

Proceedings of the International Conference on Innovation - driven Knowledge Economies and Transformation in the Global South

Proceedings of the International Conference on Innovation - driven Knowledge Economies and Transformation in the Global South

Volume X

Industry 4.0 and development : Digitalization,
automatization, productivity and employment
implications, gig economy and quality of
employment.

11-14 OCTOBER 2023

Organized by

GLOBELICS | Gulati Institute of Finance and Taxation | Research and
Information System for Developing | Indian Institute of Management
Bangalore | Kerala Development and Innovation Strategic Council



Preface / A word of welcome and thanks

It is a matter of great pride that the 20th-anniversary Globelics International Conference is being held in Thiruvananthapuram. This conference deliberates on an issue of much contemporary relevance – Innovation-driven Knowledge Economies and Transformation in the Global South, which is at the centre stage of development strategy both at the national and that of the government of Kerala. In 2006, when inclusive development was a major issue of global concern, Globelics conference was held in Trivandrum and deliberated on Innovation Systems for International Competitiveness and Shared Prosperity.

This Conference intends to bring together scholars across disciplines from over 50 countries to pool together experiences and expertise to reflect on varied issues involved in the process of transforming developing economies into sustainable, resilient, and inclusive innovation-driven knowledge economies.

In the conference, to be inaugurated by Shri Pinaryi Vijayan, the Hon'ble Chief Minister of Kerala, scholars of eminence will join as keynote speakers and panelists in plenary sessions and special sessions. The conference, organized in nine parallel tracks, will also have book presentations, excursions, and cultural events. As part of the conference, about 20 leading scholars from Globelics will visit different universities and colleges in Kerala to deliver lectures and interact with faculty and students, which is facilitated by Kerala Higher Education Council and Kerala Economic Association.

We welcome with great appreciation the Honourable Chief Minister Shri Pinarayi Vijayan and all the dignitaries who joined us for the inaugural address and eminent scholars delivering keynotes and addressing in different plenary sessions, special sessions, and parallel sessions. While the invisible hand of our honourable Chief Minister is there in all the important initiatives of GIFT, this program has been ably guided and steered by our Chairperson Shri K N Balagopal, Hon'ble Finance Minister of Kerala.

This important event is the outcome of the collective effort of all those involved, especially the Co-organising institutions: RIS, New Delhi, who inter alia managed the political clearance from the Ministry of External Affairs, IIM Bangalore, K DISC and the Digital University of Kerala. Our partnering institutions, ICSSR, Kerala State Higher Education Council, Centre for

Development Studies, Kerala Economic Association, and the Centre for Latin American Studies also have been helpful in organising this conference.

With much appreciation for all of you for being with us, we welcome you

K J Joseph
For the local organising Committee

About the 20th Anniversary Conference

The central theme of the 20th Anniversary Globelics International conference is Innovation-driven Knowledge Economies and Transformation in the Global South.

This conference intends to bring together scholars across disciplines from over 50 countries to pool together experiences and expertise to reflect on varied issues involved in the process of transforming developing economies into sustainable, resilient, and inclusive innovation-driven knowledge economies. Episodes of development in economic history highlight the successful experience of countries that have managed to harness Learning, Innovation and Competence-building Systems at the global, national, subnational and sectoral levels in order to transform their social and economic structures for achieving prosperity. In the current times, the constellation of innovations driven by rapid advances in digital technologies has given rise to a new technological revolution, commonly referred to as the fourth industrial revolution, or Industry 4.0, with profound impacts on all sectors of the economies and sections of societies. New and persisting developmental challenges have driven a renewed interest in different types of innovations that highlight their impact on society and the natural environment, such as inclusive innovation, responsible innovation, transformative innovation, grassroots innovation and Jugaad. At the core of all these approaches is how knowledge, both science-based and experience-based, is harnessed for addressing development challenges, the basic premise being “the knowledge divide is at the root of the development divide”

At the current juncture, wherein globalization is at the crossroads, the economies across the world, ravaged by the once-in-a-century pandemic and external shocks, consider strengthening the innovation system as a strategic approach towards building resilience. Evidently, there is an across-the-board drive among developing countries to transform their economies into innovation-driven knowledge economies. Such a transformation is often viewed as capable of building resilience by addressing many of the downsides of globalization-induced high growth, like excessive exploitation of exhaustible resources and accentuated divides within and between countries. Such

initiatives in India are manifested, among others, in Atal Innovation Mission (AIM), Digital India, Skill India, and Make in India and other concerted efforts to transform India into an innovation-driven knowledge economy. It is also in sync with the national agenda of harnessing Science, Technology, and Innovation (STI) to achieve sustainable development. The choice of the location of the conference is inspired by Kerala's strategy towards transforming the state into a knowledge economy, as explicitly stated in the budgets starting with 2020-2021 and the subsequent concerted efforts that followed. It is in this background that Globelics revisits India for the 20th Globelics International conference with the chosen central theme: 'Innovation-Driven Knowledge-Economies and Transformation in the Global South'.

The 20th Anniversary Globelics is also an occasion to celebrate the journey of Globelics from Brazil 2003 to India 2023 and its varied achievements. At the same time, it is also a forum for critically reflecting on our perspective and the approaches *inter alia* in terms of its achievements and limits. This is especially important at the current juncture wherein there is growing disenchantment with the development experience under globalisation on the one hand and the need on the other hand, for reimagining innovation systems in the COVID and Post COVID world towards building the much needed resilience in the South.

Globelics-India 2023 coincides with India's Presidency of G-20 that *inter alia* aims at influencing the development dynamics of Global South and it is an appropriate time for translating the insights of the conference with respect to strengthening Learning, Innovation and Competence-building Systems (LICS) into policy action. For the past two decades, we have observed an increasing influence of Global South in shaping the paradigms and trajectories of the global development agenda—sustainability, climate change, pandemic, and energy security issues—with India playing a key role; all being closely linked to the research domains of the Globelics.

World-leading scholars are set to give this year's keynote address on innovation and development. It is planned that the conference will be inaugurated by the Chief Minister of Kerala and there will be a special keynote on the conference theme along with a Globelics Lecture and Freeman Lecture. The conference will combine plenary sessions, presentations of research papers in parallel tracks, thematic panel sessions or special sessions, poster presentations, book presentation sessions, excursions, and cultural events.

Conference Tracks

While submissions on the central theme of the conference have been especially encouraged, submissions on all the issues conventionally deliberated in the Globelics international conferences are also included. Submissions have been organized around the following themes:

1. Science, technology, innovation: Theory and policy for a knowledge economy
2. Transformative innovation, responsible innovation and mission-oriented innovation
3. Knowledge-driven development of national, regional, local and sectoral innovation systems, including Agricultural innovation systems and rural development
4. Measurement of the knowledge economy: Indicators, data requirements, different approaches and methodologies
5. Skilling for the knowledge economy: Industry-academy interaction, Intellectual property rights, open innovation and development
6. Industry 4.0 and development: Digitalization and automatization, productivity and employment implications, gig economy and quality of employment
7. Economic and social upgrading for sustainable Catch-up: trade policies, FDI, value chains and innovation networks in a knowledge-driven economy
8. Innovation for inclusive development: Indigenous knowledge, Grassroot innovations, Jugaad, informal economy, micro and small enterprises
9. Entrepreneurship, employability and gender dimension in innovation and development
10. The green economy and financing innovation: Low carbon innovations, environmental technologies and renewable energy

Volume X

Industry 4.0 and development : Digitalization, automatization, productivity and employment implications, gig economy and quality of employment.

Contents

	Page No
Preface	v
About the 20th Anniversary Conference	vii
Conference Tracks	ix
1 Digital infrastructure and innovation in Africa: Does human capital mediate the effect? Dennis Boahene Osei	1
2 Role of telemedicine in healthcare delivery of India: A proposal Anitha V	25
3 Does adoption of digital marketing platforms impact household welfare? Evidence from gender disaggregated household data in Nigeria Abdulrazaq Kamal Daudu, Oyedola Waheed Kareem, Latifat Kehind Olatinwo	74
4 Developing entrepreneurial universities under the evolution of innovation system and strategy of innovation-driven development, the case of Shenzhen, China Weijia Ding	114
5 4IR technology adoption in the South African airline industry: Drivers and labour market effects Alexis Habiyaemye, Lorenza Monaco	148
6 Do digital payments enhance tax revenue? Evidence from India Jerome Joseph, N Ramalingam	163
7 Divergence and convergence in technological and industrial revolutions Mark Stephen Knell	183
8 Industry 4.0 base and front-end manufacturing technology adoption: An empirical study of Australian manufacturing firms Dilupa Nakandala	206

9	Cybercrimes, internet penetration and banking sector performance: Empirical evidence from Nigeria Faruq Umar Quadri, John O Adeoti	216
10	The financing dynamics of niche innovation in energy transitions: A South African perspective Amanda-Leigh O'Connell	239
11	Achieving sustainable industrial path development through territorial valuation dynamics – An analytical framework and case study of urban water management in Bengaluru Johan Miörner, Christian Binz, Huiwen Gong	278
12	Trade liberalisation and productivity growth: role of resource allocation and catching up - Evidence from Indian manufacturing industry Parameswaran Mavannoor	306
13	Digital trade and wealth inequality: Evidence from Asia Pacific region Sanjeev Vasudevan, Mini Thomas	356
14	On the resiliency of incubator supported start-ups: Case studies from a developing country context Phumuza Langa , Joe Amadi-Echendu	379
15	Transcending the formal- informal divide in gig work on digital platforms: The evidence from India Balaji Parthasarathy, Janaki Srinivasan, Mounika Neerukonda, Bilahari M, Pradyumna Taduri, Amruta Mahuli, Funda Ustek-Spilda, Kavita Dattani, Mark Graham	395
16	Engaging with the fourth industrial revolution: Implications for south economies Sujit Bhattacharya	416

Digital infrastructure and innovation in Africa: Does human capital mediate the effect?

Dennis Boahene Osei

DSI/NRF/Newton Fund Trilateral Chair in Transformative Innovation,
the Fourth Industrial Revolution and Sustainable Development,
College of Business and Economics, University of Johannesburg, South Africa

Abstract

While both theory and empirical evidence have established the critical role of digital infrastructure in promoting innovation, there needs to be more understanding of factors mediating this nexus, especially from the African context, which is understudied. This paper argues that digital infrastructure spurs innovation, but the overall effect depends on countries' level of human capital. Relying on data from 28 African countries spanning 2011-2019 and the system Generalized Method of Moment (GMM) estimator, this study departs from the extant literature by examining the synergistic effect of digital infrastructure and human capital on innovation while placing much emphasis on the multidimensional measurement approach of digital infrastructure. The results show a positive relationship between digital infrastructure and innovation in Africa. Furthermore, the interactive effect of digital infrastructure and human capital are significantly related to innovation, implying that digital infrastructure can indirectly enhance innovation through human capital accumulation. To promote innovation, the study recommends that policies which improve digital infrastructure should be persuaded simultaneously with policies that enhance human capital.

Keywords: Digital infrastructure, Human capital, Innovation, Africa

1. Introduction

African countries must catch up in digital revolutions. In catching up with the rest of the world, the continent must act swiftly to address the problems in the digital economy (Adam, 2019; World Economic Forum, 2022). An essential part of the digital economy that requires critical attention is the gap in digital infrastructure.¹ Closing this gap in Africa is imperative to reap the potential benefits of the Fourth Industrial Revolution and achieve the Sustainable Development Goal (SDG) 9 of “building resilient infrastructures to foster innovation within an inclusive and sustainable industry.”

Digital Council (2021) emphasised the historical deficit of digital infrastructure in Africa and the need to close the widening gap in several indicators with developed countries. To put this infrastructure gap in perspective, data from the International Telecommunication Union in 2022 suggests that the number of internet connections in Africa is the lowest globally. Thus, only 39.7% of individuals per 100 inhabitants were using the internet. This is less compared to other Telecommunication Development Bureau (BDT) regions such as the Americas (83.2%), Europe (89.5%), Arab States (70.3%), Asia-Pacific (64.3%) and Commonwealth of Independent States (83.7%). In addition, Africa is still behind other BDT regions in terms of fixed broadband (0.7%), fixed telephone (0.5%), active mobile broadband (42%), and mobile cellular telephone (86.3%). While the population covered by at least a 3G mobile network was 82.4%, only 49.6% was covered by at least a 4G LTE mobile network in 2022. Global System for Mobile Communications Association [GSMA] (2022) asserts the urgency of migrating the remaining 2G and 3G customers to 4G networks because the adoption of 4G is still less than a fifth of total connections. Focusing on new technologies, including 5G, has been problematic due to service reach and device affordability challenges. In 2020, the mobile phone market share (year-on-year) plummeted by 10%, with smartphone and feature phone shipments reduced by 5.5% and 8.0%, respectively. There is also evidence of much disparity in Africa’s data centre infrastructure. The capacity of South Africa is more significant than two-thirds of the total capacity in Africa, valued at 140,000 square meters (200 MW) (Digital Council, 2021).

However, there are some encouraging signals in the digital infrastructure space in Africa. For instance, GSMA (2022) projects global mobile subscribers to increase by almost 400 million in 2025, with sub-Saharan Africa (SSA) among the region with more new subscribers (98 million) after the Asia Pacific region (198 million). It is also estimated that SSA will reach 38 million 5G connections by 2025. As highlighted in the African Union’s Digital Transformation Strategy report (2020-2030), an extra 300 million people would be expected to access the internet by 2025.² These projections, when realised, present opportunities to bolster digital infrastructure and further improve innovation which has

historically been low in Africa. Evidence from Global Innovation Index (2011-2022) suggests that no African country has appeared in the top 40 innovative countries ranked globally. In 2022, the highest-ranked innovative country on the continent, Mauritius, was in the 45th position out of 132 countries.³ In fact, out of the bottom ten countries in the GII ranking, 8 are African countries, including Mozambique, Benin, Niger, Mali, Angola, Mauritania, Burundi, and Guinea. Therefore, policymakers must use strategies to enhance innovation and the development of digital infrastructure.

The level of African innovation can be improved with digital infrastructure leading the way. Many studies in the literature have confirmed technology infrastructure as the engine driving innovation (see Jabbouri et al., 2016; Verspagen, 1991; Wang et al., 2022). Digital infrastructure using broadband service (fixed and mobile), especially mobile, enables households and businesses to access information and communicate, potentially spurring innovation. Wang et al. (2021) assert that internet development can promote interpersonal communication and innovation by reducing information asymmetry. It is also emphasised by Zaballos et al. (2019) that digital technology investment can stimulate domestic technology, research, and innovation in developing economies.

Digital infrastructure cannot promote innovation in isolation. It needs to be supported by sound development conditions. It is widely acknowledged that high levels of human capital can complement digital infrastructure to promote innovation (Cornell University/INSEAD/WIPO, 2014). What is important to note is that digital infrastructure's potential to influence innovation would remain untapped when education and skill training is low (Shenglin et al. 2017). In recent studies, there is a great emphasis on the role of human capital in enhancing innovation (see Dakhli & De Clercq, 2004; Xu et al., 2023). However, these studies have not considered the role of human capital in serving as a conduit through which digital infrastructure promotes innovation. Against this background, this study answers the following salient questions: (i) what is the direct/unconditional effect of digital infrastructure on innovation in Africa? (ii) Does human capital have a mediating effect on the digital infrastructure-innovation nexus in Africa? An empirical understanding of these issues would help provide the basis for effective policies for digital infrastructure and innovation growth.

In responding to the above questions, this study improves literature and contributes to policy discussion in three main ways. First, this study accounts for the multidimensional context of digital infrastructure in relation to innovation. Generally, extant studies have focused on some aspects of digital infrastructure (e.g., Broadband infrastructure, internet development), ignoring the multidimensional nature of digital infrastructure, opined by Shenglin et al.

(2017). For instance, studies by Chen and Wang (2023) and Lee et al. (2016) focused on analysing the impact of broadband infrastructure on innovation, while that of Xu et al. (2023) is limited to the relation between internet infrastructure and innovation. This study departs from existing studies by capturing the multidimensional facet of digital infrastructure. This is achieved by constructing a composite index for digital infrastructure using the principal component analysis (PCA) involving indicators such as fixed telephone subscription, fixed broadband subscription, mobile cellular subscription, and internet usage. This measure has previously been used by Ndubuisi et al. (2021) to investigate the effect of digital infrastructure on employment in the service sector in sub-Saharan Africa.

Second, the study contributes to the literature by assessing how human capital complements digital infrastructure to improve innovation. Human capital developed through education and formal training raises innovation, but its importance in the digital infrastructure–innovation nexus is yet to be captured empirically in the literature. The paper focuses on human capital because it has been recognised as a panacea for achieving the SDGs (Babasanya et al., 2018; Shen et al., 2021). The 2014 Global Innovation Index report identified the human factor as a fundamental driver of innovation and, hence, emphasised the need to understand better the role of the human factor together with capital and technology in stimulating innovation (Cornell University/INSEAD/WIPO, 2014).

Third, this paper is the initial attempt to examine the relationship between digital infrastructure and innovation in Africa at the cross-country level. The closest study examined the nexus at the firm level (see Gaglio et al., 2022). Focusing on Africa is relevant in addressing the concerns of raising infrastructure and innovations to achieve SDG 9, as highlighted above. It is also important to note that Africa still has low levels of digital infrastructure and innovation relative to developed countries and, hence, needs catch-up strategies.

Following this section, the paper is structured as follows. Section 2 presents a literature review, while Section 3 discusses data and methodology. The results and their discussion are presented in Section 4. Section 5 concludes the study and highlights implications for policy and future research.

2. Literature review

2.1 Theoretical link between digital infrastructure, human capital, and innovation

Infrastructure development is an excellent way of fostering economic progress (Aschauer, 1989; Easterly and Rebelo, 1993; Calderon & Servén, 2010), and its role in key sectors of the economy cannot be overemphasized. In theory,

scholars have argued that infrastructure can play direct and indirect roles in stimulating economic growth.⁴ Of course, there are arguments that infrastructure can influence economic growth through innovation. Pan et al. (2021) stressed that technology infrastructure indirectly affects economic growth by improving innovation capabilities. This implies that innovation is central to economic growth, as espoused by Schumpeter 1934; Solow, 1956; Kraemer-Mbula and Wamae (2010). Romer (1990), relying on endogenous growth theory, emphasised the outstanding role of innovation in economic growth. This relation is conditioned on the technology infrastructure's information production and sharing role (Dretske, 1981). As a result, evolving theoretical arguments have expounded on the critical link between technology infrastructure and innovation. Lee et al. (2016) contend that technology infrastructure is necessary for innovation. Technology infrastructure in the form of the internet is critical in accumulating and agglomerating innovation elements to drive innovation by serving as a carrier of information and knowledge capital (Wang et al., 2022). Tang et al. (2021) also argue that broadband infrastructure is linked to economic activities to overcome limitations (spatial and temporal) on transmitting information, leading to knowledge accumulation, and promoting innovation.

Some scholars have also recognised the role of education in the link between technology and innovation. Nelson and Phelps (1966) indicate that “educated people make good innovators so that education speeds the process of technological diffusion”. This theoretical argument suggests that countries with higher levels of human capital are more likely to have large numbers of people adopting and using technology to access information, gain knowledge, and improve innovation.

2.2 Empirical literature review

A considerable number of studies have empirically analysed the effect of technology infrastructure on innovation. More of these studies have examined the impact of broadband infrastructure on innovation. For instance, Lee et al. (2016) studied the drivers of technological innovation in the ICT industry using data from 40 countries between 1999 and 2013. The findings based on the fixed effect model indicate broadband network infrastructure as one of the determinants. Its positive impact on innovation implies that higher levels of broadband infrastructure are necessary for ICT innovation. Zaballos et al. (2019) analysed the effect of digital infrastructure on innovation in Latin America and Caribbean countries. Evidence suggests a significant positive relationship between digital infrastructure (as measured by telecom sector investment) and innovation (as proxied by trademark applications). Nonetheless, no significant relationship between digital infrastructure and innovation was observed when the trademark applications indicator was employed. Using city-level data from

2008 to 2018, Tang et al. (2021) realised a positive relationship between telecommunication infrastructure and green technology innovation in China. The authors further observed a greater positive impact of telecommunication infrastructure on low-quality green technology innovation than on high-quality green technology innovation. Similarly, Wang et al. (2022) found a positive relationship between internet development and green innovation efficiency, relying on data from 285 prefecture-level cities in China from 2003 to 2018. Their study also established that internet development indirectly affects green innovation efficiency by enhancing financial development and product services agglomeration and reducing resource dependence. Pan et al. (2021) studied the impact of technology infrastructure on technological innovation capability using data from 30 Chinese provinces spanning 2009 to 2015. The results point to an inverted U-shaped relationship between technology infrastructure and technological innovation, implying that the technology infrastructure impact of technological innovation could be altered from promotion to inhibition.

Regarding the nexus between human capital and innovation, Furman et al. (2002), in a seminal study, examined the drivers of national innovative capacity and found human capital to be significantly related to the level of patents. Dakhli and Clercq (2004) found a positive relationship between human capital and innovation by utilising a global dataset of 59 countries. Using the GMM estimator and data for the 2000-2010 period, Nordin and Nordin (2016) point out human capital as one of the determinants of innovations in developing countries, including trade openness and regulations. Recently, Diebolt and Hippe (2019) found evidence suggesting a positive long-term relationship between human capital and innovation in European regions. Focusing on Asian countries, Malik (2022) observed the relationship between human capital and innovation as significant and positive.

Prior studies have examined the impact of digital infrastructure variables and human capital on innovation separately without considering their simultaneous effect. In addition, the multidimensional context of digital infrastructure needs to be addressed to understand the relationship between the underlying variables. There is also no existing study on the digital infrastructure-innovation nexus at the cross-country in Africa. Few existing studies have examined this relationship at the firm level. For example, Gaglio et al. (2022) found that using digital technologies (social media and internet surfing) promotes innovation among micro and small manufacturing firms in South Africa. Limited evidence exists at both firm-and-country levels in Africa on the nexus between human capital and innovation. For instance, Van Uden et al. (2017) studied the effect of various human capital elements (level of schooling, employee slack time, and training) on innovation at the firm level in sub-Saharan Africa. Their results suggest that giving employees formal training and slack time will likely improve firm innovation. Conversely, employee

schooling was found to decrease the likelihood of a firm being innovative, and hence, employee schooling serves as a minor determinant of innovation. At the cross-country level, Oluwatobi et al. (2016) demonstrate that human capital and institutions positively impact innovation in sub-Saharan Africa. A similar result was obtained in a study by Bekana (2019), who found that human capital and political institutions are significant drivers of innovation in 35 sub-Saharan African countries. Based on the above-mentioned gaps, this study provides additional insights by analysing the following hypotheses.

Hypothesis 1(H1): Digital infrastructure positively influences innovation in Africa.

Hypothesis 2 (H2): The effect of digital infrastructure on innovation is conditional on human capital development.

3. Methodology

3.1 Data

This study relied on five data sources to construct a panel of 28 African countries⁵ spanning 2011-2019 to examine the impact of digital infrastructure and human capital on innovation (see Table 1 for the list of variables, its definition, and sources). These variables were chosen exclusively based on the literature and data availability over the sample period.

Table 1: Data description and source

Variable	Measurement	Source
Dependent variable		
Innovation	Innovation output sub-index (score 0-100)	World Intellectual Property Organization
Independent variables		
Digital infrastructure index	Mobile cellular subscriptions (per 100 people)	World DevelopmentIndicators
	Individuals using the Internet (% of the population)	World DevelopmentIndicators
	Fixed telephone subscriptions (per 100 people)	World DevelopmentIndicators
	Fixed broadband subscriptions (per 100 people)	World Development Indicators
Human capital	School enrollment, tertiary (% gross)	World DevelopmentIndicators
Control variables		
Trade openness	Sum of imports and exports (% of GDP)	World DevelopmentIndicators
Financial development	Financial development index	International Monetary Fund
Economic growth	GDP per capita (constant 2015US\$)	World DevelopmentIndicators
Foreign direct investment	Foreign direct investment, net inflows (% of GDP)	World Development Indicators

Institutional quality index	Regulatory quality Government effectiveness Rule of Law Control of corruption Voice and accountability Political stability and lack of violence	World Governance Indicators
Inflation	Consumer price index (2010=100)	World Development Indicators

3.2 Justification of variables

Dependent variable: The study employed the innovation output sub-index of the Global Innovation Index (GII) as the dependent variable. This measure is appropriate because the study is interested in understanding the drivers of innovation outcomes, specifically focusing on digital infrastructure and human capital. The innovation output index captures the results of innovation activities within an economy, and it is computed based on the weighted average of indicators such as knowledge and technology outputs (knowledge creation, knowledge impact and knowledge diffusion) as well as creative outputs (intangible assets, creative goods and services, and online creativity).

⁶ This variable is normalised to range from 0 to 100. The use of the overall GI index as a proxy for innovation outcome is unsuitable because of its composition of average indicators of both innovation input and output sub-indices. The former accounts for factors that enable innovative activities such as institutions, human capital and research, infrastructure, market sophistication, and business sophistication. This paper selects the innovation output index as the measure of innovation following the study of Omid et al. (2020). The authors emphasised that the innovation output measure aligns with the OECD/Eurostat (2005) definition of product and process innovation.

Independent variables: The core independent variables employed in this study include digital infrastructure and human capital. Digital infrastructure is a multidimensional variable measured using indicators, such as mobile cellular subscriptions per 100 people, individuals using the internet as a percentage of the population, fixed telephone subscriptions per 100 people, and fixed broadband subscriptions per 100 people. These indicators were transformed using the principal component analysis to generate a digital infrastructure composite index. The higher the index score, the more improved the digital infrastructure of the economies under study. This index is richer and more comprehensive than the single variable measure used in studies examining the relationship between digital infrastructure and innovation (see Chen & Wang, 2023; Lee et al., 2016). The measure of digital infrastructure used in this study has previously been employed by Ndubuisi et al. (2021), and it is expected to correlate with innovation positively.

Human capital is used as a mediating variable and is proxied by gross tertiary education enrolment. This variable has been widely used as a determinant of innovation (see Furman et al., 2002; Omid et al., 2020). As human capital improves, innovation increases through productivity and economic growth (Cannon, 2000). Higher levels of human capital also mean more individuals can use digital infrastructure to improve innovation within an economy.

Control variables: Informed by the innovation literature, the study controlled for trade openness, financial development, economic growth, foreign direct investment, inflation, and institutional quality in the regression models. The use of these variables is in line with studies on the determinants of innovation, such as Lee et al. (2016), Omid et al. (2020), Pan et al. (2021), Saldanha et al. (2021) and Qureshi et al. (2021). Trade openness is expressed as the ratio of the summation of exports and imports to GDP. This variable is used to measure the level of openness of an economy to international trade, which is critical to promote innovation through knowledge absorption and diffusion. A priori, trade openness is supposed to impact innovation positively. The financial development index is used to proxy financial sector development. This index, which is normalised between 0 (low) and 1 (high), is composed of access, efficiency, and depth indicators of financial markets (stocks and bonds) and financial institutions (banks and non-banks) (Sviridzenka, 2016). This variable is expected to impact innovation positively. The study also controlled economic growth as an indicator of economy size measured by real GDP per capita. It is expected to have a positive relationship with innovation.

To capture the effect of macroeconomic (in) stability on innovation, this study included inflation in the models and gauged it as a consumer price index (2010=100). Higher inflation is expected to hinder innovation. Foreign direct investment is measured as net inflows of foreign direct investment as a percentage of GDP. It is expected to have a positive nexus with innovation. Institutional quality is measured as a composite indicator, which consists of variables such as regulatory quality, government effectiveness, the rule of law, control of corruption, voice and accountability, and political stability and lack of violence, computed using the principal component analysis. Institutions create the necessary environment for innovation to thrive; hence, it is expected to have a positive relationship with innovation.

3.3 Empirical strategy

In examining the unconditional effect of digital infrastructure on innovation, this study relies on the econometric model specified as follows:

$$INNOV_{it} = \gamma_0 + \gamma_1 INNOV_{it-1} + \gamma_2 DI_{it} + \gamma_3 HC_{it} + \gamma_5 X_{it} + \varepsilon_{it} \quad (1)$$

$$\varepsilon_{it} = \varphi_i + \rho_t + \mu_{it}$$

where $i = 1, 2, 3, \dots, N$ is the cross-sectional dimension of countries; $t = 1, 2, 3, \dots, T$ represents time; $INNOV_{it}$ denotes innovation; $INNOV_{it-1}$ denotes lagged of innovation; DI_{it} represents digital infrastructure and X_{it} denotes control variables such as trade openness, economic growth, foreign direct investment, financial development, inflation, and institutional quality. ε_{it} stands for error term, while μ_{it} denotes idiosyncratic error term. φ_i and ρ_t represent unobserved country-specific and time-specific effects, respectively.

The empirical analysis of equation (1) relies on the system Generalized Method of Moments (GMM) estimator inspired by Arellano and Bover (1995) and Blundell and Bond (1998). This estimation technique is appropriate to address the potential endogeneity problems prevalent in the model. It is very likely that endogeneity will occur due to the correlation between the lagged dependent variable and the error term. Arellano and Bond (1991) suggested using lagged levels of the regressors as instruments in dealing with endogeneity. However, Blundell and Bond (1998) contended that the lagged levels of the explanatory variables could be better instruments for the first difference when the series over time is persistent and the time dimension is short. Arellano and Bover (1995) proposed using differenced and first-level equations. These equations increase efficiency by reducing biases and impressions associated with the difference GMM. Additionally, system GMM is employed because of its capability to provide robust results when there is autocorrelation and heteroscedasticity. It is also the best technique when the sampled countries (N) are more significant than the number of periods (T). In this study, the panel data structure is such that number of sampled countries (28) exceeds the number of periods (9).

To assess whether human capital mediates the relationship between digital infrastructure and innovation, the study extends equation (1) by including an interaction term of digital infrastructure and human capital. This is shown in equation (2), as specified below:

$$INNOV_{it} = \omega_0 + \omega_1 INNOV_{it-1} + \omega_2 DI_{it} + \omega_3 HC_{it} + \omega_4 (DI_{it} \times HC_{it}) + \omega_5 X_{it} + \varepsilon_{it} \quad (2)$$

where $DI_{it} \times HC_{it}$ is the interaction term of digital infrastructure and human capital. All other variables denote their usual meaning.

Having estimated the interactive effects, the study further computes the marginal effect of digital infrastructure and human capital for stimulating technological innovation using a partial differentiation approach. This is specified in equation 3 as follows:

$$\frac{\partial INNOV_{it}}{\partial DI_{it}} = \omega_2 + \omega_4 HC_{it} \quad (3)$$

Two formal tests were applied to check the reliability of the results. First, the Hansen test was used to assess the appropriateness (exogeneity) of the instruments. The null hypothesis for this test states that over-identifying restrictions are valid. Second, a serial correlation test with the null hypothesis that there is no serial correlation was also performed.

4 Results and discussion

4.1 Descriptive analysis

This section presents the summary statistics for all variables employed in this study from 2011 –2019.

Table 2: Summary statistics

Variable	Mean	Std. Dev.	Min	Max
Innovation	20.223	5.675	7	35
Digital infrastructure index	0	1	-0.609	5.101
Fixed Broadband subscription	1.694	3.417	0.006	23.699
Fixed telephone subscription	3.682	6.133	0	35.386
Internet usage	21.677	18.003	0.9	84.12
Mobile cellular subscription	86.227	33.837	15.386	166.943
Human capital	15.441	11.941	0.819	52.620
Foreign direct investment	3.765	5.16	-2.545	39.456
Economic growth	2198.051	2206.728	357.826	10959.3
Trade openness	64.414	23.821	20.723	127.204
Financial development	0.189	0.136	0.062	0.643
Inflation	141.83	49.894	100.907	418.344
Institutional quality index	0	2.145	-3.347	5.947

Evidence from Table 2 suggests that innovation in Africa is very low, with an average score of 20.22 over the sample period. The minimum and maximum values indicate a wide disparity in innovation among countries on the continent. This is confirmed by the high standard deviation of 5.68. The lower values of most digital infrastructure indicators confirm why African countries must catch up in digital infrastructure. While fixed broadband subscriptions per 100 people averaged 1.69, the average value of fixed telephone subscriptions per 100 people also stood at 3.68. The average number of internet users (21.67%) is less than half depicting low internet usage among sampled countries, although the maximum percentage of individuals using the internet stood at 84.12%. This was observed in Morocco in 2019. Relatively, mobile cellular subscriptions per 100 people are quite high, averaging 86.23, with a standard

4.2 Empirical results and discussion

This section analyses the empirical results on the relationship between digital infrastructure, human capital, and innovation. Table 3 shows the results with ten models. The first two models include the main variables of interest (digital infrastructure and human capital). Models 3-9 have all control variables included sequentially. The interaction term of digital infrastructure and human capital is added to Model 10.

From Table 3, the coefficient of the lagged innovation is positive and significant irrespective of the model specification. This suggests that a country's level of innovation is positively affected by its one-period lag. In most of the models, it is realised that digital infrastructure is mainly positive and significant. This variable remains highly significant in model 10, which happens to be the main model of interest. Specifically, digital infrastructure has a positive relationship with innovation in the sampled countries, and it is statistically significant at 1% (see Models 7-10). This suggests that developing digital infrastructure could be a viable strategy to improve African innovations. The reason is that it has the potency to increase information accessibility leading to healthy research conditions to enhance innovations. It is also reasonable to expect collaborative innovation once the digital infrastructure is developed and there are no limitations in research collaborations and sharing of research findings (Tang et al., 2021; Wang & Cen, 2022). The result aligns with studies such as Lee et al. (2016) and Zaballos et al. (2019).

With regard to human capital, it is found to have a direct and positive effect on innovation. The positive coefficient, which is statistically significant at 1%, indicates that human capital accumulation spurs innovation. This corresponds to theory as one expects highly educated individuals to manifest knowledge and the capacity to contribute to innovation (Lenihan et al., 2019; Mariz-Perez et al., 2012). It is widely argued that human capital in the form of tertiary education plays an indispensable role in generating knowledge and exchanging ideas and skills necessary for innovation (Cornell University/INSEAD/WIPO, 2020). The finding of a positive relationship between human capital and innovation is consistent with Dakhli and Clercq (2004), and Oluwatobi et al. (2016).

Table 3: Digital infrastructure, human capital, and innovation in Africa (System GMM results)

Dependent variable: Innovation										
	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10
Lagged innovation	0.931***	0.814***	0.778***	0.771***	0.767***	0.709***	0.787***	0.493***	0.543***	0.392***
	-0.083	-0.042	-0.053	-0.051	-0.044	-0.086	-0.081	-0.109	-0.087	-0.083
Digital infrastructure	0.314**	0.012	-0.062	-0.067	-0.03	0.066	0.505***	0.411***	0.285***	1.339***
	-0.156	-0.053	-0.063	-0.06	-0.155	-0.171	-0.178	-0.159	-0.1	-0.438
Human capital		0.053	0.36**	0.247	0.24	1.018***	1.128***	0.930***	0.688***	0.745***
		-0.101	-0.166	-0.237	-0.168	-0.197	-0.216	-0.234	-0.199	-0.172
Foreign direct investment			0.152***	0.146***	0.138***	0.077**	0.051	0.034	0.057**	0.066***
			-0.032	-0.035	-0.026	-0.033	-0.034	-0.024	-0.022	-0.021
Economic growth				0.129	0.683	1.869	9.486***	11.947***	8.397***	6.217*
				-0.334	-2.795	-1.871	-3.463	-3.283	-3.234	-3.194
Economic growth squared					-0.038	-0.23*	-0.72***	-0.900***	-0.610***	-0.471**
					-0.191	-0.138	-0.243	-0.228	-0.22	-0.222
Trade openness						0.735***	0.476*	0.592***	0.519**	0.202
						-0.214	-0.252	-0.205	-0.228	-0.191
Financial development							-3.126***	-1.334	-1.593*	-1.288**
							-0.979	-1.138	-0.891	-0.561
Inflation								-0.519**	-0.496***	-0.397***
								-0.202	-0.156	-0.098
Institutional quality									0.063	0.086
									-0.069	-0.071

Dependent variable: Innovation										
	-1	-2	-3	-4	-5	-6	-7	-8	-9	-10
Digital infrastructure× Human capital										-0.291**
										-0.121
Constant	0.18	0.413	-0.364	-1.012	-2.976	-5.811	-33.769***	-38.985***	-28.083**	-18.518
	-0.244	-0.314	-0.532	-2.043	-10.086	-6.398	-12.332	-11.631	-11.736	-11.634
Net /marginal effect										0.605***
										-0.159
Country and time effects	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Observations	206	175	169	169	169	166	166	166	166	166
Number of countries	28	28	28	28	28	28	28	28	28	28
Number of instruments	8	16	17	18	19	20	21	22	23	23
Wald	80084.1	130074.3	52496	54862.3	39666.7	7533.48	5567.51	10875.03	21448.6	10718.2
[p-value]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Hansen test	7.14	14.51	15.1	15.04	15.67	10.48	13.24	10.39	11.21	12.31
[p-value]	[0.211]	[0.269]	[0.236]	[0.239]	[0.207]	[0.574]	[0.352]	[0.582]	[0.511]	[0.341]
AR(2)	-0.21	-0.25	0.56	0.54	0.53	0.21	0.53	0.12	0.17	0.28
[p-value]	[0.832]	[0.799]	[0.577]	[0.591]	[0.598]	[0.834]	[0.595]	[0.907]	[0.862]	[0.776]

Note: Standard errors are in parentheses. All variables are in natural logs except digital infrastructure, financial development, and institutional quality indexes.

*** p<.01, ** p<.05, * p<.1

Given that human capital improves innovation, the study further examines the mediating role of human capital in the digital infrastructure-innovation nexus. The multiplicative interaction term is included in model 10 and evaluated at the mean of human capital using partial differentiation to determine the marginal effect of digital infrastructure on innovation. It is observed that the coefficient of the net/marginal effect (0.605) is positive and highly significant (at a 1% level). This means that when human capital improves, digital infrastructure substantially promotes innovation. In other words, it suggests that digital infrastructure positively impacts innovations at higher levels of human capital. By implication, a higher level of human capital is required to adopt and exploit technology for innovative purposes. As Van Uden et al. (2017) put it, "human capital is a salient factor to spur innovation". Highly educated and skilled individuals are more likely to use digital infrastructure to yield maximum innovation returns.

Regarding the control variables, the study shows that foreign direct investment is positive and significant in most of the models, except models 7 and 8, when financial development and inflation are controlled. This evidence suggests that foreign direct investment encourages innovation. A plausible explanation is that inflows from foreign direct investment potentially serve as a source of cheaper financial capital, which can be used to support innovation activities. This result coincides with Pan et al. (2021). Similarly, the results reveal that trade openness positively impacts innovation. The positive effect of trade openness on innovation is insensitive to the model specifications and largely significant except in model 10. This result suggests that when countries open their markets to international trade, there is a greater likelihood that domestic firms will be exposed to global best practices through knowledge absorption and diffusion, which is a prerequisite for innovation (Cornell University/INSEAD/WIPO, 2020). Relatedly, trade openness allows technology spillover and increases innovation activities. This result is consistent with Lee et al. (2016).

Economic growth is another significant determinant of innovation. Specifically, a non-linear relationship (inverted u-shaped) is established between economic growth and innovation. Thus, economic growth and its squared term are significantly positive and negative related to innovation. What is clear from this finding is that innovation sufficiently improves at the initial stages of economic growth; nonetheless, innovation reduces at later stages of economic growth. The result of a positive effect of economic growth on innovation is similar to Saldanha et al. (2021). The coefficient of financial development is significantly negative, indicating that financial development undermines innovation. This finding contradicts the theory (Dabla-Norris et al., 2012) and the empirical study of Kapidani and Luci (2019), who underscored the importance of the financial sector in facilitating technological investments to support innovation. Similarly, the finding does not conform with Qureshi et al. (2021), who found

financial development insignificantly related to innovation performance in Asia and the Pacific regions. Overall, the evidence of a negative effect of financial development on innovation echoes the kind of financial system expended in Africa. Most African countries have a weak financial system, making it difficult for small innovative firms to access financial capital to support their innovative ideas and outputs.

Inflation has a significant negative effect on innovation. Thus, higher inflation is not conducive to innovation activities. This result suggests that if governments and Central Banks neglect policies that curtail the high level of inflation, then rising inflation will worsen innovation. Institutional quality did not significantly impact innovation, albeit the coefficient is positive. This result emphasised the need for African countries to improve their fragile institutional environment, which is critical for innovation. According to Roger et al. (2022), acquiring a patent and trademark for innovation outputs in many sub-Saharan African countries is wearisome due to bureaucratic requirements. The authors cited Nigeria as an example where innovative ideas and outputs get stolen and outdated before registration and certification because the procedure needs to be more bureaucratic.

The results presented in this study are reliable. Table 3 shows the results of the two formal tests (Hansen and Serial correlation tests) used to assess the validity of the system GMM estimates. While the null hypothesis of the Hansen test states that instruments are over-identified, the serial correlation test analyses the null hypothesis of no serial correlation [AR (2)]. Evidence at the bottom of Table 3 indicates a rejection of the two hypotheses since the p-values associated with the tests are insignificant.

Conclusion and implications

This study has empirically examined whether digital infrastructure and human capital simultaneously influence innovation in Africa over the 2011 –2019 period. It departs from the extant literature that has focused on the direct effect of digital infrastructure on innovation, ignoring the indirect/conditional effects. In doing so, the study contributes to the theoretical literature as well by highlighting that human capital complements digital infrastructure to improve innovation. The empirical results show that digital infrastructure proxied by a composite index involving fixed telephone subscription, fixed broadband subscription, mobile cellular subscription, and internet usage positively impacts innovation. This notwithstanding, it is observed that digital infrastructure can indirectly enhance innovation through human capital accumulation. Conclusively, digital infrastructure and human capital are complementary factors in promoting African innovation. Furthermore, the study also found foreign direct investment, economic growth,

trade openness, and inflation as significant determinants of innovation in Africa. Based on these findings, the study recommends that policies which improve digital infrastructure should be persuaded simultaneously with policies that enhance human capital. Governments need to build human capital with a key focus on digital capabilities to encourage innovations in Africa. Additionally, countries in Africa need to break the jinx of over-relying on foreign funding sources to develop their digital infrastructure by building the capacities of relevant stakeholders to raise funds locally. In terms of future research, this study can be extended to examine the multivariate causality between digital infrastructure, human capital and innovation using novel econometric techniques.

Appendix A1: Pairwise correlations

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(1) Innovation	1.000								
(2) Digital infrastructure	0.286***	1.000							
(3) Human capital	0.208***	0.632***	1.000						
(4) Foreign direct investment	0.019	-0.062	-0.283***	1.000					
(5) Economic growth	0.345***	0.732***	0.841***	-0.281***	1.000				
(6) Trade openness	0.041	0.404***	0.364***	0.273***	0.275***	1.000			
(7) Financial development	0.294***	0.600***	0.585***	-0.118*	0.796***	0.310***	1.000		
(8) Inflation	-0.076	-0.165*	-0.027	-0.077	-0.097	-0.241***	-0.067	1.000	
(9) Institutional quality	0.265***	0.598***	0.389***	0.060	0.585***	0.481***	0.651***	-0.141**	1.000

References

- Adam, H. (2019). The Digital Revolution in Africa: Opportunities and Hurdles. In *Proceedings of 10th International Conference on Digital Strategies for Organizational Success*.
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277-297.
- Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68(1), 29-51.
- Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, 23(2), 177-200.
- Aschauer, D. A. (1993). Genuine economic returns to infrastructure investment. *Policy Studies Journal*, 21(2), 380-391.
- Babasanya, A. O., Oseni, I. O., & Awode, S. S. (2018). Human capital development: A catalyst for achieving SDGs in Nigeria. *Acta Universitatis Danubius. Oeconomica*, 14(4).
- Barro, R. J. (1990). Government spending in a simple model of endogenous growth. *Journal of Political Economy*, 98(5, Part 2), S103-S125.
- Bekana, D. M. (2019). Political institutions, human capital, and innovation: evidence from sub-Saharan Africa. *Democratization*, 26(4), 666-708.
- Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
- Calderón, C., & Servén, L. (2010). Infrastructure and economic development in Sub-Saharan Africa. *Journal of African Economies*, 19(suppl_1), i13-i87.
- Cannon, E. (2000). Human capital: level versus growth effects. *Oxford Economic Papers*, 52(4), 670-676.
- Chen, J., & Wang, J. (2023). The impact of broadband speed on innovation: City-level evidence from China. *Heliyon*, e12692.
- Cornell University, INSEAD, & WIPO. (2020). The Global Innovation Index 2020: Who Will Finance Innovation? Ithaca, Fontainebleau, and Geneva. https://www.wipo.int/edocs/pubdocs/en/wipo_pub_gii_2020.pdf
- Cornell University, INSEAD, & WIPO. (2014). The Global Innovation Index 2014: The human factor in innovation. Ithaca, Fontainebleau, and Geneva. https://www.wipo.int/edocs/pubdocs/en/economics/gii/gii_2014.pdf
- Dabla-Norris, E., Kersting, E. K., & Verdier, G. (2012). Firm productivity, innovation, and financial development. *Southern Economic Journal*, 79(2), 422-449.
- Dakhli, M., & De Clercq, D. (2004). Human capital, social capital, and innovation: a multi-country study. *Entrepreneurship & Regional Development*, 16(2), 107-128.
- Diebolt, C., & Hippe, R. (2019). The long-run impact of human capital on innovation and economic development in the regions of Europe. *Applied Economics*, 51(5), 542-563.

- Digital Council (2021). Africa digital infrastructure market analysis 2021report.<https://www.wcoesarpsg.org/wp-content/uploads/2021/11/juanita-clark-africas-evolving-digital-landscape.pdf>
- Dretske, F. (1981). Knowledge and the Flow of Information, MIT Press, Cambridge, Massachusetts
- Easterly, W., & Rebelo, S. (1993). Fiscal policy and economic growth. *Journal of Monetary Economics*, 32(3), 417-458.
- Furman, J. L., Porter, M. E., & Stern, S. (2002). The determinants of national innovative capacity. *Research Policy*, 31(6), 899-933.
- Gaglio, C., Kraemer-Mbula, E., & Lorenz, E. (2022). The effects of digital transformation on innovation and productivity: Firm-level evidence of South African manufacturing micro and small enterprises. *Technological Forecasting and Social Change*, 182, 121785.
- Gramlich, E. M. (1994). Infrastructure investment: A review essay. *Journal of Economic Literature*, 32(3), 1176-1196.
- Global System for Mobile Communications Association (2022). The mobile economy.<https://data.gsmainelligence.com/research/research/research-2022/the-mobile-economy-2022>
- Hoechle, D. (2007). Robust standard errors for panel regressions with cross-sectional dependence. *The Stata Journal*, 7(3), 281-312.
- Jabbouri, N. I., Siron, R., Zahari, I., & Khalid, M. (2016). Impact of information technology infrastructure on innovation performance: An empirical study on private universities in Iraq. *Procedia Economics and Finance*, 39, 861-869.
- Kapidani, M., & Luci, E. (2019). The Effects on innovation from financial sector development: evidence from developing countries. *Journal of Competitiveness*, 11(2), 84.
- Lee, S., Nam, Y., Lee, S., & Son, H. (2016). Determinants of ICT innovations: A cross-country empirical study. *Technological Forecasting and Social Change*, 110, 71-77.
- Lenihan, H., McGuirk, H., & Murphy, K. R. (2019). Driving innovation: Public policy and human capital. *Research Policy*, 48(9), 103791.
- Malik, S. (2023). Macroeconomic determinants of innovation: Evidence from Asian countries. *Global Business Review*, 24(1), 137-151.
- Mariz-Perez, R.M., Teijeiro-Alvarez, M.M., Garcia-Alvarez, T.M. (2012), The relevance of human capital as a driver of innovation. Elsevier Economy Notebooks, 35(98), 68-76.
- Ndubuisi, G., Otioma, C., & Tetteh, G. K. (2021). Digital infrastructure and employment in services: Evidence from Sub-Saharan African countries. *Telecommunications Policy*, 45(8), 102153.
- Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. *The American Economic Review*, 56(1/2), 69-75.

- Nurnaddia Nordin, and Nurhaiza Nordin, (2016). Determinants of innovation in developing countries: a panel Generalized Method of Moments analysis. *Jurnal Ekonomi Malaysia*, 50 (2). 93-105.
- OECD/Eurostat (2005). Oslo Manual: Guidelines for Collecting and Interpreting Innovation Data, 3rd Edition, The Measurement of Scientific and Technological Activities. OECD Publishing, Paris. <https://doi.org/10.1787/9789264013100-en>
- Oluwatobi, S., Ola-David, O., Olurinola, I., Alege, P., & Ogundipe, A. (2016). Human capital, institutions and innovation in Sub-Saharan Africa. *International Journal of Economics and Financial Issues*, 6(4), 1507-1514.
- Omidi, V., Shahabadi, A., & Mehregan, N. (2020). Innovation drivers in developing countries. *Journal of the Knowledge Economy*, 11, 707-720.
- Pan, X., Guo, S., Li, M., & Song, J. (2021). The effect of technology infrastructure investment on technological innovation—A study based on spatial durbin model. *Technovation*, 107, 102315.
- Qureshi, I., Park, D., Crespi, G. A., & Benavente, J. M. (2021). Trends and determinants of innovation in Asia and the Pacific vs. Latin America and the Caribbean. *Journal of Policy Modeling*, 43(6), 1287-1309.
- Roger, M., Shulin, L., & Sesay, B. (2022). ICT Development, Innovation Diffusion and Sustainable Growth in Sub-Saharan Africa. *Sage Open*, 12(4), 21582440221123894.
- Romer, P. M. (1990). Endogenous technological change. *Journal of political Economy*, 98(5, Part2), S71-S102.
- Saldanha, T. J., John-Mariadoss, B., Wu, M. X., & Mithas, S. (2021). How information and communication technology shapes the influence of culture on innovation: a country-level analysis. *Journal of Management Information Systems*, 38(1), 108-139.
- Schumpeter, J. A. (1934). The theory of economic development: An inquiry into profits, capital, credit, interest, and the business cycle.
- Shen, H., Ali, S. A., Alharthi, M., Shah, A. S., Basit Khan, A., Abbas, Q., & ur Rahman, S. (2021). Carbon-free energy and sustainable environment: The role of human capital and technological revolutions in attaining SDGs. *Sustainability*, 13(5), 2636.
- Shenglin, B., Bosc, R., Jiao, J., Li, W., Simonelli, F., & Zhang, R. (2017). Digital infrastructure: overcoming the Digital divide in China and the European Union.
- Solow, R. M. (1956). A contribution to the theory of economic growth. *The quarterly journal of economics*, 70(1), 65-94.
- Svirydzenka, K., (2016). Introducing a New Broad-Based Index of Financial Development (IMF Working Paper, WP/16/05).
- Tang, C., Xu, Y., Hao, Y., Wu, H., & Xue, Y. (2021). What is the role of telecommunications infrastructure construction in green technology innovation? A firm-level analysis for China. *Energy Economics*, 103, 105576.

- Van Uden, A., Knoben, J., & Vermeulen, P. (2017). Human capital and innovation in Sub-Saharan countries: A firm-level study. *Innovation*, 19(2), 103-124.
- Verspagen, B. (1991). A new empirical approach to catching up or falling behind. *Structural change and economic dynamics*, 2(2), 359-380.
- Kraemer-Mbula, E., & Wamae, W. (2010). Key issues for innovation and development.
- Wang, P., & Cen, C. (2022). Does digital economy development promote innovation efficiency? A spatial econometric approach for Chinese regions. *Technology Analysis & Strategic Management*, 1-15.
- Wang, Y., Li, C., Khan, M. A., Li, N., & Yuan, R. (2021). Firm information disclosure environment and R&D investment: Evidence from Internet penetration. *PLoS One*, 16(3), e0247549.
- Wang, K. L., Sun, T. T., Xu, R. Y., Miao, Z., & Cheng, Y. H. (2022). How does internet development promote urban green innovation efficiency? Evidence from China. *Technological Forecasting and Social Change*, 184, 122017.
- World Economic Forum (2022). Regional Action Group for Africa: Attracting Investment and Accelerating Fourth Industrial Revolution Approach in Africa.
https://www3.weforum.org/docs/WEF_Attracting_Investment_and_Accelerating_Fourth_Industrial_Revolution_Adoption_in_Africa_2022.pdf
- World Intellectual Property Organization (2022). Global Innovation Index 2022: What is the future of innovation-driven growth?
<https://www.wipo.int/edocs/pubdocs/en/wipo-pub-2000-2022-en-main-report-global-innovation-index-2022-15th-edition.pdf>
- Xu, K., Loh, L., Liang, L., & Mei, R. (2023). Heterogeneous effects of influencing factors on innovation performance: evidence from European Union countries. *Technology Analysis & Strategic Management*, 1-16.
- Zaballos, A. G., Rodríguez, E. I., & Adamowicz, A. (2019). *The impact of digital infrastructure on the sustainable development goals: a study for selected Latin American and Caribbean countries* (Vol. 701). Inter-American Development Bank.

Notes

- ¹ Following Shenglin et al. (2017), this study broadly defines digital infrastructure as ‘the fundamental underlying structures – material (e.g., IT equipment) or immaterial/virtual (e.g., skills, spectrum) – that connect software, content, data, and devices to allow distributed data processing, storage, transmission and sharing.’
- ² See <https://au.int/sites/default/files/documents/38507-doc-dts-english.pdf>
- ³ This was followed by South Africa (61st), Morocco (67th), Tunisia (73rd), and Botswana (86th) as the top five innovation leaders in Africa. These countries had their innovation scores (on a scale of 0-100) ranged from 23.9 to 34.4 (see World Intellectual Property Organization [WIPO], 2022, p.19).
- ⁴ The direct effect of infrastructure on economic growth stems from the fact that infrastructure is considered as part of physical stock of a country, serving as a factor of production (Aschauer, 1993; Gramlich, 1994). On the other hand, the indirect effect implies that infrastructure complements other factors. Barro (1990) explained that this can happen by improving total factor productivity through cost reduction, and investment opportunities.
- ⁵ These countries include Algeria, Benin, Botswana, Burkina Faso, Cameroon, Cote d’Ivoire, Egypt, Ethiopia, Ghana, Guinea, Kenya, Madagascar, Malawi, Mali, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, South Africa, Tanzania, Togo, Tunisia, Uganda, and Zambia.
- ⁶ For details on the indicators of the innovation output sub-index of the GII and their measurements, see Cornell University/INSEAD/ WIPO, 2020, p207-208).

Role of telemedicine in healthcare delivery of India: A proposal

Anitha V

Department of Economics, University of Kerala, India

Abstract

Telemedicine is a recently popular and innovative method in the healthcare sector, especially for treatment. It helps us to reduce the issues in healthcare delivery, by reducing the economic burden on the patient and thereby achieving Universal healthcare. Yet, its application in developing countries is relatively new and the utilization rate is also low. India is also experiencing the same situation except in the case of COVID-19 period and a few states. Studies showed that telemedicine decreases the burden of the healthcare system to a large extent. Even though the healthcare system in India is improved a lot, still the facilities at the public healthcare centres are not equitably distributed across regions and the out-of-pocket expenditure is also high. The situation is worsened when the supply side of health care facilities not much increased based on the demand for the services. Hence, the paper presents the role of telemedicine in healthcare delivery and the need for strengthening telemedicine in India. Besides, it proposes a number of questions for further research to make the system socially viable. Secondary data and empirical works are used to establish the scope of telemedicine in India. The paper found that telemedicine is an alternative treatment option for India, which suffers from high out-of-pocket expenditure, limited healthcare expenditure and high healthcare demand.

Keywords: Telemedicine, Health care delivery, e-Sanjeevani, Universal Health Care, e-Health, Equity in health

1. Introduction

The use of Information and Communication Technology (ICT) brings revolutionary change in almost all fields and its impacts on healthcare delivery have been visible recently. ICT is widely used in digital health to improve health outcomes and telemedicine is one of the items to deliver healthcare services through mobile or desktop apps. The concept of telemedicine, healing at a distance, is defined by WHO(2010) as *“the delivery of healthcare services, where distance is a critical factor, by all healthcare professionals using information and communication technologies for the exchange of valid information for diagnosis, treatment and prevention of disease and injuries, research and evaluation, and for the continuing education of healthcare providers, all in the interests of advancing the health of individuals and their communities.”* The advantage of using telemedicine is visible in different parts of the world in various categories. It is observed that e-health provides sustainable health care to older adults in terms of prevention and cure (Morato, J, et al., 2021). The computerised reminder system helps to improve the vaccination rate (Dexheimer JW et al., 2011.) and it is a viable method for reducing tobacco use among patients (Graham, A.L.; et al., 2017).

Telemedicine/e- health has gained immense popularity in healthcare delivery in recent years (may add references) and its use reaches its peak level during COVID-19, especially for chronic disease treatment (may add references). It is widely accepted due to its availability, accessibility and affordability. Patients use telemedicine for a variety of reasons, such as no travelling cost, treatment at the comfort of home, immediate doubt clearance, easy access with functional disability, no fear of spreading diseases through hospital visits and escape from consultation fees, etc. Besides, it is helpful to create scientific data regarding the episodes of illness, (which can be used for research and development by applying epidemiological analysis and can be used for the supply side of health care management including pharmaceuticals) and can deliver health care in all areas, including remote rural areas. All these will be helpful to achieve universal health care coverage in the true sense.

Despite its merits, telemedicine remains relatively new in healthcare delivery, especially in developing countries and its disposition does not utilize its potential. At the same time, the advanced version of telemedicine remains and networks are absent in many countries, which is a challenge in healthcare delivery in an easy-to-access way. Given the above-mentioned background, this paper aims to discuss the rationale for telemedicine and the need for it in Indian healthcare delivery. As a whole, this paper supports the policymakers to think appropriately in decision-making related to healthcare delivery in the public sector.

With the support of secondary data and empirical works, the study made an attempt to answer the following questions in the context of India and proposes the same question for further detailed research. In order to achieve the objective equity in health care services and Universal Health Care coverage, can we use telemedicine/eSanjeevani in India? If so what are the different possibilities to achieve? Whether the current system is sufficient to achieve universal health care coverage? Whether the people-centred telemedicine system ensures quality healthcare services? Can we use telemedicine services to reduce the out-of-pocket expenditure and thereby the economic burden of the

patient? How can we increase the satisfaction of the patients with the support of telemedicine service? These types of studies will be benefited both the government and society, that is patient centred health care system and thereby strengthen the capacity of the public sector. By suggesting measures to reduce health inequalities, optimise telemedicine in India, and ensure the health of the future generation and sustainable development. It is accepted that intersectoral and multi-disciplinary approaches are necessary for strengthening the healthcare delivery system to achieve universal health coverage (Garzón-Orjuela, N., et al., 2022). All these will be able to create a sustainable mechanism for the healthcare system. The positive externalities and equity aspects of investing in telemedicine are ample (Dash S, et al., 2021).

2. Telemedicine in India

The discussion of telemedicine started in the 1960s, which shows that telemedicine is not a new concept. Yet, the wide use of telemedicine is fairly recent, as seen during the COVID-19 pandemic. It is notable that the availability and use of the internet and mobile phone have made accessing telemedicine in recent times.

The implementation of e-health in European Region in 2016 showed that there is progress in public health and healthcare delivery (World Health Organization, 2016). In Hungary, both doctors and patients are satisfied with telemedicine services (Békási, S., Girasek, E. & Györffy, Z., 2022). In the US telemedicine in eye care treatment improved its quality as well as its inclusive nature, especially for the disadvantaged community (Owsley, C., Rhodes, L.A., McGwin, G. et al., 2015). ehealth solutions are often deployed in the context of facilitating this transition and benefits attributed to a patient-centred approach such as safety, efficacy, available treatment options and variable delivery modes all contribute to individuals' ability to manage their health and the health of their community (WHO). Even though, telemedicine was available, it is not accessible to patients and a majority (76%) of patients faced challenges with electronic prescriptions (Kruse, M.H., Durstine, A. & Evans, D.P. , 2022). In India, the service of telemedicine is to be extended in a wide variety of forms due to ease and availability by keeping the in-person consultation or emergency medicine (Agarwal, N. et al., 2020).

eSanjeevani - National Telemedicine Service of India is a step towards digital health equity to achieve Universal Health Coverage (UHC) and can access it from smartphones or by visiting the nearest Ayushman Bharat Health & Wellness Centre (<https://esanjeevani.mohfw.gov.in/>). In India, eSanjeevani is introduced with the objective of reducing the rural-urban gap and improving the accessibility of health care services to the marginalised and thereby achieving Universal Health Coverage envisaged by the WHO. The UN's agenda for Sustainable Development Goals (SDGs) also stresses the target of achieving equity in health care services. It is also declared that health care is a fundamental right of human beings. But the quality of health care services depends on a number of factors which are already challenging in a developing country coupled with a number of new challenges like moving from a conventional system of treatment to skill needed telemedicine system. Even though, there are several successful cases of telemedicine/e-health services the widening of such practices in India is very less and the drawbacks of the existing services are not documented properly.

ISRO (Indian Space Research Organization) launched the first telemedicine services in India in 2001 for rural-urban linking in treatment (Mishra S, et al., 2009). Later, India launched its flagship telemedicine treatment in 2019 as part of the Digital India initiative for the Ministry of Health & Family Welfare under the Ayushman Bharat Scheme. It became popular and widely practised during the COVID-19 period. The two variants of the health services are eSanjeevani Ayushman Bharat-Health and Wellness Centre (AB-HWC), a Doctor-to-Doctor telemedicine service under the Ayushman Bharat-Health and Wellness Centres scheme of the Government of India, to provide general and specialised health services in rural areas and isolated communities and eSanjeevani OPD, a patient-to-doctor telemedicine service to enable people to get outpatient services in the confines of their homes. It is noted that it has crossed 3 crores of teleconsultation and 1.7 lakh consultations in a day (Ministry of Health and Family Welfare, 2022). Even though it is widely practised in the country during the COVID period, the current utilization of the service is continuously declining, particularly in some states and backward areas. A prima-face observation shows that people are not satisfied with the current services. Several barriers are there in the current services including lack of specialised care, and quality of health care services, network issues, lack of awareness of this among the rural and backward communities, etc. All these are challenges to achieve the major aim of the Ayushman Bharat scheme, Universal Health Coverage by providing universal, free and quality healthcare closer to people's homes. Hence the questions that arise here are, how can we strengthen the existing system of telemedicine in India? What are the actual problems faced by the different stakeholders (Government, doctors, patients and technical service providers) in India with respect to telemedicine delivery? What are the different ways out of the current challenges? Can we scale up the existing e-health delivery system to the rural, marginalised and backward communities? What are the policy suggestions for socially viable and sustainable telemedicine health services in India?

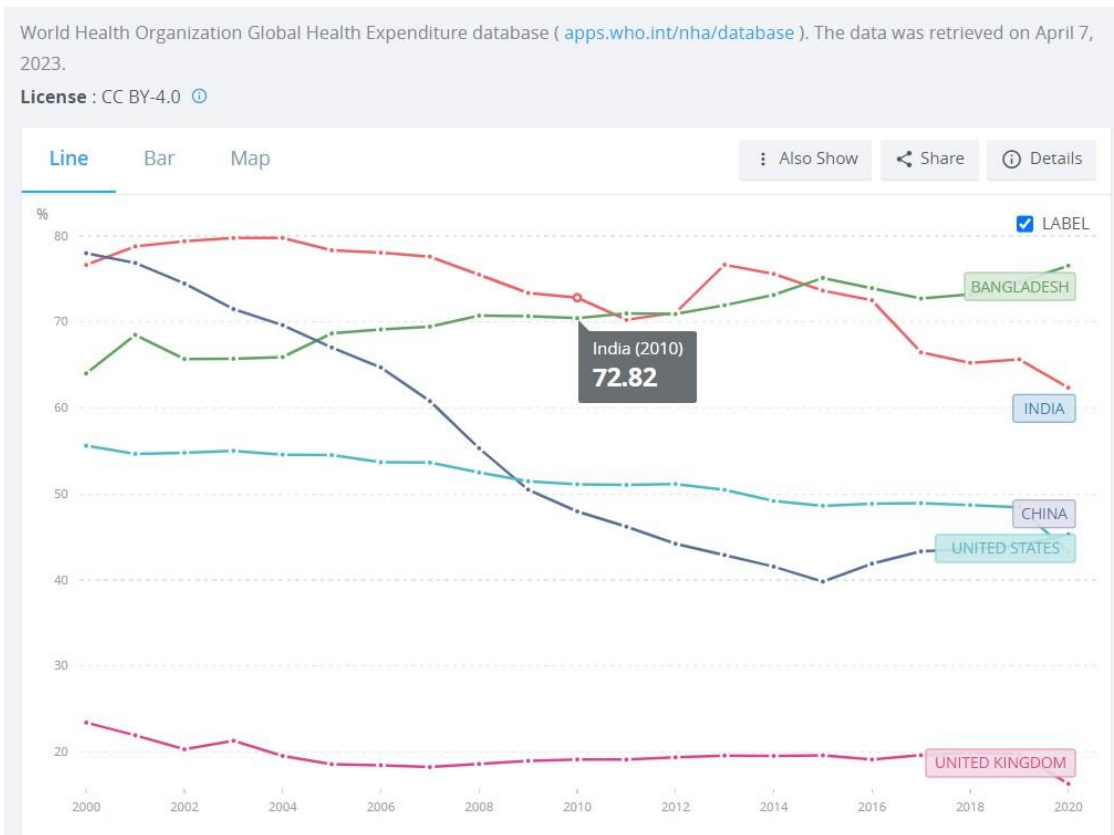
3. Importance of telemedicine in India

3.1. High health expenditure

Even though the healthcare system in India has improved a lot compared to a decade before, still the facilities at the public healthcare centres are not equitably distributed across regions and the out-of-pocket expenditure is also very high. Even though there is a decline in private health expenditure in India in 2020 (62.36 per cent) compared to 2000 (76.64 per cent), it is still very high compared to US, UK and China (Figure 1). It is due to the low share of public health expenditure in India compared to these countries and other OECD countries and is affecting the quality of healthcare. The United States is an outlier in healthcare spending and it's visible in the quality of healthcare and one of the reasons is the high income of the nation (www.oecd.org/health). India's public health expenditure as a percentage of GDP was only 2.96 per cent in 2020 whereas it was 18.2 per cent in the USA, even before COVID-19, it was 17 per cent in 2019. China spent 5.5 per cent of GDP in 2020 whereas the UK's is 11.9 per cent. This private sector caters for the needs of around 80 per cent of the people (Jilani et al., 2009). As a result of this low public health expenditure private sector is flourishing by charging a higher cost of treatment. The situation is worsened when the supply side of health care facilities not much increased based on the demand for the services. In this context, the introduction of

telemedicine/e-health services/eSanjeevani is promising and the effective utilization of that will reduce the demand for healthcare services from a private hospital to some extent and thereby the cost of healthcare. Studies showed that telemedicine decreases the burden of the healthcare system to a large extent in the case of service providers and patients (Chellaiyan, V.G., Nirupama, A. Y., & Taneja, N, 2019; Hailegebreal, S., Sedi, T.T., Belete, S. et al.,2022; Békási, S., Girasek, E. & Györffy, Z, 2022). Researchers argued that the role of ICT in patientcentred health care is very important, especially by reducing the cost (Hailegebreal, S., et al., 2022).

Figure 1: Comparison of India's domestic private health expenditure (% of current health expenditure) - India, Bangladesh, China, United Kingdom, United States



3.2 Presence of communicable and non-communicable diseases

India is experiencing a situation of epidemiological transition, that is from communicable to non-communicable diseases (Initiative Collaborators, D. B. (2017). At the same time, the presence of both diseases is high compared to other OECD countries. Besides, there are huge variations in the presence of disease across the country, especially in the case of non-communicable diseases. About 67 per cent of the disease, disability-adjusted life years (DALY) burden is associated with non-communicable diseases (Dandona L, Dandona R, Kumar GA, et al. 2017). The treatment of non-communicable diseases needs continuous follow-up actions over a long period of time and hence leads to catastrophic health expenditure. In such cases, telemedicine is the

most appropriate option and ensures the right to health for patients with chronic diseases. It is noted that telecare visits are relevant in the case of chronic care (Békási, S., et al., 2022). After the first direct visit to the doctor, telemedicine can be used for chronic diseases such as bronchial asthma, hypertension, and diabetes mellitus (Portnoy J, et al., 2020).

4. Conclusion

In conclusion, the paper proved that telemedicine/ e-health is a scientific method that can be useful for monitoring preventive and curative medicine at all ages of life. Telemedicine is a recently popular and innovative method in the healthcare sector, especially for treatment. It is also shown that telemedicine has gained immense popularity during the pandemic period COVID-19 and the expansion of internet facilities. It is clear that this alternative treatment method is needed to reduce healthcare expenditure in India. Studies showed that telemedicine decreases the burden of the healthcare system to a large extent. The paper found that telemedicine is an alternative treatment option for India, which suffers from high out-of-pocket expenditure, limited healthcare expenditure and high healthcare demand. In this context, the introduction of telemedicine is relevant and the effective utilization of that will reduce the demand for health care services from a centre and thereby sustainably manage the health care expenditure of the State.

However, it is important to note that though telemedicine is an important method it is not without limitations. First, the establishment cost in the case of specialized care is very high and safe storing of mass data needs that much specialization. If there are network issues, communication between the doctors and the patient is very difficult. Strengthen the telemedicine system by adding specialized care and diseases. Provide awareness classes to school students on using the telemedicine app and related procedures. Suggest measures to upgrade the existing technology, based on the patient-centred procedure.

References

- Agarwal, N., Jain, P., Pathak, R., & Gupta, R. (2020). Telemedicine in India: A tool for transforming health care in the era of COVID-19 pandemic. *Journal of Education and Health Promotion*, 9. https://doi.org/10.4103/jehp.jehp_472_20
- Békási, S., et al., (2022) Telemedicine in community shelters: possibilities to improve chronic care among people experiencing homelessness in Hungary. *Int J Equity Health* 21, 181. <https://doi.org/10.1186/s12939-022-01803-4>.
- Chellaiyan, V. G., Nirupama, A. Y., & Taneja, N. (2019). Telemedicine in India: Where do we stand? *Journal of Family Medicine and Primary Care*, 8(6), 1872-1876. https://doi.org/10.4103/jfmprc.jfmprc_264_19.
- Dandona L, Dandona R, Kumar GA, et al.. Nations within a nation: variations in epidemiological transition across the states of India, 1990–2016 in the global burden of disease study. *The Lancet* 2017;390:2437–60. 10.1016/S0140-6736(17)32804-0 [PMCFree article] [PubMed] [CrossRef] [Google Scholar] [Ref list]
- Dash, S., Aarthy, R. & Mohan, V. Telemedicine during COVID-19 in India—a new policy and its challenges. *J Public Health Pol* 42, 501–509 (2021). <https://doi.org/10.1057/s41271-021-00287-w>
- Dexheimer JW, et al., Computerized pneumococcal vaccination reminder system in the adult emergency department. *Vaccine*. 2011 Sep 16;29(40):7035-41. doi: 10.1016/j.vaccine.2011.07.032. Epub 2011 Jul 23. PMID: 21784117; PMCID: PMC3168965.
- Garzón-Orjuela, N., Samacá-Samacá, D.F., Luque Angulo, S.C. et al.(2020)An overview of reviews on strategies to reduce health inequalities. *Int J Equity Health* 19,192. <https://doi.org/10.1186/s12939-020-01299-w>
- Graham, A.L.; et al. An Integrated Digital/Clinical Approach to Smoking Cessation in Lung Cancer Screening: Study Protocol for a Randomized Controlled Trial. *Trials* 2017, 18, 1–14.
- Hailegebreal, S., Sedi, T.T., Belete, S. et al. Utilization of information and communication technology (ICT) among undergraduate health science students: a cross-sectional study. *BMC Med Educ* 22, 215 (2022). <https://doi.org/10.1186/s12909-022-03296-9>
- Initiative Collaborators, D. B. (2017). Nations within a nation: Variations in epidemiological transition across the states of India, 1990–2016 in the Global Burden of Disease Study. *Lancet* (London, England), 390(10111), 2437-2460. [https://doi.org/10.1016/S0140-6736\(17\)32804-0](https://doi.org/10.1016/S0140-6736(17)32804-0)
- Jilani, A., Azhar, G., Jilani, N., & Siddiqui, A. (2009). Private providers of healthcare in India: A policy analysis. *The Internet Journal of Third World Medicine*, 8, 1–5.
- Kruse, M.H., Durstine, A. & Evans, D.P.(2022) Effect of COVID-19 on patient access to health services for noncommunicable diseases in Latin America: a perspective from patient advocacy organizations. *Int J Equity Health* 21, 45. <https://doi.org/10.1186/s12939-022-01648-x>.
- Mishra S, Kapoor L, Singh I. Telemedicine in India: Current scenario and the future. *Telemed J E Health*. 2009;15:568–75. [PubMed] [Google Scholar] [Ref list].
- Morato, J., Iglesias, A., & Campillo, A. (2021). Sustainable Technologies for Older Adults. *Sustainability*, 13(15), 8465. <https://doi.org/10.3390/su13158465>.

- Owsley, C., Rhodes, L.A., McGwin, G. et al. Eye Care Quality and Accessibility Improvement in the Community (EQUALITY) for Adults at risk for Glaucoma: study rationale and design. *Int J Equity Health* 14, 135 (2015). <https://doi.org/10.1186/s12939-015-0213-8>
- Owsley, C., Rhodes, L.A., McGwin, G. et al., (2015) Eye Care Quality and Accessibility Improvement in the Community (EQUALITY) for adults at risk for glaucoma: study rationale and design. *Int J Equity Health* 14, 135 <https://doi.org/10.1186/s12939-015-0213-8>.
- Portnoy J, Waller M, Elliott T. Telemedicine in the Era of COVID-19. *J Allergy Clin Immunol Pract*. 2020;8:1489–91. [PMC free article] [PubMed] [Google Scholar] [Ref list]
- Telemedicine-Opportunities and developments in member states [Internet] 2nd ed. Geneva, Switzerland: WHO press; 2010. [cited 2019 Feb 1]. Available from: https://www.who.int/goe/publications/goe_telemedicine_2010.pdf .
- Understanding differences in health expenditure between the United States and OECD countries, September 2022, <https://www.oecd.org/health/Health-expenditure-differences-USA-OECD-countries-Brief-July-2022.pdf>
- World Health Organization. Regional Office for Europe. (2016). From innovation to implementation: eHealth in the WHO European Region. World Health Organization. Regional Office for Europe. <https://apps.who.int/iris/handle/10665/326317>.

Does adoption of digital marketing platforms impact household welfare? Evidence from gender disaggregated household data in Nigeria

Abdulrazaq Kamal Daudu, Oyedola Waheed Kareem
University of Ilorin, Ilorin, Nigeria;

Latifat Kehind Olatinwo
Kwara State University, Malete, Nigeria

Abstract:

Digital technologies have been developing at a rapid pace, especially in Africa south of Sahara as evidenced by the expansion of mobile network coverage and internet access which offers an opportunity to link farmers to markets, selling their products at remunerative prices and improving their welfare status. But many of the farmers are still not aware or adopt digital marketing to sell their products. This study seeks to investigate if adoption of digital marketing platforms has led to improved farming household welfare and helped narrow gender gap using a cross-sectional gender disaggregated survey data collected from 640 farming households across three states in northcentral Nigeria. The impact of adoption was measured in terms of household total farm sales and per capita food consumption expenditure. We used different matching techniques and endogenous switching regression model which controls for selection bias and unobserved heterogeneity, a commonly used method in adoption impact analysis. Results show that total farm size, marital status, membership of association, access to extension services, and credit, awareness, access to internet, and digital marketing training have a positive and statistically significant effect on digital marketing adoption and household welfare among male and female farming households. The outcomes indicate that adopters fare better than non- adopters in terms of their welfare status. Our econometric

results show that adoption of digital marketing platforms has significantly increased farm sales and per capita food consumption expenditure among male farming households compared to female farming households. Therefore, it is necessary to take effective measures that are aimed at improving digital marketing adoption among rural farming households and close gender gap in households' welfare may actually have significant farm sales and higher food consumption expenditure for households. Also, the significant difference of the identified factors should be upgraded to increase the assessment of each identified variable.

Keywords: adoption, digital marketing platforms, impact, rural farming household welfare, propensity score matching, endogenous regression, northcentral Nigeria,

Introduction

Agriculture is a pivotal sector of the economy of sub-Saharan Africa (SSA) and dominated by smallholder farmers who contribute about 80% to the domestic food production (Ricciardi et al., 2018 and Lowder et al., 2016). The sector is bedeviled by inefficiencies and however, smallholder farmers in most evolving nations are beset with challenges in accessing market information, knowledge and proficiency that could enhance their income (Chapagain and Raizada, 2017; Misaki et al., 2018; Petcho et al., 2019; Diaz et al., 2021). Despite these challenges, there's widespread confidence that the farming sector has the potential to feed the world's growing population, and lift hundreds of millions out of poverty. Important explanations for this positive perspective lie in smallholder agriculture and digital innovation. Digital innovation in agriculture has a great opportunity to annihilate poverty, hunger, and linking farmers to markets. Digital innovation needs to shift beyond large- scale agriculture operations to avail more smallholder farmers as well. In recent time, the rapid advancement of digital technologies is gradually present in rural areas and has remodelled the way that farming is operated and marketed. This digital transformation has a key part in agribusiness because it leads to further informed decisions, improved efficiency, and easier knowledge sharing. Some studies suggest there has been an informational revolution in the farming sector (Dyer, 2016), with the acceptance of digital technologies by farmers expanding for a decade. This trend will probably continue at an accelerated pace, given the increased venture capital allocated to developing these technologies (Boehlje and Langemeier, 2021).

Agricultural marketing system is gradually evolving over the last past few decades in Nigeria and this is due to digital transformation. Nigeria has the highest population in the region and accounts for around 47% of the West African population (Lixi et al., 2019). In 2018, the World Bank collection of development indices reported that about 49.66% of the Nigerian population lives in rural areas. In these rural areas, the farmers' access to urban markets is constrained by bad road networks, price change, lack of market information, and other market inefficiencies, which makes digital innovation an indispensable way of linking farmers to markets (Kapfudzaruwa et al., 2018). Even so, the affordability and long-range content of most cell phone providers have made it possible for nearly all rural inhabitants to be cell phone holders (Alliance for Affordability Internet, 2019; Trading Economics, 2020). In the same manner, more than 100 million citizens in Nigeria enjoy access to the internet, with over 250,000 new- fangled subscribers logging onto the internet in the terminal part of 2019 (Russon, 2020). With this growth, Nigeria has witnessed a radical breakthrough in the information communication technology sector (Ogbeide-Osaretin and Ebhote, 2020). Telecommunication and information technology contributed 77% of the entire sector's contribution to the GDP. Altogether, the mobile telecommunication sub-sector contributed 7.4% to the country's GDP in 2018, compared to 5.5% in 2017. Presently, 44% of mobile subscribers in Nigeria are harnessing 3G and 4% are exploiting 4G technology (Mobile Report, 2019). In the era of global digitalization and the introduction of new technologies, most of the activities are converted from conventional marketing to digital marketing due to the fairly lower cost of studying and the processing speed of big data has enhanced product promotion (Robul, 2020).

Digital marketing has popped in popularity in recent times, offering both advantages and challenges for small agribusinesses. All marketing enterprise that employs an electronic device and the internet are referred to as digital marketing. Digital marketing is largely recognized in the commercial world, but its application in small agribusinesses varies greatly (Upadhyay, 2022). It's expected that Digital Marketing will upheave agricultural output and growth of the agricultural market. Earlier the smallholder farmers trade with only a smaller number of buyers who pickup directly. Nowadays the buyers are getting increased in the digital agriculture marketing mode. The digital and internet- based operations in the android phones provides the market information to the farmers which can downscale the market distortions, logistics transportation losses & products' wastage & damage. In the national & transnational digital marketing scenario, android mobile and internet based digital usages can help in locating the farmers, agri-input dealers, agri-out put buyers, storage storages, recognized financial institutions for investment set care effortlessly linked and hence the farmers are getting the improved profits to their yields in the digital agriculture markets. Early research on the adoption of digital marketing offers insight into why decision makers adopt or reject new technologies (Alford and Page, 2015). Empirical study has established a number

of characteristics that are thought to affect small business adoption of digital marketing technology. therefore, it's imperative to explore factors influencing farmers' technology adoption, particularly digital based marketing, to apply the "Internet plus agriculture" strategy. Although technology adoption research has been conducted among numerous types of users, studies concentrating on farmers are quite insufficient (Venkatesh et al., 2016).

One of the major challenges in digital world is the digital divide across various parts of society such as the rich and the poor, the urban and the rural, the educated and the illiterate as well as the gender divide which is veritably vividly mirrored in the use and access of technology. Numerous women farmers fail to profit from digital marketing platforms because they lack severely in the awareness, exposure, access and proficiency in the use of digital marketing services. Women's control and access to affordable digital marketing platforms is rather low because of several socioeconomic and cultural factors similar as power relations in societies. The heavy burden of household responsibilities on women works against their exploiting digital services for selling their farms. therefore, incorporating the gender dimension is a necessary prerequisite for the achievement of a globally equitable information society (UNCTAD, 2014). likewise, for women who are new or novice mobile internet users, low mobile knowledge and a lack of digital proficiency are major barriers to employing the full potential of the internet. Studies by GSMA (2015) and other organizations have highlighted how a lack of knowledge and digital proficiency discourage new users from accessing mobile internet, and inhibit existing users from using mobile internet on their own. Also, till now there has been no comprehensive investigation on the impact of digital marketing platforms which studies a population from a gender viewpoint, similar as the rural household heads in Nigeria, where access and exposure to digital marketing platforms are veritably low. The measures of digital knowledge and marketing platforms available at present, thus, aren't suitable for such an environment with low access and low exposure. Some of these point to a significant gap in the literature that our study aims to fill. Hence, this study fills this vacuum by giving empirical data and provides a first attempt to research the impact of the adoption of digital marketing platforms on household welfare in Nigeria by gender using quantitative evidence from a cross- sectional dataset of farming households. Such empirical documentation could inform policies in sub-Saharan Africa aimed at rural digital transformation and rural economic development.

Lately, studies like those conducted by Fasakin et al. (2022), Oyetunde- Usman, et al. (2021), and Ahmed et al. (2017), have employed an economic model like Propensity Score Matching (PSM) to estimate adoption impact on household revenue and welfare. PSM is the best option when differences between adopters and non-adopters can only be explained by observable characteristics. Nevertheless, PSM results are subject to bias, especially if there's

misspecification in the propensity score model (Wossen, et al., 2019 and Wooldridge, 2010). We employ a robust analytical approach, notably the Endogenous Switching Regression (ESR) model (Teklewold et al., 2019; Di Falco, 2014; Ahmed et al., 2017; Teklewold et al., 2013; Kassie et al., 2017) to address the misspecification and heterogeneity issues that afflicted PSM and better the majority of preceding studies. This ESR model accounts for both observable and unobservable confounders that affect the intention of rural household to adopt digital marketing platforms. The following are the immediate policy issues addressed in this study (1) What factors influence rural farming household's decision to adopt digital marketing platforms? (2) Is adoption of digital marketing platforms improving rural household welfare? Answering these questions is central to linking smallholder farmers to output markets through digital transformation in rural areas, increase farm income, and household welfare in Nigeria and sub-Saharan Africa. Deichmann et al. (2016) note that the rapid digital transformation and expansion of mobile network coverage and internet access in SSA presents an opportunity to link farmers to markets amidst subsisting and arising constraints. Also, it's anticipated that the measure of impact digital marketing platform on household welfare emphasized in this study will be the earliest of its kind to address this gap in the literature for a systematic study on digital marketing which will inform and conduct further improvement in digital service uptake. The evidence on the drivers of the intention to adopt digital marketing platforms among smallholder farmers will guide the government and policy makers on the significance of upgrading factors impacting or impede adoption of digital marketing platforms in agriculture in order to enhance its relinquishment in other sectors and its overall contribution to the economy.

The remainder of the paper is structured as follows. Section 2 discusses relevant literature on digital marketing in agriculture. Section 3 presents the theoretical framework. Section 4 covers the study area, data and sampling methods, and empirical estimation strategy. The results and discussion of major findings are reported in Section 5. Finally, Section 6 draws conclusion and recommendations from the results.

Review of relevant literature

With the rapid digital transformation across the globe, numerous literatures have conducted empirical studies on the adoption of digital marketing among smallholder farming households. For instance, Seyed et al. (2016) surveyed the adoption of agricultural e-marketing employing the planned behaviour theory (TPB). Employing direct data from 146 professionals drawn from open and private sectors, results showed that the model demonstrated about 94 percent and 71 percent of the differences in behaviour towards intent for both public and private professionals collectively. The study recommended speeding up e-marketing perpetration, the easiness of application and recompenses needs to

be vindicated employing divergent programs like video displays and stores. Thomas et al. (2016) estimated the applicability of social media as a means of agricultural information using Kesses District farmers researched in Kenya. They made use of a descriptive survey on primary data. They recommended the establishment of information centres in the areas where farmers can be readily sourced for veritable information that will help enhance their farming activities. The adoption of marketing strategies predicated on social media has extensively bettered the interaction of the enterprises with consumers; enhance the marketing strategies based on consumer feedback as well as helping in gaining marketing information (Galati et al., 2017). Based on study conducted by Sudhakar Reddy (2021) point out that digital marketing platforms are useful to the farmers as they increase the selling price and reduce the marketing cost of their agriculture output like Cotton Lint, Maize Grains, Paddy, Vegetables etc. Digital marketing is an important business and information tool. He similarly states that digital innovation helps the marketing of high volumes of agriculture. Mbagwu et al. (2018) looked at the problems of meeting the information conditions of agriculturists in the villages via internet services. The explanation of the research was to ascertain the multitudinous hassles connected with meeting information conditions of rural farmers through online-based services with a specific focus on developing countries. The finding of the study shows that forming an online farmers' dialogue occasion, establishing ICT and agricultural information knowledge, adequate research to know the need of the rural farmers, repackaging of information using social media platform was among the major recommendation for the study. Indeed so, internet and web-based usages are lengthily harnessed in distribution and spreading of agricultural awareness, marketing of goods and services. The study conducted by (Ramli et al., 2015) has discovered the proposition that online campaigns are an effective answer to complications that bring about the advancement of the agricultural industry, similar as frail marketing associations, poor information dissemination, abridged productivity, low earnings and absence of diversity. Singh et al. (2015), opined that the significance of online marketing in agriculture can be seen in the area of distributing agricultural information networks to farmers at all situations (Ogbeide-Osaretin and Ebhote, 2020). Similar openness of information could support farmers to make better farming decisions which could enhance increased yields, lessen environmental effects, and raise livelihoods (Ogbeide-Osaretin and Ebhote, 2020). While Balu (2020) stated that inadequate extension services and poor access to information leads the gap in the adoption of technology and lead to poor productivity levels in Agricultural sector. He also pointed out that right information at right time plays a pivotal role for development of Indian agriculture. A digital technology tool helps youthful farmers to involve agribusiness in effective manner. Also, Fainos et al. (2014) investigated the role of agricultural marketing on the empowerment of rural farmers in Masvingo Province, Zimbabwe. Data collected from 361 farmers who are specialists in maize, groundnuts, small grain crops, and vegetable product

showed that the farmers require access to market information given their use of traditional means of gathering information. They, thus, recommended the government's enhancement of training centers to give knowledge to rural farmers as well as the establishment of a marketing board to harmonize marketing activities. Eventually, Fecke et al. (2018) have lately stated that there's a lack of studies concentrating on this field and Zeng et al. (2017) pointed out that further research is useful to better clarify which are the factors that largely impact farmers' intention to adopt, in order to support the design of further adapted policy strategies. Albeit the openness of internet access is significant (Hennessy et al., 2016), Ma et al. (2019) assert that it is not sufficient. In adoption research, the farmer's gender has become a significant socioeconomic indicator. Ragasa (2012) establish a gender difference in agricultural technology use. So far, evidence suggests that women's gender has a negative and considerable impact on the adoption of new agricultural technologies.

Theoretical framework

This study employs Technology Acceptance Model (TAM), which was proposed by Davis (1989) and has been extensively used and empirically tested across studies to inform the technology implementers whether or not targeted individuals would undertake the new technology (Kabbiri et al. 2018). It was developed from the psychological Theory of Reasoned Action (TRA) and Theory of Planned Behaviour (TPB) and has evolved to become a key model in understanding predictors of human actions toward possible acceptance or rejection of the technology. It has been extensively embraced due to its simplicity and robustness and offered by several research scholars in the field of technology uptake (Diaz et al., 2021). According to Kabbiri (2018), TAM is veritably useful in predicting an individual's acceptance of technology. However, despite, its robustness TAM (Venkatesh et al., 2003) has its own limitations, and the consequence of social influence on the implementation of the information systems wasn't included in the theory (Davis, 2007). Kabbiri et al. (2018) observed that the main two constructs of TAM may not completely explain users' behavioural intention toward the use of mobile phones. latterly, Venkatesh & Davis (2000) extended the model as the Unified Theory of Acceptance and Use of Technology (UTAUT) and gave a comprehensive note on technology acceptance and its actual use agreeing to this theory adoption of innovation is affected by four crucial factors namely effort expectancy, performance expectancy, social influence, and facilitating conditions. Several research scholars have extended the TAM and added constructs like trust, perceived playfulness, cognitive absorption, product involvement, and perceived enjoyment (Diaz et al., 2021). Jeong and Yoon (2013) offered the model by adding perceived financial cost, self-efficiency, and credibility in a mobile banking environment. Along with trust (Munoz-Leiva et al., 2017; Thanabordeekij et al., 2020), social image and perceived risk (Munoz-Leiva et al., 2017) have also been added as new constructs of TAM. UTAUT has been

extensively adopted to explain individuals' behaviour in various climes, like mobile payment services (Slade et al., 2015), e-learning services (Liu et al., 2019), and learning management systems (Buabeng-Andoh & Baah, 2020). Okumus et al. (2018) stated UTAUT has high predictive power than any other theories of human behaviour. UTAUT explains how an individual use new technology after a series of multifaceted mental processes, in which intention leads to user behaviour. latterly, this study integrates self- efficiency and sources of information as further determinants of farmers' intention to use digital marketing platforms. The subsisting studies have established that an individual's self- efficacy is linked to a positive attitude to the usage of technology (Engotoit, 2016; Okumus et al., 2018). Therefore, individuals' self- efficacy is an essential determining factor of farmers' behavioural intention to use digital marketing platforms.

Methodology

Study Area: The study was conducted in the northcentral region of Nigeria. The choice of northcentral Nigeria for this study was predicated on the active participation in arable crop farming. The region is made up of six states namely Benue, Kwara, Niger, Plateau, Nasarawa, Kogi and Abuja the Federal Capital Territory and positioned geographically in the middle belt region of the country, scaling from the west, around the convergence of the River Niger and the River Benue. It lies roughly between 3⁰ and 14⁰E and latitude 70 and 100N. It constitutes the food basket of Nigeria covering about, 730,000 km². The region itself is rich in natural land features, and boasts some of Nigeria's most exciting decor. The region is also home to numerous historical and colonial relics. The total population of north central Nigeria is over 24 million (National Bureau of Statistic, 2020). The climate of the region is halfway impacted by climates in the northern and southern region of Nigeria while the major vegetation is principally the Guinea Savannah Zone which occupies same 90% of the land mass. The tropical savannah climate defined by wet and dry condition affects most parts of north central Nigeria. Rainy seasons de-cline similarly in length as one move northward. Temperature is generally high in the region. Its rich agricultural products include as yam, rice, beans, cassava, sweet potato, corn, soybean, sorghum, millet, sesame, cocoyam etc.

Data and Sampling Methods: We used data from a cross-section of rural farming household survey which was conducted between August to December 2022. The data collection process in the survey involved a multi-level stratified random sampling technique from a cross-section of rural farming households who grow arable crops either as a primary or secondary occupation. The first stage involved a random selection of 50% of the states in northcentral Nigeria (Benue, Kwara, and Niger). In the second stage, there are 64 local government areas (LGAs) in the 3 selected states (Benue 23; Kwara 16; Niger 25), we randomly select 25% of the LGAs (Benue 6, Kwara 4, and Niger 6) making a

total of approximately 16 LGAs. The list of rural farming communities was obtained from the Department of Agriculture of each selected LGA. In stage three; 2 rural farming communities from each of the sampled LGAs were randomly selected, a total of 32 rural farming communities for the study. Finally, the fourth stage involved a random selection of 20 farming households (10 male and 10 female), from each of the selected farming communities, resulting in a total of 640 respondents for the study. However, from the sampled households, 8 adopters of digital marketing platforms and 12 non-adopters digital marketing platforms were eventually selected from each household, totalling 256 adopters and 384 non-adopters of digital marketing platforms. Effort to get equal number of male and female household respondents prove abortive, therefore, 141 and 115 adopters, 223 and 161 non-adopters of digital marketing platforms of male and female respondents respectively, were finally interviewed for the study.

We collected data using a structured questionnaire delivered on an Android electronic tablet software (Surveybe). The survey questionnaire was segmented based on the objectives of the study. Information was collected from respondents on the socioeconomic characteristics of farm households, crop production, access to information, credit and extension, expenditure on food and other items, income from farm sales, and the decision to adopt digital marketing platforms. Table 1 present the description of all variables used in our model. Consent from respondents was obtained prior to answering the enumerators' questions. All respondents were asked not to participate in the study if they felt uncomfortable and to withdraw at any time during the survey. We adopted two important welfare outcome variables in this study and these include total income from farm sales per production and food consumption expenditure per capita. The total income from farm sales represents the crop income per production and was measured by in actual sales per year in Naira (₦) preceding the survey (2022). Following similar impact studies (Wossen et al., 2017; Abdoulaye et al., 2018, Ahimbisibwe, et al., 2020), we adopt the food consumption expenditure per capita. The consumption expenditure was measured by asking the sample households on food expenditure for the preceding year covering 12 months consistent with the World Bank's LSMSISA standard module. Our study has one treatment variable which is the adoption of digital marketing platforms by rural farming households measured by a measured by a dummy that takes a value of 1 if a rural farming household adopt digital marketing platforms, and 0 otherwise.

Table 1: Description of variables included in the model

Variables	Description
Outcome variables	
Total income from crop sales	Total value of crop sales per hectare in 2022 cropping season in Nigeria naira (₦/ha)
Food consumption expenditure per capita	Total expenditures spent on food in Nigeria naira (₦) divided by the total number of household members
Treatment variables	
Adopt Digital marketing platforms	Dummy = 1 if household adopted digital marketing platforms; 0 otherwise
Control variables	
Age	Age of household head (years)
Marital status	Dummy = 1 if married; 0 otherwise
Education	Years of formal education
Household size	Total number of people in the household
Total farm size	Total farm size cultivated in hectare (ha)
Farming experience	Number of years of farming (years)
Membership of farmers' union	Dummy = 1 if yes; 0 otherwise
Extension service	Dummy = 1 if access to extension service; 0 otherwise
Credit access	Dummy = 1 if access to credit service; 0 otherwise
Awareness of digital marketing	Dummy = 1 if aware of digital marketing; 0 otherwise
Access to internet service	Dummy = 1 if yes; 0 otherwise
Access to digital marketing training	Dummy = 1 if attended any digital marketing training; 0 otherwise
Northcentral region	Dummy = 1 if the household is in the northcentral agroecological zone, 0 otherwise

Source: Field survey (2022)

Estimation strategies

Farming household's decision to adopt technology can be considered under utility maximization framework. According to Foster and Rosenzweig (2010), agricultural technology adoption and inputs use are the outcomes of optimization by heterogeneous agents. This optimization takes place in the presence of constraints on the decision-maker's budget, information and credit access, and on the availability of both the technology and other inputs. Therefore, households are assumed to maximize their utility function subject to these limitations (Asfaw, Shiferaw, Simtowe, & Lipper, 2012). Following McFadden (1974), we analyzed the decision of farming household to adopt digital marketing platforms as a question of preference within the utility context. We assumed that if the predicted net utility from adopting digital marketing platforms MiA and the utility from not adopting the digital

marketing platforms MiN may be denoted as M^* , such that a utility maximizing farming household, i , will adopt digital marketing platforms if the utility gained from adopting digital marketing platforms exceeds utility from not adopting ($M^* = MiA - MiN > 0$). The unobserved net utility is represented as a function of observable elements represented in the latent variable model as shown in Equation (1). Following Asfaw et al. (2012) and Becerril and Abdulai (2010), the adoption decision can subsist modelled in a random utility framework as follows:

$$M_i = X_i' \beta + \varepsilon_i \text{ with } M_i = \begin{cases} 1 & \text{if } M_i^* > 0 \\ 0 & \text{otherwise} \end{cases} \dots \dots \dots (1)$$

where Mi is a binary indicator variable that equals 1 if a farming household adopt digital marketing platforms and 0 otherwise. The term $X'\beta$ provides an estimate of the difference in utility from adopting the digital marketing platforms ($MiA - MiN$), using the household and farm-level characteristics such as age, education, household size, farm size, farming experience, etc., X' , as explanatory variables. The term β is a vector of parameters to be estimated and ε_i is an error term. However, adoption of digital marketing platforms among farming households might not be voluntary/random but could be a function of possible self-selection bias problem due to the observable and unobservable characteristics of the farming household that needs to be addressed to prevent producing biased estimates of how adoption of digital marketing platforms can impact household welfare.

Several methods have been adopted in the literature to resolve the issue of selection bias in impact evaluation. Such studies include Oyetunde-Usman et al. (2021); Ahimbisibwe et al. (2020); Mandate et al. (2019); Tufa et al. (2019); Abdoulaye et al. (2018); Wossen et al. (2017); Feleke et al. (2016); Khonje, et al. (2015); and Shiferaw et al. (2014). In this study, we used Propensity Score Matching (PSM) and Endogenous Switching Regression Model (ESR) approaches that take into account the potential self-selection problem. The latter is used as a robustness check for the results of the former. According to its observable characteristics, the PSM technique estimates propensity values, which are estimates of farming household's propensity to adopt digital marketing platforms. This strategy eliminates selective bias by mimicking the randomized control strategy. The proficiency of the PSM to control for the differences in observable covariates that might influence the decision of the farming household in adopting digital marketing platforms is based on the conditional independence assumption. This states that conditional on observables characteristic of household and farm-level (X), welfare outcomes are independent in the adoption of digital marketing platforms specified as: ($TY1, Y0 \perp T | X$). Another vital assumption is the common support or overlap condition: $0 < (T = 1 | X) < 1$. According to Heckman et al. (1999), this condition

ensures that the treatment observations (adopters) have comparison observations (control groups) “nearby” in the propensity score distribution. Only in areas of common support can inferences be made about causality. It is also very important to conduct a balancing test, that is, to ascertain if: $E * (X/T = 1) = E * (X/T = 0)$. However, it is possible by matching methods to determine the parameters of the average treatment effect (ATE) and the average treatment on treated (ATT). The ATE and ATT differ in that the former estimates the mean improvement across all study population, while the latter estimates the impact of differences in the subpopulation of rural farming households adopted digital marketing platform. There is a need to search for the appropriate counterfactuals that match each adopter depending on its propensity score. The effect of adoption on household welfare outcomes was estimated using the two different algorithms most commonly adopted matching methods such as nearest neighbor (NNM), and kernel-based matching (KBM) in the literature (Caliendo and Kopeinig, 2008). In the NNM, the untreated individuals are selected as partners who closely match the propensity score of the treated individuals. Kernel matching takes information from all non-participants and constructs the counterfactual results using a weighting function and reduces the variance (Gautam et al., 2017). The parameters of interest in this study are the ATT, which allows us to assess household welfare (outcome) of rural farming households adopting digital marketing platforms if they had chosen not to be in the treatment group. Following Imbens and Wooldridge (2009), the ATT is expressed as follows:

$$ATT = E[Y(1) - Y(0) | T = 1] \dots \dots \dots (2)$$

where Y (1) and Y (0) are outcome indicators (measures as total farm income in Naira and food consumption expenditure per capita). T represents the treatment indicator, which is equal to 1 for household adopting digital marketing platforms. However, given the satisfaction of the conditional independence assumption, represented as $TY1, Y0 \perp T | X$, meaning that their household welfare outcomes are dependent on observable characteristics of rural farming households independent of adoption of digital marketing platforms, and the overlap condition, which ensures that the treatment observation has adjacent comparison observations in the propensity score distribution and inferences about causality can only be made in areas of common support, then the ATT can be calculated as follows:

$$ATT = E[Y(1)|T = 1, p(x)] - E[Y(0)|T = 0, p(x)] \dots \dots \dots (3)$$

However, the PSM is limited only to evaluating observational data from rural farming households in generating ATT estimates, which may be biased due to the assumption that there are no differences in the unobservable characteristics of rural farming households. In reality, differences in household welfare

outcomes may depend on unobservable innate ability and motivations that might influence rural household's decision to adopt digital marketing platforms. For a more robust approach and consistent estimation of the impact of adoption of digital marketing platforms on the outcomes of interest (household welfare), we considered using the linear endogenous switching regression (ESR) model that accounts for the other bias related to unobservable characteristics of the growers, which cannot be controlled using the PSM technique alone (Wossen et al., 2017; Ma and Abdulai, 2016; Awotide et al. 2015; Shiferaw et al., 2014; Lokshin and Sajaia, 2004). A given endogenous switching regression (ESR) model consists of one treatment selection equation and two separate outcome equations for the outcome variable of interest that are conditional on the selection criterion. The treatment selection equation is defined by a probit model, and the two outcome equations are linear that addresses the endogeneity problem by simultaneously estimating the selection and outcome equations using Full Information Maximum Likelihood (FIML) (Wossen et al., 2017; Ma and Abdulai, 2016; Lokshin and Sajaia, 2004). Following Ahimbisibwe et al. (2020), in the context of our study, the ESR model consists of a probit model of digital marketing platforms adoption and linear models of welfare outcome variables. Here we derived the selection equation by supposing that rural farming household make important decisions about whether or not to adopt digital marketing platform based on expected benefits (in terms of improve household welfare). Suppose a rational household decide to adopt digital marketing platforms enter treatment if welfare outcomes are significantly reasonable, i.e., $Y_i^* = \mu_1 - \mu_0$. Since Y_i^* cannot be observed directly, it can be expressed as a latent variable model as follows:

$$Y_i^* = \alpha_1 + \beta H_i + \mu_i \text{ with } T_i = \begin{cases} 1 & \text{if } Y_i^* > 0 \\ 0 & \text{if } Y_i^* \leq 0 \end{cases} \dots \dots \dots (4)$$

whereby Y_i^* is the latent variable indexing the likelihood of rural farming household adopting digital marketing platforms; H_i represents variables affecting expected benefits from adoption of digital marketing platforms; β is a vector of parameters to be estimated, and μ_i is a random error related to household adoption of digital marketing platforms. The result function dependent on Y_i^* , where rural farming households are faced with two regimes (1) adoption and (2) non-adoption, and then can be specified as an ESR model in the following way.

$$\text{Regime 1 (adopters): } Y_{1i} = \beta_1 S_{1i} + \varepsilon_{1i} \text{ if } T_i = 1 \dots \dots \dots (5)$$

$$\text{Regime 2 (non-adopters): } Y_{2i} = \beta_2 S_{2i} + \varepsilon_{2i} \text{ if } T_i = 0 \dots \dots \dots (6)$$

where Y_{1i} and Y_{2i} represents the outcome indicator of treated vs. untreated farming households;

S_i depicts a vector of exogenous variables affecting outcome variables; β is a vector of parameters to be estimated; s_{1i} and s_{2i} are the error terms linked to the outcome equations. Following Maddala (1983), the error terms in the selection equation (4) and the outcome equations (5) and (6) are assumed to have a triumvirate normal distribution with a mean of zero and a non-singular covariance matrix is given as follows:

$$Cov(\mu_i, \varepsilon_{1i}, \varepsilon_{2i}) = \begin{bmatrix} \sigma_\mu^2 & \sigma_{\mu\varepsilon_1} & \sigma_{\mu\varepsilon_2} \\ \sigma_{\mu\varepsilon_1} & \sigma_{\varepsilon_1}^2 & \sigma_{\varepsilon_1\varepsilon_2} \\ \sigma_{\mu\varepsilon_2} & \sigma_{\varepsilon_2\varepsilon_1} & \sigma_{\varepsilon_2}^2 \end{bmatrix} \dots \dots \dots (7)$$

where σ_μ^2 = variance of the error term in the selection equation, which is assumed to be 1; $\sigma_{\varepsilon_1}^2$ and $\sigma_{\varepsilon_2}^2$ = variance of the error terms in the result equations, $\sigma_{\mu\varepsilon_1}$ and $\sigma_{\mu\varepsilon_2}$ = covariance of μ_i, ε_{1i} and ε_{2i} , which measure the direction and degree of non-random selection. The fact that the outcome variables Y_{1i} and Y_{2i} cannot be determined simultaneously (Maddala, 1983), therefore no covariances between the error terms in the result equations $\sigma_{\varepsilon_1\varepsilon_2}$ and $\sigma_{\varepsilon_2\varepsilon_1}$ could be defined. Hence the expected values of the error terms are ε_{1i} and ε_{2i} due to the household's participation criterion is non-zero because of the possible correlation between the error term in the adoption equation and the error terms in the outcome equations.

$$E(\varepsilon_{i1}|Y_i = 1) = \sigma_{\mu\varepsilon_1} \left\{ \frac{\theta(\hat{Y})}{\varphi(\hat{Y})} \right\} \dots \dots \dots (8)$$

$$E(\varepsilon_{i2}|Y_i = 0) = -\sigma_{\mu\varepsilon_2} \left\{ \frac{\theta(\hat{Y})}{1 - \varphi(\hat{Y})} \right\} \dots \dots \dots (9)$$

where $\theta(\cdot)$ is the standard normal probability density function, $\varphi(\cdot)$ is the standard normal cumulative function; $-\frac{\theta(\hat{Y})}{\varphi(\hat{Y})}$ and $\frac{\theta(\hat{Y})}{1-\varphi(\hat{Y})}$ are the endogenous choice terms or inverse Mill's ratio weighted at $Y^\wedge = Hi\beta$ in the participation equation where $Y^\wedge$ is the predicted probability of farming household adopting digital marketing platforms, Y_i . Because the ESR model treats the problem of selection bias as a missing variable problem, the inverse Mill's ratio terms from the probit model are included in the linear outcome equations to correct for

potential selection bias, which is given as:

$$Y_{1i} = \beta_1 S_{1i} + \sigma_{\mu\epsilon_1} \left\{ \frac{\theta(\hat{Y})}{\varphi(\hat{Y})} \right\} + \epsilon_{1i}, \text{ if } T_i = 1 \dots \dots \dots (10)$$

$$Y_{2i} = \beta_2 S_{2i} - \sigma_{\mu\epsilon_2} \left\{ \frac{\theta(\hat{Y})}{1 - \varphi(\hat{Y})} \right\} + \epsilon_{2i}, \text{ if } T_i = 0 \dots \dots \dots (11)$$

In addition to the above equations, which can be estimated in a two-stage procedure, the simultaneous estimation of the adoption of digital marketing platforms among rural farming households and outcome equations using the full information maximum likelihood (FIML) are therefore based on Lokshin & Sajaia (2004) could be considered an efficient approach. Therefore, FIML can be implemented in Stata using the movestay command. A statistically significant estimate of $\sigma_{\mu\epsilon_1}$ and $\sigma_{\mu\epsilon_2}$ indicates endogenous conversion. However, the model discussed is used to estimate the average treatment effect on the treated and the untreated, i.e., the ATT and ATU. The equations are given as follows:

For digital marketing adopters with adoption:

$$E(Y_{1i}|T_i = 1) = S_{1i}\beta_1 + \sigma_{\mu\epsilon_1} \left\{ \frac{\theta(\hat{Y})}{\varphi(\hat{Y})} \right\} \dots \dots \dots (12)$$

For digital marketing non-adopters with adoption:

$$E(Y_{2i}|T_i = 1) = S_{1i}\beta_2 + \sigma_{\mu\epsilon_2} \left\{ \frac{\theta(\hat{Y})}{\varphi(\hat{Y})} \right\} \dots \dots \dots (13)$$

The mean treatment effect on those treated (ATT) is then defined as the difference in expected outcomes between equations (12) and (13), which reflect the impact of adoption of digital marketing platforms on welfare of adopters. Following Ahimbisibwe et al. (2020), the expected outcomes for digital marketing platform adopters under observed conditions and counterfactual conditions (i.e., had they not participated) will be computed using Equation (14) and Equation (15), respectively given as:

Digital marketing adopters, had they not adopted:

$$E(Y_{2i}|T_i = 0) = S_{2i}\beta_2 + \sigma_{\mu\epsilon_2} \left\{ \frac{\theta(\hat{Y})}{1 - \varphi(\hat{Y})} \right\} \dots \dots \dots (14)$$

Digital marketing non-adopters, had they decided to adopt:

$$E(Y_{1i}|T_i = 0) = S_{2i}\beta_1 + \sigma_{\mu\varepsilon_1} \left\{ \frac{\theta(\hat{Y})}{1 - \varphi(\hat{Y})} \right\} \dots \dots \dots (15)$$

Finally, the difference in the expected results from Equation (14) and Equation (15) is called the Average Treatment Effect on the Untreated (ATU), which is the potential impact of adoption of digital marketing platforms on welfare outcome of non-adopters.

Results and Discussion

Summary of descriptive statistics

The results in Table 2 compares the descriptive statistics for the household characteristics between digital marketing platforms' adopters and non-adopters using mean difference and t- test. Figure 1-4 compares the density estimates of the outcome (total farm sales and food consumption expenditure) variables used as indicators for household welfare between digital marketing platforms' adopters and non-adopters among male and female farming households, indicating higher estimates for digital marketing platforms' adopters in both categories of gender. The total farm sales and food consumption expenditure were NGN391,833.69 and NGN12,441.98 times higher among digital marketing platforms' adopters compared with non- adopters in male farming household sub-samples. In female farming households, the total farm sales and food consumption expenditure were NGN284,031.06 and NGN3,855.15 higher among digital marketing platforms' adopters compared to the non-adopters' counterparts. The difference between male and female farming household in terms of total farm sales and food consumption expenditure were NGN107,802.63 and NGN8,586.83 respectively. While digital marketing platforms' adoption might have contributed to the observed mean differences in the total farm sales and food consumption expenditure among male and female farming households, it will be misleading to attribute the entire difference to digital marketing platforms' adoption without controlling for all the differences in household characteristics as well as any unmeasured heterogeneity.

In addition to the outcome variable in Table 2, we included other covariates that assumed to influence the adoption of digital marketing platforms among the male and female farming household heads. These covariates include age, marital status, education, household size, farm size, farming experience, membership of association, access to extension, access to credit, awareness of digital marketing platforms, access to internet service, and digital marketing training. The results show statistically significant differences in most of the covariates between digital marketing platforms' adopters and non-adopters of

male and female farming household heads' characteristics. The digital marketing platforms' adopters are relatively older, married, better educated, larger household and farm size with more farming experience among both male and female farming households. They also have more access to institutions such as extension, credit, awareness, internet services, and digital marketing training programs compared with non-adopters of digital marketing platforms. For example, about 91% of the digital marketing platforms' adopters are members of farmers association compared with 54% of non-adopters among male farming households. In female category, 81% of adopters are members of farmers' union while only 49% of non-adopters belongs to farmers associations. Also, 76% of adopters of digital marketing platforms among male farming households have extension services compared with about 46% of non-adopters. Likewise, among female farming households only 43% of digital marketing platforms adopters have access to extension services compared with just 32% of non-adopters. Results show that only 41% and 30% of digital marketing platforms adopters in both male and female farming households have access to credit facilities compared with about 23% and 19% of non-adopters respectively. About 100% and nearly 99% of digital marketing platforms adopters are aware of various online marketing platforms compared with 76% and 29% of non-adopters among male and female farming households. Similarly, about 100% of the digital marketing platforms adopters in both categories of gender have access to affordable internet services compared with 35% and 53% of non-adopters of their counterparts. Our findings are in line with Deichmann et al. (2016) who found that the rapid digital transformation and expansion of mobile network coverage and internet access in SSA presents an opportunity to link farmers to markets amidst subsisting and arising constraints. Moreover, all digital marketing platforms adopters in male farming households received training on digital marketing compared with just 68% of the adopters among female farming households. However, our findings suggest that adopters and non- adopters of digital marketing platforms among male and female farming households are systematically different.

Table 2. Summary descriptive statistics by treatment

Variables	Male farming households					Female farming household				
	Full sample (N=364)	Adopters (n=141)	Non-adopters (n=223)	Mean difference (2-3)	t-Test (p- value)	Full sample (N=276)	Adopters (n=115)	N on- adopters (n=161)	Mean difference (3-4)	t-Test (p- value)
		1	2				3	4		
Outcome variables										
Total income from crop sales	625,763.74	865,815.60	473,982.06	391,833.69	19.369***	576,228.26	741,913.04	457,881.99	284,031.06	42.869***
	-267,935.49	-290,367.15	-66,887.63			-150,378.15	-76,108.71	-30,262.47		
Food consumption expenditure per capita	29,358.41	36,980.84	24,538.86	12,441.98	19.87***	21,723.40	23,972.23	20,117.09	3,855.15	15.80***
	-8,535.67	-5,476.87	-6,323.02			-2,757.65	-1,813.29	-2,120.48		
Other covariates										
Age	49.45(14.018)	51.81(13.309)	47.96(14.2778)	3.844	2.568***	47.21(11.219)	49.39(11.227)	45.65(10.985)	3.739	2.762***
Marital status	0.81(0.395)	0.83(0.377)	0.79 9(0.406)	0.036	0.849	0.84(0.359)	0.77(0.420)	0.90(0.300)	-1.267	-2.923***
Education	7.94(5.148)	8.96(5.295)	7.29(0.4958)	1.666	3.042***	6.95(4.949)	7.52(5.314)	6.55(4.646)	0.975	1.618*
Household size	11.28(3.662)	10.38(3.856)	11.84(3.423)	-1.467	-3.791***	8.24(2.191)	9.32(2.631)	7.47(1.369)	1.856	7.624***
Total farm size	6.21(3.232)	9.23(2.796)	4.31(1.652)	4.921	21.104***	3.37(1.123)	4.15(1.226)	2.81(0.588)	1.339	12.064***
Farming experience	28.64(11.579)	31.11(11.145)	27.08(11.601)	4.037	3.284***	21.97(7.036)	23.06(7.694)	21.19(6.439)	1.862	2.182***
Membership of farmers' union	0.68(0.466)	0.91(0.280)	0.54(0.450)	0.377	8.176***	0.62(0.485)	0.81(0.395)	0.49(0.501)	0.318	5.660***
Extension service	0.58(0.495)	0.76(0.429)	0.46(0.499)	0.297	5.827***	0.37(0.482)	0.43(0.497)	0.32(0.469)	0.103	1.757***
Credit access	0.30(0.459)	0.41(0.494)	0.23(0.421)	0.183	3.768***	0.24(0.427)	0.30(0.462)	0.19(0.396)	0.112	2.157***
Awareness of digital marketing	0.60(0.490)	1.00(0.000)	0.76(0.429)	0.651	16.145***	0.58(0.494)	0.99(0.093)	0.29(0.456)	0.699	16.197***
Access to internet service	0.85(0.356)	1.00(0.000)	0.35(0.478)	0.242	6.694***	0.73(0.448)	1.00(0.000)	0.53(0.501)	0.472	10.103***
Access to digital marketing training	0.55(0.499)	1.00(0.000)	0.75(0.429)	0.416	0.847***	0.43(0.496)	0.68(0.472)	0.25(0.437)	0.415	7.517***

The t-test was carried out to test for difference in outcome and other covariates between of adopters and non-adopters of digital marketing platforms;

*, **, ***represent statistically significant at 10, 5 and 1%, respectively.

Source: Field survey (2022)

In light of such systematic differences, it will be difficult to causally attribute the observed differences in the outcome variables as shown in Figures 1-4 to adoption of digital marketing. The differences in outcome variables (total farm sales and food consumption expenditure) within and between them could well be due to the differences in the observed covariates included in the model (Table 2).

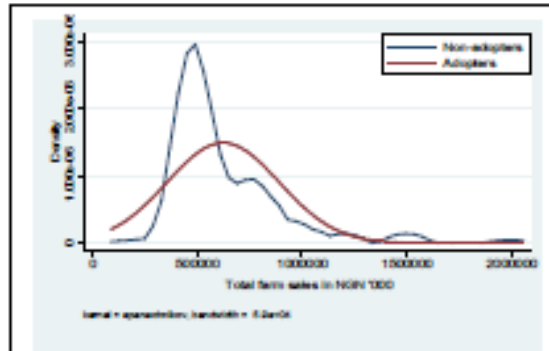


Figure 1: Kernel density estimate of the log of total farm sales in NGN '000 of adopters and non-adopters among male farming households

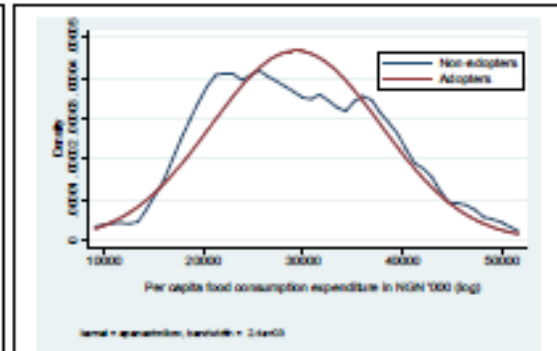


Figure 2: Kernel density estimate of the log of per capita food consumption expenditure for adopters and non-adopters among male farming households

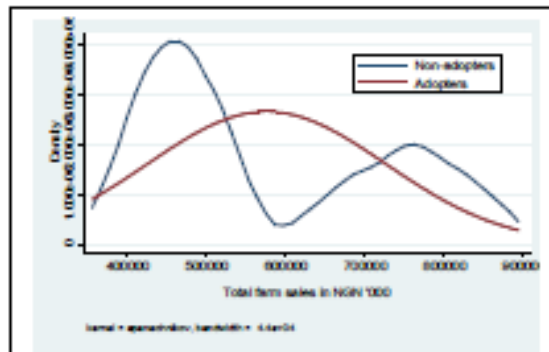


Figure 3: Kernel density estimate of the log of total farm sales in NGN '000 of adopters and non-adopters among female farming households

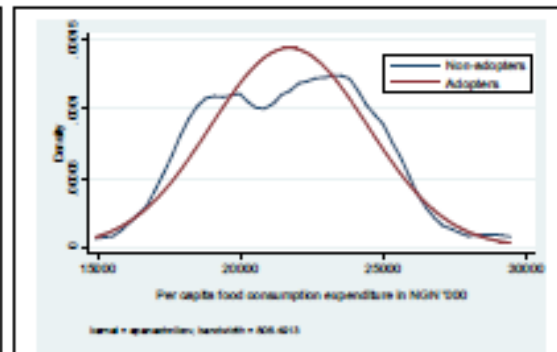


Figure 4: Kernel density estimate of the log of per capita food consumption expenditure for adopters and non-adopters among female farming households

Determinants of adoption of digital marketing platforms among farming households

Using the Probit model, the results of the main determinants of adoption of digital marketing platforms among male and female farming households are presented in Table 3. The coefficient and marginal effect from the maximum likelihood estimates of the probit model were used in explaining the influence of the socioeconomic variables of male and female farming households on their decision to adopt or not the digital marketing platforms. The marginal effect, on the other hand, is more effective than the coefficient to characterize the magnitude of a probabilistic model. This is because the sign and magnitude of the marginal effect determine the direction and magnitude of the likelihood impact of socioeconomic characteristics on their decision to adopt or not (Lokshin, 2004). The marginal effect is therefore, used in explaining the results

of the model. The Pseudo-R², statistical significance of the log-likelihood (LR) ratio, and chi-square probability in Table 3 showed that the specified model is compatible and best fit for the model used in the study. Table 3 show that seven of the twelve covariates included in the model are statistically significant in explaining the decision of male farming household heads to adopt digital marketing platforms. These include total farm size, membership of association, access to extension services and credit, awareness of online digital marketing, access to internet, and training. While among female farming households sampled only six of the twelve covariates are statistically significant in explaining the decision of household to adopt or not. Some of these variables are marital status, total farm size, membership of association, awareness of online marketing, access to internet and training.

The total farm size coefficient is positively and statistically significant and explains the likelihood of digital marketing adoption among the farming households. The marginal effect of a unit increase in total farm size on the conditional probability of digital marketing platforms adoption are 0.232 and 0.346 for male and female farming households, which implies that an additional increase in total farm size will lead to a 23.2 and 34.6 percent increase in digital marketing platforms adoption among both genders. One possible explanation is that most households either male or female with large farm size are more like to adopt digital marketing platforms than those with small farm size and which may be due to other institutional supports getting from government. Similar to our findings, size of cultivated land was found to positively affect agricultural technology use in Ghana (Nata et al., 2014), and Makate et al. (2019) opined that size of land owned by a farmer is also found to have a positive effect on adoption of multiple Climate Smart Agricultural (CSA) innovations in Southern Africa. Membership of farmers association increases farming households' likelihood of adopting digital marketing platforms. Similarly, membership of association is positive and significantly influenced the likelihood of both male and female households' decision to adopt digital marketing platforms. The marginal effect shows that being a member of association increases the probability of digital marketing platforms adoption by 34.7 and 27.1 percent respectively, in male and female farming households. A plausible reason for this could be improved access to information and knowledge sharing among members in these associations. These findings are in line with previous studies (Daadi and Latacz-Lohmann, 2021).

Table 3: Determinants of adoption of digital marketing platforms

Variable	Male farming households				Female farming households			
	Probit regression		Marginal effects		Probit regression		Marginal effects	
	Coefficient	Std. error	dy/dx	Std. error	Coefficient	Std. error	dy/dx	Std. error
Age	-0.011	0.019	-0.003	0.005	0.026	0.022	0.008	0.007
Marital status	-0.827	0.596	-0.271	0.215	1.647***	0.679	0.582***	0.215
Education	-0.028	0.035	-0.008	0.01	-0.006	0.033	-0.002	0.01
Household size	-0.156	0.039	0.043	0.012	0.253	0.093	0.076	0.033
Total farm size	0.836***	0.121	0.232***	0.041	1.149***	0.262	0.346***	0.103
Farming experience	0.027	0.019	0.007	0.005	0.012	0.024	0.004	0.007
Membership of farmers' union	1.652***	0.415	0.347***	0.069	1.007***	0.366	0.271***	0.098
Extension service	0.533**	0.323	0.142**	0.082	0.061	0.325	0.018	0.099
Credit access	0.575*	0.337	0.174*	0.106	-0.263	0.369	-0.075	0.1
Awareness of digital marketing	1.138***	0.345	0.300***	0.086	3.113***	0.612	0.702***	0.066
Access to internet service	0.551*	0.446	0.130*	0.088	0.943***	0.373	0.248***	0.089
Access to digital marketing training	1.424***	0.354	0.367***	0.082	1.035***	0.321	0.321***	0.103
constant	-7.088	1.197			-10.253	1.768		

Male: Log likelihood = -47.10964; Pseudo R2 = 0.8061; LR chi2 (12) = 391.76; Prob>chi2 = 0.000; Observation = 364 Female: Log likelihood = -44.42821; Pseudo R2 = 0.7630; LR chi2 (12) = 286.06; Prob>chi2 = 0.000; Observation = 276 *, **, ***represent statistically significant at 10, 5 and 1%, respectively.

Source: Author's computation (2022).

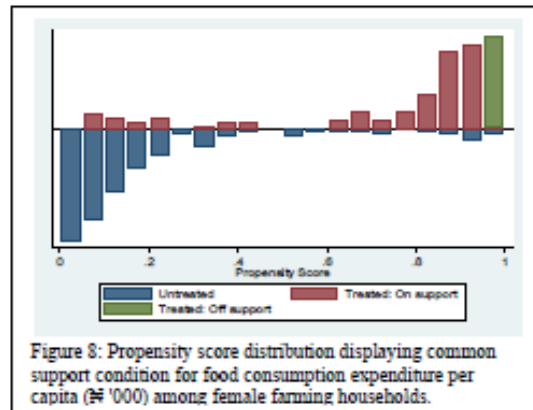
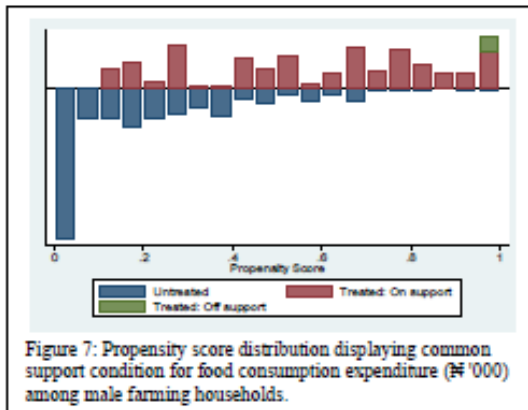
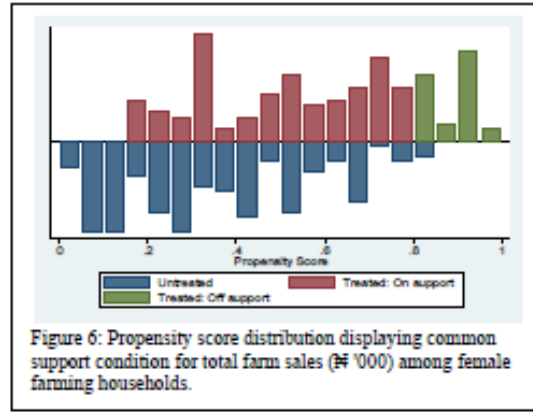
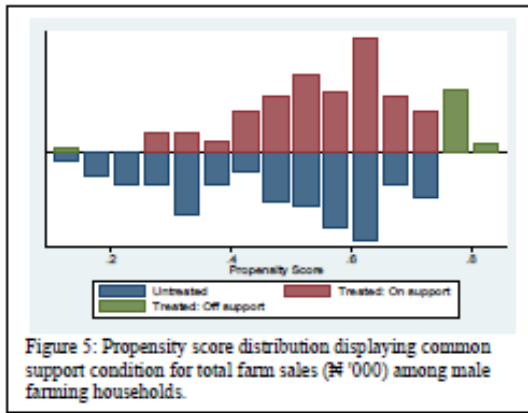
Our result showed a significant positive impact of access to extension and credit on the adoption of digital marketing platforms among male farming households and this is related to adoption studies that found extension endogenous to adoption (Wossen et al., 2017). The results suggest that households who had access to extension services and received credit were 14.1% and 17.4% more likely to adopt digital marketing platforms, respectively. Awareness had a positive and significant relationship with the decision of male and female farming household heads to adopt digital marketing platforms. This implies that an increase in enlightenment campaign and sensitization is likely to increase adoption. The marginal effects suggest that there is a probability adoption to increase by 30% and 70.2% if awareness is raised among male and female farming households respectively. This is in line with a priori expectations indicating that increased awareness is a key component driving the technology adoption decision process. Also, access to internet services and training were found to be positive and statistically significant influencing digital marketing adoption among male and female farming household heads. Having the necessary training and internet service can help boost the likelihood of digital marketing platforms adoption. This is because participation in capacity-building training pre-exposes households to opportunities in the trained online marketing.

Impact of adoption of digital marketing platforms on household welfare outcomes: Propensity score matching

In this study, we first conduct some diagnostic test before estimating the causal effect of adoption of digital marketing platforms using propensity score matching (PSM) estimator which only considers impact based on observable characteristics to ensure the quality of the propensity matches and adequacy of the PSM model specification. We checked for the extent of overall covariate balancing and overlap over the common support. By visual inspection there are substantial overlap in the distribution of the propensity scores of adopters and non-adopters. This guarantees that there are observations in the non-adopters' group that may match the group of adopters. Figures 5-8 presents the distribution of the matching propensity score and the general support condition for the propensity score estimations. The figures show that, the "treated and untreated: on support" indicates the observations in the adoption category that have an appropriate comparison, while the "treated and untreated: off support" indicates the observations in the adoption category that do not have an appropriate comparison.

We further checked for balancing property for the matching process to be sure of the common support condition quality. This was carried out to ensure that both groups of adopters and non-adopters are similar in pre-exposure characteristics after estimating the propensity score. This is in line with Austin (2009) that covariate balancing ensures whether the estimated propensity score

adequately balances observed covariates between the adopters and the comparison group. Table 4 presents the results of the balancing property check of all the covariates included in the analysis. In both male and female farming households, the results indicate that none of the covariates was significant after matching implying that the quality of matching is satisfactory for all covariates included in the model. This is consistent with previous study by Bello et al. (2021), who showed that after matching, none of the covariates included in the quality test were statistically significant. This implies that the covariate qualities of adopters and non-adopters were similar in characteristics. However, after we conditioning on the propensity scores using two conditioning methods (nearest neighbour matching, NNM and kernel-based, KBM), the test of balance in measured covariates between adopters and matched non-adopters was estimated based on three indicators (pseudo- R^2 , p-values of LR test, and mean standard bias). The results of the overall balancing test based on the three indicators are presented in Table 5. After we matched with NNM and KBM, the pseudo- R^2 reduced drastically from 80.6% before matching to 8.7% and 3.8%, respectively among male farming households, while in female farming households the pseudo- R^2 dropped from 76.3% before matching to 2.3% and 4.5% after matching. These findings show that our matching has resulted into significant reduction in systematic differences or biases in the distribution of covariates between adopters and those of matched non-adopters.



Also, the LR test for the joint significance of the covariates indicates statistically significant differences in measured covariates for unmatched samples but not in matched samples. As a result, we fail to reject our hypothesis that the distributions of the covariates are approximately the same after matching between the adopters and non-adopters among male and female farming households sampled. This implies that there is no significant difference in the distribution of covariates of matched samples of adopters and non-adopters among the two categories of farming households. From the results in Table 5, the mean standard deviation is drastically dropped from about 85.6% with the NNM to 87.4% with the KBM, with the mean bias for the overall covariates being reduced from 285.3 to 41.1 and 36.1, respectively among male farming households while among their female counterparts, the mean standard deviation reduced from about 92.2% with the NNM to 94.3% with the KBM, and the mean bias for the overall covariates dropped from 294.8 to 23.1 and 22.4. From the findings, we observed consistent results between the two matching algorithms, concluding that the matching quality is satisfactory, which invariably satisfying the conditional independence assumption and the need for proper model specification. This assumption states that, after controlling for observable covariates, the assignment of farming households to digital marketing platforms adoption is thus "as good as random," so that the possible outcomes are independent of the adoption status.

Table 4. Test of equality of means of variables before and after matching

Variable	Male Farming Households								Female Farming Households							
	Unmatched Sample				Matched Sample				Unmatched Sample				Unmatched Sample			
	Treated (mean)	Untreated (mean)	% bias	p>t	Treated (mean)	Untreated (mean)	% bias	p>t	Treated (mean)	Untreated (mean)	% bias	p>t	Treated (mean)	Untreated (mean)	% bias	p>t
Age	51.81	47.96	27.9	0.011***	51.87	46.03	42.30	0.080	49.39	45.65	33.70	0.006***	49.81	49.44	3.30	0.828
Marital status	0.83	0.79	9.2	0.054**	0.86	0.68	48.30	0.396	0.77	0.9	-34.70	0.004***	0.93	0.98	-13.40	0.311
Education	8.96	7.29	32.5	0.003***	8.08	9.38	-25.30	0.246	7.52	6.55	19.50	0.007*	7.17	9.93	-55.20	0.17
Household size	10.38	11.84	-40.2	0.000***	10.87	11.08	-5.90	0.834	9.32	7.47	88.50	0.000***	8.34	7.66	32.60	0.65
Total farm size	9.23	4.31	214.3	0.000***	6.6	6.49	5.10	0.734	4.15	2.81	139.30	0.000***	3.28	3.2	8.10	0.579
Farming experience	31.11	27.08	35.5	0.001***	30.65	26.14	39.70	0.391	23.06	21.19	26.20	0.030***	23.09	28.34	-73.90	0.201
Membership of farmers' union	0.91	0.54	93	0.000***	0.78	0.86	-20.00	0.366	0.81	0.49	70.40	0.000***	0.63	0.78	-32.40	0.149
Extension service	0.76	0.46	63.8	0.000***	0.73	0.54	40.60	0.093	0.43	0.32	21.30	0.080**	0.39	0.54	-30.30	0.188
Credit access	0.41	0.23	39.8	0.000***	0.41	0.38	5.90	0.815	0.3	0.19	26.00	0.032***	0.29	0.15	34.00	0.112
Awareness of digital marketing	0.84	0.35	116.3	0.000***	0.78	0.78	0.00	1.000	0.99	0.29	212.50	0.000***	0.98	0.98	0.00	1
Access to internet service	0.9	0.76	38.6	0.001***	0.89	0.89	0.00	1.000	0.85	0.52	74.60	0.000***	0.78	0.83	-11.20	0.583
Access to digital marketing training	0.8	0.39	93.2	0.000***	0.7	0.76	-12.10	0.606	0.67	0.25	91.20	0.000***	0.51	0.61	-21.40	0.38

Source: Author's computation (2022). Note: *, ** and *** represent, respectively, statistical significance at the 0.1, 0.05, and 0.01 level.

Table 5. Overall quality test of the propensity scores under different matching algorithms

Status	Matching method	Male Farming Households						Female Farming Households					
		Pseudo R ²	LR χ^2	p>(χ^2)	Mean Standard bias	Bias	Total % mean bias reduction	Pseudo R ²	LR χ^2	p>(χ^2)	Mean Standard bias	Bias	Total % mean bias reduction
Unmatched		0.806	391.76	0.000***	67.00	285.30		0.763	286.06	0.000***	69.80	294.80	
Matched	NNM	0.087	8.92	0.709	20.40	41.10	85.60	0.023	6.13	0.291	26.30	23.10	92.20
	KBM	0.038	6.55	0.319	29.50	36.10	87.40	0.045	13.72	0.336	22.40	16.72	94.30

Source: Author's computation (2022). Note: ***significance level at 1%.

The estimates of the average adoption effects estimated using the PSM based on the nearest neighbor (NN) and kernel (KB) matching algorithms are presented in Table 6. Based on the NN matching algorithm, the causal effect of digital marketing platforms adoption on farm productivity is statistically significant and equal to 398,540.541 and 266,902.439 for male and female farming households, respectively indicating the average difference in total farm sales per hectare between adopters and non-adopters.

This indicates that 398,540.541 and 266,902.439 ₦/ha are the average treatment effect on treated (ATT) among male and female farming households sampled. Assuming that a household is selected at random, the average treatment effect (ATE) is 346,877.700 and 286,163.100 for male and female farming households, respectively, implying that the total farm sales increases by 346,877.700 and 286,163.100 ₦/ha. In term of food consumption expenditure, the average treatment effect on treated (ATT) for male is ₦11,113.022 and ₦5,119.195 for female farming households. This shows a significant difference in food consumption expenditure per capita between male and female farming households. For instance, if a household is chosen randomly, the average treatment effect (ATE) would be ₦12,736.210 and ₦3,958.790 for male and female farming households' per capita food expenditure, suggesting that the food consumption expenditure per capita would increase by ₦12,736.210 and ₦3,958.790 respectively. Using the KB matching algorithm, results in Table 6 also show a positive and statistically significant effect on outcome variables. The results indicate that the ATT for the adopters is 535,789.230 for male farming households and 269,250.128 female farming households, which signifies that their adoption of digital marketing platforms contributed substantially to total farm sales with 535,789.230 and 269,250.128 ₦/ha. This result shows a significant difference in total farm sales between the two groups of respondents. Meanwhile, the ATE are 441,379.900 and 255,397.000 ₦/ha total farm sales per capita for a male and female household. For food consumption expenditure, the average treatment effect on treated (ATT) are 6,252.363 and 4,973.529 for male and female farming households, respectively suggesting a significant average difference in per capita food consumption expenditure between them. Supposing a household male or female is chosen at random the ATE is 7,880.792 for male and 5,361.338 for female farming households, which indicating a significant contribution to their per capita food consumption expenditure.

Estimates of endogenous switching regression (ESR) model

We employed Endogenous Switching Regression (ESR) model to complement the propensity score matching (PSM) to verify the robust estimate of the causal effect of the digital marketing platforms adoption. The results of the PSM could

be biased due to its inability to account for unobserved factors that may influence household decision to adopt digital marketing platforms. Therefore, ESR is best fit for controlling unobserved selection bias in this study. Following Di Falco et al. (2011) and Ahimbisibwe et al. (2020), we added two information-related instruments in the selection equation. These instruments include (1) access to extension services and (2) access to digital marketing training. A household decision to adopt or not digital marketing platforms could depend on information received through adequate extension services and knowledge acquired through training. Therefore, we assumed that these instruments tend to influence households' decision to adopt digital marketing platforms (relevance criterion) but are unlikely to have a direct effect on welfare outcomes (exogeneity criterion). With a simple falsification test according to Falco et al. (2011) and Ahimbisibwe et al. (2020), the validity of these instruments can be assessed. Table 7 presents the results of full information maximum likelihood estimates of ESR model.

Table 6. Propensity score algorithms: impact of adoption of digital marketing platforms on welfare outcomes

Variable	Parameters	Male Farming Households					Female Farming Households				
		Adopters	Non-adopters	Difference	S.E.	T-stat	Adopters	Non-adopters	Difference	S.E.	T-stat
Nearest Neighbour Matching (NNM)											
Total farm sales per hectare (₦'000/ha)	Unmatched	865,815.603	473,982.063	391,833.540	20,229.70	19.37***	741,913.043	457,881.988	284,031.06	6,625.55	42.87***
	ATT	867,837.838	469,297.297	398,540.541	61,086.39	6.52***	738,439.024	471,536.585	266,902.44	17,060.84	15.64***
	ATU	473,982.063	356,231.40	117,750.663			457,881.988	287,083.700	170,798.29		
	ATE			346,877.700					286,163.10		
Food consumption expenditure per capita (₦ '000)	Unmatched	36,980.836	24,538.86	12,441.979	646.633	19.24***	23,972.238	20,117.090	3,855.15	243.994	15.80***
	ATT	36,807.414	25,694.39	11,113.022	2135.928	5.20***	24,095.237	18,976.042	15952.662	810.176	6.32***
	ATU	24,538.857	13,825.07	10,713.787			20,117.090	4,164.428	3,958.79		
	ATE			12,736.210							
Kernel-based Matching (KBM)											
Total farm sales per hectare (₦/ha)	Unmatched	865,815.603	473,982.063	391,833.540	20,229.695	19.37***	741,913.043	457,881.988	284,031.056	6,625.55	42.87***
	ATT	845,857.143	310,067.913	535,789.230	46,976.699	11.41***	740,270.270	471,020.142	269,250.128	20,361.29	13.22***
	ATU	639,725.600	316,157.100	323,568.500			270,459.200	244,825.400	25,633.800		
	ATE			441,379.900					255,397.000		
Food consumption expenditure per capita (₦ '000)	Unmatched	36,980.836	24,538.857	12,441.979	646.6330	19.24***	23,972.238	20,117.090	3,855.148	243.994	15.80***
	ATT	37,044.719	30,792.356	6,252.363	3,562.1980	1.76***	23,940.858	18,967.329	4,973.529	1,405.04	3.54***
	ATU	10,715.490	3,390.804	7,324.686			5,549.156	5,093.738	455.418		
	ATE			7,880.792					5,361.338		

Source: Authors, 2022. ***represent statistical significance at 1%. S.E denote standard errors

Table 8. Endogenous treatment-effects estimation: ATE and ATT

Outcome variables	ATE			ATT			ATE			ATT		
	Coefficient	Standard Error	Estimated t-value	Coefficient	Standard Error	Estimated t-value	Coefficient	Standard Error	Estimated t-value	Coefficient	Standard Error	Estimated t-value
Total farm sales per hectare (₦/ha)	6137.487***	1310.117	4.68	5909.842***	1639.261	3.610	1324.25***	1502.459	0.88	1242.716***	1417.9	0.381
Food consumption expenditure per capita (₦ '000)	9509.299***	4027.711	2.36	8736.639***	3032.639	2.880	4054.365***	3330.097	1.22	1960.343***	4997.157	0.39

Source: Author's computation, 2022. ***represent statistical significance at 1%.

Among the male farming households, the results show that the instruments are jointly statistically significant in the selection equation ($\chi^2 = 1.88$; $p < 0.01$) and ($\chi^2 = 0.46$; $p < 0.1$), for total farm sales and food consumption expenditure per capita, respectively. While among female farming households, results also indicate that the instruments are jointly statistically significant in the selection equation ($\chi^2 = 6.11$; $p < 0.01$) and ($\chi^2 = 0.84$; $p < 0.01$), for total farm sales and food consumption expenditure per capita. We further checked by examining the statistical significance of the instruments in the selection equation, also known as the first-stage equation of the ESR model in both categories of sampled respondents. Hence, our results show that the parameter estimates of these two instruments included in the model are both statistically significant in the selection equation (Table 7), implying that the assumption of the relevance of the instruments hold. Following the exogeneity hypothesis, the effect of the instruments on the probability of farming households adopting digital marketing platforms has an indirect impact on welfare outcomes. Since households' responses to the questions about the instruments were only collected through surveys asking household heads if they had access to extension services and digital marketing training in the year preceding the survey, there is less chance of endogenous development, even though this hypothesis cannot be tested generally.

Further results of the model diagnostics in Table 7 show that the correlation between the error terms of the outcome (total farm sales and food consumption expenditure) equations for non-adopters and the selection equation is statistically different from zero among both male and female farming households. This finding means that unobservable characteristics affecting the outcome of total farm sales and food consumption expenditure are correlated with those influencing digital marketing platforms adoption. While the statistically significant negative correlation coefficient of the error terms of the digital marketing platforms adoption and that of total farm sales and food consumption expenditure equations for non-adopters implies that the exercise of self-selection among non-adopters in both categories of farming households. Non-adopters were likely to have self-selected themselves out of adoption because they may not have perceived to benefit from adoption. This result suggests that non-adopters have increase farm sales and higher food consumption expenditure than it would have been the case for a population of households who are assigned at random to non-adoption status. The Wald test of independent equations rejects the null hypothesis of joint independence, suggesting joint dependence between the selection, total farm sales, and food consumption expenditure equations for adopters and non-adopters. In line with our specification, we have two distinct regimes rather than one, justifying the use of the ESR model by providing evidence of the appropriateness of the assumption that the effects of covariates between the two regimes (with and without adoption) are significantly different. This is apparent in the variation in the second-stage parameter estimates of the control variables in the outcome

equations of the ESR model between digital marketing platforms adopters and non-adopters of male and female farming households presented in the Table 7. The results show noticeable differences in some of the parameter estimates of the outcomes equations between the digital marketing platforms adopters and non-adopters, indicating the presence of some heterogeneity in the sample. For instance, the age of the farming household heads, and access to credit among male farming households and age, household size, and access to internet service among female farming households are significantly associated with total farm sales for adopters but not for non-adopters among male farming households. In contrast, marital status, household size, membership of association, awareness, and access to internet service among male farming households and only total farm size among female farming are statistically associated with total farm sales for non-adopters but, not for adopters.

Table 7: Full information maximum likelihood (FIML) estimates of the endogenous switching regression (ESR) model of adoption of digital marketing platforms and welfare outcomes

Variable	Male farming households						Female farming households					
	Selection Equation of Adoption	Outcome equation of Total Farm sales		Selection Equation of Adoption	Outcome equation of Food consumption expenditure per capita		Selection Equation of Adoption	Outcome equation of Total Farm sales		Selection Equation of Adoption	Outcome equation of Food consumption expenditure per capita	
		Adopters	Non-adopters		Adopters	Non-adopters		Adopters	Non-adopters		Adopters	Non-adopters
Age	-0.006	-4655.39**	19.30 (565.85)	-0.011 (0.019)	-3957.96	18.07 (20.69)	0.024 (0.021)	2079.76**	-52.68	0.015 (0.022)	127.93	115.97
	-0.019	-2579.04			-0.313			-949.28	-381.22		-183.13	-182.46
Marital status	-0.879	84644.17	26024.24*	-0.833 (0.599)	1112.09	1143.99**	-1.689***	-7656.07	-447.81	-1.529**	1003.19	608.33
	-0.601	-79207.24	-14567.02		-1510.29	-533.33	-0.644	-22402.48	-9421.93	-0.652	-4911.26	-4879.6
Education	-0.016	2567.85	399.96 (1010.56)	-0.031 (0.036)	-71.05	70.36**	-0.010 (0.034)	879.34	175.23	-0.008	-193.54	-200.49
	-0.036	-4687.16			-89.31	-37.27		-1445.29	-540.59	-0.032	-302.9	-302.05
Household size	-0.152***	3886.86	-5666.09***	-0.157***	-17.78	-12.19	0.260***	6946.20**	-368.46	0.249***	218.68*	296.71
	-0.041	-6612.95	-1254.5	-0.04	-126.15	-45.92	-0.093	-2877.92	-1822.28	-0.094	-754.33	-747.49
Total farm size	0.864***	6249.1	1200.80 (2794.89)	0.841*** (0.121)	-117.5	204.96*	1.293***	8049.03	-7537.95**	1.142***	434.4	750.07
	-0.124	-9946.99			-182.98	-115.23	-0.268	-5965.25	-4228.39	-0.262	-1768.57	-1729.37
Farming experience	0.029 (0.020)	18.23 (2500.45)	57.81 (595.56)	0.027 (0.020)	3.77(47.6)	-15.21(21.92)	0.022 (0.06)	541.99 (993.4)	8.07 (644.79)	0.016 (0.024)	69.7 (252.6)	70.1 (252.1)
Membership of farmers' union	1.506***	-106282.5	20741.63**	1.663*** (0.419)	2026.51	-368.95	1.121***	-7385.95	2674.8	0.945**	2754.14**	3022.74
	-0.423	-83569.88	-9927.21		-1604.77	-364.73	-0.361	-17509.39	-5117.27	-0.369	-3176.31	-3154.5
Extension service	0.496(0.333)	54681.95	3050.79 (8751.23)	0.532 (0.327)	2170.98**	41.26	-0.023 (0.322)	5255.02	1553.52	-0.006	-278.33	0.209**
		-54242.22			-1034.59	-320.34		-13635.8	-5123.76	-0.326	-2903.43	-2103.06
Credit access	0.553*	167469.95***	-8822.02	0.586* (0.337)	-910.79	354.03	-0.566 (0.384)	-10817.36	3844.18	-0.235	-4696.88	3644.31**
	-0.337	-46042.45	-10824.11		-877.94	-397.35		-14877.52	-6217.33	-0.362	-3289.65	-2272.38

Variable	Male farming households						Female farming households					
	Selection Equation of Adoption	Outcome equation of Total Farm sales		Selection Equation of Adoption	Outcome equation of Food consumption expenditure per capita		Selection Equation of Adoption	Outcome equation of Total Farm sales		Selection Equation of Adoption	Outcome equation of Food consumption expenditure per capita	
		Adopters	Non-adopters		Adopters	Non-adopters		Adopters	Non-adopters		Adopters	Non-adopters
Awareness of digital marketing	1.032*** -0.354	-17724.82 -74644.87	15642.16* -9256.29	1.099*** (0.348)	-1204.88 -1411.65	110.52 -351.7	3.144*** -0.586	51513.45 -73518.84	422.63 -6422.2	3.091*** -0.617	4079.52 -0.27	5144.07* -5017.67
Access to internet service	0.710(0.469)	-127568.8 -86146.9	-19199.92** -9941.66	0.624 (0.454)	1298.22 -1643.04	683.58** -367.38	0.898** -0.371	57663.05*** -18952.66	-6322.7 -4985.8	0.904** -0.367	1506.63** -3234.57	1635.62 -3107.11
Access to digital marketing training	1.42*** -0.354	63866.17 -57423.69	328.65 (9072.41)	1.449*** (0.360)	1002.1 -1097.15	474.65 -340.8	0.926*** -0.314	21203.63 -15575.3	-8788.06 -5736.27	1.009*** -0.319	-3695.3 -3415.24	3174.77* -3357.87
constant	-7.601 -1.294	1085572 -213892.4	510593.7 -28078.11	-7.140 (1.245)	39939.93 -3960	18159.78 -1088.48	10.686 (1.734)	413020.3 -95743.12	483517.6 -19149.22	-10.1 -1.756	142395.2 -10935.39	140523.1 -10710.36
rho1		-0.051(0.304)			0.24(0.33)			0.491** (0.25)			-0.29(0.30)	
rho0			0.325 (0.219)			-0.003(0.277)			0.595** (0.22)			0.23* (0.27)
LR chi2 (2)	1.88			0.46			6.11			0.84		
Wald chi2(12)	35.34***			19.06*			7.69***			7.47***		
Log likelihood	-4782.6496			-3487.694			-3372.4347			-3203.25		
Number of observations	364	141	223	364	141	223	276	115	161	276	115	161

Source: Author's computation (2022). Note: *, ** and *** represent, respectively, statistical significance at the 0.1, 0.05, and 0.01 level. Robust standard errors in parentheses.

Similarly, only access to extension services among male farming households and household size, membership of association, and access to internet service among female farming households are significantly influencing per capital food consumption expenditure for adopters but not for non-adopters. Contrarily, marital status, education level, total farm size, and access to internet service among male farming households and access to extension service, credit, awareness, and access to digital marketing training among female farming households are statistically associated with per capita food consumption expenditure for non-adopters but not for adopters. However, variables such as household size and access to internet service among male farming households and total farm size among female counterparts has a significantly negative impact on non-adopters than on adopters when it comes to per capita food consumption expenditure. Our findings corroborate previous studies by Tambo and Wünscher (2017) and Ahimbisibwe et al. (2020), who find that household size significantly lowers food consumption expenditure for both innovators and non-innovators, with the effects on innovators being more pronounced than on non-innovators. Also, the results from the first-stage estimation of the ESR model presented in Table 7 for all outcome variables show that marital status, household size, total farm size, membership of association, access to credit, awareness and digital marketing training among male and female farming households are significantly associated with digital marketing platforms adoption. Specifically, households who have access to information through awareness and credit, internet facilities and access to digital training are more likely to adopt digital marketing platforms.

Impact of adoption of digital marketing platforms on welfare outcomes using Endogenous Switching Regression (ESR) model

As shown in Table 8 the ATE and ATT are nearly the same. The results of ATE shows that return of adoption of digital marketing platforms significantly ($p < 0.01$) increased by ₦6137.487 and ₦9509.299, respectively, for outcomes of total farm sales and food consumption expenditure per capita among male farming households while, among their female counterparts the ATE indicates that impact of digital marketing platforms adoption statistically significant ($p < 0.01$) and increased by ₦1324.25 and ₦4054.365 for total farm sales and food consumption expenditure per capita among female farming households. The ATT shows that the impact on welfare variables of the treated farming households has increased significantly by ₦5909.842 and ₦8736.639, respectively, for the outcome variables of total farm sales and food consumption expenditure per capita among male farming households. However, among the female farming households, the ATT also indicates that the welfare effect of treated households significantly increased by ₦1417.9 and ₦4997.157 for total farm sales and food consumption expenditure per capita. This result is consistent with previous studies on the adoption of innovation platforms; for instance, Ahimbisibwe et al. (2020), shows that the adoption of agricultural innovation platforms positively influences households'

welfare in Uganda.

Conclusion

Using a cross-sectional farming household survey conducted in northcentral Nigeria, the study has sought to investigate if adoption of digital marketing platforms has led to improved farming household welfare (measured by total farm sales and food consumption expenditure) and close gender welfare gap in Nigeria. We first assessed the determinants of adopting digital marketing platforms and found that total farm size, membership of association, access to extension services, and credit, awareness, access to internet, and digital marketing training tend to positively and significantly influence digital marketing platforms adoption among male farming households, while marital status, total farm size, membership of association, awareness, access to internet service and digital marketing training are statistically associated with female farming households decision to adopt digital marketing platforms. Similarly, using a robust ESR model, we found that household size, total farm size, membership of association, access to credit, awareness, and access to digital marketing training are significantly associated with the welfare outcomes of both adopters and non-adopters among male farming households. While among their female counterparts, marital status, household size, membership of association, awareness, access to internet and digital training are significantly associated with welfare outcomes of adopters and non-adopters. The significant difference of the identified factors should be upgraded to increase the assessment of each identified variable. Therefore, interventions that are aimed at improving digital marketing platforms adoption among rural farming households and close gender gap in households' welfare may actually have significant farm sales and higher food consumption expenditure for households. In terms of impact, we found consistently positive and statistically significant effect of adoption of digital marketing platforms on household welfare. Specifically, adoption of digital marketing platforms has increased farm sales and per capita food consumption expenditure among male farming households compared to female farming households. This indicates that, if all the male and female farming households in a given region adopted digital marketing platforms, their total farm sales and per capita food consumption expenditure would increase significantly. This result suggests the potential of digital marketing platforms to achieve improved households' welfare and close gender gap in standard of living outcomes, pointing to the need for reaching the current non-adopters. Finally, the main limitation of this study concerns the sample representing one out of six regions in Nigeria that does not reflect the country farming households' population, thus preventing the generalizability of the results. However, due to relevancy of this topic nowadays and deserves further analysis, we hope that this study can be of inspiration for future investigation that will reasonably focus on a representative sample of farming households in entire regions of Nigeria.

References

- Abdoulaye, T.; Wossen, T.; Awotide, B. (2018). Impacts of improved maize varieties in Nigeria: Ex-post assessment of productivity and welfare outcomes. *Food Security*, 10, 369–379.
- Ahmed, M.H.; Geleta, K.M.; Tazeze, A.; Andualem, E. (2017). The impact of improved maize varieties on farm productivity and wellbeing: evidence from the east hararghe zone of Ethiopia. *Develop Stud Res.* 2017;4(1):9–21.
- Ahimbisibwe, B.P., Morton, J.F., Feleke, S, et al. (2020). Household welfare impacts of an agricultural innovation platform in Uganda. *Food Energy Security.* 2020; 9: e225. <https://doi.org/10.1002/fes3.225>
- Alford, P.; Page, S.J. (2015). “Marketing technology for adoption by small business”. *The Service Industries Journal* 35.11–12 (2015):655– 669. [10.1080/02642069.2015.1062884](https://doi.org/10.1080/02642069.2015.1062884).
- Alliance for Affordability Internet (2019). *It helps me escape-2019 affordability report Nigeria impact story*. Retrieved from <https://a4ai.org/it-helps-me-escape-2019-affordability-reportnigeria-impact-story/>.
- Asfaw, S., Shiferaw, B., Simtowe, F., & Lipper, L. (2012). Impact of modern agricultural technologies on smallholder welfare: Evidence from Tanzania and Ethiopia. *Food Policy*, 37(3), 283–295. doi:10.1016/j.foodpol.2012.02.013
- Awotide, B. A.; Alene, A. D.; Abdoulaye, T.; Manyong, V. M. (2015). Impact of agricultural technology adoption on asset ownership: the case of improved cassava varieties in Nigeria. *Food Sec.* 7, 1239–1258. doi: 10.1007/s12571-015-0500-7
- Balu. (2020). Awareness Level of Farm Youth on Information and Communication Technology (ICT) Tools, *International Journal of Inclusive Development* Vol: 6 (02): pp.87-90.
- Becerril, J., & Abdulai, A. (2010). The impact of improved maize varieties on poverty in Mexico: A propensity score-matching approach. *World Development*, 38(7), 1024– 1035. doi:10.1016/j.worlddev.2009.11.017
- Boehlje, M.; Langemeier, M. (2021). The Role of Information in Today’s Uncertain Business Climate. 2021. Available online: <https://ag.purdue.edu/commercialag/home/resource/2021/02/the-value-of-data-information-and-the-payoff-of-precision-farming/> (accessed on 15 December 2021).
- Buabeng-Andoh, C.; Baah, C. (2020). Pre-service teachers’ intention to use learning management system: an integration of UTAUT and TAM. *Interactive Technology and Smart Education*.
- Caliendo, M.; Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *J. Econ. Surv.*, 22, 31–72.
- Chapagain, T.; Raizada, M.N. (2017). Agronomic challenges and opportunities for smallholder terrace agriculture in developing countries. *Front. Plant Sci.* 8, 1–15. <https://doi.org/10.3389/fpls.2017.00331>.

- Daadi, B.E.; Latacz-Lohmann, U. (2021). Organic Fertilizer Adoption, Household Food Access, and Gender-Based Farm Labor Use: Empirical Insights from Northern Ghana. *J. Agric. Appl. Econ.*, 53, 435–458.
- Davis, F.D., (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quart.* 319–340.
- Davis, F. D. (2007, December). Acceptance of Information Technology: Research Progress, Current Controversies, and Emerging Paradigms. In Workshop on HCI Research in MIS.
- Deichmann, U.; Goyal, A.; Mishra, D. (2016). *Will digital technologies transform agriculture in developing countries?* The World Bank.
- Diaz, A.C.; Sasaki, N.; Tsusaka, T.W.; Szabo, S. (2021). Factors affecting farmers' willingness to adopt a mobile app in the marketing of bamboo products. *Resources, Conservation & Recycling Advances* 11 (2021) 200056. <https://doi.org/10.1016/j.rcradv.2021.200056>
- Di Falco, S. (2014). Adaptation to climate change in Sub-Saharan agriculture: assessing the evidence and rethinking the drivers. *European Rev Agric Econ.* 2014;41(3):405–30.
- Dyer, J. (2016). *The Data Farm. An Investigation of the Implications of Collecting Data on Farm*; n 1506; Nuffield Australia Farming Scholars: North Sydney, Australia, September 2016.
- Engotoit, B.; Kituyi, G. M.; Moya, M. B. (2016). Influence of performance expectancy on commercial farmers' intention to use mobile-based communication technologies for agricultural market information dissemination in Uganda. *Journal of Systems and Information Technology*.
- Fasakin, I.J.; Ogunniyi, A.I.; Bello, L.O.; Mignouna, D.; Adeoti, R.; Bamba, Z.; Abdoulaye, T.; Awotide, B.A. (2022). Impact of Intensive Youth Participation in Agriculture on Rural Households' Revenue: Evidence from Rice Farming Households in Nigeria. *Agriculture*, 12, 584. <https://doi.org/10.3390/agriculture12050584>
- Fainos, C.; Tendai, N.; Munodawafa, N. (2014). The role of agricultural marketing on empowering rural farmers in Masvingo Province, Zimbabwe. *European Journal of Business and Management*, 6(3), 1-11.
- Fecke W.; Danne M.; Musshoff, O. (2018): E-commerce in agriculture: the case of crop protection product purchases in a discrete choice experiment. *Computers and Electronics in Agriculture*, 151: 126–135.
- Feleke, S., Manyong, V., Abdoulaye, T., & Alene, A. (2016). Assessing the impacts of cassava technology on poverty reduction in Africa. *Studies in Agricultural Economics*, 118, 101–111. <https://doi.org/10.7896/j.1612>
- Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of technology adoption. *Annual Review of Economics*, 2(1), 395–424.
- Galati, A.; Crescimanno, M.; Tinervia, et al. (2017). Social media as a strategic marketing tool in the Sicilian wine industry: Evidence from Facebook. *Wine Economics and Policy*, 6 (1), 40-47.

- Gautam, S.; Schreinemachers, P.; Uddin, N. (2017). Impact of training vegetable farmers in Bangladesh in integrated pest management (IPM). *Crop Prot.*, 102, 161–169.
- GSMA. (2015). “Bridging the Gender Gap: Mobile access and usage in low- and middle income countries”; Intel, 2012, “Women and the Web”.
- Heckman, J.; Lalonde, R.; Smith, J. (1999). The economics and econometrics of active labour market programs. In O. Ashenfelter & D. Card (Eds.) *Handbook of Labor Economics*, Vol. III, Amsterdam: Elsevier.
- Hennessey T.; Läpple D.; Moran, B. (2016): The digital divide in farming: a problem of access or engagement? *Applied Economic Perspectives and Policy*, 38: 474–491.
- Imbens, G. W., & Wooldridge, J. W. (2009). Recent developments in the econometrics of program development. *Journal of Economic Literature*, 47(1), 5–86.
- Jeong, B.K., Yoon, T.E., (2013). An empirical investigation on consumer acceptance of mobile banking services. *Bus. Manag. Res.* 2, 31–40. <https://doi.org/10.5430/bmr.v2n1p31>.
- Kabbiri, R.; Dora, M.; Kumar, V.; Elepu, G.; Gellynyk, X. (2018). Mobile Phone Adoption in Agri-Food Sector: Are Farmers in Sub-Saharan Africa Connected? Available at <https://www.sciencedirect.com/science/article/pii/S0040162517317894>.
- Kapfudzaruwa, F.; Kudo, S.; Mfuno, O.; Hansen, M.; Nyerere, J. (2018). *Rural-Urban Linkages and Sustainable Development in Africa*, Spears Media Press.
- Kassie M, et al. (2017). Land Use Policy Push–pull farming system in Kenya: implications for economic and social welfare. *Land*. 2017;77: 186–98.
- Khonje, M., Manda, J., Alene, A. D., & Kassie, M. (2015). Analysis of adoption and impacts of improved maize varieties in Eastern Zambia. *World Development*, 66, 695–706. <https://doi.org/10.1016/j.worlddev.2014.09.008>
- Lixi, M. J. Y., Zottel, S. & Neto, M. I. A. (2019). Nigeria Digital Economy Diagnostic Report. The World Bank.
- Lokshin, M. and Sajaia, Z. (2004). Maximum Likelihood Estimation of Endogenous Switching Regression Models. *Stata Journal* 4(3): 282–289.
- Lowder, S.K.; Scoet, J.; Raney, T. (2016). The number, size, and distribution of farms, smallholder farms, and family farms worldwide. *World Dev.* 87,16–29. <https://doi.org/10.1016/j.worlddev.2015.10.041>.
- Ma, Wangling, & Abdulai, A. (2016). Does cooperative membership improve household welfare? Evidence from apple farmers in China. *Food Policy* 58, 94–102.
- Maddala, G. S. (1983). *Limited-dependent and qualitative variables in econometrics*. Cambridge, UK: CUP. <https://doi.org/10.1017/CBO9780511810176>
- Ma W.; Zhou, X.; Liu, M. (2019): What drives farmers’ willingness to adopt e-commerce in rural China? The role of Internet use. *Agribusiness*, 36: 1–5.
- Makate, C.; Makate, M.; Mango, N; Siziba, S. (2019). Increasing resilience of smallholder farmers to climate change through multiple adoption of proven climate-smart agriculture innovations. Lessons from Southern

- Africa. *J Environ Manage.* 231:858– 68.
- Manda, J., Alene, A., Tufa, A., Abdoulaye, T., Wossen, T., Chikoye, D., & Manyong, V. (2019). The poverty impacts of improved cowpea varieties in Nigeria: A counterfactual analysis. *World Development*, 122, 261–271. <https://doi.org/10.1016/j.worlddev.2019.05.027>
- Mbagwu, F. C; Benson, V.; Onuoha, C. O. (2018). *Challenges of meeting information needs of rural farmers through internet-based services: experiences from developing countries in Africa*. A conference paper presented at the World Library and Information Congress: 84th International Federation of Library Association (IFLA) conference, Malaysia. Retrieved from https://www.researchgate.net/publication/333718265_Challenges_of_meeting_information_needs_of_rural_farmers_through_internetbased_services_experiences_from_developing_countries_in_Africa.
- McFadden, D. (1974). The measurement of urban travel demand. *J. Public Econ.*, 3, 303–328.
- Misaki, E.; Apiola, M.; Gaiani, S.; Tedre, M. (2018). Challenges facing sub-Saharan smallscale farmers in accessing farming information through mobile phones: A systematic literature review. *Electron.J. Inf.Syst.Dev. Ctries*.84,1–12. <https://doi.org/10.1002/isd2.12034>.
- Mobile Report (2019). *Key Nigerian facts country's economy showed positive growth*. Retrieved from <https://www.jumia.com.ng/mobile-report/>.
- Munoz-Leiva, F.; Climent-Climent, S.; Li´ ebana-Cabanillas, F. (2017). Determinantes de la intencion ´ de uso de las aplicaciones de banca para moviles:´una extension ´ del modelo TAM clasico. *Spanish J.Mark*.21,25–38. <https://doi.org/10.1016/j.sjme.2016.12.001>.
- Nata,J.T.,Mjelde,J.W., Boadu, F.O (2014). Household adoption of soil-improving practices and food insecurity in Ghana. *Agric Food Security*. 3(1):17.
- Ogbeide-Osaretin, E.N.; Ebhote, O. (2020). Does digital marketing enhance rural agricultural transformation in Nigeria? An empirical investigation. *Asian Journal of Agriculture and Rural Development*, 10(1), 450-462.
- Okumus, B.; Ali, F.; Bilgihan, A.; Ozturk, A. B. (2018). Psychological factors influencing customers’ acceptance of smartphone diet apps when ordering food at restaurants. *International Journal of Hospitality Management*, 72, 67-77.
- Oyetunde-Usman, Z., Ogunpaimo, O.R., Olagunju, K.O., Ambali, O.I., & Ashagidigbi, W.M. (2021). Welfare Impact of Organic Fertilizer Adoption: Empirical Evidence From Nigeria. *Front. Sustain. Food Syst*. 5:691667. doi: 10.3389/fsufs.2021.691667
- Petcho, W.; Szabo, S.; Kusakabe, K.; Yukongdi, V. (2019). Farmers’ perception and drivers of membership in rice production community enterprises: evidence from the central region, Thailand. *Sustainability* 11, 5445. <https://doi.org/10.3390/su11195445>.
- Ragasa, C. (2012). *Gender and Institutional Dimensions of Agricultural Technology Adoption: A Review of Literature and Synthesis of 35 Case Studies*. Retrieved from <http://purl.umn.edu/126747>

- Ramli, S. A.; Samah, B. A.; Hassan, M. S.; Omar, S. Z.; Bolong, J.; Shaffri, H. A. M. (2015). Potential benefits of ICT for youth agro-based entrepreneurs in Malaysia. *Journal of Applied Sciences*, 15(3), 411-414.
- Ricciardi, V.; Ramankutty, N.; Mehrabi, Z.; Jarvis, I.; Chookolingo, B. (2018). How much of the world's food do smallholders produce? *Global Food Security*, 17, 64-72.
- Robul, Y. (2020). Structure and constituents of digital marketing systems, *Olsztyn Economic Journal*, 15(1), 53-62.
- Russon, M. (2020). *How internet access is improving in Nigeria*. Business reporter, BBC News. Retrieved from <https://www.bbc.com/news/business-51377955>.
- Seyed, J. A.; Mohammad, S. A.; Ahmad, S. A.; Jhalukpreya, S. R. (2016). Adoption of agricultural e-marketing: application of the theory of planned behavior. *Journal of International Food & Agribusiness Marketing*, <http://dx.doi.org/10.1080/08974438.2016.1229242>.
- Shiferaw, B.; Tesfaye, K.; Kassie, M.; Abate, T.; Prasanna, B.; Menkir, A. (2014). Managing vulnerability to drought and enhancing livelihood resilience in sub-Saharan Africa: Technological, institutional and policy options. *Weather and Climate Extremes*, 3, 67– 79.
- Singh, V.; Sankhwar, S.; Pandey, D. (2015). The Role of Information Communication Technology (ICT) in Agriculture. *Global Journal of Multidisciplinary Studies*, 3(4), 2- 10.
- Slade, E. L.; Dwivedi, Y. K.; Piercy, N. C.; Williams, M. D. (2015). Modeling consumers' adoption intentions of remote mobile payments in the United Kingdom: extending UTAUT with innovativeness, risk, and trust. *Psychology & Marketing*, 32(8), 860-873.
- Sudhakar Reddy (2021). *Nat. Volatiles & Essent. Oils*, 2021; 8(4) pp.426-437.
- Tambo, J.; Wünscher, T. (2017). Farmer-led innovations and rural household welfare: Evidence from Ghana. *Journal of Rural Studies*, 55(2017), 263-274. <https://doi.org/10.1016/j.jrurstud.2017.08.018>
- Teklewold, H., Kassie, M., Shiferaw, B., & Köhlin, G. (2013). Cropping system diversification, conservation tillage and modern seed adoption in Ethiopia: Impacts on household income, agrochemical use and demand for labour. *Ecol. Econ.* 2013; 93:85–93.
- Teklewold, H., Gebrehiwot, T., & Bezabih, M. (2019). Climate smart agricultural practices and gender differentiated nutrition outcome: empirical evidence from Ethiopia. *World Dev.* 2019; 122:38–53.
- Thanabordeekij, P.; Sudtasan, T.; Tanamee, D. (2020). Integrating trust into the technology acceptance model: the case of mobile banking adoption in Myanmar.
- Thomas, K.; Michael, O.; Silahs, C. (2016). Impact of social media on agricultural extension in Kenya: a case of Kesses district. *International Journal of Agricultural Extension and Rural Development Studies*, 3(1), 30-36.

- Trading Economics (2020). *Nigeria-rural population*. Retrieved from <https://tradingeconomics.com/nigeria/rural-population-percent-of-total-population-wbdata.html>.
- Tufa, A.; Alene, A.; Manda, J.; Akinwale, M. G.; Chikoye, D.; Feleke, S.; Manyong, V. (2019). The productivity and income effects of adoption of improved soybean varieties and agronomic practices in Malawi. *World Development*, 124(2019), 104631.
- Upadhyay, J. (2022). "Digital Marketing Practices Adoption in Micro, Small and Medium Businesses for Inclusive Growth." *International Research Journal on Advanced Science Hub* 04.04 April (2022): 75–80. <http://dx.doi.org/10.47392/irjash.2022.021>
- Venkatesh, V.; Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, 46(2), 186-204.
- Venkatesh, V.; Morris, M. G.; Davis, G. B.; Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS quarterly*, 425-478.
- Venkatesh, V.; Thong, J.Y.; Xu, X. (2016). Unified theory of acceptance and use of technology: a synthesis and the road ahead. *J. Assoc. Inf. Syst.* 17:328–376.
- Wossen, T.; Abdoulaye, T.; Alene, A.; Haile, M.; Feleke, S.; Olanrewaju, A.; Manyong, V. (2017). Impacts of extension access and cooperative membership on technology adoption and household welfare. *Journal of Rural Studies*, 54, 223–233. <https://doi.org/10.1016/j.jrurstud.2017.06.022>
- Wossen, T.; Alene, A.; Abdoulaye, T. (2019). Poverty Reduction Effects of Agricultural Technology Adoption: The Case of Poverty Reduction Effects of Agricultural Technology Adoption: The Case of Improved Cassava Varieties in Nigeria. *J. Agric. Econ.*, 70, 392–407.
- Wooldridge, J.M. (2010). *Econometric Analysis of Cross Section and Panel Data*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2010.
- Zeng Y.; Jia, F.; Wan, L.; Guo, H. (2017): E-commerce in agrifood sector: a systematic literature review. *International Food and Agribusiness Management Review*, 20: 439– 460.

Developing entrepreneurial universities under the evolution of innovation system and strategy of innovation-driven development, the case of Shenzhen, China

Weijia Ding

University College London, United Kingdom

Abstract

As the world economy is increasingly driven by knowledge creation, the innovation system aiming to development and competitiveness becomes a central goal of most leading economies (World Bank, 2019). In a knowledge-based economy, universities are taking on more third mission activities, such as entrepreneurship and innovation, to transfer knowledge and technology to industries and society. While more than 200 universities globally are identified as leadership of innovation, there lacks empirical and comparative research in different contexts. This study focuses on China's emerging regional innovation system and innovation-driven development policies, with an emphasis on the development of entrepreneurial universities.

This study utilizes the Triple, Quadruple, and Quintuple Helix framework to investigate the entrepreneurial environment and activities, focusing on analyzing interactions and partnerships among actors in the Triple Helix model. The research takes place in Shenzhen, a renowned hub of entrepreneurship and innovation known as the "Silicon Valley of China." A mixed-methods approach will be employed in the study, combining a literature review with

expert semi-structured interviews that include structured surveys and open-ended questions. The research finding will shed light on the pattern of innovation partnerships and describe the entrepreneurial university as a dynamic ecosystem that can mobilize its resources, capabilities, and strengths. A proposed input-process-output-outcome model that encompasses inputs (such as resources from different stakeholders), processes (such as funding selection, research activities, and incubation), and outputs (such as inventions, networks, and innovation centers) could help explain why some entrepreneurial universities are more successful and provide implications for innovation-driven development policies.

This study systematically explores how universities become more entrepreneurial and how innovation partnership of Triple Helix actors is forged within a different economic context. Theoretically, the study contributes to empirical case in an eastern developing context. The study's findings offer a better understanding of the role of universities in socio-economic transformation by elucidating how universities become entrepreneurial in a developing context. Unpacking novel approaches and practices to orchestrating innovation partnership, this study also underpins implications for science, technology and innovation policies, university practices, and R&D collaborations. The findings will inform better science, technology and innovation policies and offer practical insights into successful multi- collaboration projects. Furthermore, the study will provide implications for university development strategies and approaches to collaboration with external stakeholders by examining the current role, opportunities, and challenges faced by universities.

Keywords: Entrepreneurial University, Innovation-driven development, Triple Helix Model, Hybrid institutions, tech entrepreneurship

Acronyms and abbreviations

TM	Third Mission
EU	Entrepreneurial University
STI	Science, Technology, and Innovation
S&T.	Science and Technology
R&D	Research & Development
E&I	Entrepreneurship & Innovation
HEI	Higher education institutes
NIS	National Innovation System
FDI	Foreign Direct Investment
WTO	World Trade Organization
PCT	Patent Cooperation Treaty
IP	Intellectual property
WIPO	World Intellectual Property Organization
SOEs	State-owned Enterprises
SME	Small and Medium Enterprise
TTO	Technology Transfer Offices
TM	Third Mission

1. Introduction

1.1. Research background

As the world economy is increasingly driven by knowledge creation and commercialization, a National Innovation system (NIS) to promote development and competitiveness becomes a central goal of most leaders of economies (Lundvall 1992, World Bank, 2019). Innovation systems mobilize different resources and create dynamic interactions among various entities, including government, enterprises, universities, and hybrid institutes, to jointly promote knowledge creation, diffusion, and transfer which are cornerstones of economic growth and competitiveness (OECD, 2019).

In a knowledge-based economy, universities involved in Third Mission (TM) activities like university-industry labs and spinning-off enterprises are becoming engines that contribute to the social, economic, and cultural development of the region where they operate by transferring knowledge and technologies to industry and society at large (Secundo et al., 2017a; Agasisti et al., 2019) and the importance of TM has been widely recognized among policymakers and universities (Benneworth et al., 2015; Compagnucci & Spigarelli, 2021). Universities are increasingly proactive in TM such as technology transfer and entrepreneurship education. More than 200 universities worldwide have well-established reputations in entrepreneurship in the knowledge-based economy and have been identified as leaders in innovation (OECD, 2017; Frolund et al., 2017). The concept of entrepreneurial university (EU) mainly refers to universities that are actively participating in or incubating entrepreneurship activities and have mission statements as well as strategies, teaching & learning structures for its entrepreneurship development (Clark, 1998; Tijssen, 2006; Guerrero & Urbano, 2010; Ruth, 2014).

A triple helix perspective has also been utilized to understand how innovation emerges from and is sustained by government-industry-university collaboration and how entrepreneurial universities are emerging in developed economies in recent years (Etzkowitz & Leydesdorff, 1995, 2000; Klofsten et al., 2000, 2010; Lawton Smith et al., 2013; Etzkowitz & Zhou, 2018). However, whether the triple helix model of innovation applies to developing economies are under-explored and empirical studies of entrepreneurial universities in developing contexts are still lacking (OECD, 2017; Newman et al., 2018, 2019).

1.2. Objectives and structures of the study

This study mainly attempts to tackle the following research questions:

- 1) Is the Triple, Quadruple or Quintuple Helix model of innovation relevant to the Chinese context?

- 2) How do the three spheres of innovation characterised by the Triple Helix model operate in the Chinese context?
- 3) What are interactions among the three actors in the innovation system and how do interactions vary across regional and local contexts?

This study utilizes the Triple Helix framework to analyze China's entrepreneurship environment and activities with a focus on how interactions between the three actors. The study further analyzes whether regional institutional context will result in different entrepreneurship patterns. This study chooses Shenzhen, one of China's most successful economic zones famous for innovation and entrepreneurship, as the pilot study area. The following analysis examines the local examples of proactive collaboration in Shenzhen that involved government, universities, research, research institutions, and industry sectors in promoting and implementing integrated innovation policies. The analysis starts with S&T rise in China, including its growing attractiveness as a desired knowledge-based innovation-driven development. It also points to a limitation of innovation policy making and potential tensions during the interactions among three entities that curb innovation. This study contributes to the existing literature from theoretical, empirical, and practical aspects.

Theoretically, the study adds to the development of the Triple Helix system thinking and innovation studies concepts by untangling the evolutionary process of the Triple Helix system in promoting innovation and entrepreneurship in the Chinese context. The study systematically explores interactions among actors to uncover Triple Helix relationships, e.g., trilateral collaboration, networking, and complementary synergies. Empirically, analysis by utilizing first-hand survey data collected from industrial, academic, and governmental institutions in a representative Chinese city of entrepreneurship provides additional assessments and analyses of the framework of the Triple Helix system in a different context. Most literature has developed conceptual models make conclusions based on case studies in developed economy. This study provides additional assessment and analysis of the emerging Triple Helix system in a developing economy context. It further provides implications for innovation policymaking, university practices, and industry-government-university collaborations. This section introduces the study and sets its crucial research questions in context.

The remaining report is structured as follows: Section 2 reviews National Innovation System (NIS) literature and interactions between actors in the context of China and Shenzhen. It also presents a review of the theoretical framework of Triple Helix. Section 3 analyzes the entrepreneurial environment and science, technology, and innovation (STI) policies in the context of Shenzhen. Section 4 explains the research methodology that underpins the case

study with implications for further investigations. Section 5 presents the study's findings, followed by a discussion in Section 6. The final section presents the conclusion and implications of this study.

2. Literature review

2.1. The emergence of the triple helix concept

The NIS approach reflects the rise of systemic approaches to the study of technology development as opposed to the linear model of Innovation (Godi, 2009). The systematic approach focuses on knowledge flows and interactions among actors in a NIS, including interactions among enterprises, 2) interactions among enterprises, universities, and research institutions and 3) diffusion of knowledge and technology to firms, and 4) personnel mobility. (OECD, 1999; 2006) The synergy among industry, universities, and government is known as the Triple Helix (Etzkowitz & Leydesdorff, 1998, 2000). Focusing on the interaction among main actors within a NIS, the model considers these factors the key to improving conditions conducive to innovation. Etzkowitz & Leydesdorff (2000) argue that university, industry, and government are the main pillars of many innovation systems theories, including NIS.

The "triple helix" scholars provide a general framework for exploring and informing complex innovation dynamics in societies and have been demonstrated as a successful model for fostering innovation and entrepreneurship (e.g., in the US: Leydesdorff, 2012; Etzkowitz & Zhou, 2018; in Europe: Lawton.S et al., 2013; Geuna & Rossi, 2016). It reveals that trilateral collaborations among universities, industries, and the government in fostering innovation and entrepreneurship works under an interactive rather than a linear model of Innovation (Triple Helix Research Group, 2017). According to this framework, innovation processes and outcomes are not only driven by mutual coordination between university-industry, industry-government, and government-university pairs but also by three-way interactions and collaborations among the three entities. Collaborations among the three types of organisations in the system help foster a more favorable entrepreneurial environment comprised of research-centered institutions, university spin-offs, strategic alliances among private enterprises, financial intermediaries, public-sponsored infrastructure organizations, etc. (Etzkowitz and Leydesdorff, 2000). This paper introduces the Triple Helix systems as an analytical concept that systematizes the critical features of university-industry- government interactions.

Each component (government, university, industry) in the Triple Helix framework evolves to adopt characteristics of other institutions, which leads to hybrid institutions. Therefore, the framework can characterize the development of hybrid institutions. Entrepreneurial university--for instance, increasingly being recognized with a hybrid element--has studied with the "triple helix"

perspective and has thus been argued to contribute to transitions of science-based innovations to global society (Etzkowitz, 2006; Etzkowitz & Zhou, 2018). The Triple Helix innovation system also takes different forms, with one of the three spheres serving as the key innovation driver (Etzkowitz, 2008; Ranga & Etzkowitz, 2013). In a statist regime, the government plays a leading role in driving industry and academia, while the industry is the leading force in a laissez-faire regime. The prominent role of universities is usually recognized in a balanced model. In addition, the Triple Helix system is associated with an evolving overlay of communications, networks, and organizations among the three helices and is dynamic and complex (Leydesdorff, 2000).

2.2. The rationale and criticism of the triple helix model

The model of Triple Helix has its critics who point out limitations. Shinn (1999) argued that the Triple Helix model is theoretically vague. He is also doubtful whether the model can be accepted and defined as analytic elements or only as an evocative metaphor (Shinn, 1999). Saad and Zawdie (2005) further raised doubts about the role of the institutions within the Triple Helix system as a knowledge source in the growth of regional innovation systems (RIS). According to the main idea of the RIS, it can be applied only to effective growing regions. Based on this, it raises doubts about the application of the Triple Helix as to whether the model is suitable to use in a region where the universities do not have good infrastructure and enough knowledge capacities.

Similarly, the university is expected to play an underpinning role as an entrepreneur and develop capacities to help Triple Helix model institutions achieve a hybrid status. However, it is unclear what unique characteristics a hybrid institution system would be in different contexts. Etzkowitz (2008) has mentioned the incubator movement in countries as a direction towards a hybrid Triple Helix. It requires more description and a further conceptualisation. Another problem with the hybrid Triple Helix is that its emergence as an institutional system could be elusive, as it involves a complex process based on high levels of commitment, understanding, and trust between all three spheres (Leydesdorff & Zawide, 2010). It can also lead the university to lose sight of its critical thinking and become too business-oriented (Saad, 2004). Etzkowitz (2003, 2008) highlighted that the Triple Helix model is transitive and alternative, where the process of change and development is ongoing. The evolutionary characteristic of the Triple Helix manifests in the way that each institutional sphere starts to take the role of the others in addition to fulfilling their traditional functions.

2.3 Rationale for selecting China and Shenzhen as the pilot study area

While the Triple Helix framework was used to analyse innovation systems in developed countries such as the US and Sweden, there remains a lack of empirical analysis of how Triple Helix works in developing economies. This

study investigates whether the Triple Helix framework applies to the entrepreneurial environment in developing countries, with the provision of emerging entrepreneurial universities in China. Over the past two decades, entrepreneurship in China has grown exponentially. It is bringing forth disruptive changes not only to China but also to the rest of the world. In recent years, China has issued a call for the country to accelerate its efforts to become self-reliant in innovation and innovation-driven development in S&T. A handful of Chinese companies, such as Alibaba, an online retail giant, Tencent, online social communication and gaming ecosystem giant, and Huawei, a global leader in telecoms, have emerged as global innovators and gathered much attention.

China's vision to become an innovative nation is a critical element in the development narrative, within which the higher education sector, notably its leading universities, plays an important role (World Bank, 2019). Higher education institutes (HEIs) are regarded as a critical source of innovation in development strategy in China. Overall, China's R&D input and scientific output have risen sharply in comparison with the US, Japan, and European Union (ERC, 2021). R&D investment rose almost three-fold from 2000 to 2017, catching up with international standards and reaching a total of 2.4 trillion yuan (\$372 billion) in R&D spending with a share of 2.4 per cent of the country's aggregate GDP in 2020. Therefore, China provides an excellent sample to understand how the triple Helix model works in a developing context.

2.4 The triple helix model and innovation policies in China

As the world's largest developing country and second-largest economy, China has become an indispensable player in the global economy. The Chinese government has set its strategic priority on the development of S&T to move the nation towards a knowledge-based economy (Lu, 2008; Zhou & Leydesdorff, 2006) since the late 1990s. The promulgation and implementation of the 2006 version of the 15-year 'Medium to Long-Term Plan for the Development of S&T' set an ambitious goal to turn China into an innovation-oriented society by 2020 and a global leader in S&T by 2050 through indigenous Innovation (Liu et al., 2011). Despite a significant rise in R&D expenditure in the past two decades, China still faces considerable challenges in fostering and commercializing innovation. To further encourage innovative and entrepreneurial activities, the Chinese central government has recently depicted a vision of building a national-wide innovation system towards mass E&I in the 13th five-year plan (State government, 2019). An official guideline of popular E&I issued by the State Council in 2015 endorsed this vision by further stressing the role of the government in facilitating entrepreneurial and innovative activities and recognizing the need for close interactions between research institutions, universities, enterprises, and government agencies to achieve this strategic goal jointly. In 2017, the state Council further implemented policies and actions to

propel mass innovation and entrepreneurship and has ambitious to upgrade the NIS (World Bank, 2019).

Distinguished from the balanced mode of Triple Helix Innovation (Table 1), where universities underpin a conspicuous role in the innovation system (ERC, 2021), it is apparent that the government plays a dominant role in the Chinese context. The government leads the direction of innovation by owning most research-led universities and institutions, leading strategic decisions of state-owned enterprises, and establishing innovative infrastructures such as high-tech development zones, science parks, and incubators, as well as initiating and reforming regulations and policies related to the promotion of innovation and entrepreneurship. Zhou (2008, 2018) thus argues that the Chinese NIS, a mixture of government-pulled, corporate-led, and university-pushed models, could also be classified as a statist regime based on Etzkowitz (2008)'s typology.

Models	Literature	Characteristics	Leading Force
Statist regime	SponUOOl Etzkowitz 2008	The government plays a leading role in a statist regime to drive academia and industry. The government directs research resources at specific mission-driven or technological priorities.	The government
1 ait*c/- (airr regime)	Mowcry. 2010 Ranga and Etzkowitz. 2013	The industry is the leading force in a laissez-faire regime. There is a strong business focus on R&D with a balance of investment between large firms and SMEs.	The industries
Balanced regime	Etzkowitz. 2004 Etzkowitz. 2008	The university plays a prominent role in a balanced regime. R&D are more associated with curiosity driven activities and prototype attempt	The universities

While China's dramatic economic rise and formidable scientific output are well known, the innovation model demonstrates significant regional differences. Regarding technology development and innovation, although the central government enacts a universal S&T policy for all provinces, different regions differ significantly in the implementation of national policies and the launch of local programs to supplement national policies (OECD, 2017; World Bank, 2019). Regional and local governments are administratively and economically independent in many aspects. Chinese provincial governments have the autonomy to formulate regional economic and social development policies, set local tax rates, grant business licenses, coordinate local business development, design provincial innovation plans, and manage their budgets to foster regional entrepreneurship. As a result, Chinese regions' performances on innovation

pathways and entrepreneurship outcomes vary, attributed to various innovation and entrepreneurship environments (Gu & Lundvall, 2006; Wadhwa et al., 2012 ; Manson& Brown, 2014).

In general, eastern coastal areas are far more developed than central and western provinces. The number of annual patent applications among coastal regions, for instance, is significantly higher than in other regions due to international trade, better infrastructure, and more advanced technology transfer markets. The efficiency of innovation input-output thus is significantly higher in the eastern Chinese regions than in other regions. Li (2010) argues that such a regional gap is caused by variations in provincial understanding and attributes variations in innovation performance of Chinese provinces to the regional difference in inputs by innovation participants, including governments' financial subsidies, the effectiveness of financial intermediary and R&D expenditure as well as the local environment. A recent study (Prodi et al., 2017) further indicates that the disparity in regional innovation infrastructures, such as knowledge production, access, knowledge transfer, and collaborations, results in distinct innovation performance.

The current stratified higher education system invests heavily in its top-tier HEIs to compete at the highest levels internationally. Leading universities such as Tsinghua University and Peking university usually win the most considerable funding boost, as well as top students, professors, and staff from around the world. International collaboration and facilities standards were all part of China's innovation investment, making it a leading knowledge partner for other nations, including Australia, Germany, the UK, and Canada (Welch, 2018). Also, science parks affiliated with well-known universities, e.g., Tsinghua's well-known Tsinghua QIDI Park in Beijing and others in Hangzhou and Shenzhen, are other marks of this spirit of innovation and entrepreneurialism (Rhoads et al., 2014).

Notably, unlike Beijing or Shanghai owning abundant higher education resources, Shenzhen's selected pilot study area is a young city lacking top-tier universities or research institutes. Shenzhen has experienced a high speed of development from a small village to an international entrepreneurial zone and one of the super tier1 cities in China in recent 40 years. However, the city only has two universities- Shenzhen university and Southern University of science and technology until 2017. Despite the lack of top-tier research resources, both universities perform outstandingly in E&I.

In the next section, the significant transformation of Shenzhen from a fish village into the world-famous Chinese Silicon Valley is discussed with the analysis of the entrepreneurial approach adopted by the local government and multiple collaborations in steering innovation. Responding to a recent call for the Greater Bay Area development opportunities, the study will also depict

how innovation participants (governments, universities, industries) in Shenzhen have tried to work with external resources to capture the development opportunities.

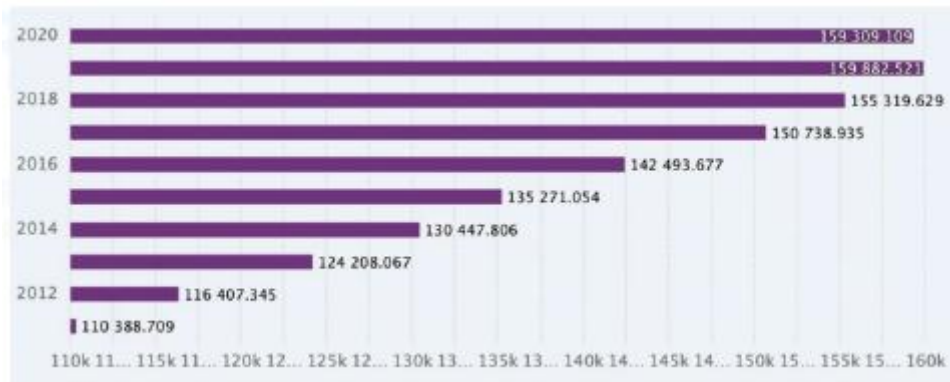
3. Study area and policy context

3.1 Shenzhen: From a small fishing village to Chinese Silicon Valley

Shenzhen has experienced a high speed of development from a small village to an international entrepreneurial zone and one of the super tier1 cities in China in recent 40 years. However, the city lacks top-tier HE resources and only has two universities- Shenzhen university and Southern University of science and technology until 2017. Despite the lack of innovative talents, the key driver of economic success is attributed to its spirit of reform and innovation. By 2019, the city has surpassed Beijing to become the city with the most active innovation pursuits. In 2019, the number of patents granted in Shenzhen exceeded 1.66 million, accounting for 6.73 percent of the total number of patents granted in China.

The Chinese economy's unparalleled growth since enacting Reform and Opening policies in 1978 started in Guangdong. When reformers decided on a socialist market economy in 1978, they feared it would cause political tension. Located 140 km from Guangzhou and with 30,000 residents, a small village was renamed Shenzhen and awarded Special Economic Zone status, granting it the autonomy to experimental trial , including accepting foreign direct investment (FDI) in 1980s. Benefiting from its proximity to Hong Kong and abundant cheap labor resources, Shenzhen soon became a global manufacturing and assembly center, with annual average GDP growth rates of around 30% (1980–2000). Seizing the opportunity of WTO membership Shenzhen's export-oriented industries soared after 2000. Since 2017, the GDP volume of Shenzhen (Figure 1) has surpassed Guangzhou, Beijing, and Shanghai, and in early 2019, it surpassed Hong Kong(Figure 2). Now it is regarded as Silicon Valley in China for its success and fame in tech innovation and entrepreneurship. Despite the shortage of higher education resources, both local universities also have the fame of outstanding performance in E&I (THE,2022; CUAA, 2020; Tao, 2019). Taking the first local university in Shenzhen- Shenzhen City, for example, the university submitted the most (201) patent cooperation treaty (PCT) application in 2018, surpassing Tsinghua University, the most renowned research-led university, with the number of 137 PCT applications in the same year.

Figure-1. Shenzhen gdp per capita

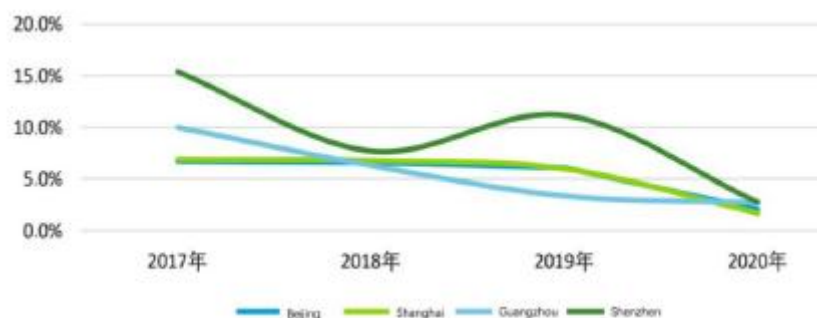


Source: Shenzhen Municipal Bureau of Statistics.2020

3.2 An adaptive pathway to promote hi-tech development and industrial upgrade

Shenzhen's success could attribute to both geographic advantage and the state's positive policy environment, visionary planning, and entrepreneurial organizational framework (Chen & Kenney, 2007). Many people view Shenzhen as a successful case of globalization, and its economic growth largely depends on international trade. However, globally resurgent nationalism hit Shenzhen's economy, especially during the US-China trade war in 2018. By late 2018, it was apparent that the 'Trade War' was a proxy for a technology war, aiming to curb China's innovation drive, most notably initiatives in its signature Made in China 2025 plan (White House, 2018). Consequently, Shenzhen's total exports dropped for four successive months since the second quarter of 2018, highlighting the massive challenge of policies and economic transformation.

Figure 2. The comparison of GDP acceleration in Tier one cities in China



Source: National Bureau of Statistics.2020

The Shenzhen government realized the urgency to transform its growth pattern far before 2018. The Shenzhen government acted early to transform its growth

pattern. In the 1980s, it recognized the limitations of relying on cheap labor and low-value exports. Shenzhen shifted focus to higher value-added industries, establishing Shenzhen University in 1983 and the S&T Industry Park in collaboration with the Chinese Academy of Science Shenzhen in 1985. The government issued numerous policies to encourage the hi-tech industry, allowing individuals to establish private hi-tech companies with intellectual property rights. In 1995, Shenzhen launched a policy to revitalize the city through science and technology, with IT, new materials, and biotechnology becoming pillar industries.

3.3 From entrepreneurial governance to innovation-driven development

At the policy level, the government has introduced institutions and measures to spur business start-ups and investment, attract various professionals, and build systematic policies to motivate market entities to engage in innovation. The city accelerates the gathering of global innovation resources and standards so that the city forms a new innovative industrial cluster that further attracts global resources and moves Shenzhen to the upper ladder of the global industrial chain.

Shenzhen is committed to Intellectual Property (IP) protection. The city issued the Strategic Guideline on Intellectual Property (2006–2010) as the first document at the city level to legislate IP protection in China. A working group headed by the Deputy Mayor contributed to The Opinion on Building Intellectual Property Index System of Shenzhen (2007). The system measured innovation ecology in five dimensions, including innovation ability, management efficiency, IP protection, IP implementation, and social-environmental support. As a result, the annual application and volume of IP in Shenzhen achieve consecutive championships in China.

The municipal government also helps solve financial difficulties to support the high-tech industry. It introduces capital to invest in patent holders from universities and technology start-ups to promote knowledge transfer. Shenzhen has cooperated with banks in providing loans to companies and has instituted measures to encourage intellectual property rights and finance pledge. The Shenzhen government started developing the online infrastructure for High-Tech Enterprise Financing in 2014, in cooperation with banks in China.

To further encourage innovation and entrepreneurship, the Shenzhen government established a strategic policy-based Venture Capital: Shenzhen Angel Investment Guiding Fund, in recent years to strengthen the capacity for venture capital investment and offer helps to start-ups. Investment in private tech firms bolstered the growth of private high-tech enterprises and further supported them to move up the market and go global. As a result of these efforts to innovation-driven development, the hi-tech industry in Shenzhen achieved an added value of ¥974.7 billion in 2020. The volume of international PCT

patents in Shenzhen rose significantly and accounted for 30.2% of Chinese international PCT patents total in 2020, achieving 17 consecutive championships nationwide and ranking secondly to Tokyo internationally. Shenzhen's share of GDP was elevated to 32.8%.

4. Research methodology

4.1 Research design

This study uses literature review techniques and semi-structured interviews to explore the emergence of entrepreneurial universities in Shenzhen, China. Before the study, I used archives and primary documents to understand the local context of E&I. These sources include investigating local policies relevant to the entrepreneurial movement in universities.

The study conducts semi-structured interviews, including structured and open-ended questions, to explore the relationship between universities and their local innovation systems and context. These methods give insight into how university staff understands and value different support mechanisms. The interview protocol contains a mix of structured and tightly defined questions (the survey) and more open-ended questions for participants (the semi-structured interviews).

The interviews took between from November 2021 to March 2022. The interview questions are prepared in English and translated into Chinese. Items will be back translated to ensure that the meaning of the original items is maintained. Because the questionnaire requires respondents to choose in the form of options, the final response can be simplified into digitized data, so a translation of the responses is not needed. From open-ended questions, phrases, emerging themes/ patterns, and different reactions are gathered and used to sort analytic leads, assumptions, and potential codes.

According to preliminary information on regional entrepreneurship environment and entrepreneurship activities from the literature review and the archival documentary analysis, this study selects participants with the following criteria: (1) Participants above the age of 18. (2) Participants have at least three years of related (practical or theoretical) experience in university entrepreneurship. (3) Participants obtain relevant research backgrounds in academic entrepreneurship and entrepreneurial universities. (3) Participants play as key stakeholders in entrepreneurial activities in the region. All participants will answer survey questions and will be invited to participate in the semi-structured interviews.

Our focus groups of interviewees consisted of 22 members, including directors in universities' technology transfer offices (TTO) or innovation centers (3), Academics in universities and institutes (3), Entrepreneurs or business

executives in hi-tech enterprises (7), Director in venture capital institutions (3) Government officials (4) Director of an innovation center in quasi-public institutions (2). We also revised our survey questions based on gathered suggestions from these focus group members. We randomly asked questions on various constructs to reduce common method bias in the survey design.

4.2. Data collection and analysis

A series of semi-structured interviews with representatives from universities, enterprises, and the government to test the reliability and clarity of our survey questions was then conducted. The in-person interviews then deeply explore different opinions and situations of entrepreneurial universities in practice. The open-ended questions are best characterized as semi-structured because in answering open-ended questions, participants will be encouraged to share their thinking fully, which may involve asking follow-up questions.

Data analysis began with reading the transcripts of the interview. In this pilot study, we interpreted the qualitative and quantitative data on the basis of the definition of qualitative research, which involves collecting and analyzing non-numerical data (e.g., text, video, or audio) to examine concepts in terms of their meaning and interpretation in specific contexts of inquiry and quantitative research that mainly refers to an approach that examines concepts in terms of numbers, statistical analysis, intensity, or frequency (Aspers & Corte, 2019). It can be used to gather in-depth insights into a problem or generate new ideas for research.

Themes, patterns, and dimensions (e.g., university and city entrepreneurial culture, funding and policies, spin-off experiences, self-efficacy) deduced from the research questions become the codes and labels that are assigned units of meaning to the information gathered during this study. A coding scheme based on an entrepreneurial university's conceptual framework and paradigm is designed for the initial data sorting and analysis. Thus, codes are categorized into different groups to identify the dimensions, elements, and variables. Each category follows with more detailed descriptions of different aspects, contributing to a potential pattern of an innovation system. Also, attention is paid to any meaningful units of data that can stimulate the researcher and readers to think beyond the information and reveal some pieces of evidence for the study questions.

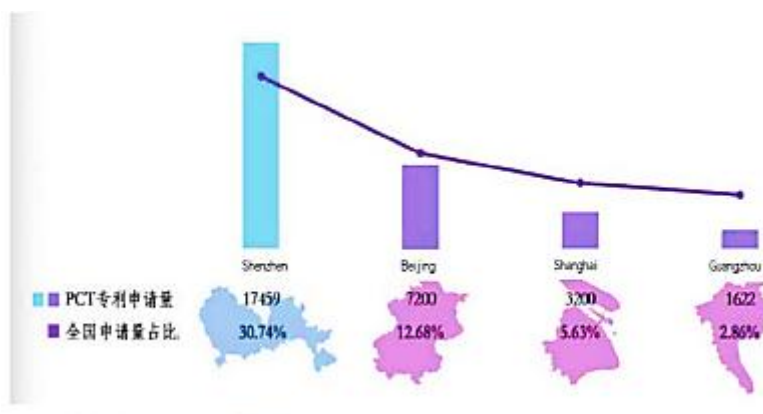
The above themes and variables are verified through the analysis of empirical data, and Nvivo software, auto-coding, and Word cloud analysis are illustrated and translated in Figures 3&4.

5. Findings

5.1 Industry-oriented innovation and upgraded global supply chain

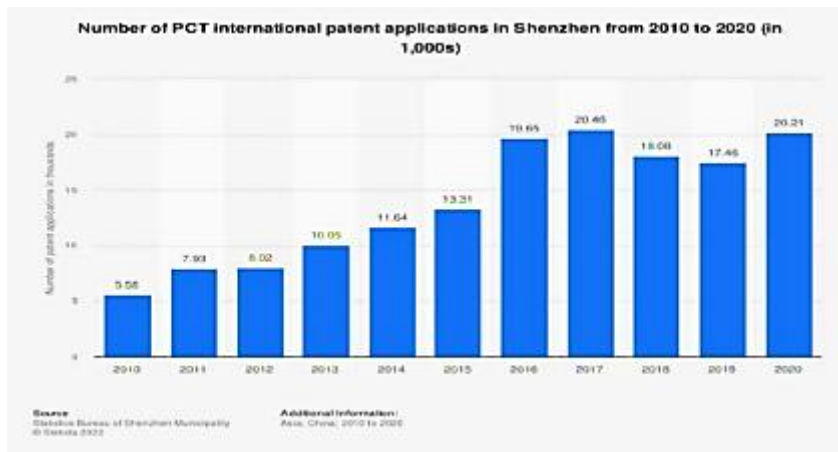
The economic reform and transformation since the 1980s have made Shenzhen the most ‘entrepreneurial city’ in China. In Shenzhen, different drivers can be identified in the innovation system that promotes the emergence of the innovation system. Overall, the critical driver of innovation in Shenzhen is industry and entrepreneurs in the local or regional economy. Enterprises trigger in-house innovation to stimulate new growth points and broader revenues, increase market competitiveness and broaden opportunities for R&D and new products. According to the “National Innovative City Innovation Capability Monitoring Report 2020” released by the national technology department in 2021, Shenzhen’s innovation capability ranks first in the country, with the number of core intellectual property indicators ranking first. By 2020, Shenzhen’s high-tech industries will achieve an added value of ¥974.7 billion, accounting for more than 35% of GDP. It is the first pillar of Shenzhen’s four pillar industries and a significant economic growth point. In 2020, Shenzhen will rank first in the number of patent applications and authorizations among first-tier cities (Figures5 & 6). Among them, the number of patent applications from an enterprise account for 89.58%, much higher than the national average of 65%. In 2020, the number of PCT applications reached 20,209, accounting for 30.2% of the national total (Shenzhen S&T committee,2021), ranking as 17 consecutive championships in China and performing significantly when compared to international cities ranking after Tokyo and ahead of Silicon Valley, New York, and Israel (WIPO,2021).

Figure 5 – statistical analysis report of Shenzhen intellectual property data



Source: Shenzhen Bureau of Statistics, 2020

Figure 6 – Number of PCT international patent application Shenzhen Data



Source: World International Property Organization, 2020

The entrepreneurial spirit underpins the core culture in Shenzhen. Driven by young grassroots, alumni, and global talents, Shenzhen is the city with the highest density of start-ups and the most active small and medium-sized enterprises. Among the 3.6546 million commercial entities, small and medium-sized enterprises account for 63.49%, and there are 131.5 people owning enterprises per 1,000 people in the city. Notably, compared to most Mainland cities in China, where much of the local economy is dominated by state-owned enterprises (SOEs) with an average of 56% of gross capital (National Bureau of Statistics, 2018), Shenzhen's state-owned sector is much smaller (State Department, 2017). Private commercial entities account for 97.65% of the economic entities in Shenzhen, contributing to 67% of gross capital in the same year (Shenzhen Government, 2018). In the high-tech industry sector, private companies account for 93% (Figure 7). Established as an Economic Specific Zone close to Hong Kong, Shenzhen acted as a base for market reform. Consequently, entrepreneurship has become a popular culture in Shenzhen, resulting in active SMEs and an entrepreneurial ecosystem.

Figure 7 – Type of enterprises in Shenzhen high industry, Shenzhen Municipal Bureau of Statistics, 2018



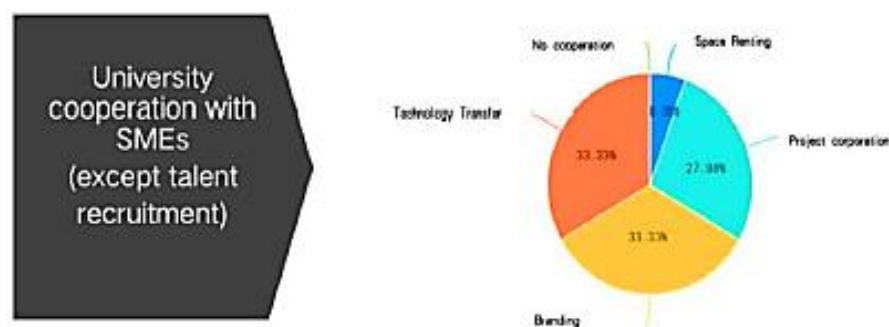
5.2 Towards entrepreneurial and multiple network governance for innovation

The local government in Shenzhen takes an entrepreneurial and gradual approach to policymaking, aiming for long-term success in the economic-specific zone. It acts as a platform constructor and policy-oriented investor in the innovation model. Instead of focusing solely on state-owned enterprises (SOEs), the government issues strategic guidelines and policies to encourage the development of private high-tech enterprises, promoting a shift to a high-tech industry. Notable policies include the Technology Innovation Regulations for the Shenzhen Special Economic Zone, the first local legislation in China for the technology innovation ecosystem. The government invests in the hi-tech industry, both financially and through infrastructure development and software environment improvement, to facilitate knowledge transfer. Various innovation centers, such as national infrastructure and engineering platforms, key laboratories, innovation incubators, and public service platforms, support industries and attract talent resources. The government also acts as a strategic policy-oriented investor, indirectly guiding industrial upgrading by owning policy-oriented capitals and investment companies while attracting social capital. Recognizing the importance of higher education resources for innovation, the Shenzhen government aims to introduce global higher education resources to support its innovation-driven strategy. However, the city has historically lacked higher education resources and research-led universities. To address this, a developmental strategy was launched in 2018 to establish around 20 higher education institutions by 2025 and attract global talents to build internationally competitive universities. Most universities in Shenzhen prioritize innovation and entrepreneurship as key development strategies. The next section will delve into the role of universities in Shenzhen's innovation system and analyze the process of institutional transformation toward a more entrepreneurial model.

5.3 University : Knowledge resources for local innovation and grassroots entrepreneurship

Unlike research-led universities like Tsinghua universities, where entrepreneurship is typically triggered by the desire to realize income from university research, the university's strategic development agenda is driven by and focused on a solid and ambitious technology transfer (Ruth,2014). Often building on established university research strengths, this model of university entrepreneurship offers a robust and fully institutionalised approach. In comparison, university entrepreneurship is much more bottom-up and community-led, catalysed by students, alumni, and entrepreneurs in Shenzhen. Technology transfer and entrepreneurship are triggered by a desire to cooperate with surrounding corporations and broader avenues for university support, stimulating local economic growth and creating graduate career and entrepreneurship opportunities. Driven by the university's grassroots, students, and alumni, a dynamic and inclusive ecosystem is gradually created through solid partnerships of trust between the entrepreneurial community and the university. This type of entrepreneurship focuses more on external influence rather than institutional capacity. Universities also downplay the importance of IP ownership and many technology transfer activities outside the university in the early stages. Enjoying a successful entrepreneurship image through grassroots start-ups and spin-offs, local universities in Shenzhen gradually realize that entrepreneurship is a differentiation advantage that could strengthen competitive advantage among universities. Therefore,

Figure 8 – University cooperation with SMEs (except talent recruitment)



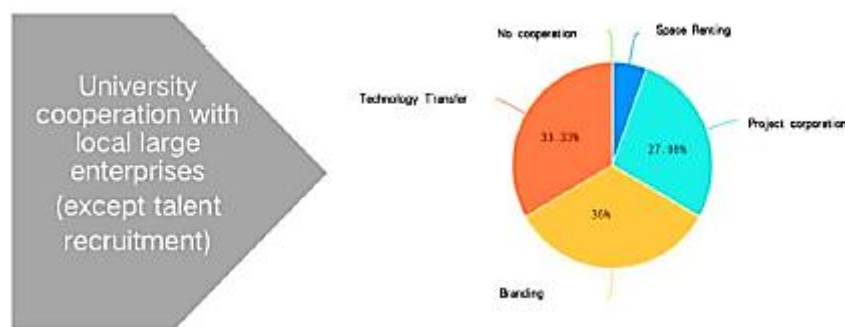
Source: Survey, 2021-10

universities in Shenzhen actively involve innovation and entrepreneurship in the universities' strategic development and seek to regulate and institutionalize their entrepreneurship profile.

The primary role of universities in the innovation system in Shenzhen is regarded as a pool of talent resources and the first seed of entrepreneurship. As seen in the survey, 90% of entrepreneurs and managers in different institutions approve that an essential role of the university is providing talents. For 80% of

SMEs, talent recruitment is the most frequent corporation with universities in the region. For SMEs, other types of corporations with the university include technology corporation & transfer (33.33%), project corporation in a specific program (24.98%), branding (33.33%), and entrepreneurial space renting (8.36%)(Figure 8). For a large corporation, talent recruitment is a regular corporation with universities. Other types of a corporation with the university include technology corporation & transfer on a university level (32%), technology corporation on department and individual level(63%), corporate in a specific program(50%), branding (65%) and co-spin off (23%) (Figure 9).

Figure 9 – University cooperation with local large enterprises



Source: Survey, 2021-10

Higher education serves as a catalyst for entrepreneurship, with more than 50% of young entrepreneurs attributing their initial ideas to university and 20% starting their first entrepreneurial ventures while in college. The exposure to new trends and technology during university studies can spark entrepreneurial aspirations. For example, a student attending a Tsinghua-Stanford joint summer entrepreneurial camp changed their career plans after developing a real business plan and pitching it to investors and industry leaders. This experience led them to pursue entrepreneurship outside academia. Collaboration between research labs and hi-tech industries extends beyond the university, with over 60% of academics and research staff at Shenzhen University engaging in partnerships with companies. The Media Lab at Shenzhen University, for instance, provides advanced equipment to local game and entertainment companies, reducing their costs and creating internship opportunities for students. Similarly, hi-tech companies involved in VR device development find valuable business opportunities by collaborating with university laboratories. Universities actively encourage entrepreneurship on campus, leveraging the surrounding industrial environment and support from alumni. Shenzhen University, for instance, embraces a tradition of fostering entrepreneurship due to the success of its notable alumnus, Pony Ma, the founder of Tencent. The university organizes entrepreneurship and innovation (E&I) competitions at various levels, simulating real business environments. Winning programs in these competitions may receive funding and support to transition into real

businesses. The university also invites external funding resources and facilitates connections with the industry. The entrepreneurial ecosystem in the vicinity further facilitates collaboration between the university and industry.

5.4 A tendency from bottom-up entrepreneurship to top-down innovation branding

Recently, E&I has further become a brand for universities in a regional area to win the competition among universities. Driven by successful bottom-up entrepreneurship stories and the external environment, the university not only actively encourages E&I on campus with a package of policies but also involves E&I in the universities' strategic development plan. Within the fierce competition among universities in the regional area, Shenzhen university recently has the ambition to involve entrepreneurship in the strategic development plan at the university leadership level to promote the whole university as a brand of innovation and entrepreneurship.

Therefore, a package of top-down policies to encourage E&I has been designed and delivered. The top-down model of entrepreneurship is usually involved in the strategic agenda of the university leadership team with a strong ambition of developing an entrepreneurial university to win the competition in the higher education sector, triggered by the desire to augment income from university research (Ruth, 2014; Philpott et al., 2011). This model offers a robust and fully institutionalised approach (Etzkowitz, 2018). The top-down approach to encouraging entrepreneurship includes incentive policies, connecting external resources (including funding and industry), providing relevant entrepreneurial training skills, and cultivating an entrepreneurial culture.

Almost all universities in Shenzhen now actively formulate policies and a prominent agenda to promote an alliance for the university's innovative output and technology transfer, which is responsive to regional circumstances and opportunities. Shenzhen University, for instance, has released a series of policies to promote scientific and technology transfer from universities since 2014. The policies provide strong incentives that 85% of innovation outputs and technology transfer rewards are given to academics or research groups. The university and department respectively charge 5% of profit to replenish revenue. The remaining 5% of benefits will go to the Technology transfer office for Intelligent Property rights management and compliance. The universities also actively provide academics and students a platform to approach external industrial and funding resources. For instance, the grassroots community-led effort to build E&I engagement is strengthened by universities through annual entrepreneurship competitions. Usually executed by the innovation or student center, the university organizes the first round of E&I competition at the collegiate level to select E&I programs with tremendous potential from March. Providing entrepreneurial skill training and guidance in the following months,

50-100 E&I projects will be selected and enrolled in the second round of competition. The selected E&I projects will be exposed to external commercial and investment opportunities in the second round of entrepreneurship competition across the university campuses. In the annual competition, the university, as a local higher education and research institution, will invite and integrate social resources of seed funding, venture capital, large companies, and industry associations to help like-minded entrepreneurs from the university pitch opportunities. Some competitors also give out hefty sums-- ¥2,000,000 at the university level and ¥30,000,000 investment from the venture or social capital. "The entrepreneurial experience on the campus will help future aspiring entrepreneurs and change-makers be aware of different business ideas, running start-ups, and access to seed funding and social ventures. For venture capital firms and angel investors, it is also a good opportunity to follow cutting-edge technological development and access to entrepreneurial opportunity in the very early stage." (Sui, 2022, Jan 10th, personal interview).

6. Discussion

6.1 Current model of E&I in Shenzhen

The study argues that the current innovation model in Shenzhen is driven by industries and supported by local government and universities. In this innovation model, E&I mainly derives from the industrial upgrade and grass-root entrepreneurship culture. Meanwhile, the government has a strategy of innovation-driven development. It provides various support from policy incentives and resources to strategic platforms and institutional innovation to prompt the construction of an innovation ecology. Universities are critical in innovative talent resources as public higher education institutions. Experimental innovation and technology transfer from university also accelerate innovation from initial thought to a demo product. Due to the support of the industry and the entrepreneurial culture in the region, although Shenzhen is a young city that lacks higher education resources, universities cooperate closely with the industry and have the most potential to develop an entrepreneurial university. The universities also start a strategic plan to combine the entrepreneurial genes in the city to develop an entrepreneurial university as core competitiveness in terms of long-term development.

6.2 Interactions of actors and trilateral collaborations in an innovation system

In our research, trilateral collaboration relationships exist among three entities (government, industry, and universities) in the Shenzhen area. The Triple Helix model of innovation and trilateral collaborations plays a role in the development of Shenzhen, yet the model is still evolving. At the institutional level, the city has started a higher level of innovation since 2013 by implementing trial measures and practices in four aspects of establishing a high-level innovation system to support the development goal of a socialist

market economy system, service-oriented government, institutions for independent innovation and energy-saving and environmental-friendly society. Since states' initiative for economic reform and marketization from 1980s, Shenzhen's innovation system now is more market-oriented, and closely integrated with government, industry, R&D sectors. The objective of innovation system is to maintain Shenzhen's continued innovation and development.

For industries, this is the inexhaustible motive force for innovation. Establishing an innovation system and trilateral collaborations have several essential roles in the sustainable development of high-tech industries and market vitality in Shenzhen. Primarily, it could systematically reduce the uncertainty and risks of innovation in a business context, thereby overcoming market and organizational failures to enhance investment and long-term competitiveness. Secondly, establishing technological transformation, enterprise derivative, and incubator is a testing trial to convert R&D achievement to profitable products for market growth. As the first pillar of industry in Shenzhen, the government introduced various policies and measures to spur business start-ups and investment, attract various professionals, and build systematic policies for technological innovation, as discussed above.

If entrepreneurship and the market initiated the motor of Shenzhen's STI, the objective of the local government is to sustain the long-term success and competitiveness of the city in technological innovation. Technological innovation is the talent resources of scientific and technical personnel and professionals. The number of people engaged in S&T activities in large and medium-sized enterprises increased from 57,200 in 2005 to 257,000 in 2012, an increase of 3.5 times in Shenzhen (UN Habitat,2019). By the end of 2020, there will be about 428,500 people working in the R&D field in Shenzhen, including 392,100 professional and technical personnel with intermediate professional qualifications above.

While enterprises are still key innovators in Shenzhen's city innovation ecosystem that generates 90% of R&D patents, including personnel and funding, and has close collaboration with local government, universities and research institutes have gradually become pioneering innovators since 2008, when Shenzhen became China's first National Innovative City. Realizing the importance of talents for STI and long-term competitiveness, Shenzhen has implemented several measures to attract talents and introduce HEI resources.

By strengthening local universities' development and improving incentives and services to introduce world-class higher education resources, the local government gradually builds a variety of talent policies to attract and retain global talents. By the end of 2021, the "Peacock Plan"(A well-known municipal talent policy to recruit oversea talents) had attracted 2954 people. In 2021 alone, the plan attracted 19 R&D teams in biology, life sciences, pharmaceuticals,

software, telecommunications, microelectronics, and new energy, which are pillar industries of the local economy.

In recent years, more resources have been allocated to tertiary education and research institutes to compensate for the lack of higher education and research institutes. Besides the measures to support local universities, Shenzhen has introduced the Research Institute of Tsinghua University in Shenzhen, the PKU-HKUST Shenzhen- Hong Kong Institution, the Shenzhen Virtual University Park, and other research organizations. These have become a continuous source of knowledge and technology. According to the Shenzhen Municipal Government plan, the total number of universities will be 20 by 2025.

The spontaneous partnership between the universities research lab and surrounding hi-tech industries is comprehensive. Due to its industrial system, Shenzhen also introduced and attracted the establishment of a national R&D infrastructure, including national research labs and engineering research centres. Equipped with ambitious national strategic R&D projects and newly developed R&D infrastructure, Shenzhen promotes self-reliant innovation in its six pillar industries and innovation-driven development strategy. Besides the rising trilateral collaborations among industry governments and universities, innovation intermediaries have become another player in Shenzhen's innovation ecosystem. These innovation intermediaries include incubators, professional services firms, financial consultancies, technology consultancies, trade associations, and recruitment agencies. Services provided by innovation intermediaries—including information exchange, decision-making advice, resource allocation, technical service, and technology evaluation—enable innovators and entrepreneurs to mitigate risks and commercialize technology faster. The city has become an important innovation hub in China by establishing an independent market-driven STI system and improving collaborations in the innovation system. It has ambitions to sustain its competitiveness in E&I. In the current innovation system, the demand for market and enterprise is the leading force. In contrast, the university's role becomes prominent in the evolving process. The model is still dynamic and evolving to the government's ambition to promote innovation-driven development and university attempts to put entrepreneurship as a brand and play a prior role in innovative talent resources and technology pioneering.

6.3 Limitations of the current model of E&I in Shenzhen

Despite a high degree of innovative vitality and proactive interactions in the innovation system, the Triple-helix model in Shenzhen still shows the area's characteristics and faces certain challenges and uncertainty. Government, even though it adopts an entrepreneurial approach and service-oriented stance and provides strong incentives to encourage innovation, lacks the follow-up work to know the situation on the ground. One main issue in Shenzhen is the rational

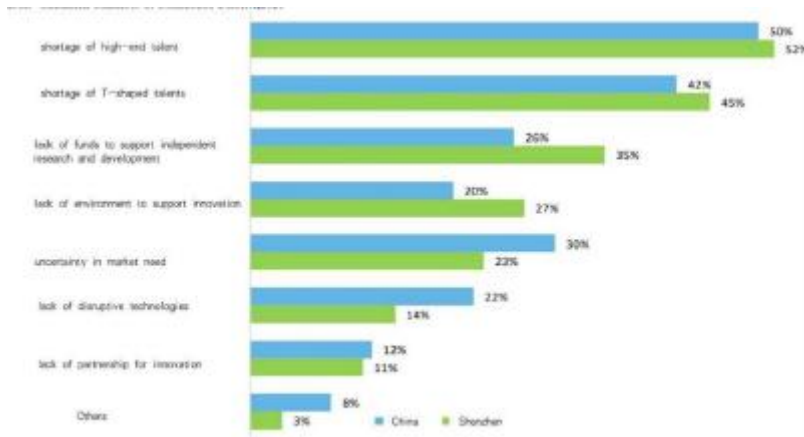
use of public resources that government initially wants to support E&I. Taking the innovation center, for instance, the whole city has 2660 spaces for innovation acceleration. For the development, except for 37% of the public resources, developers can obtain land at a meager cost than land for commercial or residential use and renovations. However, many innovation spaces, tiny ones, whose primary function is not to provide space for early enterprises at lower prices but to let for profit. Some well-known innovation spaces with geographical advantages want to attract well-known high-tech enterprises or scientific research institutions. Innovative companies with high growth potentials, however, cannot afford the rent. Therefore, while the government invests massive amount of public resources and introduces a series of policies to build an environment that supports innovation, how much of the resources are flowing into technological innovation and its output is difficult to measure and evaluate. A group of recent literature examining innovation centre or labs have echoed the chaos (Aitken, et al., 2021; Simone 2019).

Another main challenge faced by the government is that the multiple-departmental management model leads to a lack of awareness of who should be responsible for policy implementation. The development & reform department, the S&T innovation commission, and the SMEs promotion department have relevant policies to support innovation. However, if the municipal leadership team does not lead the project at the implementation level, many innovative projects may face barriers. Although the Shenzhen government is efficient and pragmatic, how cooperating across departments remains an intractable problem when it comes to the collaboration of different departments. some R&D projects are forced to suspend or take prolonged procedures when it faces the collaborations and responsibilities across the department, which can exemplify the lack of collaboration across institutions and policy support consistence(Wei et al., personal Interviews, 2022). Considering the balance of different stakeholders and potential competition and political balance in the different management departments, not every department is willing to promote it actively. Paradoxically, advancing innovation needs the persevering supports from public sectors and ensure stability in policies at the same time (Rainer et al, 2022).

Due to the policy uncertainty and continuity, universities, as quasi-public institutions, also face the challenge of developmental policies to consistently support E&I on the leadership level and the policy implementation across the campus. Although university in the Shenzhen area has ambitions to develop entrepreneurial university as a competitive image, the balance between the traditional function as a higher education institution and the third mission, e.g., entrepreneurship & innovation, remains an issue. For one thing, regarding providing talent resources as the university's primary function, voices from entrepreneurs argue that the current talent development of universities cannot satisfy their needs (Peng, 2021). While R&D talents are considered the first

driving force for the development of high-tech enterprises, 52% of high-tech enterprises in the survey list the lack of talents as the most crucial obstacle to enterprise innovation (Figure 10).

Figure 10. Obstacles that have a greater impact on the scientific and technological innovation in hi-tech enterprises in Shenzhen



Source: Shenzhen Chamber of Commerce-Deloitte.2021

Comparatively, 63% of companies list R&D talent recruitment and development as the biggest challenges in human resource development. It takes at least one year for university graduates to become qualified employees, and innovative talents with both technical and market sense are even rarer (Wang, 2021, Dec 10th, Personal interview). Venture capital institutions in the survey also state that at least 80% of E&I programs from universities are immature for universities may be good at technology innovation and knowledge production in a specific discipline, and they have less sensitivity to market or enough experience in running an enterprise (Sui& Xu, 2022 Jan 10th, personal interview). Currently, feasible cooperation is regarding the university as a testing ground for innovation. Some cutting-edge and early-stage entrepreneurial projects can experiment at a relatively small cost.

Unlike the operation mechanism in the industry, universities face the challenge of enough market sense and follow-up work to turn a single technology into a competitive market product systematically. Many E&I programs from universities are too idealized to apply to the real market. Due to the fundamental function of universities as research-led institutions and emphasis on knowledge production, most technological innovations from universities concern less about their applications to the industry and costs. However, for high-tech enterprises, especially SMEs, whether the R&D can quickly transfer into competitive products and profits is crucial for the survival of enterprises. One of the most apparent barriers in current university-industry collaborations is that for enterprises that need R&D for competitiveness, a cutting-edge technology from the university may be too advanced or expensive, while the

practical value of the technology in industry is still a doubt. A venture capitalist also indicates that some R&D companies emerging from university laboratories are overly dependent on capital investment rather than their capability of marketing profit.

Meanwhile, the cost and investment in R&D are sometimes too ideal to fail. Once the capital investment exit or business financing fails, the enterprise without enough profitability faces the risk of bankruptcy. Meanwhile, for those companies that could afford the R&D, a common approach to collaboration is to build private cooperation with the research team by buying out patents directly or investing in the research team to co-found a new company with the joint partnership. Patent transformation, therefore, does not need to go through the TTO (Huang, Dec 14th 2021, personal interview).

Another important barrier (58%) is the dual pressure of academics to balance academic research and innovation application (Figure 9, Appendix). While most academics have patents in their research field, compared to patent transfer or industrialization, they prefer to spend more time and attention on their academic research, especially those young professors who are pressured to obtain tenure position (Cui & Huang, Dec 14th 2021, personal interview). In the current higher education system, 80% of academic evaluation metrics focus on attaining research funding, research outputs, and publications rather than industrialization. Therefore, even though young professionals have the patent to translate ideas of entrepreneurship, they cannot devote enough time to it.

7. Conclusion and implications of the research

7.1 Conclusion

The purpose of this study is to investigate whether the Triple Helix model applies to the Chinese context and how interactions between the three spheres in the Triple Helix model may vary across the Chinese context in the emergence of entrepreneurial universities. The study analyzes whether regional institutional context will result in different patterns of the E&I model.

The Triple Helix model of university-industry-government collaboration provides an analytical framework to explain the dynamic innovation processes and complex network relationships among universities, industries, and the government. This pilot study applies this framework to examine university entrepreneurship in the Shenzhen area. Although the Triple Helix is a widely recognized heuristic principle for studying interactions among the three innovation spheres, how the model promotes the emergence of hybrid institutions like the entrepreneurial university still lacks the empirical analysis in developing regions. Theoretically, the study attempts to add to the development of the Triple Helix system thinking and innovation studies concepts by untangling the underlying evolutionary process of the Triple Helix

system in promoting E&I in the Chinese context. We examine each innovation agent's function separately with the following analysis of the trilateral collaboration relationships among these entities. The Triple Helix innovation model applies in the Shenzhen area, yet the model is still ongoing for forge and change. Neither is it a pure statist regime where government plays a leading role in driving industry and academia nor a total laissez-faire regime where the industry is a dominant force. The industry and enterprise are the leading force in the current innovation model, whereas the university's role becomes prominent in the evolving process.

Empirically, by utilizing first-hand survey data collected from industrial, academic, and governmental institutions in a representative Chinese city of entrepreneurship to provide additional assessments and analyses of the framework of the Triple Helix system, the study finds that the university-industry-government collaboration positively influences the entrepreneurial environment. The study highlights the prominent role of regional contexts and governance stance in influencing the pattern of innovation and entrepreneurship. We demonstrate that variations in regional context and governance stance affect the innovation system's functioning as well as interaction patterns. By analyzing the individual function of each entity in the innovation system, the study argues that the current key driver of innovation in the pilot study area is industry and is supported by governments and universities.

However, the model is dynamic and evolving for the government's strategy of promoting innovation-driven development and university attempts to put entrepreneurship as a brand and play a prior role in innovative talent resources and technology pioneering. By analyzing the current innovation model's current limitations, the study also stresses the crucial role of synergies in the innovation system. Promoting information exchange and interactions, strengthening trust between partners (e.g. individuals, universities, and industries), and encouraging long-term cooperation and joint development between universities, industries, and governmental agencies can vigorously promote the E&I system.

7.2 Policy implications

This pilot study provides evidence-based insights into STI policies and interactions of actors in the innovation system and the emergence of hybrid institutions such as entrepreneurial universities in the Chinese context. The study has highlighted the gap in policymaking, implementation, monitoring, and evaluation. While government adopts an entrepreneurial and service-oriented stance to provide strong incentives and a series of policies to encourage innovation, the lacks the follow-up monitoring and evaluation of STIs may lead to the misuse and the waste of public resources. Therefore, the

recommendations include establishing a successful policy framework and policy-making process to improve policy alignment, coherence, and evaluation. The missing policy evaluation instrument has hindered the implementation of STI policies. Such hindrances not only hamper innovation policy's effectiveness but also hamper the achievement of overall policy goals.

Another main challenge identified by the study is the lack of awareness of who should be responsible for policy implementation, intra-actors interaction, and corporation. The recommendation, therefore, proposes the establishment of multi- departmental policy instruments to stimulate intra- department interactions, together with policy alignment and coherence in overall development strategies. There is still a lack of tools to tackle tacit knowledge and asymmetric information from different actors in the dynamic innovation system. Therefore, the knowledge exchange and capacity of different actors are critical for sustainable and effective policy-making for development.

The policy instruments also apply to universities with a strategic agenda to develop entrepreneurial universities where university, college, and departmental- level knowledge exchange and intra-actors' interactions are critical to the coherence and implementation of universities' development strategy. Therefore, an unequivocal commitment on the part of the university leadership team, academics as well as students to promoting and implementing capacity building of E&I will be critical. Similarly, understanding the needs and tactic knowledge of other actors(government, industries) in the innovation system will help the university provide intellectual resources to develop to be an entrepreneurial university together with the needs of local or regional development strategies and a pioneering actor in the local innovation system.

Reference

- Bellini, E., Capaldo, G., Edström, A., Kaulio, M., Raffa, M., Ricciardi, M., & Zollo, G. (1999). Strategic Paths of Academic Spin-Offs: A Comparative Analysis of Italian and Swedish Cases. Proceedings from the 44th ICSB Conference, Naples, June :20-23.
- Birley, S. (2002). "Universities, Academics, and Spinout Companies: Lessons from the Imperial". *International Journal of Entrepreneurship Education*, 1(1): 133-53.
- Charmaz.K.(2001).Grounded Theory: Methodology and Theory Construction *International Encyclopedia of the Social & Behavioral Sciences*,6396-6399
- Clark, B.R. (1998). *Creating entrepreneurial universities: Organizational pathways of transformation*. Kidlington, Oxford: Pergamon.
- Debackere K. and Veugelers R. (2005). The role of academic technology transfer organizations in improving industry-science links. *Research policy* 34 (3), 321-342
- Djokovic.D and Souitaris. V.(2008). Spinouts from academic institutions: A literature review with suggestions for further research, *The Journal of Technology Transfer* 33(3):225-247
- EC & OECD (2012). *A Guiding Framework for Entrepreneurial Universities*. European Commission & the Organization for Economic Co-operation and Development
- Etzkowitz, H. & Leydesdorff, L. (1995)*The Triple Helix: University - Industry - Government Relations A Laboratory for Knowledge Based Economic Development* , *EASST Review* 14 (1).
- Etzkowitz, H. (2003). Research groups as quasi-firms': the invention of the entrepreneurial university. *Research policy* 32(1): 109-121.
- Etzkowitz, H. (2008). *The triple helix: University-industry-government innovation in action*. New York, N.Y.: Routledge.
- Etzkowitz,H and Chunyan Zhou.(2018) *The Triple Helix: University-Industry-Government Innovation and Entrepreneurship* (second edition) New York, N.Y.: Routledge.
- Geuna, Aldo & Rossi, Federica,(2011). Changes to university IPR regulations in Europe and the impact on academic patenting, *Research Policy*, Elsevier, vol. 40(8),pages 1068-1076.
- Gibb, A. (2012). Exploring the synergistic potential in entrepreneurial university development: Towards the building of a strategic framework. *Annals of Innovation & Entrepreneurship*, 3(1), 1-21.
- Guerrero, M., and D. Urbano.(2010). The development of an entrepreneurial university. *The Journal of Technology Transfer*: 1-32
- Guerrero, M. & Urbano, D. (2012a). *The creation and development of entrepreneurial universities in Spain: An institutional approach*. New York: Nova Science Publishers.

- K.Chen & M.Kenney.(2007). Universities/Research Institutes and Regional Innovation Systems: The Cases of Beijing and Shenzhen. *World Development* (35) 6: 1056-1074.
- Kirby, D.A., Guerrero, M. & Urbano, D. (2011). Making universities more entrepreneurial: Development of a model. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, 28(3), 302-316.
- Klofsten, M., Heydebreck, P. & Jones-Evans, D. (2010). Transferring good practice beyond organizational borders: Lessons from transferring an entrepreneurship programme. *Regional Studies*, 44(6), 791-799.
- Lawton Smith, H. ,S. Romeo., R. Waters. & Chadwick, A. (2013). Entrepreneurial regions: Evidence from Oxfordshire and Cambridgeshire. *Social Science Information*, 52(4), 653-673.
- Leydesdorff, L. (2012). The Triple Helix, Quadruple Helix , and an N-Tuple of Helices: Explanatory models for analyzing the knowledge-based economy? *Journal of the Knowledge Economy*, 3(1), 25-35.
- Mason, C.&Brown, R.(2014) Entrepreneurial Ecosystems and Growth-Oriented Enterprises: Background Paper Prepared for the Workshop Organised by the OECD LEED Programme and the Dutch Ministry of Economic Affairs. Discussion Paper.
- Markuerkiaga, L., Errasti, N. & Igartua, J.I. (2014). Success factors for managing an entrepreneurial university: Developing an integrative framework. *Industry and Higher Education*, 28(4), 233-244.
- Meira Soares, V. A, Amaral, A. (1999). The Entrepreneurial University: a Fine Answer to Difficult Problem?' In *The Entrepreneurial University: How to survive and Prosper in an Era of Global Competition*, Higher Education in Europe, Vol XXIV, 1.
- National Bureau of Statistics (2018) Approach to annual review from National bureau of Statistics, from <http://www.stats.gov.cn/tjsj/ndsj/2018/indexeh.html>
- Newman.A, Obschonka.M, Schwarz.S,&Cohen.M(2019) Entrepreneurial self-efficacy: A systematic review of the literature on its theoretical foundations, measurement, antecedents, and outcomes, and an agenda for future research, *Journal of Vocational Behavior* 110:403-419
- North, D.C. (2005). *Understanding the process of economic change*. Princeton: University Press
- Ó Shea, R. and Allen, T.J. (2005). "Determinants and consequences of University spin-off activity: A conceptual framework". Working Paper presented at Babson-Kauffman Entrepreneurship Conference
- OECD (2005). *Higher Education Management and Policy*. Special Issue: Entrepreneurship. Volume 17, No. 3. Paris: OECD Publishing.
- OECD (2007), *Eurostat-OECD Manual on Business Demography Statistics*, Paris: OECD Publishing.

- O'Shea, R.P., Allen, T.J., Morse, K.P., O'Gorman, C. & Roche, F. (2007). Delineating the anatomy of an entrepreneurial university: The Massachusetts Institute of Technology experience. *R&D Management*, 37(1), 1-16.
- Pruett, M., Shinnar, R., Toney, B., Lopis, F., Fox, J. (2009). Explaining entrepreneurial intentions of university students: a cross-cultural study. *International Journal of Entrepreneurial Behaviour & Research*, 15(6), 571-594.
- Rainer, K., Wolfgang, D., Erkki, K. (2022). How to make an entrepreneurial state, Why Innovation Needs Bureaucracy, New Haven: Yale university press.
- Ruth, G. (2014). Creating University-based Entrepreneurial Ecosystem: Evidence from Emerging World Leaders, Cambridge, MA: MIT Skoltech entrepreneurial ecosystem report.
- Scott, W.R. (2008) *Institutions and Organizations: Ideas and Interests*. Los Angeles, CA: Sage Publications.
- Scott, W.R (2008) Approaching adulthood: the maturing of institutional theory. *Theory and Society*, 37:427
- Siegel, D.S., Waldman, D., Atwater, L., Link, A. (2004) Toward a Model of the Effective Transfer of Scientific Knowledge from Academicians to Practitioners: Qualitative Evidence from the Commercialization of University Technologies *Journal of Engineering and Technology Management* 21(1-2):115-142.
- Shane, S. (2004). "Encouraging university entrepreneurship?. The effect of the Bayh-Dole on university patenting in the United States". *Journal of Business Venturing*, 19: 127- 151.
- Shane, S. (2004a.) *Academic Entrepreneurship: University Spin-offs and Wealth Creation*, Edward Elgar, UK.
- Shane, S. and Stuart, T. (2002), Organizational endowments and the performance of university start-ups, *Management Science*, 48(1): 154-170.
- Shenzhen Government. (2018) approach to Shenzhen Science & Technology committee annual review of hi-tech industry, from <https://opendata.sz.gov.cn/>
- Shenzhen science & technology committee. (2019), The annual statistic database of Hi-tech corporation, <http://stic.sz.gov.cn/xxgk/sjfb/>
- Smilor, R.W., Gibson, D.V. and Dietrich, G.B. (1990). University spin-out companies: Technology start-ups from UT-Austin. *Journal of Business Venturing*, 5(1): 63-76.
- Triple Helix Research Group website: <https://triplehelix.stanford.edu/>
- UN Habitat (2019). THE STORY OF SHENZHEN, Its Economic, Social and Environmental Transformation, UN Habitat Publishing.
- Wolfe D.A (2005) The role of universities in regional development and cluster formation, In Glen Alan Jones, Patricia L. McCarney & Michael L. Skolnik (eds.), *Creating Knowledge, Strengthening Nations: The Changing Role of Higher Education*. University of Toronto Press. 167--94

- World Economic Forum (2009). Educating the next wave of entrepreneurs: Unlocking entrepreneurial capabilities to meet the global challenges of the 21st century. Geneva: Switzerland: World Economic Forum
- WIPO(2021).Patent Cooperation Treaty Yearly Review 2021,The International Patent System, World Intellectual Property Organization 34, Switzerland.
- Wadhwa,A., Phelps C, Kotha S(2016), Corporate venture capital portfolios and firm Innovation, *Journal of business venturing*, 31:95-112.
- Zhou, C., and L. Lu. 2008. Emergence of the Entrepreneurial University in Evolution of the Triple Helix: The Case of Northeastern University in China. *Journal of Technology Management in China* 3 (1): 109-126. doi:10.1108/17468770810851539

4IR technology adoption in the South African airline industry: Drivers and labour market effects

Alexis Habiyaremye

University of Johannesburg, South Africa;

Lorenza Monaco

University College London

Abstract

The emergence of 4IR technologies has been presented as holding the promise of drastic transformations both in the structures of productive systems and in the concomitant skills composition of the coming decades. While high expectations have been placed on the potential socio-economic benefits of these transformations for the South African economy, empirical evidence of the patterns and the effects of 4IR technology adoption is still scarce. This study aims to delve into the adoption of 4IR technologies in the South African airline industry in order to shed light on its drivers, patterns, constraints and associated effects on labour market dynamics. Using technology organisation environment (TOE) framework we analysed qualitative data collected from interviews with airlines executives and found the adoption of 4IR technologies to be slowly taking shape but lagging behind industry leaders in advanced countries. The main drivers of technological change include the post-COVID-19 strive for safety standards compliance, cost reduction and international competitiveness. The identified main constraints are: availability of commensurate infrastructure, low scale of operations and an unfavourable trade-off between the required capital investments and the low labour costs in a context of high unemployment rates. Regarding labour market effects, job displacement has been limited, but is expected to increase as the transition to 4IR technology intensifies. Growth and environ-

mental sustainability of the airline industry require a technology transition combining a more intensive use of digital technologies, green energy sources and the stimulation of alternatives to carbon-intensive air travel.

Key words: 4IR technologies, labour-augmenting technical change, productive skills, labour displacement, airline industry **JEL classification codes:** J23; L93; O33; Q42

1. Introduction

The emergence of the fourth industrial revolution (4IR) technologies appears to form the peak of economic changes experienced in the course of successive industrial revolutions that the world has gone through (Carvalho et al., 2018). The expected shifts in production and interaction paradigms are estimated to be even more transformative than those produced by previous industrial revolutions. The 4IR involves a combination of connection and collaboration, data and high- efficiency analysis and a fusion of various technologies and human capital in a production system driven by frontier innovations such as artificial intelligence, 3D printing, robotics, biotechnology, blockchain technology, digital devices, smart factories, autonomous vehicles, intelligent machines, big data analytics and virtual reality. The new social environment enabled by the advent of the 4IR is also expected to create dilemmas amongst economic stakeholders and require drastic changes in productive skills, with considerable implications for labour market dynamics. Of particular concern is the threat of work displacement, especially for low-skilled workers, as an increasingly important fraction of human labour might be replaced by machines and automation. It is namely feared that changes in production systems may entail a bias towards highly technical and digital skills, thereby displacing the demand for low-skilled and mid-level skilled occupations (World Economic Forum [WEF], 2020). On the other hand, some analysts expect the expansion of the 4IR to increase production output and shore up economic growth, thereby creating employment and new job opportunities (e.g. Fukuyama, 2018; Keidanren, 2017; Norton, 2017).

Despite high stakes placed on 4IR technologies by numerous observers and analysts, there is hardly any empirical evidence of the effects of their adoption in South Africa. The purpose of this study is precisely to bridge this gap by garnering empirical data through discussions with business enterprises that are involved in the adoption of these technologies in order to gain evidence- based insights into the implications of 4IR adoption for

firm performance, skills intensification, employment growth and sustainability. The current study focuses on the dynamics and impact of 4IR technologies in the airline industry in South Africa. The main questions this case study on 4IR productive skills in the SA airline industry attempts to answer are the following:

1. What are the main drivers of 4IR technology adoption in the airline industry and how does cost-benefit analysis affect the adoption choice?
2. What are the main obstacles/constraints to the adoption of 4IR technologies in the airline industry?
3. In which ways are the adopted technologies transforming the operations in the industry and what are the implications for skills requirement, firm performance, employment, operational & environmental sustainability and social inclusion?

We structure our technology adoption analysis around a customised version of the technology organisation environment (TOE) framework proposed by Tornatsky and Fleischer (1990). Our research approach makes use of rigorous analysis of data collected from first-hand sources in order to shed light on the patterns, drivers, constraints, costs and potential benefits that characterise the adoption of 4IR technologies in the South African airline industry. Implications for employment and labour productivity, industry performance, skills development, energy & carbon footprint, as well as the overall structural transformation is equally explored to derive potential policy recommendations. The research also aims to highlight the role of industrial innovation policies in resolving technology lags & infrastructure bottlenecks as well as providing the necessary support for the adoption of technologies that can foster structural transformation and improve the overall performance of the South African economy.

The rest of this paper is structured as follows: the next section presents key theoretical considerations on technology adoption and diffusion in order to provide a context to the 4IR adoption process. Section 3 explains our methodological approach and study design in more details. The fourth section provides the findings on the drivers and constraints of 4IR technology adoption. Section 5 discusses the views of respondents on the impact of 4IR technology adoption on the labour market dynamics, shifts in productive skills requirements, income distribution and environment sustainability. The final section concludes the discussions and provides key policy recommendations for industry players and policymakers.

2. Theoretical considerations on technology adoption and diffusion

2.1 Skills and capabilities for 4IR adoption in the airline industry

The fourth industrial revolution is defined by Schwab and Samans (2016) as the advent of “cyber- physical systems” that require novel human and machinery competencies in the production processes. For the announced potential of the 4IR technologies to be transformed into actual economic opportunities in South Africa, the different players in the airline industry need adequate technological capabilities that enable them to absorb technologies developed outside the country (Cohen and Levinthal, 1990; Narula, 2004). Those capabilities are usually dependent on the supply of adequate levels and quality of human capital, which enables firms to identify, select and acquire externally developed complex technologies, which require adaptive competence to internalise them and put them to profitable use (Cohen and Levinthal, 1990).

Factors related to the size of the firm as well as its organisational structure and corporate culture can also contribute to influencing its ability to successfully adopt new technologies (Lall, 1992). Similarly, the attitude adopted by firms towards learning is also a critical indicator of technological capabilities (Bell, 1993; Kim, 1997). Likewise, the ability to learn is part of organisational capabilities, as it reflects companies’ attitude towards new ideas, as well as the eagerness to adopt them (Rousseva, 2006). Human capital related capabilities include managerial skills, technical competence of workers as well as collective impact skills and efficiency of employees.

Requisite capabilities also include the ability to marshal the financial resources needed to finance the acquisition (Lall, 1992; Habiyaemye and Ziesemer, 2006). Operational or production capabilities refer to the efficiency with which existing personnel can expedite the operational tasks with current production methods. They include skills in quality control, equipment maintenance as well as adaptive competence enabling the company to manage adversity and respond to external shocks (Habiyaemye, 2021). Linkage capabilities refer to the skills needed to exchange information, resources and technology with subcontractors, consultants, suppliers and technology institutions in order to strengthen the productive efficiency of the firm and enhance its capacity to manage and diffuse its productive technologies (Lall, 1992).

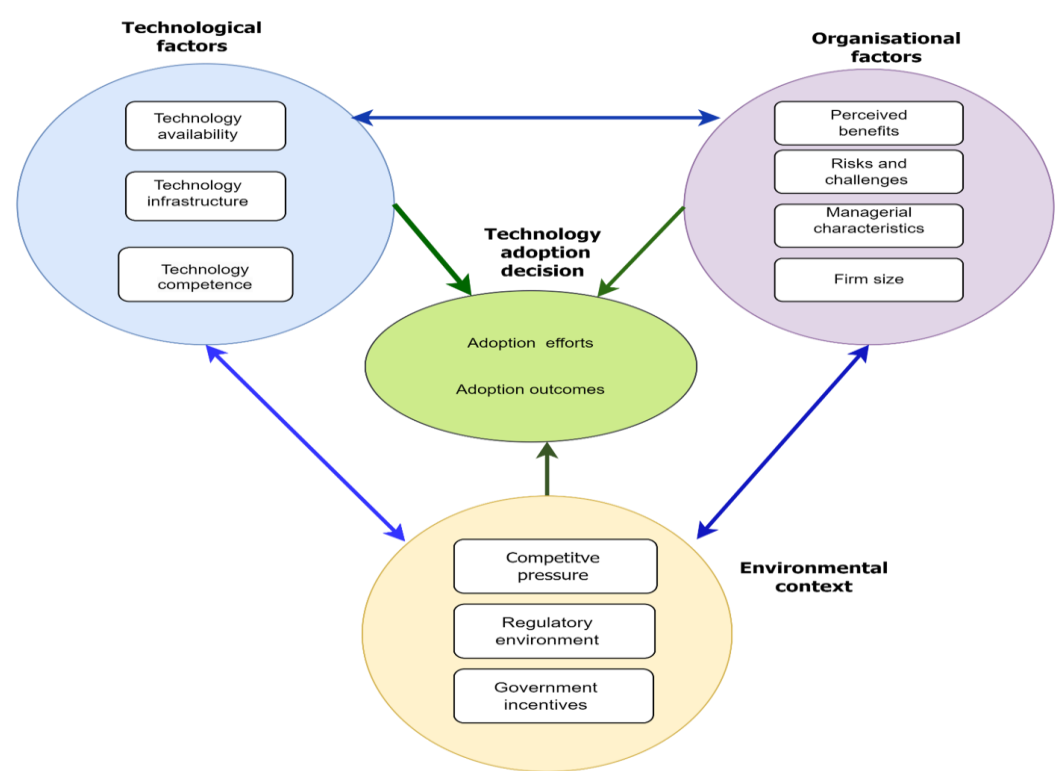
2.2 Technology adoption and diffusion: drivers and obstacles

The literature on innovation diffusion has suggested a number of microeconomic factors that influence the adoption decision at the firm and individual levels and therefore affect both

the rate of adoption and the speed of diffusion. Rogers (1995) proposed five of such factors having direct implications for the adoption and diffusion dynamics: 1: relative advantage over existing alternatives; 2: compatibility with established social norms; 3: implementation complexity of the new technology; 4: trialability (opportunity to experiment with the technology before adopting it); 5: observability (opportunity to see the technology in action in other firms). Other important factors in the adoption decision are financing and adoption costs as well as availability of complementary resources and skills (Hall and Khan, 2003). On the other hand, uncertainty, insufficient political support, weak adoption networks and resistance of already established firms can constitute diffusion obstacles (Foxon and Pearson, 2008). For newly developed technologies to successfully diffuse, the capacity of potential adopters to absorb them in their operations and production systems also plays a crucial role (Cohen and Levinthal, 1990, Narula, 2004, Habiyaemye and Ziesemer 2006). Institutional factors, such as regulatory and market-related policies can equally hinder or catalyse innovation diffusion (Katz and Allen, 1982; North, 1990; Hofstede, 1997; Talegeta, 2014). Another useful approach to examining the patterns, the drivers and the constraints to technology adoption is the technology-organisation-environment (TOE) framework proposed by Tornatsky and Fleicher (1990). It offers a taxonomy of factors affecting the adoption and implementation of new technologies by categorising them into three main contextual dimensions, namely the technological dimension (e.g. technological infrastructures and practices), organisational characteristics (e.g. organisational and managerial attributes), and the environmental context (e.g. competitive pressure, industry characteristics, regulatory frameworks). As represented in Figure 1, these main contextual domains exert a mutual influence in a dynamic process, while their different elements can either foster or hinder the technology adoption process (Aboelmaged, 2014). The TOE framework does not prescribe a fixed set of factors to use for an empirical analysis of the adoption and deployment of new technologies; its strength as an empirical tool resides mainly in its flexibility and adaptability to different disciplines (Aboelmaged, 2014). The present paper applies the TOE framework to investigate the drivers and patterns of 4IR technology adoption in the airline industry in South Africa. In the technological dimension, we base our analysis on two main considerations: technology infrastructure (Weil and Broadbent, 1998) and technological competence (Ifinedo, 2011; Lin and Lin, 2008; Abidinour-Helm et al., 2003). The availability of a common infrastructural platform plays a critical role in fostering the adoption and implementation process (Weil and Broadbent, 1998), while the capacity of an organisation to understand, adopt, use and modify externally developed technological knowledge increases its propensity to adopt and implement new technologies in their operations (Cohen and Levinthal, 1990; Lin and Lin, 2008). Regarding the organizational dimensions, our focus is on

perceived benefits of 4IR technology adoption (enhancing competitiveness, efficiency of operations, cost reduction and operational flexibility, system integration), expected adoption risks and challenges (such as implementation costs, organisation restructuring, compatibility with other organisational processes) and managerial characteristics (such as optimism, innovativeness and attitude towards risk) (Rogers, 1995, Aboelmaged, 2014).(?) While perceived benefits provide the motivation for technology adoption, expected risks and challenges tend to dampen the adoption enthusiasm, while managerial attributes such as innovativeness and risk aversion also play a key role in the adoption decision, especially in its timing relative to their competitors (Awa et al., 1017, Van den Berg and van der Lingen, 2019; Matsepe and van der Lingen, 2022). Less emphasis is put

Figure 1: TOE framework for 4IR technology adoption in the airline industry in South Africa (Source: Author)



on firm size in this analysis since many of the technologies considered for the airline industry are in common use. As for the environmental contextual dimension, we focus our attention on the role of competitive pressure, the regulatory framework and government incentives for 4IR related innovations. Competitive pressure often serves as the impetus for technology adoption, especially if the new technology is perceived as having the potential to increase operational efficiency or enhance the quality of the provided service (Matsepe and van der Lingen, 2022). Likewise

favourable regulations and government incentives can provide a conducive contextual drive to adopt the new technologies, while bureaucratic complexities and lack of government support can act as adoption impediments (Tornatsky and Fleischer 1990, Aboelmaged, 2014). These different influences are schematically summarised in Figure 1 as a customised TOE framework.

2.2 Nature of technological change and effects on labour input

Recent discussions on the potential impact of 4IR technologies have portrayed them as being sufficiently disruptive for their adoption and diffusion to herald a technological paradigm shift (Philbeck and Davis, 2018). This implies that the combination of such technologies is expected to modify, not only the technological and production ecosystems of the coming decades, but also the way our societies shape and manage social interactions as well as socio-economic values. The introduction and diffusion of new products or new production methods in an economy paves the way for technical progress, which constitutes the primary source of long-term economic growth (Schumpeter, 1947; Romer, 1990; Aghion and Howitt, 1992). Technical change leads to output growth through its effects on increasing of factor productivity, i.e. by enabling firms to produce more with the same inputs, or keep the production level constant with reduced input level (Jackson, 1982). Technical progress can therefore be either capital-augmenting or labour augmenting in its short-term manifestations, but in the long run, its asymptotic direction points towards a purely labour-augmenting balanced growth path (Acemoglu, 2003). The associated increase in labour productivity has complex implications on labour-market dynamics as well as the skills requirements to match the new tasks made necessary by new production techniques (Acemoglu and Restrepo, 2018). While the supply shock provoked by labour saving automation systems can often lead to technological unemployment if other sectors fail to absorb surplus labour (see. e.g. Acemoglu and Restrepo, 2017), some theoretical models in this framework, such as the one put forward by Ziesemer (2001), postulate that increased adoption and diffusion of digital and automation technologies have the potential to lead to higher wages and lower unemployment rates in the long run. This study framed the adoption of 4IR technologies in the perspective of labour-augmenting technological change and analyses its labour market effects within the structural transformation framework in the context of a developing economy (see e.g. Schlogl & Sumner, 2020; Andreoni et al., 2021). It is important to highlight that new technology adoption does not function in isolation and only works effectively within a conducive institutional and industrial policy framework (Lall, 1992). Likewise, for the benefits of any technology adoption to materialise and generate tangible socio-economic impact, significant diffusion of the new technology is necessary, as pointed out by Hall and

Khan (2003) in their theorisation of technology adoption, while isolated cases of technology adoption are more likely to lead to increase in income inequality, as illustrated by Lahiri and Ratnasiri (2007).

3. Methodological approach

3.1 Study design

Gaining in-depth insights into the effects of technology adoption on the dynamics of skills requirements at firm level requires a close observation of the actual process and a detailed dialogue of the researcher with the people involved in the process at different levels of decision making and operation of the involved companies. The pertinence of our analysis is therefore dependent on how deep we can delve into the different layers of technology adoption processes in each of the organisation we analyse, rather than relying on the quantitative analysis of statistical data of a large number of firms. For this reason, we took the case study approach and attempted to gather as much pertinent information as possible from each of the firms with which we interacted. Semi- structured interview guides were designed and used to garner key elements of this information with the view to analyse it using a qualitative approach. The analysis of interview data was carried out in light of the background information on technology adoption in general and the dynamics of the airline industry in a global context.

3.2 Sampling

As the study aimed to garner as much information as possible on the 4IR technology adoption process, the sampling targeted the whole population of domestic and international airline companies operating on the main South African airports. We also included Airports Company South Africa (ACSA) as the primary infrastructure provider on most domestic and international airports in the country, because of the crucial role that it plays as the enabler of technologies. Domestic airlines are interconnected through their business association Airlines Association of Southern Africa (AASA), while international airlines coordinate their collaboration through the Board of Airline Representatives on South Africa (BARSA). Interviews were also targeted to the representative of the tourism industry because of their close proximity with the airline industry and the potential influence that the tourism association may have on technology adoption in the airline industry, due to shared customers (tourists travellers form the core of passengers for most airline companies operating in South Africa, according to the board of Tourism Business Council of South Africa[Interview: August 2022]).

3.3 Data collection

Contact information on domestic and international airline companies operating in South Africa including their telephone numbers and email address was obtained from the ACSA portal. After sending emails and making phone calls, preparatory meetings were organised with the head of research and innovation at ACSA in December 2021, which enabled the preparation for accessing the different airline companies that we wished to approach to collect data on their 4IR technology adoption experience. Several preparatory meetings were necessary to get the endorsement and support of the two main business associations for data collection among their members.

Interviews were conducted with local and regional managers of domestic and international airline companies. We also held an interview with representatives of ACSA to get a limpid picture of the technological infrastructure deployed at the disposal of airline companies as a common platform for their own technological systems. Our primary aim was to hold as many face-to-face interviews as possible, also for the sake of making beneficial use of *in situ* observations to gain a deeper understanding of how the adoption of technology is handled in practice. Because of the still ongoing COVID-19 related restrictions at the time of the inception of the data collection process, many interviews had to be conducted via online connection (Zoom and MS Teams applications). Interviews with some of airlines executives were conducted in face-to face conversations when conditions allowed to hold talks at their offices.

4. Drivers and patterns and dynamics of 4IR technology adoption in SA airline industry

Table 1 provides a summary of our findings regarding drivers and constraints to 4IR technology adoption in the airline industry in South Africa, grouped according to the TOE framework dimensions. Those findings are subsequently described in more details following the same taxonomy.

4.1 Technological infrastructure for the SA airline industry

With the technology adoption theoretical frame as a background, we interviewed a dozen executives and managers representing various operators in the South African air transport industry, regarding the expected effects of adopting the newest technologies in their operations. The primary thrust for adopting advanced airport technologies is aimed to increase operational efficiency by minimizing the time that each aircraft remains in the hangar, i.e. out of active operation, and to allow for

Table 1: Summary TOE drivers and constraints of 4IR technology adoption in SA airline industry

Contextual dimension	Favorable influence	Constraint	Unknown
<i>Technological</i>	-Technology availability -Some technologies are fairly developed and already in use overseas -Technological competence: sufficient skills base	-Absence of common advanced infrastructure -Absence of compatible technologies -Some technologies still in pre-deployment maturation stage -Scarcity of technological competence	-New technological infrastructure -Training of local skills to run and maintain highly sophisticated digital systems
<i>Organisational</i>	- Expected benefits -Cost reduction -Operational efficiency -international competitiveness -Sustainable & health friendly solutions -Connection to overseas assets	-Investment costs -Scale of operations insufficient to recoup investments -Adaptive restructuring costs -Financial distress -Financial leverage structure -Centralised decision making	-Investment prudence -Attitude towards risks
<i>Environmental</i>	-Government support -Clear regulations -Post-Covid-19 recovery -Collaboration with external operators -External risks	-Privacy protection -Ownership of digital data -high unemployment \& low labour costs -Immigration laws impeding sourcing of foreign skills	-Labour displacement effects -Environment protection -Development of alternative transport (e.g. HST) to optimise CO2

smoother passengers flows. Major airports in South Africa are equipped with the most modern technological infrastructure for world class navigation and flight control systems to ensure a safe traffic for aircrafts of any modern size, ranging from regional jets to the superjumbo A380. The infrastructure provider (ACSA)'s ambition is to provide world class airport technologies to ensure a seamless air travel to passengers, but remarks from airline operators and ground handling service providers suggest that the technologies they deploy closely follow global leaders in the domain, but cannot afford to be at the cutting edge of the technological frontier because of the limited scale of operations (interview with overseas airline executives, July 2022). Below, we provide a short overview of technologies deployed at the main South African airports along with examples of leading technologies already in use at some leading airports overseas:

1. Booking system

Most respondents we interviewed indicated that their airlines use the AMADEUS global distribution system. Such a system enables user airline companies and travel agents to manage their flight inventory and offer the best deals to their customers. AMADEUS is one of the preferred network-based reservation systems globally.

2. Smart check-in

All major airports in South Africa have shared self-service check-in kiosks allowing passengers to print their own boarding passes and avoid long queues. The airline industry is evolving towards the intensification of smart check-in facilities that incorporate facial recognition and self-service bag drop services. Smart check-in system is designed to allow passengers to drop their bags without having to go to the baggage-drop counter, which can help reduce congestion and delays during busy times. During the covid pandemic, ACSA had intensified the use of touchless screening services with mobile devices to ensure a secure flow for passengers with minimal contagion risks. This kind of technology is expected to save money in the long run through increased check in efficiency and accuracy.

3. Security screening

South African airports have adopted advanced security technologies to help keep airports safe for passengers and public. The use of sophisticated X-ray machines enables airport security personnel to spot any potentially dangerous objects on passengers or in their cabin luggage, while the use of biometrics and facial recognition help to ensure that only authorised passengers and personnel get access to restricted areas. For enhanced security check, the CT scanning technology uses X-rays to create a three-dimensional image of a person's body, which enables security personnel to

easily identify any objects that may be concealed on the body. The expected adoption of advanced scanning technologies could also enable airport security control to relax the restrictions on 100ml liquids and enhance the efficiency of security screening.

4. Baggage handling and tracking system

The automated baggage handling system (BHS) used on South African airports is based on automated departure control systems (DCS) using baggage reconciliation system (BRS) and hold baggage screening (HBS) technologies. Baggage transfer with conveyor belts to and from aircrafts relies on HBS and a sort-allocation computer (SAC) system with tilt trays, which automatically directs luggage to the correct dump on the basis of the allocation obtained from the flight planning module. Actual baggage loading on and unloading from aircraft remains a manual handling operation primarily based on human labour.

The current state-of-the art in baggage tracking systems includes the use of radio-frequency identification (RFID) tracking technology. RFID tags are small computer chips that use radio waves to transmit data, thereby enable the tracking of objects by corresponding readers. They provide an optimal tracking system for passenger luggage on long distance flights, even when it gets lost or stolen in transit. The safety and efficiency of baggage handling can also be considerably increased by the use of robotic loaders as already demonstrated by its utilisation by Saga Airport (Japan).

5. Immigration control & boarding

The current immigration processing system at South African airports relies mainly on immigration officers to control passengers' documents with the help of a centralised computer system. Expansion in the adoption of biometrics and facial recognition technology could enable immigration control to reduce the clearance process time. Such systems use artificial intelligence to identify passengers by matching their photos against records in airport databases. The processing of passengers' documents is thereby rendered more accurate and can be executed faster. As of November 2022, the Department of Home Affairs had launched a pilot project aimed to capture biometric data at South Africa's international airports in order to implement its Enhanced Movement Control System (EMCS) at all ports of entry. This pilot will initially run at Lanseria International Airport, OR Tambo International Airport, King Shaka International Airport and Cape Town International Airport to test its functionality before extending it to other airports.

4.2 Technological competence: Skills and capabilities

A successful transition to the 4IR technology-driven operating environment requires a commensurate adaptation in skills supply in order to meet the needs of such a technological setting. Observations from our key informant interview respondents suggest that for the needs of the air-line industry, training for the new skills needs can be done without disrupting ongoing operations. The set of technical skills required to run the new technological systems became more complex as departure control system and security check moved to digitalised systems. Some of the required skills had to be imported from abroad owing to the limitations of highly technical skills supply from the South African vocational and technical training institutions. The inflow of foreign skilled workers has however remained rather limited because of the complexities of the work permit processing requirements and red tape. As a consequence, some operators report frustrations about their inability to source the needed talents from abroad, while they cannot be easily sourced from local supply either. Technical experts for running and maintaining the AI systems, for instance, are readily available abroad, but find it difficult to obtain work visas even when they are critically needed to run such systems in South Africa. While the issue of importing foreign workers into a country with an effective unemployment rate of more than 40% remains contentious, for the needs of post covid recovery, the discussion needs to shift to marshalling all resources and skills to give a new impulse to the industry. Respondents envision a recovery that will create more jobs, and therefore consider that bureaucratic simplification is needed to enable the inflow of foreign talents required to implement the adoption of complex technological systems, such as AI and big data analytics.

4.3 Perceived benefits and challenges of 4IR technology adoption in the SA airline industry

All international and domestic airline operators that we interviewed indicated to see a significant potential in the adoption of 4IR technologies and singularly in increasing digitalisation to improve travelling experience and increase the efficiency of ground-handling operations. While some of the most advanced technologies have already been tested in South Africa (for example, e-gates at Cape Town International Airport) or in the main hubs of foreign companies operating in South Africa (ex. Schiphol in the Netherlands), the most sophisticated technologies remain on a future agenda. Ultimately, they consider such technology adoption as an essential move to gain international competitiveness. Amongst the technologies estimated to have high transformative potential, a synthesis of insights from our interviews with industry players indicates the following:

- Expanding the use of biometrics, e-gates and face-recognition systems to improve immigration procedures and streamline incoming passenger flows: the three main hubs connecting South Africa to international destinations are currently in the process of implanting the biometrics and e-gates systems, but it is not yet functional. The system is expected to be put in operation in the coming months, but no precise date cannot be given.
- The self-baggage-drop-in system is expected to be implemented at all major South African air- ports, but the migration to fully automated systems will take several years as a result of technology maturation and cost advantage considerations in comparison to existing labourintensive systems.
- The use of big data to archive passengers' profiles and to improve the travelling experience: some major international airlines have indicated to regularly make use of big data analytics but its application is primarily driven by data accumulated in overseas operations.
- The adoption of autonomous vehicles at the airports, for which demand has significantly in- creased during Covid-19 as they present a clear potential advantage for reducing the contagion risk. Effective implementation requires significant investments in appropriate infrastructure, which is not yet available, but is expected to be part of the technological landscape at South African airports if the projected conditions evolve according to ACSA's vison.
- Digitalisation and automation of sanitary checks required since the Covid-19 pandemic started are also expected to contribute to reducing the risk of airport congestions.
- The digitalisation and the increasing automation of cargo handling operations, with particular interest for the digitalisation of goods classification and the potential adoption of 3D scanners to optimise cargo loads: digitalisation in air cargo operations requires related but not identical technologies in comparison to passenger air transport. Increased digitalisation combined with 3D scanning can be expected to yield considerable gains in efficiency and drive logistic costs down in the medium to long term.
- Technologies to drive this digitalisation are mostly embodied in the corresponding machines and equipment, while its implementation requires an alignment with integrated logistics systems. The use of drones for cargo handling is another technological development that

airlines operators expect to intensify in the future. Drones are expected to increase accuracy and efficiency of cargo handling, for example by speeding up the collection of light packages for delivery to or from cargo warehouses.

- The deployment of AI for predictive modelling and management decision support systems for light operations makes it possible to connect all available data and propose centralised solutions. AI and machine learning tools are also deployed in chatbots that provide customised passenger information and can respond to individual queries to help customers solve their flight-related problems.

4.4 Organisational and managerial attributes

Our discussions with industry executives and managers probed their attitudes towards innovation and 4IR technology adoption in their operations. They also sought to gain a deeper understanding of their general expectations about the transformative outcomes of the 4IR technology adoption for their respective companies and for the South African economy. All airline managers and executives involved in this study expressed a high degree of inclination towards innovation and technology adoption. They mostly feel constrained by the limitations of relatively obsolete infrastructural platforms currently in use in South Africa. For some of the airlines based overseas, the local executive we spoke to also indicated the limitations they face in local technology adoption as a result of centralised decision making. The representatives of the state-owned infrastructure provider as well as the executives of the ground handling operations have similarly very favourable attitudes towards innovation and are only limited in their propensity to innovate by the degree of compatibility with local infrastructural platforms. With respect to attitude towards risks, the industry is characterised by a high priority attached to security and the necessity to comply with all relevant aviation regulations and prescriptions., many interview respondents among industry executives expressed their enthusiasm about the new operational possibilities expected to emerge from the mainstreaming of 4IR technologies. Most of them appear to have an inclination for prudence when it comes to new investments in 4IR-related technological and physical infrastructure by making conditional on rational expectation of the long-term evolution of the travel market within regulatory and environmental constraints. In sum, observed managerial attributes are expected to support the adoption of 4IR technologies in a coordinated and environmentally friendly rhythm, which balances the demands of the country's labour market structure and the estimated changes in efficiency of business operations.

4.5 Industry structure and competitive pressure

Before the COVID-19 pandemic, the airline industry in South Africa was characterised by a size-able number of domestic and international operators, with regular scheduled services to and from 3 main international hubs (Johannesburg, Cape Town and Durban), in addition to 6 main domestic service airports (Port Elizabeth, East London, George, Bloemfontein, Kimberley and Upington). Lanseria Airport is a privately owned alternative infrastructure that also offers frequent scheduled domestic flight services. The air transport infrastructure comprises other public and private airports with international standards, but they currently offer less frequent services. The grounding and subsequent liquidation of Comair in June 2022 (which accounted for 40 percent of the country's domestic seat capacity before its financial distress), as well as the financial restructuring of South African Airways (SAA) meant a considerable reshuffle of the domestic air transport industry. In the wake of the crippling restrictions, the industry counts 4 main domestic operators and 36 international operators. Domestic airline companies coordinate their operations through their Airlines Association of Southern Africa, (AASA), while international operators cooperate through the Board of Airline Representatives of South Africa (BARSA). The most important hub for domestic and international flight services remains O. R. Tambo International Airport (JNB) with a total of 40 operators, while Cape Town International Airport (CPT) accommodates all 4 main domestic operators and 13 international operators. As for King Shaka International Airport (DUR), it currently offers services for only 3 of the 4 domestic and 3 international operators. The current recovery trajectory of the airline industry in South Africa reflects post-pandemic trends of global air transport in its ramifications to the country's major international hubs. While the global airline industry experienced a considerable drop in passenger flights and 20% drop in cargo transport when drastic pandemic restrictions were first imposed in April 2020, the cargo transport has recovered most of its lost ground, while domestic passenger transport revenues are still 18.6% below their pre-pandemic levels, according to data from Airport Companies South Africa (ACSA, 2023). Revenues from international passengers have even more ground to recover, being at 35% below their pre-pandemic levels. The envisaged recovery strategies include increased digitisation of airlines and airport service systems, increased interconnectivity and standardisation of digital platforms, resource optimisation through the expansion of self-service tools, green airports that adopt renewable energy sources and optimise energy consumption, more digitalised security systems and transformation of airport into social hubs by diversifying the array of services

offered to passengers. The reduction of the operational capacity of South African Airlines remains a considerable challenge in the effort to restore international connectivity to its pre-pandemic levels, but ACSA's collaboration with international operators is betting on route expansion to achieve the kind of economies of scale that will make it possible to recoup the required investments in commensurate infrastructure.

4.6 Regulatory and institutional framework

The various operations and processes involved in running the airline industry require a coherent regulatory framework because of the security implications they represent for travelers, for the personnel and for the general public. Air transport and airport ground operations fall under the regulatory authority of the Department of Transport (DoT), while the control of emigration and immigration on international travelers fall under the regulatory authority of the Department of Home Affairs (DHA). For the use of biometrics, for example, ethical considerations of who owns and controls access to personal information are critical to the implementation of related technologies while guaranteeing optimal protection of privacy for the users. The implementation of biometrics, face recognition systems as well as e-gates in security control at various airports has to rely on databases controlled by the DHA, with the understanding that the state control guarantees the necessary privacy protection. As an autonomous state-owned enterprise in charge of managing the main country airports, ACSA also plays a pivotal role in the regulatory framework. Any technology change introduced at SA airports has to comply with the stipulations and the procedures of this complex regulatory arrangement. All new technological systems need to obtain the required approval by the regulatory authorities before they can be put into use. Collaboration and coordination between these various entities is therefore essential for the success of 4IR technology adoption. Some of our respondents expressed frustrations about the slow responsiveness and the complexity of bureaucratic procedures in their dealings with the DoT to implement new technologies in their operations (interview with airline executive, August 2022). The institutional setting is generally favourable to business and technological development, with different entities around the Department of Science and Innovation and the Department of Trade and Industry readily equipped to facilitate and support various innovation initiatives in the national economy. As underscored in the introduction, 4IR related technological innovations are particularly encouraged as recommended by the presidential advisory council. Airline operators introducing innovations that contribute to the technological transformation can therefore be confident to find a conducive and supportive institutional environment.

4.7 Exogenous risks

The airline industry in South Africa is currently confronted with various challenges ranging from the slump caused by the Covid-19 pandemic to the threat of high energy prices amid sanctions against Russia for its military operations in Ukraine, the growing social polarisation, the threat of expanding poverty, as well as soaring unemployment rates. Through its innovative technologies, the 4IR is perceived as having much potential to offer solutions to many of these challenges by driving efficiency in production (Fukuyama, 2018; Keidanren, 2017; Norton, A., 2017). However, its differentiated adoption throughout the globe also increases the potential to drive the kind of structural change that may widen the gap between technologically advanced countries and technological laggards, which may lead to job destruction for low-skilled workers and a subsequent worsening of income inequality (Norton, 2017; Schwab, 2017; Springer and Schnelzer, 2019). Changes in production systems may entail a bias towards high technical and digital skills, thereby displace the demand for low-skilled and mid-level skilled occupations. The supply of needed technical skills cannot be assumed to follow automatically as some technical specialisations are critically low and need to be sourced from abroad. This has the potential for employment displacement if no measures are taken to reorient productive skills and adapt them to the changing requirements of the new technological environment. To counter the undesired effects of such shifts in South Africa, where the unemployment rates are very high and the skills distribution still biased towards low digital capabilities, preventive and corrective policies need to be put in place to ensure that the transition to 4IR is not only economically efficient, but also socially inclusive and sustainable.

5. Estimated effects of 4IR technology adoption in the SA Airline industry

5.1 Vision for technology-enabled transformation in the airline industry

One of the most important impacts of 4IR technology adoption, though intangible and not readily easy to measure is undoubtedly its contribution to the societal shifts brought about by the advent of a new era of automated and digitally connected networks of production systems. These systems are estimated to generate positive social changes by transforming the ways and the speed with which firms and households interact, not only in the workplace, but also in the daily behaviour of the private sphere. Most of our respondents envision the adoption of security scanning technology capable of providing 3-dimensional images of the content of luggage as an indispensable evolution, implying a much quicker detection of security threats. Improved efficiency in security control operations is expected to play a key role in accelerating the technology-enabled transformation,

namely through behavioural changes resulting from an improved general sense of security. An integrated security information management system must be built to ensure that the operations run smoothly with minimum human intervention (interview with an airline company executive, august 2022). This in turn will streamline the security control and enable to relax some of the restrictions put on liquids and on electronics equipment in cabin luggage for example (Amies, 2022). A partnership has been concluded with a US company to create a library of images and build AI-based algorithms that will use such a library of images for a quicker identification of threats. The capacity of integrated system to constantly connect the baggage information to the corresponding passenger also reduces the risk of baggage loss. Different airline operators estimate that when the use of the IOT connectivity becomes more widespread with a standardised system, the resulting operating environment will enable the airline industry to enhance security around airports at a time when the use of drones and balloons is increasing the risk of unexpected incidents (interviews with airline executives, July 2022). Blockchain technology could be used to manage the biometrics systems and enable passenger flow to be more efficient and more secure. The vision of the future configuration powered by automation includes the introduction and expansion of self-driving airside car mobility system. Supporting infrastructure will require intensive investments and adequate personnel training to ensure a well-resourced capacity to control the functioning and the maintenance of this infrastructure as well as the associated digital systems. Airline operators are of the opinion that a backup system in the form of traditional mobility and baggage loading equipment with manual handling personnel will remain needed for a long period before the operating environment moves to full automation. To mitigate the risk of fraudulent attempts of those who seek to take undue advantage of automated systems, human control will remain crucial for the success of technological change.

5.2 Operations and production costs

The operational changes that came about with the increased digitisation have been incremental. For instance, some operators find it challenging to adopt the most advanced computer systems for fear of system incompatibility because ACSA airport management computers are still running on legacy systems. In certain cases, they manage to build workaround systems that enable them to bypass some of the inhibiting technical challenges associated with the older systems that they have to interact with in the airport. Such compatibility issues are an inhibitor of technological modernisation, which limits the speed at which operators can move to fully integrated information systems. Since the technological shifts that were made necessary by covid restrictions, the trend has been to strive to an app-based system relaying on information access through tablets and

phones, whereby passengers can more easily access most of the information they need through such apps. Transit and boarding information can be easily shared through such systems to streamline passenger fluxes. In the absence of integrated information systems, transit can remain challenging, especially as regards the transfer of baggage between airline operators, which may sometimes require passengers to retrieve their luggage at transit airports and have to check-in again. In such cases, visa issues may bring about additional complications as passengers may need to obtain a visa to get to the luggage collection points. The introduction of integrated systems not only helps reduce such problems, but can also result in significant cost reduction as the people and resources devoted to dealing with the evoked complication can be redirected to other operational tasks (interview with airline company executives, July 2022). One airline operator asserted that the shift to digital and virtual revenue management system has enabled the company to halve their costs. Artificial intelligence systems centralised in the main (overseas) operation control centres of some international airlines have been successfully deployed in management decision support to optimise operations and minimise costs. Some international operators have also conducted pilot projects to use robots to guide passengers in their overseas hubs. Such technology is also being tried for loading baggage on aircraft with the purpose of using it to make up for expected shortfalls in human resources as the baby boom generation exits the labour market (interview with overseas airline company executive, July 2022).

5.3 Industry recovery and growth

The post-Covid-19 recovery is slowly picking up, but the industry expresses the need for structural measures to support the necessary efforts towards the transition to a technology-enabled operational environment. Such a transition is seen as an important component of the technological innovation strategy aimed at adapting the operating structures to the requirements of the 4IR technological environment. Industry growth is projected to build on advances in high-speed connectivity, a more intensive application of AI in operation management, as well as the diffusion of cloud and quantum computing services (Mc Kinsey, 2019). Immersive reality techniques and machine learning will equally play an important role in the future growth of the sector. Technical constraints still make it difficult for the aviation industry to move to fully electric engines, while the use of renewable fuel sources in jet engines is still in its infancy (Carroll, 2021). The infrastructure provider and ground handling companies as well as airlines executives are however considering an increased use of clean energy in the ground operation fleet (including aircraft towing trucks) to limit the carbon footprint of their operations. The recovery of the airline industry is intimately linked to the recovery of the tourism sector. In July

2022, data from ACSA statistics indicate a recovery rate of 87% in the rate of passengers arriving through Cape Town International Airport in comparison to the pre-covid numbers (Damiyani, 2022). Comparable data for O.R. Tambo International Airport show that the passenger volumes, both international and domestic, are still below their pre-pandemic levels (ACSA, 2023). This sluggish recovery implies a slow adjustment to 4IR technologies as new investments need to be recouped through growth in revenues from higher passenger volumes.

5.4 Change in skills requirements

The introduction of common use self-service terminal equipment has facilitated the efforts to tackle long queues, but has not led to a dramatic change in productive skills requirements among respondent airline operators. A large portion of the operating tasks is still carried out using traditional check-in counters, especially since the self-baggage drop-in system has not been made operational. While future technology uptake is expected to require the acquisition of more advanced skills to ensure an optimal system administration and maintenance, airline and ground handling operation executives expect the current skills mix to persist in the foreseeable future, especially since the labour market conditions with high unemployment rates make automation less profitable than it would be if labour costs were high. Together with government policies that favour labour-absorbing production systems, high investment requirement and low profitability explain why many do not feel the drive to invest a considerable capital in automated systems.

5.5 Labour-market outcomes and implications for income distribution

The introduction and widespread access to common use terminal equipment for self-service kiosks has contributed to increased efficiency of the passenger service and led to cost savings at the various airports where they were put in service. Even as many passengers are able to make effective use these kiosks for self-check-in, some number of airport personnel will remain needed to provide back-up assistance for passengers who may not be very familiar with the digitalised system or in case of malfunction. The expected intensification of biometrics and self-bag drop-off kiosk, if the corresponding use diffuse widely, could lead to shifts in employment and a reorientation of skills requirement towards a higher demand for advanced skills for the control of digital systems, while the demand for traditional skills is expected to slump. Airline and ground handling operators expect that by the time AI is widely adopted to facilitate passenger flow and security management the demand for skills will also move towards more sophisticated capacity to control and maintain such systems, which will take place at the expense of some of personnel

within the current skills mix. Considering the labour market characterised by cheap labour costs and a high unemployment rate in South Africa, robotisation has not been prioritised in local operations. Because of the incremental approach to the introduction of new systems, the different instances of new technology adoption in the airline industry have been structured in a way to enable employees to adapt to operating methods with adequate skills training with only limited employment loss (interview with airline company executive, July 2022). More significant employment loss was due to the heavy impact of Covid-19 on air travel, especially with the resulting collapse of several domestic operators. A change in skills demand that favours more advanced technical skills is likely to result in exacerbating income inequality as the handful of highly skilled workers land in good paying jobs, while the demand for low skilled work falls. Without a reskilling and well-structured labour redeployment programme, the intensification of 4IR technologies in SA would threaten to increase income polarisation in an economy that is already among the most unequal in the world.

5.6 Operational sustainability and environment protection

The air transport is among the most carbon-intensive activities, with an estimated average of 90 kilogram of CO₂ emission per passenger per hour of flight. Emission data for 2020 suggest that the aviation industry accounts for 2.5 percent of global carbon dioxide emissions (Ritchie, 2020). To mitigate the consequences of this intensity, efforts have been invested in the development of fuel-efficient engines as well as aircraft designs. Moreover, investments have been committed for the development of biofuels that can be used on existing aircraft engines without modification. Airline companies renewing their fleet with more fuel-efficient aircrafts or using fuel produced from renewable energy sources are signalling their commitment to reducing their environmental impact. Even though there are obvious limitations in what airline operators can use as alternative forms of propulsion, especially for large, long-haul airliners, the use of clean energy such as green hydrogen batteries and electric engines can still be intensified for airport fleet of vehicles and towing (pushback) trucks. Interview respondents also reflected on the possibility of using a different propulsion system with clean energy for taxiing, which would still contribute to reducing the carbon footprint. They estimate the related fuel consumption to be high enough for significant gains to be made with such an alternative propulsion system. Because of the high costs involved in investing in such new systems, transition is likely to be slow in the absence of a good coordination because the required investments represent a first-mover disadvantage. Increased collaboration among airline operators and coordination with the infrastructure provider is necessary to overcome the first-mover disadvantage hurdle because the transition is estimated to have

positive payoff in the long run. Some respondents indicated to be in collaboration with airport management partners in Germany to examine how some of the techniques used to optimise operation and minimise carbon emissions can be adopted in South Africa.

5.7 Impact of Covid-19 and other threats

Air travel was one of the most profoundly affected activities from the time Covid-restrictions were put in place in the first quarter of 2020. The years of the pandemic have been particularly difficult for the airline industry in general, but South Africa has been one of the hardest-hit as it witnessed the financial distress and the collapse of several of its main domestic airlines companies, including the already ailing national flag carrier, South African Airways. According to IATA estimates, the global airline industry was at risk of losing USD 9.6 billion in revenues and 44 million airline- related jobs for the year 2021-2022 as a result of the Covid-19-related slump. The measures imposed to curb the spread of the virus have brought about profound changes in travel behaviour and have necessitated technological adaptations. Technological adaptation helped mitigate some of the adverse effects of the imposed restrictions. While health authority information system on South African airports was primarily paper-based and caused considerable challenges to airline operators before the pandemic, adaptation measures put in place enabled a more intensive use of touchless passenger control and digital tablets for boarding operations. Whereas hygiene and physical protection measures have been extensively used to compensate for the lack of touchless passenger flow control systems, the progressive adoption of such systems has provided the impetus for an increased digitalisation, which will remain useful long after the covid restriction have been lifted.

6. Conclusion and recommendations

6.1 Concluding discussion

Air transport is one of the industries in which the adoption of 4IR technologies creates great opportunities for considerable efficiency gains and transformative social changes. In the airline industry in South Africa, the pattern and rate of 4IR technology adoption was found to be positively driven by the technological lead observed in advanced countries, the favourable attitude of industry executives towards innovation and the capacity to source the requisite technically skilled personnel either locally or from abroad. However, while the overall distribution of digital skills in the sector is fairly good, high-skilled professionals such as aviation engineers, supply chain managers and specialised data analysts are rather scarce and bureaucratic simplification in critical skills visa processing is desirable to facilitate a seamless sourcing and deployment of the necessary skills for the

technology adoption process. Despite considerable efforts by the infrastructure provider to invest in advanced physical and digital infrastructure platforms, technological infrastructure compatibility still constitutes a non-negligible limitation on the rate of 4IR adoption. While most airline and handling operations executives who participated in the study appear to have a clear picture of the potential benefits offered by such technologies and have seen them in action at major air transport hubs outside South Africa, they mostly consider them as an aspirational goal for their operations in South Africa in the near future. ACSA, the main infrastructure provider, however, has indicated that considerable budgets have already been reserved for infrastructure modernisation, including the deployment of digitalisation and biometrics systems to ensure higher passenger and cargo flow efficiency as well as enhanced security. The cost of investment needed and the centralisation of investment decision making at some airline companies are some of the other obstacles to the adoption of 4IR technologies, considering the large scale of operation needed to recoup such investment. The regulatory environment is not perceived as constituting a sizeable constraint to adoption, although the slow responsiveness of the bureaucratic apparatus is perceived as causing unnecessary delays. As for the environmental factors, practical concerns such as fuel and staff shortages (particularly aggravated by the lay-offs caused by the Covid-19 crisis) financial difficulties (such as the financial distress that crippled SAA and led to the liquidation of Comair) and institutional fragmentation (for example, the lack of procedural harmonisation and data integration between airport authorities and the Department of Home Affairs on immigration procedures). Structural and ethical considerations are an equally important obstacle (ex. data availability following the application of the POPI act). Regarding the labour market effects of 4IR technology adoption, whereas technological change can be expected to lead to shifts in skills requirements and jobs displacement, this study has not uncovered any significant labour displacement resulting from the gradual adoption of 4IR technologies in the airline industry so far. However, various interview respondents expressed the concern that job displacement may occur in the future as technology adoption intensifies. The smooth transition to digital and automated systems was made possible by adapting to the changing skills mix requirement through new skills training and reconversion of existing employees. The most job losses have come from the heavy restrictions imposed on air travel during the Covid-19 pandemic: employment recovery to the pre-Covid-19 levels will be conditional on the post-pandemic rebound of the travel and tourism industries. Although job losses due to the Covid-19 pandemic worsened the already highly skewed income distribution (with low skilled being disproportionately affected), direct testimony from our key informants did not point to any significant job losses to be associated with the adoption of advanced technologies by their

companies. Finally, environmental and climate change consciousness has let may airline operators to direct more attention to the adoption of green energy both for the aircrafts and for the fleet of ground transportation, baggage handling and towing vehicles. Respondents agree that despite the limitations imposed by the requirement of the jet engines for long-haul air transportation, the future of the airline industry must and will be greener through efficiency gains and the diversification of energy sources wherever possible in their ground operations.

6.2 Policy implications and recommendations

In light of global trends in the technological advance of aviation and air transport industries, the South African airline industry will have to make strategic choices, including investments in smart technology adoption not only to overcome existing weaknesses and bottlenecks, but also to create a new impetus for overall profitability and sustainability of the sector. In particular, five areas are identified as being of crucial importance, that require careful consideration:

1. Modernisation of airport infrastructure as a platform for new technology adoption.

The technological content and the amenities of airport infrastructure play a key role in determining the set of technologies airline operators can adopt in their local operations. For the expected recovery of the South African airline industry to harness the benefits of the smartest technologies and be ready for the future of traveling, the infrastructure provider needs to strive for a more profound modernisation of existing airport facilities by adopting the best practices that enable digitalisation and automation. Without adequate technological infrastructure, airline operators could remain constrained in their ability to deploy efficient technologies in their South African operations, even as they have the means to apply them in major overseas hubs which are setting the trends. As such transformation requires considerable investments a careful analysis of the long-term benefits of this transition is necessary to ensure positive returns and spillover benefits.

2. Digitalisation and automation

This will have to cover multiple operations across the board, from self-service check-in and bag-drop to baggage handling in the airport, but also departures and transfers integration as well as cargo services. Advance in these fields may allow seamless passenger flows, avoid congestion and improve the comfort and the security of the travel experience. However, the sustainability of all future innovations will have to be examined from all perspectives, including financial feasibility, the possibility of labour-displacing

effects and the privacy implications of increasing data-handling.

3. Pre-emptive intersectoral skills planning For South Africa to participate fully in the digital revolution that accompany the 4IR, it will have to undertake some fundamental transformation in the area of local digital skills. The relatively low level of available digital skills literacy remains a daunting challenge to the country's aspiration to harness 4IR as a source of technological transformation. While the potential benefits of increased adoption of 4IR technologies are clear, intensification of digitalisation and automation may lead to labour displacement and to changes in skills requirements. In a country with extremely high rates of unemployment, industrial policy will have to anticipate the trajectory of technological change with its employment effects and provide the necessary support for strengthening labour absorption in sectors that are less prone to automation. While a successful adoption of 4IR technologies requires specific skills for which advanced training is needed, reskilling of negatively affected employees is also essential to facilitate their absorption in the so-called automation-resistant sectors. A narrow collaboration between the industry, the SETAs and the Department of Higher Education and Training (DHET) will be necessary to ensure both an adequate supply of skills required to adopt and operate these new technology systems and a reorientation towards new growth sectors.

4. Environmental sustainability

The contribution of the industry to global emissions and its environmental sustainability will be an essential consideration for the sustainability of the sector. The level of competitiveness of individual companies and national industries will be largely determined by how green their future transformations will be. In this regard, smart technological development may help achieve important steps in terms of fuel-efficiency, reduction of current emissions, improvement of waste cycle management and the wider use of clean, green energies. Achieving a greener air travel will also require a shift in travel habits to reduce the passenger mileage, because technology alone will not be enough if the fuel and emission efficiency gains are followed by a growth in the flight hours. The development of alternative travel infrastructure such as the high-speed rail service offers one the most credible alternatives to alleviate the environmental impact due to the crowding of air travel.

5. Policy and institutional coordination

Better institutional coordination and policy alignment will be necessary at several levels. For example, to improve immigration procedures, a closer collaboration between airport authorities and the DHA will become very

important. For matters related to additional sanitary measures, a closer partnership with the Department of Health is necessary to ensure that the protocols that are critical to the protection of our health do not become a drag on the competitiveness of the industry. In order to secure increasing sustainability and adopt smarter, greener technologies, a commitment of all stakeholders from all operations (airlines, ground handling and cargo) will be strictly necessary (including associations like AASA and BARSA). Matters related to infrastructure innovations will require more effective collaboration between all airlines, the Air Traffic Navigation Centre (ATNS), ACSA, and the Department of Transport. For the growth of the industry, important competition matters will have to be solved, with special focus on the issue of licenses on new routes, moving beyond exclusive bilateral agreements and reinstating an efficient licensing department (DoT).

References

- Abdin, M. (2019). Fourth Industrial Revolution and Reality of Industrialisation in Bangladesh. SSRN 3319582. Retrieved from - <https://papers.ssrn.com/sol3/papers.cfm abstractid=3319582>. Date of access: 23 Dec. 2022.
- Abdinnour-Helm, S., Lengnick-Hall, M., & Lengnick-Hall, C. (2003). Preimple- mentation attitudes and organizational readiness for implementing an enterprise resource planning system. *European Journal of Operational Research*, 146, 258–273.
- Aboelmaged, M.G. (2014). Predicting e-readiness at firm-level: An analysis of techno- logical, organizational and environmental (TOE) effects on e-maintenance readiness in manufacturing firms. *International Journal of Information Management*, 34(5), 639–651.
- Acemoglu, D. (2003). Labour- and Capital-Augmenting Technical Change. *Journal of the European Economic Association*, 1, 1-37.
- Acemoglu, D. & Restrepo, P. (2018). The Race between Man and Machine: Implications of Technology for Growth, Factor Shares, and Employment. *American Economic Review*, 108, 1488–1542. <https://doi.org/10.1257/aer.20160696>.
- Acemoglu, D. & Restrepo, P. (2017). Robots and Jobs: Evidence from US Labor Markets. NBER Working Paper No. 23285.
- ACSA [Airports Company South Africa] (2023). Passenger Statistics. <https://www.airports.co.za/airports/or-tambo-international-airport/statistics/passenger>.
- Aghion, P. & Howitt, P. (1992). A model of Economic Growth Through Creative Destruction. *Econometrica*, 60, 323–351.
- Ahn, S. H., & Lee, M. H. (2016). Fourth industrial revolution impact: How it changes jobs. *Korean Academic Society of Business Administration*, 8, 2344–2363.
- Amies, N. (2022). 3D airport scanners could lead to end of 100ml liquids rule. *The Brussels Times*, 11 December. <https://www.brusselstimes.com/333561/proposed-3d-airport-scanners-could-lead-to-end-of-100-ml-liquids-rule>.
- Andreoni, A., Mondliwa, P., Roberts, S., & Tregenna, F. (2021). Structural transformation in South Africa: The challenges of inclusive industrial development in a middle- income country. Oxford: Oxford University Press.
- Awa, H.O., Ojiabo, O.U., & Orokor, L.E. (2017). Integrated technology- Organisation- Environment (T-O-E) taxonomies for technology adoption. *Journal of Enterprise Information Management*, 30(6), 893–921. <https://doi.org/10.1108/JEIM-03-2016-0079>.

- Bell, M. & Pavitt, K. (1993). Technological Accumulation and Industrial Growth: Contrasts between Developed and Developing Countries. *Industrial and Corporate Change*, 2(2): 157-210.
- Boute, R. N., Disney, S. M., Gijsbrechts, J., & Van Mieghem, J. A. (2021). Dual Sourcing and Smoothing Under Nonstationary Demand Time Series: Reshoring with Speed Factories. *Management Science*, 68(2), 1039-1057.
- Carroll, J.P.(2021).World's first passenger flight with 100% renewable fuel takes off – thanks to biotech. Biotechnology Innovation Organization. <https://www.bio.org/blogs/worlds-first-passenger-flight-100-renewable-fuel-takes-thanks-biotech> (9 December 2022).
- Carvalho, N., Chaim, O., Cazarini, E. & Gerolamo, M. (2018). Manufacturing in the fourth industrial revolution: A positive prospect in sustainable manufacturing. *Procedia Manufacturing*, 21, 671-678.
- Cohen, W. M., & Levinthal, D. A. (1990). Absorptive capacity: A new perspective on learning and innovation. *Administrative science quarterly*, 128-152.
- Damiyani, M. (2022). Covid-19-ravaged tourism sector on its way to re-covery, with over 100% increase in domestic trips. News24, 6 September.<https://www.news24.com/news24/southafrica/news/covid-19-ravaged-tourism-sector-on-its-way-to-recovery-with-over-100-increase-in-domestic-trips-20220906:~:text=July%202022%20figures%20show%20an,the%20same%20period%20in%202019>.
- De Backer, K., DeStefano, T., Menon, C., & Suh, J. R. (2018). Industrial robotics and the global organization of production. *OECD Science, Technology and Industry Working Papers*, No. 2018/03, OECD Publishing, Paris.
- Fukuyama, M. (2018). Society 5.0: Aiming for a New Human-Centered Society. *Jpn. SPOT- LIGHT*, Special Article 2 220, 47–50.
- Get smarter (2021). The 4th Industrial Revolution: Will South Africa Be Ready for the Jobs of the Future? <https://www.getsmarter.com/blog/career-advice/the-4th-industrial-revolution-will-south-africa-be-ready-for-the-jobs-of-the-future/>.
- Graetz, G. & Michaels, G. (2018). Robots at work. *Review of Economics and Statistics*, 100(5), 753-768
- Habiyaremye, A (2020). Water innovation in South Africa. Mapping innovation success and diffusion constraints. *Environmental Science and Policy*, 114, 207-229.
- Habiyaremye, A. (2021). Co-Operative Learning and Resilience to COVID-19 in a Small- Sized South African Enterprise. *Sustainability* 13(4), 1-17.
- Habiyaremye, A & Ziesemer, T. (2006). Absorptive Capacity and Export Diversification in Sub-Saharan African Countries. *UNU-MERIT Working Paper* 2006-030.

- Hall, B. H., & Khan, B. (2003). Adoption of new technology. NBER Working Paper 9730.
- Ifinedo, P. (2011). Internet/e-business technologies acceptance in Canada's SMEs: An exploratory investigation. *Internet Research*, 21(3), 255–281.
- Jackson, D. (1982). Technical Progress: Technological Change, the Learning Curve, and Economic Growth. In: *Introduction to Economics: Theory and Data*. London: Palgrave. <https://doi.org/10.1007/978-1-349-16933-7-8>.
- Kagermann, H., Lukas, W. D., & Wahlster, W. (2011). Industrie 4.0: Mit dem Internet der Dinge auf dem Weg zur 4. industriellen Revolution. *VDI Nachrichten*, 13(1), 2-3.
- Keidanren (2018). Society 5.0 - Co-creating the future - <http://www.keidanren.or.jp/en/policy/2018/095-proposal.pdf> . Last accessed on 25 August 2022.
- Khalid, W., Akalpler, E., Khan, S., & Shah, N. H. (2021). The Relationship Between Unemployment and Economic Growth in South Africa: VAR Analysis. *Forman Journal of Economic Studies*, 17(1).
- Kim, L. (1997). *Imitation to Innovation: The Dynamics of Korea's Technological Learning*. Boston: Harvard Business School Press.
- Koch, V., Kuge, S., Geissbauer, R., Schrauf, S. (2014). *Strateg Former Booz Company*. PwC 13, 1–51.
- Lahiri, R., & Ratnasiri, S. (2007). Concerning Inequality, Technology Adoption, and Structural Change. In *Proceedings 2007 Australasian Meeting of the Econometric Society*, Brisbane.
- Lall, S. (1992). Technological Capabilities and industrialisation. *World Development* 20, 165-186.
- Lee, J. (2013). Industry 4.0 in Big Data Environment. *German Harting Magazine*, 8–10.
- Lin, H., & Lin, S. (2008). Determinants of e-business diffusion: A test of the technology diffusion perspective. *Technovation*, 28, 135–145. <https://doi.org/10.1016/j.technovation.2007.10.003>.
- Lorentz, E., Morceiro, P. & Tessarin, M. (2019). Report on the Adoption of 4th Industrial Revolution Technologies in South African Industry. DST/NRF SARChI Chair in Industrial Development, University of Johannesburg.
- Malhotra, N. K. (2010). *Marketing research: An applied orientation*. New Jersey: Pearson Education.
- Matsepe, N.T. & van der Lingen, E. (2022). Determinants of emerging technologies adoption in the South African financial sector. *South African Journal of Business Management*, 53, 2493.
- Min, J., Kim, Y., Lee, S., Jang, T. W., Kim, I., & Song, J. (2019). The fourth industrial revolution and its impact on occupational health and safety, worker's compensation and labor conditions. *Safety and health at work*, 10, 400-408.

- Monaco, L., Bell, J., & Nyamwena, J. (2019). Understanding technological competitiveness and supply chain deepening in plastic auto components in Thailand: Possible lessons for South Africa, IDTT Working Paper.
- Mosconi, F. (2015). The new European industrial policy: Global competitiveness and the manufacturing renaissance. London: Routledge.
- Norton, A. (2017). Automation and inequality. The changing world of work in the global South., IIED Issue Paper. International Institute for Environment and Development (IIED), London, UK.
- Omni (2020). The 4th Industrial Revolution in South Africa. <https://www.omniaccounts.co.za/is-sa-ready-for-4th-industrial-revolution/>. Date of access: 15 November 2022.
- Philbeck, T. & Davis, N. (2018). The fourth Industrial Revolution: Shaping a New Era. *Journal of International Affairs*, 72, 17-22.
- Ritchie, H. (2020) Climate change and flying: what share of global CO2 emissions come from aviation? Our World in Data. <https://ourworldindata.org/co2-emissions-from-aviation:text=Flying%20is%20a%20highly%20controversial,impacts%20on%20climate%20into%20acc>
- Last accessed on 20 November 2022.
- Romer, P. (1990). Endogenous Technological Change. *Journal of Political Economy*, 98, S71-S102.
- Rousseva, R. (2006). Technological Capabilities with Different Degree of Coherence: A Comparative Study of Domestic-Oriented vs. Export-Driven Bulgarian Software Companies. UNU-MERIT Working Paper.
- schloglsummer2020Schlogl, L., & Sumner, A. (2020). Automation and structural transformation in developing countries. In: *Disrupted Development and the Future of Inequality in the Age of Automation* (pp. 51-78). Palgrave Pivot, Cham.
- Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. New York: Harper and Brothers.
- Schwab, K. (2017). *The Fourth Industrial .* Random House LCC US, New York.
- Sima, V., Gheorghe, I. G., Subic', J., & Nancu, D. (2020). Influences of the industry 4.0 revolution on the human capital development and consumer behaviour: A systematic review. *Sustainability*, 12(10), 4035.
- Simchi-Levi, D. (2012). U.S. re-Shoring: A turning point (Annual re-shoring report). Cambridge(MA): MIT Forum for Supply Chain Innovation.
- Soete, L., & Verspagen, B. (2021). Niet alle innovatie hangt direct samen met hogere welvaart. *Economische en Statistische Bulletin*, 106, 67-71.

- Speringer, M. & Schnelzer, J. (2019). *Differentiation of Industry 4.0 Models: The 4th Industrial Revolution from different Regional Perspectives in Global North and Global South*. Vienna, Austria.
- Tornatzky, L. G., & Fleischer, M. (1990). *The processes of technological innovation*. Lexington, MA: Lexington Books.
- Van den Berg, J., & Van der Lingen, E. (2019). An empirical study of the factors affecting the adoption of mobile enterprise applications. *South African Journal of Industrial Engineering*, 30(1), 124–146. <https://doi.org/10.7166/30-1-1992>.
- Venkatesh, V., & Morris, M. (2000). Why don't men ever stop to ask for directions? Gender, social influence and their role in technology acceptance and usage behavior. *MIS Quarterly*, 24(1), 115–139. <https://doi.org/10.2307/3250981>.
- Weill, P., & Broadbent, M. (1998). *Leverage the new infrastructure: How market leaders capitalize on information technology*. Boston, MA: Harvard Business School Press.
- Ziesemer, T. (2001). ICT as Technical Change in the Matching and Production Functions of a Pissarides-Dixit-Stiglitz model. MERIT-Infonomics research memorandum 2001-025.

Do digital payments enhance tax revenue? Evidence from India

Jerome Joseph, N Ramalingam

Gulati Institute of Finance and Taxation, India

Introduction

Payment systems all over the world have been undergoing a radical transformation in recent years. The surveys conducted during the last few years by the central banks of various countries revealed a decline in cash usage and preferences for digital payments by consumers (See Caddy et al., 2020; Foster et al., 2020; Reserve Bank of New Zealand, 2019; Henry et al., 2018). The leading factors for this could be attributed to policy interventions supported by technological advancements and rapid innovations. Moreover, apart from the aforesaid factors, the Covid-19 pandemic too has made a momentous impact on non-cash payments all over the world. The sudden huge shift from cash to digital payments associated with the spread of the pandemic can be attributed to a number of contributing factors such as the closedown of shops, the non-availability of physical cash, and the hesitation to make physical contact (Bank for International Settlements, 2021). The usage of digital modes of payments such as cards and other non-physical modes of transfer witnessed tremendous growth during and after the pandemic period (Pietrowiak. et al., 2021). Thus, the "virus" became a catalyst for the exponential growth of digital transactions. Digitalization, initiated as a matter of convenience and then became a matter of necessity during the pandemic and a major behavioral pattern of payment thereafter.

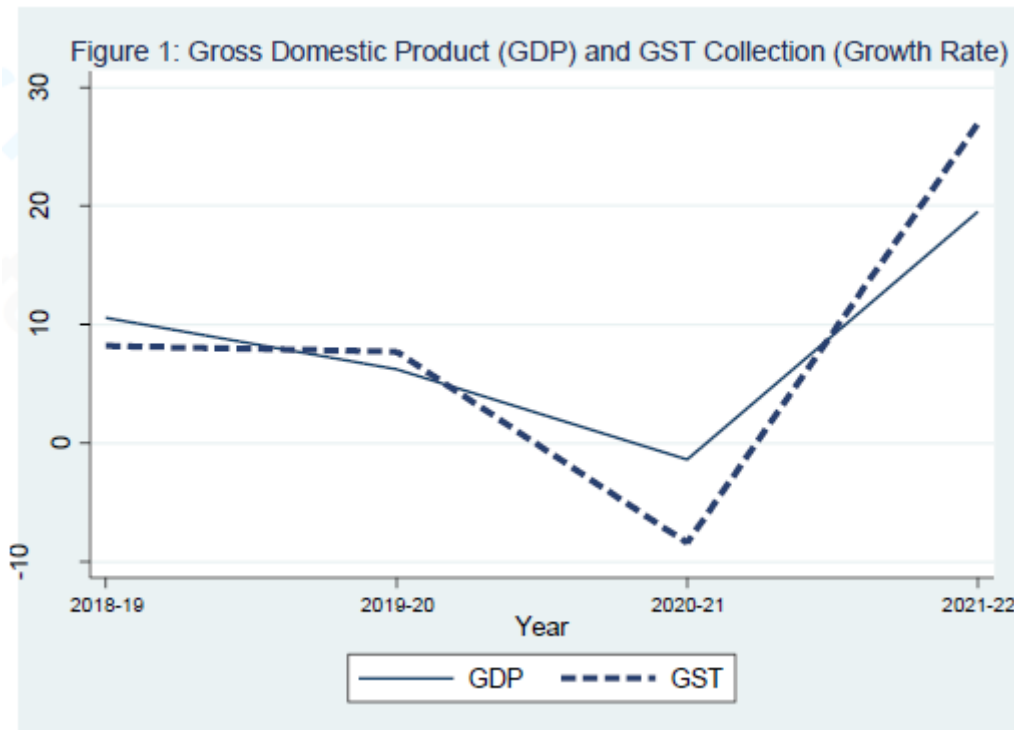
The story of India concerning digitalization is also not a different one. The affirmative action policies of the Government of India including the demonetization of 2016 have envisaged a digitally enabled cashless economy. RBI has designated the period 2010-20 as the digital payments' decade. As a result of all these factors, digital payments in India have shown exponential growth in recent years (See Annexure 1). Besides this, the positive valence of digital transactions has added fuel to this thriving. (Xavier and Zakkariya, 2021).

The pandemic, bottlenecking the normal activities of the economy, curtailed the production and income generation activities and slashed the consumption expenditures, thus, hitting hard on revenue collected from Goods and Service Tax (GST). It was interesting to witness the growth of GST revenue outpacing the growth in the Gross Domestic Product (GDP) of India after the pandemic when the pandemic started to recede (See fig 1). This exceptional growth in the GST revenue can be accredited to the combined effect of increased consumption and improved compliance. Other than consumption what contributed more to the improved GST revenue is the primary question of our inquiry. Various literature suggests a positive relationship between the digitalization of payments and improved tax revenue.

Danchev et.al., (2020) used the method of least squares to conclude that the growth of electronic payments has positively impacted tax compliance in Greece. They portrayed that out of the new nine digital transactions recorded in Greece between 2015 and 2017, at least one corresponded to a previously unrecorded economic transaction. There are also other studies observing an increase in the tax compliance as a result of an increase in the value of payment card transactions that ultimately resulted in the growth of VAT revenue in the Greek context (Hondroyiannis and Papaoikonomou, 2017). Evidences from Europe also show that VAT evasion and the use of card payments are negatively related (Immordino and Russo, 2018). The prime reason identified by them for this relationship is the non-traceability of the cash transactions. Madzharova (2014), enquired whether there is any association between the mode of payment and VAT revenue using the data from 26 European Union (EU) countries for the decade, 2000-2010. He observed a consistently negative impact of cash on VAT payments. It is also important to note that POS (Point of Sale) terminals are shown to affect the VAT-to- consumption ratio positively. Even income taxes are affected by the mode of payment of transactions. Slemrod et al. (2017) find that sales with electronic payments significantly increase the reported income of US sole proprietorships. Apart from tax revenue, the promotion of digital payment, has a strong positive correlation with economic growth, trade, and consumption. (Hassan, et al. 2012).

In the Indian context, Das et al. (2022) find out that increased electronic payments will lead to reporting higher sales to the tax authorities by the firms in West Bengal. However, to the best of our knowledge, barely any literature explores the relationship between digital payments and tax payments, at the national level, especially in the context of GST. This could be attributed to many reasons such as the primitive phase of development of the digital payment environment of India, limited availability of data, and the shorter period after the implementation of GST. In this context of the drastic change in India's transaction or payment architecture favouring the use of non-cash payments, it is evident that digitization of payments is a strong predictor of increased tax revenue. In fact, there exists a wide gap in the literature on this area and we try

to address a fraction of this gap through our analysis. This article thus focuses on the evolving relationship between digital payments and GST Revenue in India.



Source: Own calculation based on data from MOSPI and GST portal

The discussion is organized as follows. Section II will set the theoretical background for the study. Section III will provide a conceptual framework to demonstrate how digital payments, as well as other factors, can increase GST revenue in India while section IV will provide an overview of the digital payment ecosystem in India. In Section V, data and methodology will be presented, followed by the results and their interpretation in section VI. The concluding remarks are presented in section VII.

II. Theoretical framework

Researchers explain the affinity of tax revenue and digital payments in multiple ways. Among them, the traceability of transactions is the paramount one. The non-traceability of transactions, as Immordino and Russo (2018) have argued, will induce the buyer and the seller to enter into a bargain of price discount and the success of this bargain leads to tax evasion. When the transaction becomes traceable, this probability evades. Therefore, by creating digital records and leaving digital traces, digital payments pour light into the otherwise unseen financial lives of individuals and businesses. The digital payment trails make it easier for governments to measure statistics, including the tax and income, for their county.

The increased use by businesses of financial institutions to channel receipts and expenditures provides information not only on those businesses but also on their suppliers and businesses further down the economic activity chain. Similarly, the increased use of credit cards or more sophisticated electronic payment mechanisms provides not only information as to the financial capacity of the purchaser but also information that may be used to confirm the value-added tax (VAT) and income tax filings of the sellers (Bird and Zolt, 2008). In the United States, third-party organizations, including banks, are required to share the information collected by them with the revenue authority and this data is used by the authority to match merchants' sales with reported sales (Treasury Inspector General for Tax Administration, 2011). Even the knowledge of the firms that transactions are recorded by third parties can serve as a major deterrent to evasion and as a tool to diminish the tax gap (Madzhariva, 2014). The availability of checking credit card sales while preparing tax returns has helped to reduce noncompliance to a large extent in Turkey (Dogan, 2011).

Das et al., (2022) estimate that firms report total sales to the tax authorities that are substantially higher than the total electronic sales in their areas. This is consistent with the evidence in Gadenne et al. (2021) that firms report sales that are much higher than their total third-party reported sales and suggests firms pay taxes on more than just their electronic sales. We, therefore, do not expect firms to simply increase their reported sales by the amount of extra electronic transactions they make during demonetization. The increase in electronic transactions may instead increase reported sales if firms suspect that information on third-party verified sales gives tax authorities a better proxy of their true sales, or provides them with more readily available information on the basis of which to start an audit investigation. If, in addition, firms cannot adjust other aspects of their tax returns (the amount of input tax they claim back for example), this increase in reported sales will translate into higher tax payments.

When more people are brought into the purview of formal financial institutions, the tax revenue will eventually rise. Financial inclusion nourishes the interconnection between digital payments and tax revenue. The literature suggests that the broader adoption of digital payments significantly helps the financial inclusion of the disadvantaged sections (Leora and Dorothe, 2014). 'If financial inclusion is the road to bring people to the formal economy, then digital payments are the vehicle that makes the journey possible' (Alim, 2017, p.14). Consumers can demand and perform digital modes of payments only if they are financially included in its broader sense. Financial inclusion leads to better tapping of the benefits of digitalization and digitalization in turn leads to the further deepening of financial inclusion. They are reciprocating the derived benefits. Both these factors together contribute to better tax revenue.

The reduction of the shadow economy is the next factor that relates digital payments with tax revenue. In a study of Europe's shadow economy, Schneider (2013) found that increasing digital payments by an average of ten percent annually for four years can shrink the shadow economy by five percent and shift the behaviour of merchants who underreport sales. This is because the use of electronic payments produces a document of the transactions making it more difficult to participate in the shadow economy. Alim (2017) also provides shreds of evidence of digital payments help reduce the shadow economy and improve adherence to tax laws worldwide.

III. Conceptual framework

Goods and Services Tax (GST) in India was implemented in India from 1st July 2017. One of the main arguments envisaged in the implementation of GST was the self-policing feature of GST returns. It was envisaged that the taxpayers were required only to fill GSTR-1, which recorded the outward supply. GSTR 2, which recorded the inward supplies, and GSTR 3 which involves the calculation of tax payable would be calculated on the basis of filled GSTR-1. This could prevent turnover suppression and evasion to a large extent, But, this mechanism did not become a reality as expected. The technical advancements were not so adequate to make this a reality. Therefore, evasions crept into the GST system too from the old tax regimes.

Improvement of tax compliance mainly depends on the public who should always insist on genuine tax invoices for every procurement of goods and services. This compliance of ultimate consumers at large will voluntarily ensure the compliance of the immediately preceding registered person (retailer) under the supply chain. This impact continues towards the upward direction of the chain and creates a multiplier outcome in tax compliance.

A digital payment by the public will insist his suppliers to record their supply outwards of goods and services. This systemic change reduces the chances of the retailer suppressing the turnover. When the retailers' outward supply is largely accounted for, to match with the recorded outward supply, the inward supply will also be recorded. This, in turn, forces all the registered persons under the supply chain to record the transactions. Thus, the benchmark of improving tax compliance started from the public at large and percolated in the upward chain. This ensures the recording and presenting of transactions by different players in the supply chain who earlier could easily evade tax in the absence of digitalization.

As the standard theory of evasion suggests (Allingham and Sandmo, 1972), evasion has a negative relationship with the probability of being caught. The probability of being caught increases since digitalization leaves digital traces. Even when there is less probability of being caught, the digitalization ecosystem will increase the perception of the evaders that they will be caught. This results

in strengthening the negative relationship between digitalization and underreporting of supply.

The probable evasion that may creep in is the use of non-business accounts for escaping from formalizing digital payments. The intention of using non-business accounts (including bank accounts of relatives and friends) is non-recording or under-recording the transactions and thereby creating a barrier to efficient tax collection. Through cash withdrawals from such non-business accounts, the registered person may opt for unaccounted purchases from the predecessor. However, this practice will not work on a large scale since income tax law restricts cash withdrawal from bank accounts as unexplained debits.

However, there are still ways in which evasion could take place even in the digitalization era. Use of multiple QR codes, non-business accounts, etc to receive payments are a few of them. Still, this evasion is possible only among small traders or transactions. When it comes to large scale, information sharing of banks and other financial intermediaries with various tax authorities can act as a barrier for this type of evasion. The suspicious Financial Transaction (SFT) reports of banks will also act as a barrier to such evasion. Bank transactions are strong pieces of evidence of evasion for tax authorities whereas cash payments leave no such traces. Section 150 of the CGST/SGST requires reporting of the statement of accounts, periodic returns, etc. by various authorities and agencies.

The basic question is whether the taxpayer is ready to pay the tax or not. If yes, then the mode of payment does not matter. Whatever the payment mode, there will be no evasion or turnover suppression. However, this does not happen in real life. So there arises the importance of the mode of payments. The consumer who makes payment tax has two options. Either to make payments through cash or through digital modes. When he chooses digital mode, there is traceability of transactions which increases the chances of being caught (at least increases the perception that he will be caught) which will lead to lesser evasion and increase the tax revenue. Besides this, as we have discussed, digital payments, by deepening financial inclusion and reducing the shadow economy, also increase tax revenue.

However, the mode of payment is not the single factor that affects GST revenue. The literature identifies various other factors that can potentially affect the same. Based on the literature we have identified the following control variables.

1. GST Compliance

Compliance with a tax system is one of the prime elements that can significantly impact the revenue from it. It is simply the taxpayer's decision to comply with the laws and regulations that govern the tax system which results in the timely and accurate filing of returns and payment of the tax. It is tenable

to argue that the more compliant the tax, the more will be the revenue and vice versa. Improving the efficiency of tax collection and tackling tax evasion can increase government revenue and ensure that the redistributive properties of the tax system function (European Union, 2012).

2. Economic performance

When the economy is moving up, the aggregate demand goes up and the tax revenue will go up. Similarly, when the economy is in a recession, the aggregate demand goes down. Bikas and Andrskaite (2013) identified economic situation of the country affects the tax revenue.

3. Inflation

Since GST is an ad-valorem tax charged on the price of goods and services, the tax amount will also move in the same direction along the movement of price. Therefore, inflation is a factor that can potentially affect GST collection positively. Researchers like Imam and Jacobs (2014), have supported this positive influence of inflation on tax revenue. His study was conducted in 12 Middle Eastern countries from 1900 to 2003.

At the same time, some researchers argue that there is a negative relationship between tax revenue and inflation. Taxpayers respond to "tax increases" due to inflation through informal economic activities, underground economies, and tax evasion (Amin et al., 2014). Ayenew (2016) also confirms this negative relationship. Tanzi (1977) has identified the lags in tax collection as a reason for this negative relationship.

4. Exports and Imports

Literature also suggests exports and imports as important factors affecting tax revenue. Amin et al. (2014) also made a significant finding that imports raise revenue while exports reduce tax revenue. Tanzi (1992) argued that the share of imports in the GDP has a positive relationship with the tax revenue. Ahmed and Mohammed (2010) also observed that imports have a positive impact on tax revenue. Anware (2014) also established a positive relationship between tax revenue with exports and imports. The study by Stotsky and WoldeMariam (1997) concluded that the share of exports has a positive impact on tax revenue.

There are a large number of other variables that are identified to be affecting tax revenue. To list a few, GDP per capita, agricultural sector, industrial sector, capital inflow, openness, political stability, corruption, inflation, accountability, and mortality are some of the important economic determinants of tax revenue (Tanzi and Davoodi, 1997; Ghura, 1998; Piancastelli, 2001; Gupta, 2007; Bird et al., 2008; Chaudhry and Munir, 2010; Dioda, 2012; Karagoz, 2013; Castro and Camarillo, 2014; Gobachew, et al., 2017). However, the majority of the above-

said variables are used to capture the country-specific characteristics and become nearly time invariant when considering a comparatively small span of a period of fewer than five years, we only control for four variables explained above.

IV. Digital payment ecosystem in India

India offers a wide range of payment facilities for customers and RBI is considered the official authority to regulate and supervise the payment systems in India. The enactment of the Payment and Settlement Systems Act, of 2007 (Roy, 2021) and the launch of the National Payment Corporation of India Ltd (NPCI) in 2008 for operating retail payments and settlement systems in India were some of the landmarks in the Indian digital payment scenario. Due to all these efforts, the share of digital payments in India has been gradually increasing.

The retail digital payment modes in India can be classified into two categories; Fund Transfer Payment Systems (FTPS) and Merchant Payment Systems (MPS). FTPS includes National Electronic Funds Transfer (NEFT), and Immediate Payment Service (IMPS). These mainly involve money transfers, Government disbursements, etc. Meanwhile, MPS facilitates payments for availing goods or services (RBI, 2022). It includes card networks (Debit cards, Credit Cards) and Prepaid Payment Instruments (PPI). UPI is acting both as an FTPS and an MPS.

V. Data and methods

A. Data

The period of study is from August 2017 to March 2022.

We measure compliance by the percentage of registered taxpayers who have filed GST 3B

¹ on or after the due date. State-wise monthly data of all the aspects of GST collection (Central GST, State/Union Territory GST, Integrated GST, and Cess collected), which add up to the total GST collection of the country, as well as compliance for all the states are taken from the GST portal².

Regarding digital transactions, our analysis takes into consideration only the transactions relating to Merchant Payment Systems and UPI, being the most commonly used digital payment systems, related to the purchase and sale of goods/services. Even though some of the remaining digital payment modes facilitate merchant payments, these are not popular channels for the daily purchase of goods and services (RBI, 2022).

The data regarding UPI transactions were taken from NPCI and those pertaining to Debit cards and Credit cards were taken from RBI. From the Bank

wise ATM/POS/Card statistics available from the Reserve Bank of India, the individual banks' data was added up to arrive at the data for card transactions³. Due to the unavailability of data concerning Prepaid Payment Instruments (PPI), the variable has been ignored in the study.

There are sufficient reasons for including these three payment modes in the analysis. (a). The growth in tax revenue due to increased usage of digital payments is mainly attributed to traders who are small, unnoticed, and transacting in the shadow economy coming to the purview of the formal economy with the increased use of digital modes of payments. (b). The development of UPI significantly contributes to the journey toward a cashless economy by giving confidence to people to go out without taking any cash in hand. This necessitated such traders to accept digital payments to ensure sales. This not only has resulted in the growth of UPI transactions but also facilitates the growth of other digital payment modes. (c). The tremendous growth of online transactions during the past few years which substantially used card payments or UPI for making payments, where tax could not be avoided. The reduced prices and offers they receive through online shopping outweigh the advantage they gain through the bargaining made with the traders for underreporting the sales.

Monthly data with respect to Consumer Price Index (CPI), capturing the inflation level of the country is taken from the MOSPI (Ministry of Statistics and Programme Implementation).

The Gross Domestic Product of the country is the true, convenient, and conventional measure of its economic performance. But due to the limitation concerning the availability of high- frequency data on GDP, we use the Index of Industrial Production (IIP) as an indicator of economic performance that is available on a monthly basis from MOSPI.

The data pertaining to imports and exports of the country has been taken from the data published by the Department of Commerce, Ministry of Commerce and Industry. All the variables except GST compliance (which is expressed in percentages) were transformed into logarithmic form.

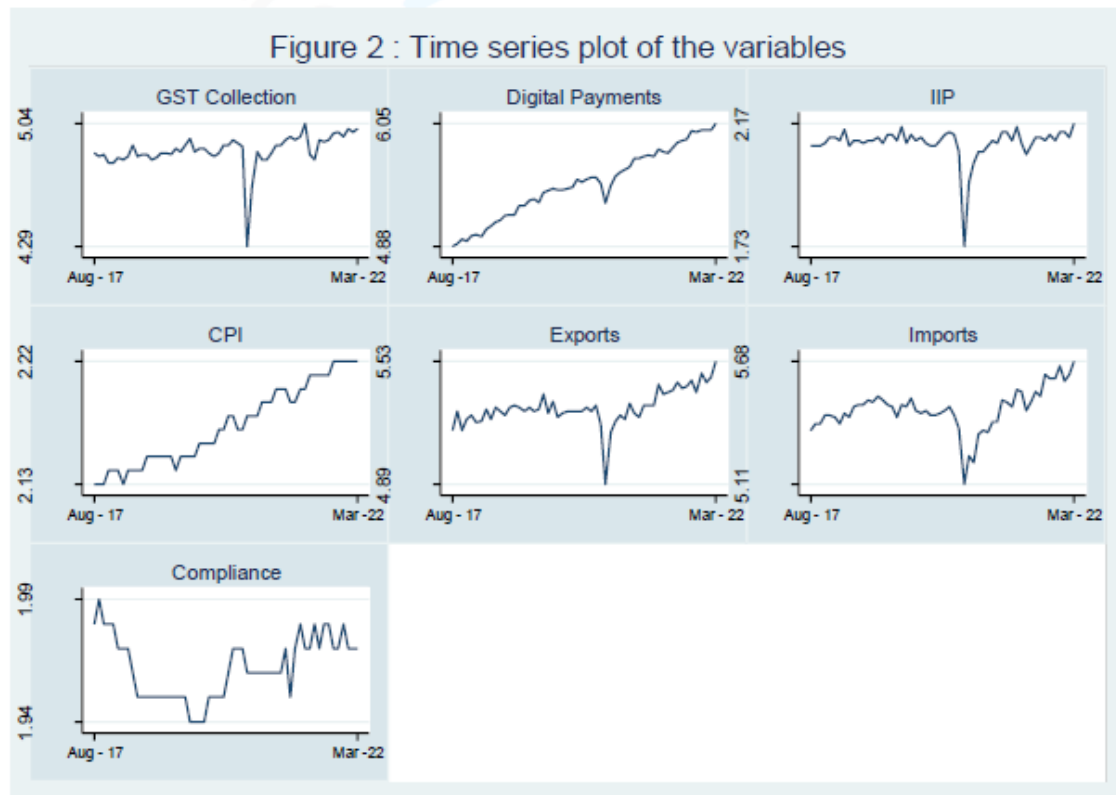
While we expect positive signs for the coefficients of digital payments, imports, compliance, and IIP. We anticipate a negative relationship between GST collection with CPI and Exports

Table 1: Variable Description

Variable Code	Variable Name	Description
gst	GST Collection	Monthly GST collections in India
digpay	Digital Payments	Monthly Value of transactions made through UPI, Debit card POS and credit card POS
Iip	IIP	Index of Industrial Production
Cpi	CPI	Consumer Price Index
Export	Exports	Total exports made by India
Import	Imports	Total imports made into India
Comp	Compliance	The number of eligible taxpayers who have filed GSTR 3B before or after the due date

Table 2: Descriptive statistics

Variables	Obs	Mean	Std. Dev	Min	Max
GST collection	56	4.87	.10	4.29	5.04
IIP	56	2.10	.06	1.73	2.17
GST compliance	56	91.82	2.66	87.65	96.63
CPI	56	2.17	.03	2.13	2.22
Digital Payments	56	5.47	.33	4.88	6.05
Export	56	5.29	.09	4.89	5.53
Import	56	5.45	.10	5.11	5.68



B. Methods

Unit root test

The Augmented Dickey-Fuller (ADF) Test done to check the stationarity of variables explain that the variables are stationary at $I(0)$ or $I(1)$ and not at $I(2)$ (See Table 3). A p-value less than percent indicates that the variables are stationary at a 5 % significance level or else non- stationary. This prevents us from doing a simple Ordinary Least Squares (OLS) regression or cointegration. Therefore, we adopt the Autoregressive Distributed Lag (ARDL) approach, which can be used to establish the long-run relationship between the variables irrespective of whether the variables are $I(0)$, $I(1)$, or a combination of both.

Table 3: Unit Root Test Results (ADF)

Variables	P- Value	
	At Level	At First Difference
gst	0.000	0.000
comp	0.232	0.037
cpi	0.987	0.000
digpay	0.854	0.000
iip	0.001	0.000
import	0.424	0.000
export	0.071	0.000

ARDL Model

Due to the time series properties of the variables and since variables are stationary at different levels (See Table 3), Autoregressive Distributed Lag (ARDL) model is applied to establish the long-run relationship between the variables. The dependent variable is GST collection and the independent variables are digital payments, the Index of Industrial Production (IIP), the Consumer Price Index (CPI), GST compliance, imports, and exports.

An Autoregressive Distributed Lag model is an ordinary Least Square (OLS) based model which could be used when the variables are stationary or of mixed order of integration. This Model takes sufficient number of lags to capture the data-generating process in a general to specific modeling framework (Pesaran and Pesaran, 1997).

The model equation is as follows

$$gst_t = \beta_0 + \beta_1 gst_{t-1} + \beta_2 gst_{t-2} + \gamma_0 comp_t + \gamma_1 comp_{t-1} + \gamma_2 comp_{t-2} + \gamma_3 comp_{t-3} + \theta_0 cpi_t + \theta_1 cpi_{t-1} + \theta_2 cpi_{t-2} + \theta_3 cpi_{t-3} + \alpha_0 digpay_t + \alpha_1 digpay_{t-1} + \pi_0 export_0 + \pi_1 export_{t-1} + \pi_2 export_{t-2} + \pi_3 export_{t-3} + \vartheta_0 iip_t + \vartheta_1 iip_{t-1} + \vartheta_2 iip_{t-2} + \vartheta_3 iip_{t-3} + \omega_0 import_t + \varepsilon_t$$

Bounds test

The bounds test is to check the long-run relationship of the variables and this is done by computing the Bound F statistic. The null hypothesis is that there is no levels relationship and the alternate hypothesis is that there is a levels relationship. When the F-statistic is above the upper band, we reject the null hypothesis that there is no levels relationship and when it is below the lower band, we fail to reject the null hypothesis that there is no levels relationship. If the F statistic is within the band, then we have to know whether the underlying variables are I(0) or I(1).

Levels equation and error correction regression

The levels equation illustrates the long-run relationship of the variables at the level while the Error correction regression explains the short-run dynamics of the underlying variables.

V. Results and interpretation

The results of the F-Bounds test are given in Table 4. Since the F statistic is above the upper bound critical value, it implies that there exists a co-integration or a long-run relationship between the dependent variable and independent variables in the model.

Since there exists a long-run relationship between the variables, the ARDL approach to co- integration is applied. The optimum lag length (k) is determined by Akaike Information Criterion (AIC).

Table 4: F - Bounds test results

Dependent variable: GST collection				
Test Statistic	Value	Significance level	I(0)	I(1)
F-Statistic	13.951	10%	2.139	3.204
Actual Sample Size	53	5%	2.49	3.658
		1%	3.33	4.708

The results of the level equation are given in Table 5.

Table 5: Levels equation

Dependent variable: GST collection				
Variable	Coefficient	Std. Error	t-Statistic	Prob
comp	0.003	0.002	1.85	0.07
cpi	-1.5	0.565	-2.65	0.01
digpay	0.31	0.046	6.63	0.00
export	-0.56	0.226	-2.46	0.02
iip	1.52	0.160	9.50	0.00
imp	0.27	0.125	2.15	0.03
c	4.38	0.929	4.71	0.00

The result of the levels equation confirms our expectations about every variable. The results suggest that digital payment has a positive and significant effect on GST revenue collection. All other control variables have also appeared significant in the model. While compliance, economic performance, and imports have a positive effect, inflation and exports have a negative effect on GST revenue.

IIP can be considered a proxy for the economic performance of the country. Manufacturing or industrial output is expected to be high during the time of economic boom and recession phase bottlenecks the same. So, the Index of Industrial production positively and significantly impact the tax revenue since increased production is accompanied by increased income and therefore increased consumption. A hike in consumption expenditure directly leads to increased tax revenue.

Further, the results also suggest a negative relationship between CPI and GST revenue. In fact, there are a large number of literature suggesting a negative relationship between them (Ayenew, 2016 ; Tanzi, 1977). A possible explanation for the same is that the reduction in demand due to the increased price may outweigh the revenue gain from the same.

Further the results also suggest a positive relationship between imports and GST collection. Import of goods and services are significant sources of GST revenue they are consumed in the country at various stages which will lead to increased tax revenue. The treatment of exports in the GST regime is unique as they are treated as zero-rated goods under which a supplier is eligible to receive back the input tax paid by him to acquire the exported goods or services. In such a way, there is no wonder in seeing a negative relationship between exports and GST revenue.

D. Error correction regression

The estimated error correction model is as follows (Table 6.)

Table 6: ECM Regression

Dependent Variable: GST collection				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(digpay)	0.17	0.19	0.89	0.38
D(digpay(-1))	0.36	0.17	2.12	0.04
D(digpay(-2))	-0.12	0.16	-0.75	0.46
D(digpay(-3))	0.88	0.15	6.01	0.00
D(cpi)	-1.29	1.87	-0.69	0.50
D(cpi(-1))	6.22	2.02	3.08	0.00
D(cpi(-2))	-0.16	2.08	-0.08	0.94
D(cpi(-3))	10.72	1.83	5.84	0.00
D(export)	0.26	0.15	1.78	0.08
D(export(-1))	0.83	0.12	6.99	0.00
D(export(-2))	0.21	0.11	1.91	0.07
D(export(-3))	-0.34	0.10	-3.53	0.00
D(iip)	0.80	0.19	4.17	0.00
D(import)	0.25	0.12	2.15	0.04
CointEq(-1)*	-1.46	0.13	-11.31	0.00

The error correction term, -1.46, is statistically significant at 1 percent significance with a negative coefficient between -1 and -2. This implies dampened fluctuations about the equilibrium path, which implies that instead of monotonically converging to the equilibrium path directly, the error correction process fluctuates around the long run value in a dampening manner. However, once this process is complete, convergence to the equilibrium path is rapid. (See Narayan & Smyth. (2006) for a similar interpretation)

VI. Conclusion

This paper analysed the nature of the relationship between digital payments and tax revenue growth in India. Our empirical analysis finds that digital payments is a strong predictor of GST revenue in India and they are positively associated in the long run. Besides this, other factors that positively influence GST revenue are GST compliance and imports. At the same time, CPI and Exports have a negative impact on GST revenue.

Tax revenue of the government can be enhanced through multiple channels including audits, check posts, account scrutiny, etc. But optimum tax revenue with least complications and effort is the desirable state of affairs. The aforesaid measures have their limitations in that regard. These methods of course will increase the revenue but create pressure on the taxpayers and demands huge work by the tax officials. But digital payments, bypassing this problem,

volunteers augmented and sustainable revenue. Further, it involves zero cost, either from the Government or the taxpayer side. Thus, digitalization proves to be the best tool for voluntary compliance by taxpayers.

At the same time, the mode of retail digital payment is also not an ever-winning solution to mitigate tax evasion; We cannot close all doors of tax evasion. However, digital mode of payment can reduce evasion to a large extent and increases tax revenue. Here again, this mechanism works efficiently only if information is shared by the third parties to the tax authorities and authorities use this information to sustain the tax revenue. Digital payments not only bring tax revenue to the Government, but digital payments at current times is, not a luxury but a necessity for a developing nation like India.

References

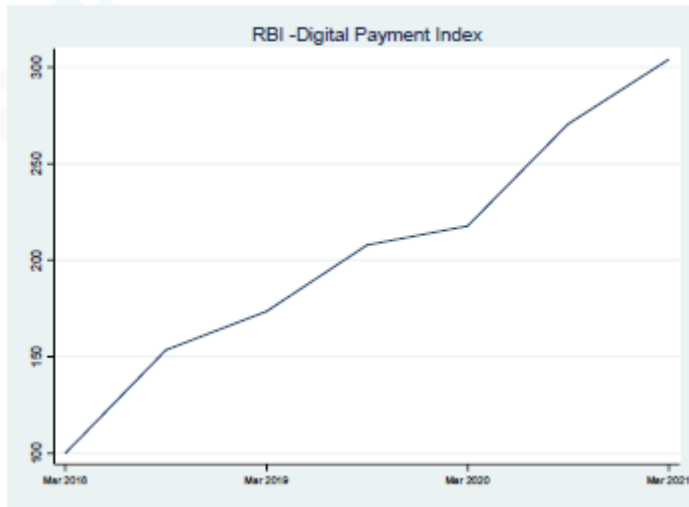
- Amin, A., Nadeem, A.M., Parveen, S., Kamran, M.A., & Anwar, S. (2014). Factors Affecting Tax Collection in Pakistan: An Empirical Investigation. *Journal of Finance and Economics*, 2(5), 149-155. <http://10.12691/jfe-2-5-3>
- Ahmed, Q.M., & Mohammed, S.D. (2010). Determinant of Tax Buoyancy: Empirical Evidence from Developing Countries. *European Journal of Social Sciences*, 13(3), 408-418. <https://ssrn.com/abstract=1583943>
- Allingham, M.G., & Sandmo, A. (1972), "Income Tax Evasion: A Theoretical Analysis," *Journal of Public Economics*, 1(3-4), 323-338
- Alim, M. (2017). Financial Inclusion, Digital Payments and Their Impact on Income and Tax Revenue Around the World. Master's thesis, Harvard Extension School. <http://nrs.harvard.edu/urn-3:HUL.InstRepos:33826588>
- Anware, N. (2014). Determinants of tax revenue performances in Ethiopia: mini research paper.
- Ayenew, W. (2016). Determinants of tax revenue in Ethiopia (Johansen co-integration approach). *International Journal of Business, Economics and Management*, 3(6), 69- 84
- Bikas, E., & Andruskaite, E. (2013). Factors affecting Value Added Tax Revenue. *European Scientific Journal*, 1. <https://doi.org/10.19044/esj.2013.v9n19p%25p>
- Bird, R. M., & Zolt, E. M. (2008). Technology and taxation in developing countries: from hand to mouse. *National Tax Journal*, 791-821.
- Bird, R.M., Martinez-Vazquez, J., Torgler, B. (2008). Tax efforts in developing countries and high-income countries: the impact of corruption, voice and accountability. *Econ. Anal. Policy*, 38(1), 55-71
- Committee on Payments and Market Infrastructures. (2021). *Covid-19 accelerated the digitalisation of payments*. Bank for International Settlement
- Caddy, J., Delaney L., Fisher, C., & Noone. C. (March 2020). *Consumer Payment Behaviour in Australia*. Reserve Bank of Australia Monthly Bulletin
- Castro, G.A., Camarillo, D.B.R. (2014). Determinants of tax revenue in OECD countries over the period 2001-2011. *Contaduria Adm.* 59(3), 35-59
- Chaudhary, I.S. & Munir, F. (2010). Determinants of low tax revenue in Pakistan. *Pakistan Journal of Social Sciences* 30(2), 439-452
- Danchev, S., Gatopoulos, G., Vettas, N. (2020). Penetration of Digital Payments in Greece after Capital Controls: Determinants and Impact on VAT Revenues. *CESifo Economic Studies*, 66(3), 198-220. <https://doi.org/10.1093/cesifo/ifz019>
- Das, S., Gadenne, L., Nandi, T., & Warwick, R. (2022). Does going cashless make you tax- rich? Evidence from India's demonetization experiment. IFS Working paper W22/03
- Dioda, L. (2012). *Structural determinants of tax revenue in Latin America and the Carribean, 1990-2009*. Economic Commission for Latin America and the Caribbean (ECLAC). LC/MEX/L1087. Press in United Nations.

- Dogan, U. (2011). Data Warehouse and Data-Mining Tools for Risk Management: The Case of Turkey In M.S. Khwaja., R.Awasthi, J. Loeprick (Eds.), *Risk based tax audits Approaches and country experiences*. <https://doi.org/10.1596/978-0-8213-8754-2>
- European Union. (2012). Tax reforms in EU Member States 2012. *European Economy*, 6
- Foster, K., Greene, C., & Stavins, J. (2020). *The 2019 Survey of Consumer Payment Choice: Summary Results*. Federal Reserve Bank of Atlanta. <https://doi.org/10.29338/rdr2020-03>
- Gadenne, L., T. Nandi, and R. Rathelot (2021): "Taxation and Supplier Networks: Evidence from India," Tech. rep., Mimeo, University of Warwick.
- Ghura, M.D. (1998). Tax revenue in Sub-Saharan Africa: effects of economic policies and corruption. *IMF Working paper no 98-135*, 1-25
- Gobachew, N., Debela, K.L., & Shibiru, W. (2017). Determinants of Tax Revenue in Ethiopia, *Economics*, 6(6), 58-64. <https://doi.org/10.11648/j.eco.20170606.11>
- Gupta, A.S. (2007). Determinants of tax revenue efforts in developing countries. *IMF Working paper No. 7-184*, 1-39
- Imam, P. A., & Jacobs, D. (2014). Effect of corruption on tax revenues in the Middle East. *Review of Middle East Economics and Finance*, 10(1), 1-24. <https://doi.org/10.1515/rmeef-2014-0001>
- Hasan, I., Renzis, T., & Schmiedel, H. (2012). Retail Payments and Economic Growth. *Bank of Finland Discussion Paper No. 19*. Helsinki, Finland
- Henry, C.S., Kim, P., Huynh., & Welte, A. (2018). 2017 Methods-of-Payment Survey Report. *Bank of Canada Staff Discussion Paper*, 2018-17
- Hondroyiannis, G., & Papaoikonomou, D. (2017). The effect of card payments on VAT revenue: New evidence from Greece. *Economic Letters*, 157(c), 17-20
- Immordino, G., & Russo, F. (2018). Cashless payments and tax evasion. *European Journal of Political Economy*, 55, 36-43. <https://doi.org/10.1016/j.ejpoleco.2017.11.001>
- Karagoz, K., (2013). Determinants of tax revenue: does sectoral composition matter? *Journal of Finance, Accounting and Management*, 4(2), 50-63
- Leora K., & Dorothe S. (2014). The Opportunities of Digitizing Payments. World Bank, <http://hdl.handle.net/10986/19917>
- Madzharova, B. (2014). The impact of cash and card transactions on VAT collection efficiency. *The usage, costs and benefits of cash-revisited*, 521-559.
- Narayan, P.K., & Smyth, R. (2006). What determines migration flows from low-income to high income countries? An empirical investigation of FIJI-US migration 1972-2001. *Contemporary Economic Policy*, 24(2), 332-342
- Pesaran, M. H., & Pesaran, B. (1997). *Working with microfit 4.0*. Camfit Data Ltd, Cambridge.

- Piancastelli, M. (2001). Measuring the tax effort of developed and developing countries: Cross country panel data analysis – 1985/95. *IPEA Discussion paper no 103*
- Pietrowiak, A., Korella, L., & Novotny, J. (2021). Payment behaviour in Germany in 2020 – making payments in the year of the coronavirus pandemic: Survey on the use of payment instruments. Deutsche Bundesbank. <https://www.bundesbank.de/paymentbehaviour>
- Reserve Bank of India. (2020). *Payment and Settlement Systems in India Journey in the Second Decade of the Millennium*. Mumbai
- Reserve Bank of India. (2021). *Monthly Bulletin – April 2021*. Mumbai
- Reserve Bank of India. (2022). *Discussion Paper on Charges in Payment Systems*. Mumbai
- Reserve Bank of New Zealand. (2019). Cash Use in New Zealand - Public Survey 2019 High Level Findings.
- Roy, P.M. (2021). Anatomy of the digital Payment Ecosystem in India. *Bimaquest*, 21(3). 40- 61
- Slemrod, J. (2016). Tax Compliance and Enforcement: New Research and Its Policy Implications. *Ross School of Business Paper No. 1302*. <https://dx.doi.org/10.2139/ssrn.2726077>
- Schneider, F., & Kearney, A.T. (2013). The shadow economy in Europe- Using payment systems to combat the shadow economy. *Johannes Kepler University of Linz*
- Stotsky, M. J. G., & WoldeMariam, M. A. (1997). Tax effort in sub-Saharan Africa. Working paper No. 97-107. International Monetary Fund.
- Tanzi, T., & Davoodi, H. (1997). Corruption, Public investment and Growth, *IMF Working paper 97/139*,
- Tanzi, V. (1992). Structural factors and tax revenue in developing countries: A decade of evidence. In I. Goldin & L.A. Winters (Eds.), *Open Economies: Structural adjustment and Agriculture*. Cambridge
- Tanzi, V. (1977) Inflation, Lags in Collection, and the Real Value of Tax Revenue, *IMF Staff Papers*, 24, 154-167.
- Xavier, P.S., & Zakkariya, K.A. (2021). Factors Predicting Consumers' Continuance Intention to Use Mobile Wallets: Evidence from Kerala, India. *Colombo Business Journal*, 12(1), 114-144. <http://doi.org/10.4038/cbj.v12i1.73>

Annexures

A1. RBI - Digital Payment Index



Source: RBI (2022)

A2. GDP and GST from 2017-2022

Year	GDP (Billion)	GST (Billion)	GDP Growth	GST Growth
2017-18	170900	8103		
2018-19	188997	8767	10.58	8.2
2019-20	200749	9444	6.21	7.71
2020-21	198009	8651	-1.36	-8.4
2021-22	236646	10975	19.51	26.87

Source: EPWRF, GST Portal

Model adequacy tests

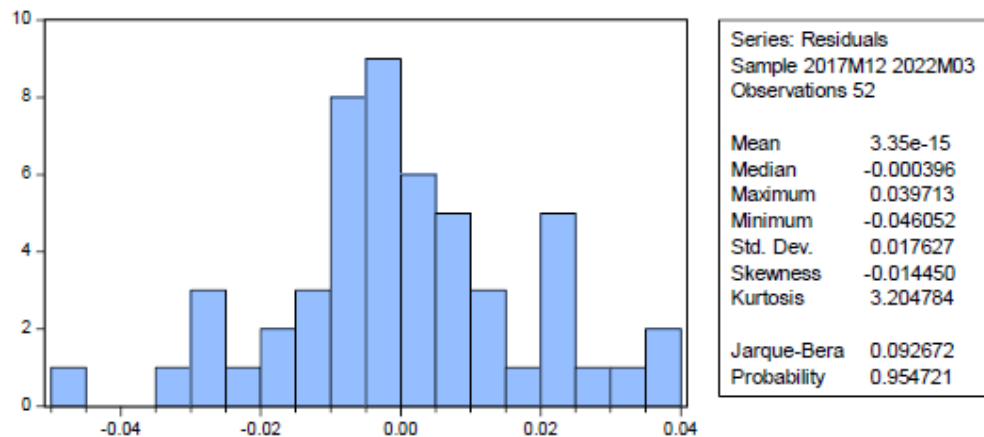
A3. Test for Heteroskedasticity

Heteroskedasticity Test: Breusch-Pagan-Godfrey			
F-statistic	1.605	Prob. F(20,32)	0.115
Obs*R-squared	27.514	Prob. Chi-Square(20)	0.154
Scaled explained SS	10.095	Prob. Chi-Square(20)	0.977

A4. Test for auto correlation

Breusch - Godfrey Serial Correlation LM Test			
F-statistic	0.655	Prob. F(2,30)	0.526
Obs*R-squared	2.326	Prob. Chi-Square(2)	0.312

A5. Test for normality



A6. Test for Multicollinearity – Variance inflation factors

Variable	Coefficient Variance	Uncentered VIF	Centered VIF
GST(-1)	0.02	57830.99	25.69
DIGPAY	0.10	298182.70	861.43
DIGPAY(-1)	0.12	360863.90	1065.13
DIGPAY(-2)	0.09	261486.70	786.22
DIGPAY(-3)	0.09	254432.00	778.45
DIGPAY(-4)	0.05	152922.10	475.71
COMP		0.00	8687.23
CPI		5.50	2561435.00
CPI(-1)		14.38	6690311.00
CPI(-2)		14.03	6518207.00
CPI(-3)		16.56	7679925.00
CPI(-4)		10.36	4797126.00
EXP01		0.05	126804.20
EXP01(-1)	0.04	107345.80	27.02
EXP01(-2)	0.03	76490.41	18.51
EXP01(-3)	0.03	72104.94	16.57
EXP01	(-4)	0.02	43961.09
IIP		0.09	40055.28
IIP(-1)		0.10	42841.85
IMP		0.03	74426.24
IMP(-1)	0.03	84076.80	24.39
C		4.06	399701.70

Notes

¹ GST Return 3B is a self-declared summary GST return filed every month. (Quarterly for QRMP scheme). It is a simplified return to declare summary GST liabilities for a tax period. Even in the absence of any business activities, GSTR 3B has to be filed (nil returns). It is on the basis of GSTR 3B, tax liability is calculated.

² <https://www.gst.gov.in/download/gststatistics>.

³ Monthly data available at Payment system indicators were not comparable from November 2019 onwards.

Divergence and convergence in technological and industrial revolutions

Mark Stephen Knell

NIFU Nordic Institute for Studies in Innovation,
Research and Education, Norway

Abstract

This paper explores the concepts of convergence and divergence within the context of technological and industrial revolutions, analyzing five successive techno-economic paradigms from the industrial revolution to the current digital revolution, with a particular emphasis on the latter. The study argues that while knowledge creation and novelty often result in divergence, international technology transfer and diffusion typically generate convergence. To remain competitive, countries must establish the necessary capabilities and competences to prevent falling behind. Drawing on a Schumpeterian approach, the paper analyzes the dynamics of economic development over time. By highlighting the importance of building necessary capabilities in rapidly changing environments, this paper contributes to our understanding of the cyclical pattern of development in technological and industrial revolutions.

1. Introduction.

Joseph Schumpeter (1939) believed that major or disruptive innovations and the diffusion of these breakthroughs produce a fundamental change in the way things are produced, the types of products being produced, how a firm is organized, and the way people transport things and communicate. These innovations appear discontinuously in groups or swarms, causing a cyclical pattern of development. Each cycle or long wave transforms the technological paradigm, but also the entire economy and society. Schumpeter highlighted the temporal nature of entrepreneurial profit and the disequilibria behavior of the entrepreneur in spreading technology over the course of the cycle.

This paper studies convergence (economic catching up) and divergence (falling behind) over the course of five technological revolutions. It does this by weaving together ideas from both the history of economic thought and economic history perspectives. Building on Schumpeter's work, Freeman and Perez (1988) identified five successive techno-economic paradigms dating back to the time of Arkwright's first cotton mill and Adam Smith's *Wealth of Nations*, up to the current digital revolution. This fifth technological revolution, which began with the creation of the first commercially available microprocessor, has been ongoing for the past fifty years.

Knowledge creation and novelty tend to produce divergence, while the international technology transfer and diffusion generates convergence. From the historical perspective, falling behind (divergence) is a much more frequent phenomenon than catching up Abramovitz (1986). Changes in techno-economic paradigm suggest countries must build the necessary capabilities and competences to remain competitive. The most prominent examples of catching up include the United States in late 18th and 19th centuries, western and central Europe in the 19th century, Nordic countries in the late 19th and early 20th centuries, and East Asia in the late 20th century and early 21st century. These countries achieved rapid rates of productivity growth by learning to use and improve existing technologies and to integrate them into their productive activities in a technology revolution. Every techno-economic paradigm has a core technology underlying the dynamics of convergence and divergence. In each paradigm, we see instances of convergence and divergence.

We organize the paper in the following way. The following section surveys divergence and convergence in economic thought. The classical economists understood that capital accumulation has negative and positive feedbacks, and endogenous growth theories are important for generating convergence and divergence over time. Section three briefly summarizes long waves in economic development. The paper then returns to the twin issue of issue of divergence and convergence as part of the discussion on technological revolutions, techno-

economic paradigms, and great surges of development. Section five gives a historical overview of the five technological revolutions, or great surges. Focus then centres on convergence and divergence in the digital revolution. A closing section presents some concluding thoughts on development and Underdevelopment in the history of economic thought.

2. Divergence and convergence in economic thought

The issue of wealth and technology distribution has been a subject of concern for economists and philosophers throughout history. Plato proposed a solution for addressing wealth disparity in the Athenian world by advocating for common ownership of all property, whereas Aristotle suggested that a significant portion of property could be privately owned but redistributed through a system of justice. Eventually, the Aristotelian perspective prevailed with the advent of capitalism. However, the subsequent rapid growth spurred by capitalism resulted in a substantial imbalance in the global distribution of wealth and technology. Between 1000 and 1820, per capita income experienced a modest 50 percent increase, while from 1820 to 2000, it grew eight-fold. This dramatic disparity led to a staggering per capita income gap of approximately 20 to 1 between the developed nations of Europe and the least developed countries (Maddison, 2006). Consequently, economists made it a primary objective to study the key issues surrounding economic growth, encompassing its origins, patterns, and consequences.

Convergence analysis has emerged as a central concern within the field of growth theory, as the phenomenon of countries falling behind economically occurs more frequently than catching up. This raises significant apprehension regarding the ability of growth theory to provide effective policies to reverse this trend. The economic literature has engaged in an extensive and ongoing debate regarding various aspects of convergence. While some models offer support for the idea of convergence, others suggest the contrary. However, all these models emphasize the crucial role of technological and institutional changes in fostering sustainable growth. As these debates persist in the growth theory literature, a substantial body of research has emerged, focusing on empirically testing the validity of these growth theories. This extensive literature, characterized by conflicting interpretations of long-term economic performance, may explain the lack of a clear connection between theoretical models and economic policy guidelines (Temple, 1999).

Another area of focus in the literature pertains to catching-up and technology gaps. It is important to note that convergence does not necessarily entail catching up, as highlighted by Abramovitz and David (1996). Barro and Sala-i-Martin (1992) define catching up as the potential to narrow the technology gap by enhancing the productivity level of a less developed nation, while convergence pertains to a decrease in the dispersion of relative productivity

levels among a group of countries over time. Empirical studies examining the technology gap often adopt a descriptive approach, encompassing a diverse range of technology indicators and frequently emphasizing historical analysis (Fagerberg, 1994).

A significant portion of these studies center around the concept of "appropriate knowledge, experience, and skills" as defined by Richardson (1972), which are crucial for firms and organizations to introduce innovations in the form of new products, processes, and organizational structures. Alternatively, they may examine the set of national characteristics that impact economic growth, as described by Abramovitz (1986), encompassing factors like the overall level of education and technical proficiency, institutional arrangements (both formal and informal), and political influences on risk and uncertainty (referred to as social capability). Interestingly, some of these studies employ similar econometric techniques employed in convergence analysis, such as the nonlinear framework introduced by Verspagen (1991).

One important implication of the neoclassical assumption of diminishing returns to capital is its tendency to lead to convergence. However, this convergence can be counteracted by the presence of increasing returns, which are often incorporated into the model through the inclusion of knowledge or human capital. These increasing returns or externalities sustain perpetual growth by preventing the marginal productivity of accumulated factors from declining to zero. The incorporation of such offsetting factors is a key characteristic of endogenous or new growth theory. In order to make growth endogenous, economists have also modified the relationship between savings and investment (Kurz and Salvadori, 1999).

The theoretical debate on convergence traces its roots back to the research conducted by Ramsey (1928) and Solow (1956, 1957, 2000). Solow (1956) developed a growth model based on standard neoclassical assumptions to demonstrate that competitive markets, characterized by substitution and flexible wages, result in the full employment of labor and capital. Using a production function with constant returns to scale that incorporates labor and capital, Solow illustrated the existence of a unique and globally stable growth path towards which labor productivity and income per person tend to converge. Policies aimed at enhancing growth have only temporary effects during the transitional phase. Productivity growth is restricted by diminishing marginal returns on capital unless there is labor-augmenting technological progress. The rate of technological progress, given exogenously, will also converge to the growth path as the model assumes that knowledge is a public good with costs associated with its acquisition but freely accessible. Consequently, the capital-labor ratio (capital intensity) and real wages will converge to a steady-state position over time if all countries share the same preferences (saving rate) and technology. Starting from diverse initial

conditions, less developed countries experience convergence due to their lower capital intensity and higher rate of profit.

If we assume that technology is a public good and the saving rate is uniform across countries, we can anticipate beta-convergence occurring between poorer and wealthier nations. Additionally, if the production function follows the Cobb-Douglas form, national income growth will unequivocally be higher. However, when we relax the assumption that all economies possess identical preferences and technology, convergence becomes conditional on the extent to which the less developed country is comparatively further from its steady-state position. The hypothesis of sigma-convergence relies on the presence of distinct saving rates between two countries, with each country gradually converging towards its own steady-state value at a rate inversely related to its distance from that steady state. According to this model, different saving rates among countries should result in varying levels of income per person.

The empirical literature employs various econometric techniques, including time series, cross- section, and panel data analysis. A significant portion of this analysis is based on the concepts of beta- convergence and sigma-convergence among countries (Solow, 2000). Beta-convergence refers to the phenomenon where less developed economies grow at a faster rate compared to wealthier ones, enabling them to narrow the gap in terms of per-capita income. On the other hand, sigma-convergence occurs when the dispersion of per capita income across a group of economies, measured, for instance, by the standard deviation of the logarithm of per capita income, decreases over time (Quah, 1996). Beta-convergence is considered a necessary condition for sigma-convergence, although it alone is not sufficient to ensure it.¹

Endogenous growth models predominantly focus on the production and utilization of new knowledge, taking into account factors like increasing returns, non-convexities, and monopoly power. These models incorporate a technology parameter that captures the influence of innovative ideas through research and development and human capital. In contrast to the exogenous neoclassical models that predict convergence, many endogenous models propose that countries follow their unique growth path. Kaldor (1957) and Arrow (1962), for instance, introduced learning effects as a source of technical progress in their models, leading to endogenous growth. The presence of increasing returns or externalities sustains perpetual growth by preventing the marginal productivity of accumulated factors from declining to zero. To achieve endogenous growth, the model must either eliminate or compensate for the assumption of diminishing marginal returns, thus allowing for sustained productivity growth.

The concept of endogenous growth can be traced back to the time of Adam Smith (1776), who recognized the presence of both negative and positive

feedback mechanisms in capital accumulation. Smith identified that competition for capital led to negative feedback through the decline in the rate of profit, while market growth and division of labor created positive feedback through increased labor productivity. Enhancements in worker skills, time-saving from task specialization, and the introduction of new products and processes all contribute to higher labor productivity and establish a competitive advantage that follows a specific path. Kurz and Salvadori (1999) note that Smith anticipated two key ideas in modern new growth theory: (1) technological change is endogenously determined by the economic system, and (2) new technical knowledge ultimately becomes a non-rival and non-excludable public good. Von Tunzelmann (1995) highlights Smith's anticipation of human capital theory and the understanding that learning by doing, learning by using, and formal scientific learning all play crucial roles in the growth process. Additionally, David Ricardo (1817) recognized that non-accumulable factors of production create negative feedbacks that limit growth, while positive feedbacks stemming from technological advancements can counteract diminishing returns on land.

Endogenous growth models explain growth by making technological progress endogenous. These models address the negative feedback resulting from diminishing returns to physical capital by incorporating externalities associated with increasing returns into the production function. They place emphasis on the generation and utilization of new knowledge in the presence of increasing returns, non-convexities, and monopoly power (Amable, 1994), recognizing both the public and private nature of technological progress. The early models introduced by Romer (1986) and Lucas (1988) proposed that varying growth rates arise from the positive feedbacks generated by external technological learning (diffusion), which can outweigh the negative feedbacks arising from diminishing marginal returns within firms. Subsequent models developed by Romer (1990), Grossman and Helpman (1991), and Aghion and Howitt (1993) explicitly incorporate the role of monopolistic market structures in the creation and dissemination of knowledge.

A separate branch of literature stemming from endogenous growth theory delves into the study of "general purpose" technologies that have substantial, widespread, and enduring effects on productivity, such as electricity, steel, and microprocessors. Researchers like Bresnahan and Trajtenberg (1995), Helpman (1998), and Lipsey, Carlaw, and Bekar (2006) explore various aspects of these technologies and their relationship to endogenous growth, particularly in models involving expanding product variety. These models expand upon the works of Lucas (1988) and Romer (1986, 1990) by explicitly examining the interplay between breakthrough technologies and complementary technologies (Faccarello and Kurz, 2016). A crucial aspect relevant to catching-up and convergence is the diffusion of these general purpose technologies, which has played a significant role in Asia (Helpman and Trajtenberg, 1998). Rosenberg

(1963) described the positive feedback loop that drives technological convergence and collective learning, exemplifying this phenomenon in the machine tool industry as a general purpose technology. Bresnahan (2019) further illustrates how this convergence and learning process applies to the digital industries. The concept of upstream firms capturing the benefits of economies of scale is readily apparent, forming the core idea behind the general purpose technology narrative.

3. Long waves in economic thought

Prior to Schumpeter, several economists such as Hyde Clarke, W.S. Jevons, Clement Juglar, John Bates Clark, Knut Wicksell, Vilfredo Pareto, and Mikhail Tugan-Baranovsky had already hinted at the existence of long-wave theories. In the 19th century, these economists explored various factors that could potentially explain the long cycle, including climate, demographics, land speculation, sunspots, famines, secular stagnation, and the likelihood of crises. Tugan-Baranovsky, for instance, identified long-term patterns within the normal business cycle, noting extended periods of rising and falling commodity prices and the expansion and contraction of trade occurring at approximately fifty-year intervals (Barnett, 2001). While Marx and Engels were not direct precursors to long wave theory, their observations of frequent periods of unrest, such as the global crisis of 1857-58, secular stagnation, and extreme inequality during their lifetimes, contributed to the evolving understanding of long-wave phenomena.

In the 1920s, Nikolai Kondratiev published several papers on long wave cycles, which garnered significant attention. The immediate impact of his work was evident through two English translations, followed by comprehensive versions published posthumously in 1979 and 1984, as well as a collection of his works in 1999. Kondratiev drew inspiration from Marshall's approach to competitive equilibrium, Marx's concept of periodic renewal of fixed capital, and Tugan-Baranovsky's endogenous theory of the business cycle. While Kondratiev used the term "long wave cycle" to emphasize their cyclical nature, others referred to them as long waves, long cycles, or long swings. However, Kondratiev faced criticism during his lifetime, particularly regarding his use of statistical and econometric techniques. In response, Kondratiev addressed some of these concerns in a methodological paper on reversible dynamic processes published in 1925. Interestingly, Trotsky expressed a belief that the downward phase of a Kondratiev cycle signaled the decline of capitalism, sharply criticizing Kondratiev's interpretation of equilibrium and self-reinforcing mechanisms in economic development (Trotsky, 1973: 273).

Kondratiev's (1979, 1984) analysis of long wave cycles identified three distinct phases: expansion, stagnation, and recession. Through the examination of various price indicators such as wages, interest rates, raw material prices, coal

and pig iron production and consumption, foreign trade, and bank deposits, Kondratiev dated two complete long cycles. The first cycle encompassed an upswing period from 1789 to 1814, characterized by the first Industrial Revolution, followed by a downward swing from 1814 to 1849. The second cycle consisted of an upswing phase from 1849 to 1873, marked by the age of railroads, steam, and steel, followed by a downswing from 1873 to 1896. The third cycle began in 1896 and ended in 1920, representing the era of electricity, chemistry, and motors. Kondratiev utilized the method of least squares, employing both linear and third-order nonlinear trends, to estimate the theoretical curve based on these data points.

The Kondratiev model distinguished between two types of cycles: normal cycles of fixed capital replacement and extraordinary waves of investment in infrastructure that could span around 50 years. These long swings represented periods of intensified reinvestment in fixed capital. Kondratiev expanded upon Marx's concept of periodic reinvestment by introducing the idea of diverse capital goods with varying degrees of durability. The theory also incorporated an endogenous mechanism, with certain variables arising as a consequence of the system's rhythm rather than being inherent to it. While radical technological change, wars, revolutions, and changes in gold production and monetary circulation were seen as exogenous factors, Kondratiev believed that technological advancements were influenced by endogenous forces within capitalism. Furthermore, he recognized that significant societal transformations occurred prior to the onset of a long upswing.

Schumpeter (1935, 1939) used the term "Kondratiev cycle" to describe the long wave in his work on business cycles. The model emphasized the clustering of radical innovations as the driving force behind the endogenous long-investment cycle (Silverberg 2003). While primarily serving as a descriptive tool, the model aimed to explain cyclical patterns observed over time (Schumpeter 1939: 177). It relied on capital-goods dynamics, investment clustering, and self-reinforcing behavior to generate oscillations. Schumpeter's historical analysis focused on key industries such as cotton textiles, railroads, steel, automobiles, and electric power in the United States, Britain, and Germany between 1790 and 1920 (McCraw 2007). The narrative highlighted the uneven distribution of innovative activities across countries, industries, and periods, with the evolving technology of the time serving as a defining characteristic of the cycle.²

4. Technological revolutions, techno-economic paradigms, and great surges of development

Freeman and Perez (1988) and Perez (1985) introduced the concept of the techno-economic paradigm as a response to the ahistorical nature of Schumpeter's theory. Their work focused on macro-level theories that linked cyclical patterns

of technological evolution with path dependence and structural and institutional changes. Perez (1985: 443) defined a techno-economic paradigm as a set of guiding principles that inform technological and investment decisions, shaping the choices made by actors in the system. Freeman (2019) further developed this idea within a historical framework that recognized the role of institutional diversity and path dependence for firms, industries, national economies, and technologies. In this context, Schumpeter's long waves are viewed as a succession of techno-economic paradigms. Freeman and Louçã (2001) associated the Kondratiev long wave with technological revolutions, while Perez (2002, 2015) referred to them as great surges of development.

Perez (1983, 1985) suggested that each major technological surge is characterized by a dominant and all-encompassing low-cost key sector. These sectors typically involve new energy sources, materials, technologies, products, and processes, as well as unique organizational structures. To qualify as a key sector within a techno-economic paradigm, it must meet four criteria: 1) exhibit low and declining relative costs; 2) possess an unlimited supply; 3) have a widespread presence of low-cost core inputs; and 4) contribute to cost reduction and quality improvement in labor, capital equipment, and intermediate goods. During the initial phase of a techno-economic paradigm, there is a growing disconnect between the techno-economic and socio-institutional realms (Perez 1983, 2002).³

Freeman and Perez (1988) introduced the term techno-economic paradigm to capture the comprehensive transformation that occurs after a technological revolution. The concept of techno-economic paradigms adds a nuanced perspective to the discussion on long waves by highlighting the intricate interplay among technology, economic structure, management practices, social institutions, and the societal adoption of technology. According to Perez (2002: 8), a technological revolution represents a significant and visible cluster of new and dynamic technologies, products, and industries that can disrupt the entire economic landscape. It involves a series of interconnected radical breakthroughs, forming a substantial constellation of interdependent technologies—a cluster of clusters or a system of systems (Perez 2010: 189).

Each technological revolution follows a distinct logic that shapes its specific techno-economic paradigm. This encompasses not only the emergence and development of core inputs and physical technologies but also a transformation in how individuals interact with the technology (Nelson 2005). Take the microprocessor, for instance, which not only revolutionized the process of economic growth but also transformed human interactions through its various applications. This pattern holds true for other significant innovations such as mechanization, the rise of the chemical industry, and mass production, all of which acted as catalysts for technological revolutions. Within each revolution, there are one or more core technologies that manifest as a cluster of

interconnected radical breakthroughs, forming a constellation of interdependent technologies. Over time, these core technologies drive down production costs. Long-term economic growth is not solely tied to major technological innovations but also to their diffusion throughout the broader economy.

Economists already understood the idea of S-shaped (sigmoid function) growth curves (Nelson and Winter 1977). The diffusion of radical technologies appears S-shaped as they spawn novel ideas and new combinations until exhaustion of the paradigm. Perez (2002) interpreted the growth curve as an S-shaped diffusion curve or technology life cycle to describe a technology revolution or great surge. Diffusion involves a learning process in which there are leaders and followers that evolves over the course of the techno-economic paradigm.⁴ Once a particular paradigm becomes dominant, the breakthrough technology will define the evolution of the technology life cycle. There are four phases in each techno-economic paradigm: (1) irruption phase, when the latest technology supplants old technology; (2) frenzy phase, or the period of intense exploration; (3) synergy phase, when the new technology is diffused throughout the economy; and (4) maturity phase, as the diffusion process becomes complete.

There is also a gestation period before the paradigm, in which a laboratory-invention phase, with prototypes, patents, and early applications develop, and a period of overlap as the new techno-economic paradigm replaces the old one (Freeman and Louçã 2001).⁵ Perez (2002, 2015) stressed the importance of the diffusion process during each technological revolution and its transformative effects on all aspects of the economy and society.

5. From the industrial revolution to the digital revolution

Freeman and Perez (1988: 50) provide a concise overview of the main characteristics of successive techno-economic paradigms or great surges. According to their account, there have been five technological revolutions since the introduction of Richard Arkwright's first mechanical spinning machine, or water-frame, in a Cromford water-powered mill in 1771. While the precise dating of the first industrial revolution may be subject to debate, Schumpeter (1939), Freeman and Louçã (2001), and Perez (2002) all acknowledge Richard Arkwright's patent on the water-frame in 1769. Perez (2002) summarizes the five great surges that have occurred over the past 250 years. The first column indicates the key energy source associated with each revolution. The second column highlights the initial breakthrough in each revolution. Columns three and four condense the essential physical and social technologies within each paradigm, drawing from the distinction made by Nelson (2005). Physical technologies encompass the core technologies and leading (or carrier) industries, while social technologies encompass the institutional arrangements

of the economy, including changing modes of division of labor and new ways of organizing and governing work (Nelson 2005: 195, e.g., von Tunzelmann 1995). Finally, the fifth column identifies the key networks associated with each revolution.

Table 1. Summary of great surges or technological revolutions

	Key energy source	Initial breakthrough	Physical technologies	Social technologies	Networks
First (1771)	<i>Water- powered industry</i>	Arkwright's water-frame	Iron, raw cotton, coal, machine tools	Factory system vs. domestic system, shift-work, entrepreneurs, interchangeable parts	Canals, roads, sailing ships, local networks
Second (1829)	<i>Steam- powered industry</i>	Manchester Liverpool railway, Brunel's steam ship	Iron, coal, Railways, Steamships,	Factory system and capital process, worker loss of control, professional management, joint stock companies	Railways, steam ships, telegraph networks, national power networks
Third (1875)	<i>Electrification of industry</i>	The Carnegie Bessemer steel plant, Edison's electric power	Steel, copper, metal alloys, internal combustion engine, electricity	American system, key machine tools, vertical integration, horizontal integration, management, R&D lab	Transport, communication, and electricity networks
Fourth (1908)	<i>Fossil fuel driven industry</i>	Ford moving assembly line; Burton oil cracking process	Oil, gas, synthetic materials, automobiles, aerospace	Mass production, managerial control (scientific management), multinational corporations	Production networks, radio, television, motorways, airports
Fifth (1971)	<i>Digitalization of energy systems</i>	Intel Micro-processor	Integrated circuits, computers, robots, software, AI, e- services, satellites	Factory and workplace automation, flexible specialization, integrative management	Global innovation networks, Internet, mobile connectivity networks

Source: Own summary based on Freeman and Perez (1988), Perez (2002), and von Tunzelmann (1995).

Steam engines based on James Watt's designs appeared during the first technological revolution, but they faced challenges such as high costs and technical infeasibility, resulting in limited adoption by entrepreneurs. Innovation in the machine tool industry and precision engineering gradually led to the development of smaller, high-pressure steam engines. The Rainhill Locomotive trials and the opening of the Liverpool-Manchester railway between 1829 and 1831 marked the Second Technological Revolution.

During this revolution, steam-powered mechanization of industries became the key technology, driven by falling resource prices. Agglomeration, standardization, and specialization played crucial roles in accelerating productivity growth. The emergence of railways, telegraphs, transatlantic steamship navigation, and a universal postal service facilitated networking of the economy, enabling greater connectivity and communication. Timesaving management practices and specialization within enterprises contributed to labor productivity growth in this paradigm. It is important to note that during this

period, there was significant repression of the working class and extreme inequality.

The United States lagged behind during the first two great surges, characterized by small markets, limited capital and labor, and a lack of technological knowledge compared to Europe (Thomson 2009: 17). To bridge this gap, the country relied on importing technical knowledge from abroad and combining it with local technological learning. However, there were promising signs of catching up to England. In the first technological revolution, the United States successfully internalized and improved upon innovative technologies in textiles, engines, and printing that were initially transferred from England. Additionally, advancements were made in clocks, woodworking, and firearms, drawing upon European technology (Thomson 2009: 59). The United States showcased its progress at the Paris Universal Exposition of 1867, where it received numerous accolades for machine tools, steam engines, telegraphs, locomotives, and other technologies.

The Third Technological Revolution, starting from 1875, marked the onset of industrial electrification and the emergence of a new electrical equipment industry. This revolution also brought advancements in steel production, making it more affordable. Steel, along with chemicals, copper, and other metal alloys, played a crucial role in the development of transportation systems, including shipping, railways, bridges, and tunnels, both in Europe and the United States. Notably, Thomas Edison established the first dedicated research and development (R&D) laboratory in New Jersey. This laboratory aimed to consistently introduce new products to the market, which included inventions such as the phonograph, microphones, electric lighting, and a system for electrical distribution, among others (Israel 1998). Edison's approach facilitated the recombination of existing ideas in innovative ways and fostered the creation of social networks, both within and outside the laboratory.

Declining oil prices combined with the moving assembly line (to produce Ford's Model T) in 1913 made it possible to produce inexpensive motorized vehicles. This application of mass production techniques was essential to the Fourth Technological Revolution, including making use of machines and presses to stamp out parts and ensure interchangeability, leading to the relative cheapness of large-scale production and the emergence of mass consumption (Hounshell 1984). Many industries appeared using mass production techniques, including automotive components, tractors, aircraft, consumer durables, and synthetic materials.

To ensure that mass consumption continued (Freeman and Louçã 2001), the rapid growth of the consumer markets generated large economies of scale. However, it also created even larger corporations that needed new ways of managing diverse operations (Chandler 1977). Speculation leading to the

financial crash of 1929 and the (Great) depression triggered the global war. A new dynamism reignited mass production and consumption after World War II, which led to extended periods of high growth, innovative product design, and intensive positive feedback effects.

The digital revolution had its roots in the fourth industrial revolution, with significant advancements in vacuum tubes (1935) and transistors (1947). Bell Labs played a crucial role in developing prototypes, securing patents, and early applications that foreshadowed the digital revolution. This progress paved the way for the emergence of a thriving electronic cluster in the Santa Clara Valley, also known as Silicon Valley (Lécuyer 2006). In 1969, Intel announced its plans to create the first commercially viable microprocessor, the Intel 4004, using silicon. This groundbreaking achievement was realized in November 1971. Concurrently, in the same year, the US Department of Defense installed the initial computers on the Advanced Research Projects Agency Network (ARPANET), the precursor to the Internet. The microprocessor, which integrated all the functions of a central processing unit (CPU) into a single integrated circuit, was a significant milestone. While Intel's announcement was overshadowed by the news of the Apollo 11 mission, it marked a pivotal moment in the development of the microprocessor (Isaacson 2014).

The microprocessor stands as a revolutionary breakthrough, with advancements in photolithography and planar techniques enabling the doubling of transistors in an integrated circuit approximately every two years, famously known as Moore's Law. Presently, Apple's Notebooks, iPads, and smartphones utilize processors boasting an astounding 11.8 billion transistors through 5nm technology. In contrast, the original Intel microprocessor contained a mere 2,250 transistors. Following this pivotal advancement, numerous new technological trajectories emerged within the digital techno-economic paradigm, giving rise to clusters of dynamic technologies, products, and industries that reverberated throughout the entire economy and society. These developments spurred the formation of new enterprises that interacted within intricate networks, ultimately leading to the establishment of a global digital telecommunications network, the internet, and various electronic services, including email (e-services).

We are now in the second phase of the digital revolution (Perez 2013, Reati and Toporowski 2009). The financial collapse of 2008 marked the turning point, as power shifted away from financiers to entrepreneurs and enterprises, and currently we should be in the period of high economic growth, or the golden age. Here diffusion is key, as is the complete digitalization of everyday objects and activities, or what we might refer to as ubiquitous computing (Weiser 1991).

A wide array of new products dependent on computing devices have appeared

in recent decades, including situated robotics, artificial intelligence (AI) technologies and smart energy networks (Perez 2013; Knell 2013). Modern robots can be autonomous or semi-autonomous, appearing human-like at times, but most often they are just a sophisticated industrial (and numerical control) machine with little guiding intelligence. These included industrial robots, warehouse robots, agricultural robots, autonomous vehicles, caring robots, medical robots, and robots in education. A few robots embed AI in the sense of (statistical) learning algorithms, but these are narrow and limited by the inability to deal with common sense solutions to everyday problems (Mitchell 2019).

6. Catching-up and falling behind in the digital revolution

Pomeranz's (2000) narrative suggests that Western Europe witnessed several great surges of development. Table 2 presents per capita GDP data from 1776 to 2018 for selected European countries, the United States, and East Asia. Interestingly, East Asia experienced a stagnation in labor productivity since the time of the industrial revolution. The United Kingdom emerged as the leading country during the first two waves but was surpassed by the United States in the third cycle. The table also reveals a distinct disparity in per capita income between Europe and Asia during the initial four surges, followed by a rapid convergence in the fifth technological revolution, known as the digital revolution. So, what contributed to East Asia's rapid convergence during this revolution? Notably, Korea demonstrated significant convergence, catching up with France and the United Kingdom. However, for now, let us shift our focus to China.

Table 2. Selected Europe, United States, and Asia, in the world economy.

Per capita GDP, 1776-2018.

	1776	1820	1870	1913	1950	1975	2000	2018
France	1,728	1,809	2,990	5,555	8,266	20,653	33,410	38,516
Norway	...	1,384	2,177	4,415	8,655	19,560	54,040	84,580
United Kingdom	2,962	3,306	5,829	8,212	11,061	18,884	31,946	38,068
United States	2,419	2,674	4,803	10,108	15,240	25,956	45,886	55,335
China	981 ^a	882	945	895	799	1,594	4,730	13,102
Korea (South)	...	816	820	1,171	998	4,362	23,108	37,928
Taiwan (China)	...	966	968	1,286	1,460	5,614	26,787	44,664
India	1,068 ^b	937	850	1,073	987	1,430	5,384	6,806

Note: Percentage of total world population and Real GDP per capita in 2011US\$. ^a1780; ^b1750

Source: Maddison Project Database, 2020. Bolt, Jutta and Jan Luiten van Zanden (2020), "Maddison style estimates of the evolution of the world economy. A new 2020 update".

China found itself in the midst of the Cultural Revolution when the digital revolution commenced. The country operated under a typical shortage economy, characterized by prevailing sellers' markets. Shortages of essential

consumer goods like sugar, soap, bicycles, sewing machines, and household coal persisted, while energy, materials, and transportation bottlenecks further constrained development.⁶ Despite some advancements in space, lasers, biotechnology, semiconductors, electronics, and computing technology during the 1970s, the disruptive impact of the Cultural Revolution significantly hampered production activities for over a decade (Brock and Wei 2013:24).

The 1978 economic reforms aimed to shift China's extensive growth model, reminiscent of the Soviet style, towards a more domestically driven intensive growth approach (Byrd 1991). These reforms paved the way for a transformation of China into a mixed economy, characterized by the emergence of buyers' markets in many sectors, while sellers' markets persisted in certain areas such as raw materials, energy, and specific consumer goods. The growing private economy, initially comprised of non-state rural, village, township, and urban enterprises, gradually alleviated shortages in most consumer goods sectors. As the private economy continued to expand, state enterprises became less influential in the overall economic landscape.

The telecommunications industry holds significant relevance in this context, particularly concerning its development in China. Prior to the Cultural Revolution, there existed a small state-owned telecommunications sector, but progress was limited until the late 1970s when Deng Xiaoping initiated upgrades in telephone, telex, and fax transmission technology. The mid-1980s witnessed the emergence of two domestic telecommunications enterprises that posed as competitors to the state sector: ZTE (Zhongxing Telecommunications Equipment Company) and Huawei Technologies. ZTE, founded in 1985, received support from the Ministry of Aerospace Industry, while Huawei, established in 1986, operated as a privately-owned enterprise.⁷

Fierce competition between these two firms in the advancement of 5G technology has yielded substantial results in terms of patents. Huawei boasts 2,570 5G patents, while ZTE has accumulated over 1,700 5G patents, as reported in their respective annual reports. The intense interfirm rivalry between Huawei and ZTE has played a significant role in driving innovation within the telecommunications industry in China.

China's digitalization strategy relies on smart specialization and the diffusion of knowledge, both of which play crucial roles in its technological upgrading. This strategy entails a gradual shift from lower-value to higher-value activities, effectively ascending the global value chain (Milberg and Wickler, 2013). Upgrading, as described by Grossman and Helpman (1991), can be visualized as climbing a quality ladder, and it can manifest in various contexts such as factories, inter-firm networks, local or national economies, and regions (Gereffi, 2018). Upgrading can take the form of adopting new production and organizational methods (process upgrading), developing more sophisticated

products (product upgrading), engaging in higher value-added activities (functional upgrading), and adapting competencies to new industries (intersectoral upgrading) (Humphrey and Schmitz, 2002).

Radošević and Yoruk (2016; 2018) have adopted a multidimensional framework that incorporates technological capabilities and competencies, institutional governance arrangements, and global knowledge flows through trade. This comprehensive framework considers the interplay of these factors in facilitating China's technological development and economic growth.

The upgrading of China's technology plays a pivotal role in achieving convergence, particularly when considering its position relative to the technology frontier (Lee, 2013). To catch up, it is essential to foster the development of new knowledge within the country while also effectively diffusing existing knowledge from external sources (Fagerberg, 1988). Cohen and Levinthal (1989) put forth the argument that research and development (R&D) activities not only generate fresh knowledge but also enhance a firm's capacity to comprehend and utilize existing knowledge, which they refer to as absorptive capacity. The first type of R&D activity establishes a positive feedback loop that drives technological convergence and collective learning. Recognizing the second role of science, technology, and innovation (STI) activities in catching up implies that national learning systems will influence both the potential for catching up (technological opportunities) and its actual realization (appropriability conditions) (Abramovitz, 1986). The convergence process hinges on the technological opportunities accessible to China and its ability to leverage specific local learning advantages.

[The final paper will include an explicit discussion of South Korea, Taiwan, and India]

7. Emerging technologies in the sixth techno-economic paradigm

Today, there is much discussion about Industry 4.0 and the Fourth Industrial Revolution. The concept of Industry 4.0 originated as a marketing-style promotion during a press conference at the *Hannover Messe* in Germany in 2011. Two years later, the *Federal Academy of Science and Engineering*, with the support of the *Federal Ministry of Research* (BMBF), presented a research agenda and implementation recommendations (Kagermann et al. 2013). This agenda became part of a specialized discourse meant to promote the digitalization of production and innovation networks that integrate machinery, warehousing systems, and production facilities through cyber-physical systems (Pfeiffer 2017). The term Industry 4.0 has since sparked ongoing debates about the future of work and society.

The slogan for the 2016 World Economic Forum meeting in Davos was "Mastering the Fourth Industrial Revolution". In his book, Schwab (2016)

associated this revolution with the diffusion of digital technologies and emerging breakthrough technologies like biotechnology, nanotechnology, artificial intelligence, robotics, 3D printing, the Internet of Things, and quantum computing. Schwab identified four revolutions that correspond to long waves of technological change: 1) water- and steam-powered mechanical manufacturing (first two long waves); 2) electricity-powered mass production (second two long waves); 3) digitalisation (fifth long wave); and 4) cyber-physical systems (sixth long wave). However, in this framework there is no discussion of nor reference to long-waves, technological revolutions, and techno-economic paradigms. The discussion resembles the crafting of a strategic narrative rather than a solid conceptual scheme to interpret techno-economic dynamics. Schiølin (2019: 4) labels the Fourth Industrial Revolution as an act of 'Future essentialism', or:

discourses, narratives or visions that (...) produce and promote an imaginary of a fixed and scripted, indeed inevitable, future, and that can be desirable if harnessed in an appropriate and timely fashion but is likewise dangerous if humanity fails to grasp its dynamics.

In contrast, the concepts and tools highlighted in this chapter are patterns derived from historical or analytical analysis, not "imaginaries."

Over the past decade, discussions on digitalization and automation technologies have re-ignited decades-old debates between techno-optimists and techno-pessimists (Mokyr et al 2015). Our concepts and tools can help to understand these debates. For example, Rifkin (2011) predicts that "green" technologies, such as wind, solar photovoltaic, hydropower, and wave and tidal power, will encourage a new industrial revolution. Brynjolfsson and McAfee (2014) believe we are in the second machine age where productivity gains from intangible assets and digitalization of the economy intensified changes in skills, organizations, and institutions. They suggest that earlier technology-driven revolutions, such as the steam engine and electrification, took longer to diffuse through the economy.

Schot and Kanger (2018) grouped the first four technological revolutions into what they call "industrial modernity", or the first deep transition. Their framework builds on the concepts we describe, as Deep Transitions describe "an evolutionary multi-level model of surges". During the first deep transition, labor and machines were complementary. The second deep transition began with the digital revolution and involves the automation of cognitive tasks (especially symbol processing) that make humans and software-driven machines substitutes rather than complements. Brynjolfsson et al. (2021) suggested that productivity growth may have "paused" in the first decades of the digital revolution but could surge again as a novel wave of digital technologies deploys its effect, resembling the way ICTs pushed productivity

growth in the Nineties after the puzzling slowdown in the Eighties. These narratives confirm Perez's (2013) idea that we are now in the golden age of the digital revolution.

Can we envisage the future and the sixth technological revolution through the tools and concepts outlined above? Possible emerging technologies include nanotechnology, biotechnology, AI and quantum computing. This is where the physical, digital, and biological worlds could converge in what Lombardi and Vannuccini (2021) call a 'cyber-physical universe'. Inspiration for the idea originates in Richard Feynman's 1959 lecture, "There's Plenty of Room at the Bottom". Here Feynman described a process in which scientists would be able to manipulate and control individual atoms and DNA molecules. The idea of the transistor and the microprocessor triggered the process of miniaturization. Feynman anticipated nanoscience and nanotechnology, which influenced other science fields, including chemistry, biology, physics, materials science, and engineering. Kaku (2023) suggests we near the end of the digital age and moving toward the post-Silicon or Quantum age.

It is possible that two or three independent technological systems could converge into one system (or a system of systems, to use Perez's expression), which would then trigger the explosive take-off into the sixth technological revolution (Roco and Bainbridge 2003, Knell 2010). Atoms, DNA, bits, and synapses will supply the basic elements and foundational tools that will make it possible to integrate several emerging technologies, including nanotechnology, biotechnology, quantum biology, information technology, and the latest cognitive technologies, into multifunctional systems. It is a long way from Kondratiev and Schumpeter's contributions, but the tools we illustrated remain the cornerstone of a useful theory of innovation and change even in our current complex socio-technical landscape.

Bibliography

- Abramovitz, M. 1986. Catching up, forging ahead, and falling behind, *The Journal of Economic History* 46:385- 406.
- Abramovitz, M. and P.A. David, 1996. Convergence and Deferred Catch-up: Productivity Leadership and the Waning of American Exceptionalism', in R. Landau, T. Taylor, and G. Wright, *The Mosaic of Economic Growth*, Stanford: Stanford University Press.
- Aghion, P. and Howitt, P. 1992. A Model of Growth Through Creative Destruction, *Econometrica* 60/2: 323-351. Amable, B. 1994. Endogenous growth theory, convergence and divergence, in G. Silverberg and L. Soete, L (eds) (1994), *The economics of growth and technical change*, Cheltenham: Edward Elgar.
- Arrow, K. 1962. The economic implications of learning by doing, *Review of Economic Studies* 29:155-173. Arthur, W.B. 2009. *The Nature of Technology: What it is and how it evolves*, London: Allen Lane.
- Barnett, V. 2001. Tugan-Baranovsky as a Pioneer of Trade Cycle Analysis, *Journal of the History of Economic Thought* 23/4: 443 – 466.
- Barro, R.J. and Sala-i-Martin, X. 1992. *Convergence*. *Journal of Political Economy*, 100:223-51.
- Beinhocker, E.D. 2006. *The Origin of Wealth: Evolution, Complexity, and the Radical Remaking of Economics*. Boston, Harvard Business School Press.
- Bresnahan, T.F. 2019. Technological change in ICT in light of ideas first learned about the machine tool industry, *Industrial and Corporate Change* 28:331–349
- Bresnahan, T.F. and M. Trajtenberg, 1995. General Purpose Technologies "Engines of Growth?", *Journal of econometrics* 65:83-108.
- Brynjolfsson E., and McAfee A. 2014. *The second machine age. Work, progress, and prosperity in a time of brilliant technologies*, W.W. Norton, New York
- Chandler, A.D. 1977. *The visible hand*. Harvard University Press, Cambridge
- Cohen, W.M. and Levinthal, D.A. 1989. The two faces of R&D, *Economic Journal* 99:569-96.
- Dosi, G., 1982. Technological paradigms and technological trajectories, *Research Policy* 11: 147-162.
- Dosi, G., C. Freeman, R.R. Nelson, G. Silverberg and L. Soete 1988. *Technical change and economic theory*, London: Pinter Publisher.
- Faccarello G.; Kurz, H.D. (eds.), *Handbook on the History of Economic Analysis*, Vol. III: Developments in Major Fields of Economics, Edward Elgar.
- Fagerberg, J. 1994. Technology and international growth rates, *Journal of Economic Literature*, 32:1147–75.
- Feynman, R.P. 1960. There's Plenty of Room at the Bottom, *Engineering and Science* 23:22–36
- Freeman, C. [1995] 2019. History, Co-Evolution and Economic Growth, *Industrial and Corporate Change*: 28/1:1–44.

- Freeman, C., and F. Louçã 2001. *As Time Goes By. From the Industrial Revolution to the Information Revolution*, Oxford: Oxford University Press.
- Freeman, C., and Perez, C. 1988. Structural crisis of adjustment, business cycles and investment behaviour. In Dosi, et al. London: Pinter.
- Gereffi G 2018. *Global value chains and development*. Cambridge University Press, Cambridge
- Grossman, G. and E. Helpman 1991. *Innovation and growth in the global economy*, Cambridge: MIT Press. Helpman, E. 2004. *The Mystery of Economic Growth*, Cambridge: Belknap Press.
- Helpman, E. ed. 1998. *General Purpose Technologies and Economic Growth*. Cambridge: MIT Press.
- Hounshell, D.A. 1984. From the American System to Mass Production, 1800–1932: The Development of Manufacturing Technology in the United States. Johns Hopkins University Press, Baltimore
- Humphrey, J. and Schmitz, H. 2002. How Does Insertion in Global Value Chains Affect Upgrading in Industrial Clusters?, *Regional Studies* 36:1017–27.
- Isaacson, W. 2014. *The innovators: how a group of hackers, geniuses, and geeks created the digital revolution*. Simon & Schuster, New York
- Isreal, P. 1998. *Edison: a life of invention*. New York, John Wiley.
- Kaku, M. 1997. *Visions: How Science will Revolutionize the 21st Century*, Anchor books.
- Kaku, M. 2023. *Quantum Supremacy: How Quantum Computers will Unlock the Mysteries of Science – and address Humanities Biggest Challenges*, Allen Lane.
- Kaldor, N. 1957. A model of economic growth, *Economic Journal* 67:591-624.
- Knell, M. 2013. Multi-source energy networks and the ICT revolution, *European Planning* 21: 1838-1852.
- Kondratiev, N.D. [1925] 1979. Long business cycles, *Voprosy kon'iunktury* I:28-79. Translated as: Long waves of economic life, *Review (Fernand Braudel Center)* 2: 519-62. Shorter version in *The Review of Economic Statistics*, 17/6: 105-115 (1935).
- Kondratiev, N.D. [1926] 1984. *The Long Wave Cycle*, (tr. Guy Daniels), New York: Richardson and Snyder. Kurz, H.D. and Salvadori, N. 1999. Theories of 'Endogenous' Growth in Historical Perspective, in M.R. Sertel, *Contemporary Economic Issues*, London: Macmillan.
- Lécuyer, C. 2006 *Making Silicon Valley: innovation and the growth of high tech, 1930-1970*. Cambridge, MIT Press.
- Lee, K. 2013. *Schumpeterian Analysis of Economic Catch-up: Knowledge, Path-Creation, and the Middle-Income Trap*. Cambridge University Press.
- Lipsey, R.G. K.I.Carlaw and C.T.Bekar, eds. 2005. *Economic Transformations: General Purpose Technologies and Long Term Growth*, Oxford: Oxford University Press.
- Louçã, F. 1999. Nikolai Kondratiev and the Early Consensus and Dissensions about History and Statistics, *History of Political Economy*, 31: 169-205.

- Lucas, R.E. 1988. On the mechanics of economic development, *Journal of Monetary Economics*, 22:3-42.
- Maddison 2006. *The World Economy*, Paris: OECD.
- McCraw T.K., 2007. *Prophet of innovation: Joseph Schumpeter and Creative Destruction*, Cambridge: Harvard University Press.
- Milberg, W and Winkler, D. 2013. *Outsourcing Economics: Global Value Chains in Capitalist Development*, Cambridge.
- Mitchell, M. 2019. *Artificial intelligence: a guide for thinking humans*. Pelican Books, London.
- Nelson, R.R. 2005. *Technology, Institutions, and Economic Growth*, Cambridge: Harvard University Press.
- Nelson, R.R. and Winter, S.G. 1982. *An Evolutionary Theory of Economic Change*, Cambridge: Harvard University Press.
- Perez, C. 1983. *Structural change and assimilation of new technologies in the economic and social systems*. *Futures*, 15(5), 357-375.
- Perez, C. 1985. Microelectronics, Long Waves and World Structural Change: New Perspectives for Developing Countries, *World Development* 13(3): 441-463.
- Perez, C. 2002. *Technological revolutions and finance capital: The dynamics of bubbles and golden ages*. Cheltenham: Edward Elgar.
- Perez, C. 2010. Technological revolutions and techno-economic paradigms, *Cambridge Journal of Economics* 34: 185-202.
- Perez, C. 2013. Unleashing a golden age after the financial collapse: Drawing lessons from history, *Environmental Innovation and Societal Transitions* 6: 9-23.
- Perez C. 2015. From Long Waves to Great Surges. *European Journal of Economic Social Systems* 27(1-2):69-79.
- Pfeiffer S. 2017. The Vision of Industrie 4.0 in the Making – a case of future told, tamed, and traded, *Nanoethics* 11:107-121
- Pomeranz, K. 2000. *The Great Divergence: China, Europe, and the making of the Modern World Economy*. Princeton, Princeton University Press.
- Quah, D. 1996. Twin peaks: growth and convergence in models of distribution dynamics, *Economic Journal* 106:1045-1055.
- Radosevic, S. and Yoruk, E. 2016. Why do we need a theory and metrics of technology upgrading?, *Asian Journal of Technology Innovation*, 24:sup1, 8-32.
- Radosevic, S. and Yoruk, E. 2018. Technology upgrading of middle income economies: A new approach and results. *Technological Forecasting and Social Change* 129, 56-75.
- Ramsey, F.P. 1928. A mathematica theory of saving, *Economic Journal*, 38:543-59.
- Reati, A. and J. Toporowski 2009. An economic policy for the fifth long wave, *PSL Quarterly Review*, 62: 147- 190.
- Ricardo, D. [1817] 1955. *The principles of political economy and taxation*. Cambridge: Cambridge University Press.
- Richardson, G. 1972. 'The Organisation of Industry', *Economic Journal*, 82:883-896.
- Rifkin J (2011) *The Third Industrial Revolution: how lateral power is transforming energy, the economy, and the world*. Palgrave Macmillan, New York.

- Roco, M.C., and Bainbridge, W.S. 2003. Converging technologies for improving human performance: nanotechnology, biotechnology, information technology, and cognitive science. Springer, Heidelberg.
- Romer, P.M. 1986. Increasing returns and long-run growth, *Journal of Political Economy* 94:1002-37.
- Romer, P.M. 1990. Endogenous technological change, *Journal of Political Economy* 98:S71-S102.
- Rosenberg, N., 1963. Technological change in the machine tool industry, 1840-1910, *Journal of economic history*, 414-443.
- Rosenberg, N. and Frischtak, C.R., 1984. Technological innovation and long waves. *Cambridge Journal of Economics*, 8(1):7-24.
- Rostow, W.W. 1975. Kondratiev, Schumpeter, and Kuznets: Trend Periods Revisited, *The Journal of Economic History* 35/4: 719-753.
- Schot J, and Kanger L 2018. Deep transitions: emergence, acceleration, stabilization, and directionality. *Research Policy* 47:1045–1059
- Schumpeter, J.A., [1912] 1934. *Theorie der wirtschaftlichen Entwicklung*, Leipzig, Verlag von Duncker & Humblot. Third edition translated by R. Opie as *The Theory of Economic Development*, Cambridge: Harvard.
- Schumpeter, J.A., 1935. The Analysis of Economic Change, *The Review of Economics and Statistics* 17/4: 2-10, Schumpeter, J.A., 1939. *Business Cycles: A Theoretical, Historical, and Statistical Analysis of the Capitalist Process*, New York: McGraw-Hill.
- Schumpeter, J.A. 1942. *Capitalism, Socialism and Democracy*. New York: Harper and Brothers.
- Schwab, K. 2018. *Shaping the Future of the Fourth Industrial Revolution*, Portfolio Penguin.
- Silverberg, G. 2003. Long waves: conceptual, empirical and modelling issues. In *Elgar Companion to neo- Schumpeterian Economics*, edited by H. Hanusch and A. Pyka, Aldershot: Edward Elgar.
- Smith, A. [1776] 1976. *An Inquire into the Nature and Causes of the Wealth of Nations*, ed. R. H. Campbell and A. S. Skinner, volumes 1 and 2 of the Glasgow Edition of the Works and Correspondence of Adam Smith, Oxford: Oxford University Press.
- Solow, R.M. 1956. A contribution to the theory of economic growth, *Quarterly Journal of Economics* 70:65- 94.
- Solow, R.M. 1957. Technical change and the aggregate production function, *Review of Economics and Statistics* 39:312-20.
- Solow, R.M. 2000. *Growth Theory: An Exposition, 2nd Edition*. Oxford: Oxford University Press.
- Temple, J. 1999. The new growth evidence, *Journal of Economic Literature*, 37(1), March, 112-156.
- Thomson, R. 2009. *Structures of Change in the Mechanical Age: Technological Innovation in the United States 1790–1865*, Baltimore: Johns Hopkins University Press.
- Trotsky. L. 1973. “The Curve of Capitalist Development” in *Problems of Everyday Life and Other Writings on Culture and Science*, New York: 273–280.

- Tunzelmann, G.N von 1995. *Technology and Industrial Progress: The Foundations of Economic Growth*, Aldershot: Edward Elgar.
- Tylecote, A. 1992. *The Long Wave in the World Economy: The Current Crisis in Historical Perspective*, London, Routledge.
- Verspagen, B. 1991. A new empirical approach to catching up or falling behind, *Structural Change and Economic Dynamics* 2:359-80.
- Weiser, M. 1991 The computer for the 21st Century, *Scientific American* 265 (3):94-104 (Sept.).

Notes

-
- ¹ Sala-i-Martin (1996) pointed out, economists often equate beta-convergence with conditional convergence and sigma- convergence with absolute convergence. Another example is that club convergence allows for multiple equilibria, while the standard neoclassical model depends on uniqueness.
 - ² Schumpeter (1935, 1939) proposed a three-cycle framework, encompassing different time scales. The first cycle, known as the 'Kitchin' inventory cycle, lasts approximately 40 months and reflects information asymmetries. The second cycle, referred to as the 'Juglar' investment cycle, operates over a period of eight to ten years and captures the dynamics of monetary and financial markets. The third cycle is the long-term 'Kondratiev' cycle, spanning fifty to sixty years, which corresponds to the evolution of major technological innovations. Within each long wave, there exist shorter cycles: six Juglars for every Kondratiev and three Kitchins for each Juglar.
 - ³ This mismatch also appears between physical and social technology (Beinhocker 2006).
 - ⁴ Diffusion is not prominent in Economic Development and only appears as a process by which firms copy, imitate, and gradually improve on the original innovation, or what he described as induced innovations.
 - ⁵ This phase challenges the dominant technology system, and can last for decades, culminating in the new paradigm.
 - ⁶ Under Mao, the economy had a centralized socialist organization of the market where money and prices are passive and where the planning bureau distributed over 90 per cent of total output through a network of supply and marketing cooperatives (Robinson 1964).
 - ⁷ In 2018, Huawei became the fifth largest global enterprise in terms of R&D activity, and ZTE was the 76th largest enterprise, according to the 2018 EU Industrial R&D Investment Scoreboard. By the end of 2018, Huawei obtained 87,805 patents, of which were 11,152 classified as core patents granted in the United States.

Industry 4.0 base and frond-end manufacturing technology adoption: An empirical study of Australian manufacturing firms

Dilupa Nakandala

Western Sydney University, Australia

Abstract

Industry 4.0 technologies have become a key focus for the manufacturing industry in recent years, particularly due to their role in driving supply chain and operations resilience during times of disruption. This study investigates how strategic orientations of firms drive the adoption of Industry 4.0 technologies, with a focus on firms' strategic focus on customer requirements, competitor behaviour, technological advancements, and their ability to create, disseminate, and use knowledge. Recognising the heterogeneity of Industry 4.0 technologies, the study differentiates between Industry 4.0 base technologies and advanced front-end manufacturing technologies, where the latter includes technologies such as AI, Machine-to- Machine communication, cyber-physical systems, big data analytics and autonomous machines and systems. Based on survey data collected from Australian manufacturing firms and using the PLS-SEM method for data analysis, the study found empirical evidence for the effects of customer, competitor, and learning orientations on technology orientations and the effects on the adoption of both Industry 4.0 base technologies and front-end manufacturing technologies through technology orientation. However, the effects of the technology orientation strategy are weaker on advanced front-end manufacturing technologies compared to Industry 4.0 base technologies. This study provides valuable insights for management practitioners by highlighting the importance of the firm's customer, competitor, and learning orientations on the technology orientation and the significant effects of technology orientation on Industry 4.0 technology adoption. It also suggests that the adoption of advanced Industry 4.0

manufacturing technologies is influenced by various other factors beyond these strategic orientations relative to Industry 4.0 base technology adoption. The study contributes to the strategic management literature on Industry 4.0 by including base and advanced manufacturing technologies as different independent constructs in model testing. It further asks for a differentiated policy response to support manufacturing firms to upgrade their advanced manufacturing technologies to be competitive in the Industry 4.0 era.

Keywords: Industry 4.0, technology adoption, strategic orientation, front-end manufacturing technologies; manufacturing firms, Australia

Introduction

The concept of Industry 4.0, also known as Industrie 4.0 and Industrial Internet of Things (IIoT), originated with a focus on the manufacturing industry and evolved from the mechanisation (second industrial revolution) and automation (third industrial revolution) eras (Lasi et al., 2014). Industry 4.0 is enabled by the Internet of Things (IoT) and is characterised by the integration of cyber-physical systems into manufacturing operations. The emergence of advanced technologies such as 3D printing, artificial intelligence, advanced robotics, autonomous systems, and big data analytics is prominent in Industry 4.0. These technologies enable mass product customisation without sacrificing economies of scale, predictive planning and maintenance, real-time monitoring for increased productivity and reduced downtime, and data-driven decision making (Oztemel and Gursev, 2020; Ghobakhloo et al., 2021).

As the adoption of Industry 4.0 technologies continues to gain increased attention, numerous research studies have emerged investigating firm-level behaviours from diverse perspectives. Many of these studies focus on the advantages of Industry 4.0 technologies, including increased productivity, resilience, product quality, and, more recently, sustainability, including the circular economy (Nakandala et al., 2023; Oztemel and Gursev, 2020; Ghobakhloo et al., 2021). The adoption of Industry 4.0 technology has also been the subject of many studies, which tend to be mainly conceptual frameworks, descriptive analyses, or literature reviews. For instance, Hopkins (2021) discovered that robotics is the most widely used Industry 4.0 technology, while big data analytics is expected to have the highest impact in the sample of manufacturing firms in Australia. Khin and Kee (2021) conducted an exploratory study of Malaysian manufacturing firms and discovered that operational benefits, market opportunities, customer requirements, labour

issues, competitive pressure, and quality image were driving factors for Industry 4.0 adoption.

The broad range of Industry 4.0 technologies and their various technological complexity levels create challenges in identifying what technologies suit different firm-level conditions and characteristics and whether all technologies should be treated the same way when we investigate adoptions of Industry 4.0 technologies by firms in research. Many research studies focus on Industry 4.0 base technologies, which are cloud computing, big data analysis, cyber-physical systems and the use of the internet to access data from remote sensors and control physical assets. The heterogeneity of Industry 4.0 technologies investigated in previous research was identified by Hopkins (2021). There is limited research that focuses on industry-specific Industry 4.0 technologies. Extending beyond base Industry 4.0 technologies to capture industry-specific Industry 4.0 technologies will make research findings more relevant to the industry in focus.

This study investigates how strategic orientations of manufacturing firms influence their Industry 4.0 technology adoption levels, focusing on differentiating base technologies from front-end manufacturing Industry 4.0 technologies. The hypothesised model shows a positive relationship between the market, competitor and learning orientations of firms with Industry 4.0 base and front-end technologies through technological orientation. Based on an anonymous online survey data from 117 manufacturing firms in Australia, it is found that market, competitor and learning orientations have direct positive effects on the technological orientation and then technological orientation leads to both Industry 4.0 base and front technology adoption levels. We also find that the effect of technological orientation is stronger on Industry 4.0 base technologies compared to their effects on front-end manufacturing technologies. It is also found that the strategic orientations explain the limited variance of the Industry 4.0 technology adoption variable where the coefficient is strong for the relationship between technological orientation and Industry 4.0 base technologies. Still, both R squared value and the coefficient are small for front-end manufacturing technologies.

The findings of the study show the importance of strategic focus on technology shows that manufacturers need to sense opportunities in the technological environment. It also finds that the strategic orientations of the manufacturing firms differently impact Industry 4.0 base and front-end technologies. Such a finding indicates the need for future researchers not to integrate all Industry 4.0 technologies (base and industry-specific) together to make conclusions on the benefits or antecedents of Industry 4.0 implementations. The paper will first provide a brief literature review leading to hypothesis development and the methodology, including the data collection method and sample characteristics. Next, data analysis and results will be presented, followed by a discussion of

results and a conclusion that will identify the limitations of the study and suggestions for future research.

Literature review and hypothesis development

The study is underpinned by the TOE framework from Tornatzky and Fleischer (1990) and takes a strategic perspective on the TOE framework. It investigates the effects of customer and competitor orientations from the environmental perspective, learning orientation from the organisation perspective, and technology orientation from the technological perspective. Previous studies on firm-level strategies identify the importance of the customer, competition, learning and technological orientations for firm performance (Gatignon & Xuereb 1997; Frambach et al. 2016). Customer orientation is “the sufficient understanding of one's target buyers to be able to create superior value for them continuously” (Narver & Slater, 1990, p. 21). Competitor orientation is the firm's understanding of competitors' strengths, weaknesses, capabilities, and strategies. Learning orientation collectively captures a firm's ability to acquire new knowledge and share knowledge within the firm and its commitment to learning (Chen et al. 2011). Customer-oriented firms are focused on technology to meet the needs of their existing markets and customers or expand to new markets. Increasing customer awareness of new technologies and the increased digital literacy of customers drive firms to focus on updating, creating, and adopting new technologies to sustain customer satisfaction. Open-mindedness and Technology orientation is the strategic focus on technological developments within the firm. Customer, competitor, learning, and technology orientations will likely have interdependent effects on firm performance (Frambach et al. 2016).

H1 Competitor orientation directly affects technology orientation.

H2 Customer orientation directly affects technology orientation.

H3 Learning Orientation directly affects technology orientation.

Technological capabilities have been a primary factor of competitiveness between firms in recent times, where those who develop, adopt, and adapt advanced technologies have advantages over the other firms. Technology orientation is more important than other strategic orientations for high firm performance in uncertain times (Gatignon & Xuereb 1997). Simply put, firms compete on their technological capabilities in an industry environment that is uncertain and dynamic. Firms' technology orientation leads to competitive advantage as customers prefer their products to be of superior quality.

Technology-oriented firms invest in new technologies and attempt to surpass the capabilities of competitor firms by investing in technologies to develop low-cost or superior products (Hakala 2010; Gatignon & Xuereb 1997). Technology-

oriented firms are receptive to new technologies and sense, seize and adapt new technologies to improve their product quality, process efficiency, planning and resilience capabilities. Manufacturing firms with a strategic orientation towards technology increasingly adopt industry 4.0 technologies for increased proactive planning, predictive maintenance and real-time monitoring for decision-making (Oztemel and Gursev 2020; Ghobakhloo et al. 2021),

H4 Technology orientation affects Industry 4.0 base technology adoption

H5 Technology orientation directly affects Industry 4.0 manufacturing front-end technology adoption

Method

The study collected data from managers and executives of manufacturing firms in Australia based on an anonymous survey. The importance of Australia as the research context for this study is supported by the recent government initiatives for advanced manufacturing, such as the \$1.5 billion investment in the Australian Modern Manufacturing Strategy and the more recent \$15 billion investment in the National Reconstruction Fund. The data were collected by working with a reputable survey agency with a large panel of manufacturing firms. Participants representing 117 manufacturing firms responded to the survey.

The four independent constructs of customer, competitor, technology and learning orientations were developed from the strategic management literature. All the survey items had a measurement scale of seven points ranging from strongly disagree to agree on a seven- point Likert scale strongly. Customer orientation was measured using a four-item scale which evaluated if the participant firm's competitive advantage was based on understanding their customers' needs and business objectives are driven primarily by customer satisfaction and if they assess customer satisfaction and pay attention to after-sales service. The measure for competitor orientation assessed the level of integration of information about the products and strategies of competitors, collection and analysis of competitors' activities, management collaborative actions on sharing competitors' information and the possession of knowledge of competitors' capabilities. Four survey items for the learning orientation were based on Chen et al. (2011) and Sigauw et al. (2006). Those items captured the firm's learning philosophy towards employees' willingness to learn and share new knowledge and that the firm encourages employees to resolve problems innovatively.

The measures of the dependent constructs of Industry 4.0 base and front-end technologies were adapted from Frank et al. (2019). The four Industry 4.0 base technologies of cloud computing for storage and data management, the use of the internet to access data from remote sensors and control physical objects in

the surrounding environment, cyber-physical systems and advanced data analytics (e.g. simulation, optimisation, regression) have been commonly used in numerous studies where the manufacturing front-end technologies had limited presence in Industry 4.0 technology measures. The base Industry 4.0 technologies were extended to include front-end manufacturing technologies following Frank et al. (2019), and respondents were asked to rate the level of implementation in their firms. Manufacturing front-end technologies were proposed as a second-order construct to differentiate conventional and advanced manufacturing technologies proposed under the smart factory concept.

Data analysis

The respondent firms had varying degrees of Industry 4.0 technology implementations. Cloud computing has the highest implementation rate in base technologies (Mean=5.368, SD=1.344). AI for production planning, virtual commissioning, robots (industrial robots, autonomous guided vehicles or similar), and AI for predictive maintenance and automatic nonconformity identification were at the lowest end of the front-end manufacturing technology implementation in the sample firms.

The dataset was tested for construct validity, where Cronbach's alpha had a minimum value of 0.769, and the average variance extracted had a minimum value of 0.591, as shown in Table 1. The discriminant validity of the items was tested by checking HTMT values where all were lower than the minimum cut-off of 0.85, with the highest value at 0.785. Following the criteria from Fornell and Larcker (1981), the discriminant validity of the main constructs in the model was confirmed by testing that the correlation between any pair of the constructs was smaller than the square roots of the AVE of them, as shown in Table 2.

Table 1. Construct validity measurements

	Cronbach's alpha	Composite reliability	Average variance extracted (AVE)
Automation	0.988	0.988	0.933
CoO	0.898	0.9	0.766
CxO	0.802	0.827	0.621
I4.0-Base	0.769	0.811	0.591
I4.0-Manuf	0.993	0.993	0.904
LeO	0.852	0.855	0.692
Smart Factory	0.984	0.984	0.912
TechO	0.877	0.879	0.731

Table 2. Discriminant validity - Fornell and Larcker (1981) criteria

	Automation	CoO		CxO		I4.0- Manuf	LeO		TechO
Automation	0.966								
CoO	0.177		0.875						
CxO	0.111		0.663		0.788				
I4.0-Base	0.112		0.617		0.506				
I4.0-Manuf	0.99		0.181		0.093	0.951			
LeO	0.125		0.645		0.647	0.124		0.832	
Smart									
Factory	0.961		0.186		0.075	0.989		0.124	
TechO	0.21		0.648		0.634	0.21		0.641	0.855

Following the measurement model testing, hypothesis testing was conducted using the PLS- SEM method. The model fitness value was tested using the standardised root mean square residual (SRMR) value at 0.055, indicating a good model fit level (Henseler et al. 2014).

Results

When the direct effects of customer, competitor, learning and technological orientations on Industry 4.0 base technologies were tested, only the relationship between technological orientation and Industry 4.0 base technologies was statistically significant. Then the hypothesised model shown in Figure 1 was tested, and the testing results are shown in Table 3.

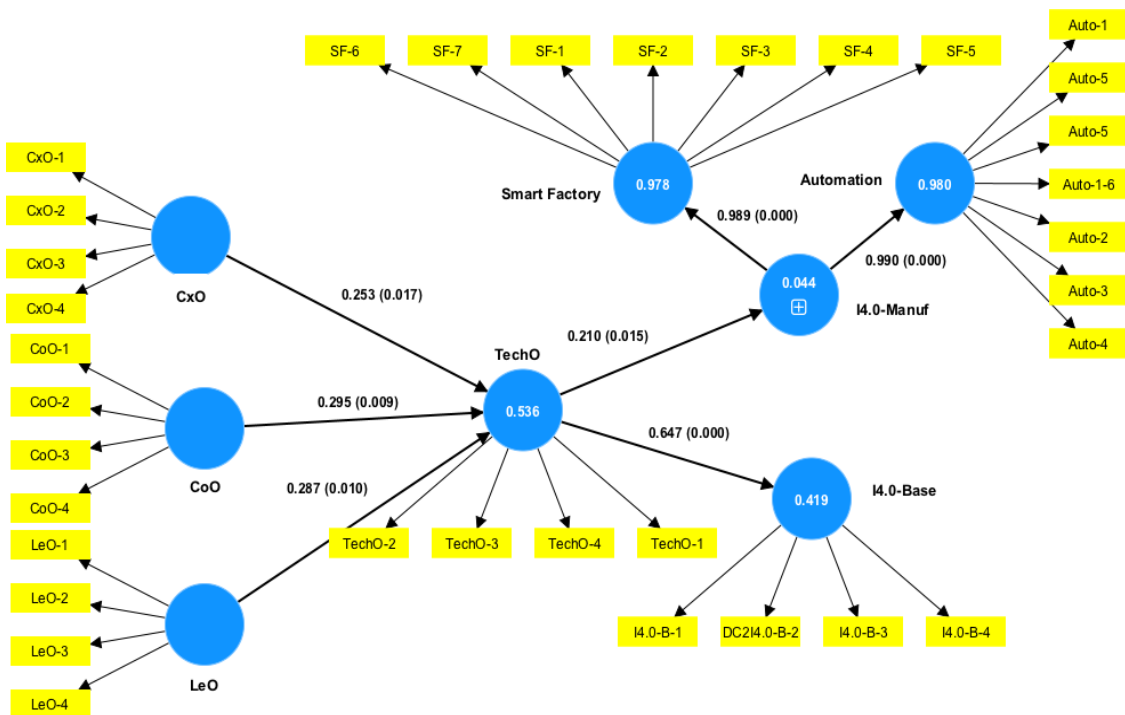
Table 3. Hypothesis testing results

	Sample	Standard deviation	T statistics	
	mean (M)	(STDEV)	(O/STDEV)	P values
CoO -> TechO	0.304	0.113	2.607	0.009
CxO -> TechO	0.252	0.106	2.38	0.017
LeO -> TechO	0.284	0.111	2.577	0.01
TechO -> I4.0-Base	0.66	0.058	11.179	0
TechO -> I4.0-Manuf	0.209	0.086	2.438	0.015

In testing the relationships between the strategic orientations of customer orientation, competitor orientation, learning orientation and technology orientation, the relationship between customer orientation and technology orientation was positive and significant, supporting H1 ($\beta=0.253$, $p<0.05$), between competitor orientation and technology orientation was positive and significant supporting H2 ($\beta=0.295$, $p<0.05$) and between learning orientation and technology orientation was positive and significant ($\beta=0.287$, $p<0.05$) supporting H3. The relationship between technology orientation and Industry 4.0 base technology adoption was positive and significant ($\beta=0.647$, $p=0.000$)

and between technology orientation and Industry 4.0 front-end manufacturing technology adoption was positive and significant ($\beta=0.210$, $p<0.05$). The path coefficient of the relationship between technology orientation and Industry 4.0 base technologies ($\beta=0.647$) was significantly higher than that between technology orientation and Industry 4.0 front-end manufacturing technologies ($\beta=0.210$). R^2 values of dependent constructs, as in Figure 2, show that the model explains 53.6% of the variance in the technology orientation construct, 41.9% of the variance in the Industry 4.0 base technology construct but only 4.4% of the variance in the Industry 4.0 front-end manufacturing technology construct.

Figure 2. Structural model testing results



The Q^2 (predictive relevance) values of technology orientation, industry 4.0 base technologies, and Industry 4.0 front-end manufacturing technologies were at 0.375, 0.217 and 0.04, respectively; thus, they were larger than zero. However, it shows that the predictive relevance of Industry 4.0 front-end manufacturing technologies was very small, even though it was >0 . The f^2 statistic values show that Industry 4.0 capabilities have the strongest effect size on Industry 4.0 base technologies with $f^2=0.722$, but a very weak effect size on Industry 4.0 front-end manufacturing technologies with $f^2=0.046$. All path coefficients were tested by controlling for the number of employees, firm age, and annual revenue, where none had significant effects on Industry 4.0 front-end manufacturing technologies. However, only the annual revenue significantly and positively affected Industry 4.0 base manufacturing technologies ($\beta=0.198$, $p<0.05$), and others had no significant impact.

Discussion and conclusion

The study finds that customer, competitor and learning orientations have positive effects on technology orientation, and the effects are at roughly equal levels indicating that they interact harmoniously. It aligns with the TOE framework (Tornatzky and Fleischer, 1990), where technological, organisational and environmental factors influence technology and innovation in firms. It finds that technology orientation has statistically significant and positive effects on both Industry 4.0 technology adoption rates and Industry 4.0 front-end manufacturing technology adoption rates.

Given that the study found no significant direct effects of customer, competitor and learning orientations on Industry 4.0 technology adoption, it argues that the focus on market dynamics, including customer requirements and competitor behaviour, as well as internal learning capabilities, will lead to Industry 4.0 technological capabilities, only through technological orientation of the firms. Hence, the findings show the importance of the technology orientation of firms for various Industry 4.0 technology adoption. The practitioners are informed of technology orientation, meaning technological innovation capabilities and the acceptance and the drive towards integrating advanced technologies for superior products are important strategic capabilities influencing Industry 4.0 adoption.

For manufacturing firms, it is important to understand that the adoption of Industry 4.0 base technologies and advanced manufacturing technologies have different antecedents. The different effects of technology orientation on Industry 4.0 base and front-end manufacturing technologies were evident in the study. That shows that various other factors influence the adoption of manufacturing-specific advanced technologies in Industry 4.0 relative to base technologies. Hence, the heterogeneity of Industry 4.0 technologies needs to be recognised, and the industry-specific technology requirements are important to be considered in research studies (Hopkins 2021; Frank et al. 2019).

Limitations of the study are known. Even though the study found empirical evidence for the relationships between strategic orientations and Industry 4.0 adoption, it did not explore why the firm focus on customers, competitors, and learning does not directly influence Industry 4.0 technology adoption and how those strategic orientations lead to technology orientation. Further, it needed to investigate how different technology development trajectories firms have taken to extend their Industry 4.0 technologies from base technologies to more advanced manufacturing technologies.

References

- FORNELL, C. & LARCKER, D. F. 1981. Evaluating structural equation models with unobservable variables and measurement error. *Journal of marketing research*, 18, 39-50.
- FRAMBACH, R. T., FISS, P. C. & INGENBLEEK, P. T. 2016. How important is customer orientation for firm performance? A fuzzy set analysis of orientations, strategies, and environments. *Journal of Business Research*, 69, 1428-1436.
- GHOBAKHLOO, M., FATHI, M., IRANMANESH, M., MAROUFKHANI, P. & MORALES, M. E. 2021. Industry 4.0 ten years on: A bibliometric and systematic review of concepts, sustainability value drivers, and success determinants. *Journal of Cleaner Production*, 302, 127052.
- GOLGECI, I. & Y. PONOMAROV, S. 2013. Does firm innovativeness enable effective responses to supply chain disruptions? An empirical study. *Supply Chain Management: An International Journal*, 18, 604-617.
- HENSELER, J., DIJKSTRA, T. K., SARSTEDT, M., RINGLE, C. M., DIAMANTOPOULOS, A., STRAUB, D. W., KETCHEN JR, D. J., HAIR, J. F., HULT, G. T. M. & CALANTONE, R.J. 2014. Common beliefs and reality about PLS: Comments on Rönkkö and Evermann (2013). *Organizational research methods*, 17, 182-209.
- HOPKINS, J. L. 2021. An investigation into emerging industry 4.0 technologies as drivers of supply chain innovation in Australia. *Computers in Industry*, 125, 103323.
- KHIN, S. & KEE, D. M. H. 2022. Factors influencing Industry 4.0 adoption. *Journal of Manufacturing Technology Management*, 33, 448-467.
- MARCUCCI, G., AN TOMARIONI, S., CIARAPICA, F. E. & BEVILACQUA, M. 2022. The impact of Operations and IT-related Industry 4.0 key technologies on organizational resilience. *Production Planning & Control*, 33, 1417-1431.
- NAKANDALA, D., YANG, R., LAU, H. & WEERABAHU, S. 2023. Industry 4.0 technology capabilities, resilience and incremental innovation in Australian manufacturing firms: a serial mediation model. *Supply Chain Management: An International Journal*.
- OZTEMEL, E. & GURSEV, S. 2020. Literature review of Industry 4.0 and related technologies. *Journal of intelligent manufacturing*, 31, 127-182.
- SIGUAW, J. A., SIMPSON, P. M. & ENZ, C. A. 2006. Conceptualizing innovation orientation: A framework for study and integration of innovation research. *Journal of product innovation management*, 23, 556-574.
- SPIESKE, A. & BIRKEL, H. 2021. Improving supply chain resilience through industry 4.0: A systematic literature review under the impressions of the COVID-19 pandemic. *Computers & Industrial Engineering*, 158, 107452.
- ZHENG, P., WANG, H., SANG, Z., ZHONG, R. Y., LIU, Y., LIU, C., MUBAROK, K., YU, S. & XU, X. 2018. Smart manufacturing systems for Industry 4.0: Conceptual framework, scenarios, and future perspectives. *Frontiers of Mechanical Engineering*, 13, 137-150.

Cybercrimes, internet penetration and banking sector performance: Empirical evidence from Nigeria

Faruq Umar Quadri, John O Adeoti

Nigerian Institute of Social and Economic Research (NISER), Nigeria

Abstract

This study examines the links between cybercrime, internet penetration and the performance of the Nigerian banking sector. Data on the two key factors that drive the study (cybercrime and banking performance) are quarterly data from 2011 to 2021 sourced from the Nigeria Deposit Insurance Corporation (NDIC), a statutory entity with the responsibility of managing the Deposit Insurance System (DIS) in Nigeria. A foundational model of a bank that maximizes performance (profit) is developed, and a Two-Stage Least Square (2SLS) technique was employed for data analysis. The findings of the study revealed that internet penetration has a positive and statistically significant impact on banking performance in Nigeria, while cybercrime negatively influences the nexus of internet penetration-banking performance. This implies that as the internet penetration increases in Nigeria, volume of cybercrime increases, and thus damps the impact of internet on banking performance, all things being equal. Since cybercrime has thus become an unintended consequence of internet penetration in Nigeria, the operators in the Nigerian banking industry should implement strict corporate rules that guarantee proper protection of customer data; and ensure the implementation of proper approval protocols for the entire banking system/operation with at least two permission needs for wire transfers.

Keywords: Cybercrimes, Internet penetration, Banking performance, Nigeria

JEL Classification: K24, L11, E71, G21, C32

1.1 Background of the study

The last two decades have brought about a tremendous change in the way banking services are provided due to the quick growth of technologies. The banking business is fast innovating across the board, from analogue banking to digital currency to blockchain applications (Lucey et al., 2018). In Nigeria, internet banking has grown quickly in recent years, just like it has in many other economies of the world. Initially, investment bankers were those predominantly made use of internet service in banking system but currently, however, a notable trend is the gradual acceptance of internet penetration by commercial banks and other non-banking financial operators, in the provision of services to the financial industry (Cummins, et al., 2020).

Even with all the economic advantages that came with internet usage in the banking industry, it also has unanticipated negative impacts on the efficiency of the industry due to illicit activities like credit card fraud (Victor and Oluwatosin, 2010). Cybercrimes are computer-related offenses committed for profits/gains, including identity theft and other offenses involving computer contents that are difficult to define legally as a whole. Fraud is said to occur when an individual or organization in an official capacity with obligations willfully defies the law for their own or their company's gain at the cost of everyone else. It is a global ailment that impacts all institutions and economic systems. Contrarily, bank fraud is the use of questionable practices to gain access to assets held by or belonging to financial institutions (Nwaeze, 2008). The financial system, and banks in particular, are seriously threatened by the rising tide of widespread fraud in recent years in credit intermediaries (Usman & Shah, 2013). Additionally, the prevalence of frauds and forgeries may, in rare circumstances, result in the shutdown of banks (Fatoki, 2015). Internal frauds and credit abuses which were greatly facilitated by the internet penetration is singled out as the cause of many distressed banks in Nigeria (Nwaeze, 2008). In its 2020 annual report for the industry, the NDIC reported that the quantum of fraud cases increased significantly from 52,754 in 2019 to 146,183 in 2020, representing a 177.10% rise. However, the total sum of monetary value involved in 2020 was N120.79 billion as against N204.65 billion in 2019, a decrease of 40.98%. The overall actual loss also decreased somewhat from N5.46 billion in 2019 to N5.33 billion in 2020 (NDIC, 2020). Regrettably, the channels and tools used to carry out these banking frauds and forgeries have gradually changed from traditional to internet/digital channels including ATM cards, mobile banking, and web-based. For instance, fraud involving ATM cards was the most prevalent, equals to 39.81% of all fraud cases, trailed by fraud involving mobile banking at 17.35% (NDIC, 2020). According to Adebisi (2009), the financial institutions face significant obstacles to their survival and viability due to the rise in fraud incidents and the comparatively huge amounts

involved. It is thus imperative that the widespread of fraud, including cyber fraud, in Nigeria's banking sector is given a deserved attention.

Concern over cybersecurity and other fraudulent activities in the banking industry is still quite high in Nigeria, despite the regulations putting in place by the highest bank in the land, the Central Bank of Nigeria (CBN), in conjunction with the roles of the NDIC in the banking cyberspace. This deep concern seems to be indirectly moderating the impact of internet penetration on the industry's performance in Nigeria. According to the NDIC's 2018 report on examinations and special investigations, the banking industry is still battling with a number of challenges, which include cybercrime. Accordingly, fraudulent practices via the Internet and automated teller machines (ATMs) remain one of the elements that has the biggest impact on the performance of the Nigerian banking sector (Akinyomi, 2012; Okpara, 2009). It is thus apparent that the advantage of internet penetration on the performance of Nigeria's banking sector is coming along with unanticipated effects of cybercrimes.

In spite of financial resources that are currently allocated to government agencies such as the Economic and Financial Crimes Commission (EFCC), the Nigerian Police, and other agencies upholding law in Nigeria to combat this menace, their efforts have not yielded a substantial result. They have not successfully tamed or significantly reduced the rate at which this plague ravages the Nigerian economy. For instance, in order to have any real influence on containing this plague, it is predicted by Gartner (2019) that global investment on information security would need to considerably increase to \$124 billion. Yet, according to some security researchers, spending on cybercrime will outstrip investment on cyber-security by nearly 16 times by the end of 2019, when it is projected that expenditures will have quadrupled from their 2015 number to about \$2.1 trillion (Fadairo-Cokers and Ibrahim, 2021). This thus implies that for these agencies to have any meaningful impact in the war against cybercrime, there is need for increase in budgetary allocation to these agencies in a bid acquire cutting-edge smart technologies to tackle this challenge. This consequently, constitutes another drain on the lean revenue of the government at all levels.

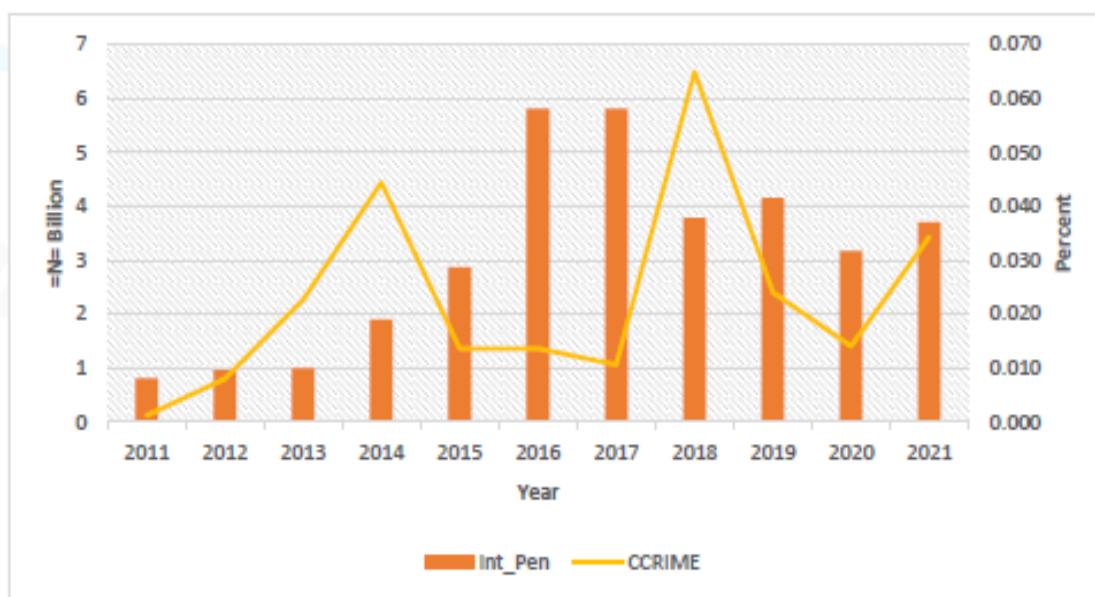
Therefore, the study's objective is to look at the moderating role of cybercrime on the nexus between internet penetration and performance of banking industry in Nigeria. Specifically, to examine the effect of internet penetration on banking performance in Nigeria. And to determine the moderating role of cybercrimes on the nexus between internet penetration and banking performance in Nigeria.

1.2 Trends of cybercrime, and internet penetration in Nigeria: 2011-2021

In figure 1, the trends of cybercrime, and internet penetration in Nigeria from

2011 to 2021 is considered. It can be seen from the Figure that as the internet subscription per 1000 inhabitants increases, cybercrime increases in the following year. For instance, in 2011, when internet sub was at its lowest value of 0.008 per 1000, cybercrime recorded was value at =N= 0.12 billion. This increased to =N= 0.79 billion in 2012 when internet sub was around 0.01 per 1000. In 2017, at the highest internet sub of 0.06 internet sub per 1000, the highest value of cybercrime of =N=6.49 billion was recorded. In summary, the trend of cybercrime seems to be proportional to the trend on internet penetration with some lags.

Figure 4.1: Trends of cybercrime, and internet penetration in Nigeria: 2011-2021



Source: Author's Computation (2023)

2.0 Literature review

2.1 Conceptual clarification

Cybercrime: These crimes include the use of cutting-edge technology. It covers crimes committed through computers and other gadgets that are interconnected to enable the conduct of offenses against computer systems, data, and other offenses (Campbell, 2014). A lot of people have talked about cybercrime from many different perspectives (Ibikunle and Eweniyi, 2013). According to Halder and Jaishankar (2011), cybercrime refers to any conduct with intent to commit a crime that is performed against an individual or collective of individuals with the goal of sullyng their identities and permanently disfiguring hardware harm to important infrastructure, including the web and mobile gadgets.

Mobile Banking: This is a type of financial transaction completed on the go employing a mobile device. Customers of banks can use mobile banking to

check their account balances, make credit card purchases, and get details on their most recent transactions.

Cyber Security: To safeguard an organization's and a user's digital assets, a variety of tools are available, including technology, training, best practices, assurance, and policies. These techniques, procedures, concepts, safeguards, rules, methods of risk management, technology, and tools make up cybersecurity.

Cyberterrorism: According to Lewis (2002), cyber-terrorism refers to the intentional exploitation of a computer network to disrupt essential national infrastructures (such as energy, transportation, and government activities), including by intimidating or coercing both public and private individuals.

Cyber Stalking: Cyber stalking involves someone using the internet facilities such as electronic devices, according to Ellison and Akdeniz (1998). Online abuse, cyberstalking, and online harassment are all interchangeable phrases. Instead of physically threatening the victim, the perpetrator gathers information about the victim's online actions, watches those behaviors, and then threatens the victim.

2.1.2 Performance indicators in the banking sector

The majority of indicators used to assess the performance of the banking system are bank profitability ratios, although other indicators include the proportion of non-performing loans or the ratio of bank capital to assets. Profitability also reflects the banks' ability to make money. It emphasizes the bank management's managerial skill. Profitability ranks as the most substantial and dependable metric since it offers a general indication of a bank's ability to enhance its earning level Dasgupta (2002). Every commercial organization continues to exist for one reason: to make money. Profitability rate and profit margin are seen as measures of how effectively banks use their resources. Indicators of profitability are often used to assess a commercial bank's overall risk posture. (Drigă, 2009). Efficiency and profitability analysis in banks dependent upon a given set of ratios:

The Profit-to-Turnover Ratio: Numerous studies have backed the profitability indicator known as the profit-to-turnover ratio (Hall, et al., 2000). Additionally, profitability remains a helpful indicator for evaluating a firm's capacity to make a earnings (Atrill & McLaney, 2006). Additionally, profitability as a ratio disregards the effect of magnitude of a firm on its capacity to generate profit, this is important since certain businesses may produce bigger revenues simply because of their magnitude. As a result, it is acceptable to emphasize profitability.

Fixed-to-Total Asset Ratio: According to Reid and Jacobsen, external financiers prefer to utilize fixed asset as collateral for loans since the fixed-to-total asset ratio is a good tangibility scale (1988). In some additional research, this ratio was also used as a tangibility indicator (Chittenden, et al., 1996; Huang & Song, 2001). On the other hand, a high tangibility does not necessarily mean that a company has more fixed assets available as loan collateral, but rather that fixed assets make up a sizable fraction of the company's overall assets. As a guarantee against default risk, banks would utilize the total value of actual assets rather than a ratio as loan collateral.

Return on Assets (RoA): Return on assets (ROA), a financial indicator, measures a firm 's performance in proportion to its total assets. By examining its ROA, corporate management, analysts, and investors can assess a company's resource usage effectiveness. The metric is frequently calculated as a percentage using the average assets and net income of a firm. A corporation can manage its balance sheet more skillfully and profitably if its ROA is higher, whereas a weaker ROA suggests there is space for progress (Aduda & Kingoo, 2012).

Return on Equity (ROE): The measure of a firm's net income in proportion to every penny of equity that shareholders invest is called return on equity (RoE). The ROE statistic, which is commonly expressed in percentage form, can be a very useful tool to evaluate a management team's capital allocation decisions and ability to support shareholder value development (see, for instance (Oyewole et al., 2013).

Debts-to-Shareholder's Equity Ratio: Gearing is a measure that quantifies the ratio of debt to equities of shareholders utilized to fund the operations of a firm, as is shown from a number of studies (Gomes & Schmid, 2010; Baturevich & Muradoglu, 2010). Companies in different industries frequently have a variety of asset structures, and the financing arrangements for these companies are based on the makeup of these asset structures. a commercial bank's position (Drigă, 2009).

2.2 Theoretical framework

The blend of Technology-Enabled Crime theory and theory of Structure-Conduct-Performance (SCP) are adopted as theoretical framework to drive the study. For the purpose of assisting society in better understanding how criminal behavior changed in tandem with computer and communications technology, the theory of technology-enabled crime is made up of numerous kinds of criminological theories. According to McQuade (1998), it is difficult to understand and manage more complex criminality at first. There is also an ongoing technological arms race between criminal organizations and law enforcement organizations. Criminal conduct must be monitored by law enforcement as it changes and becomes more inventive to avoid, manage,

prevent, and stop new kinds of crime. Certain activities, according to McQuade (2006), are part of technology-enabled crime theory. This theory offers a framework for comprehending all types of criminality, particularly those that are developing as a result of advancements and discoveries in computing and telecommunications technology.

According to the SCP model, market structure has an impact on bank conduct, which has an impact on performance. Banks are more likely to act in a *quid pro quo* manner in a highly concentrated market, and their oligopoly rents enhance profitability. In this instance, market concentration is used to infer behavior, which is an unobservable. We may utilize our fundamental profit model to generate the SCP relationship even though the SCP hypothesis lacks a formal foundation.

2.3 Empirical literature

Oluwagbemi, et al. (2011) claim that the use of IT facilities has caused significant modifications to the character and standard of the operations of the Nigerian financial market. Sreedevi & Meena (2011), opined that the presence of IT in the sector offered a variety of benefits, including simple exposure to banking services, insurance facilities, and relevant guidance or information. A nation's financial system can be made more inclusive through the combination of mobile and internet banking. The broad usage of the internet clearly encourages economic growth via the banking system.

The traditional operations of commercial banks are being impacted by the increased competition caused by internet penetration in China's financial services sector (Funk, 2019). Also, it might have an effect on the financial operation and the industry operating framework (Gonzalez & Loureiro, 2014). On the flip side, it is undeniable that the growth of the Internet has given commercial banks the push and chance to reform. The advancement of Internet technology has reduced asymmetric information and processing fees, which has helped to expand the financial market and rebuild the monetary system (Srivastava, 2014). The growth of the Internet has benefited commercial banks since it has compelled established banks to modernize their framework and boost their comprehensive organizational performance (DeYoung et al., 2007).

Mobile phone usage exhibits a favorable effect on economic performance, claim Andrianaivo & Kpodar (2012). In addition, there is a favorable and strong correlation between internet penetration rates and financial inclusion. Moreover, Mago & Chitokwindo (2014) in Zimbabwe found that mobile banking has been popular among the country's poor since it is simple to use, quick, affordable, and secure. In similar research, Kpodar & Andrianaivo (2011) find that the introduction of internet devices substantially accelerates the expansion of this sector in Africa since it promotes financial inclusion.

Moreover, Oyewole, et al., (2013) revealed that electronic banking has been of immense benefits to the growth of banking in Nigeria in terms of return on assets (ROA) and net interest margin (NIM), but not return on equity (ROE).

We can conclude that the bunch of current research reviewed suggest that internet penetration ultimately leads to increased tremendously the industry's performance in Nigeria.

Between the years of 1994 and 2015, Funso & Olaniyan (2018) examined how fraud affected the operation of Nigerian deposit money institutions. According to this study, the deposit of banks in Nigeria are negatively and significantly impacted by the quantities of cases of fraud, the loss of funds due to fraud, and the volume of staff members participated. The study's results, which were obtained using ordinary least squares regression methods, demonstrated that fraud showed a detrimental influence on the money market in Nigeria. The influence of bank cyber fraud on the expansion of Nigerian financial market was explored by Iyodo, et al., 2016. They find that the development of the industry in Nigeria is negatively and remarkably impacted by banking frauds. The proliferation of the internet has made banking procedures more complex; this has altered how bank fraud is perpetrated and other illicit activities. The volume of internet banking transactions has increased, according to Berney (2008), as more and more customers rely on the internet for their financial requirements. Cybercriminals can target the clients of banks who are only available virtually to confirm their online banking transactions, according to Gates and Jacob (2009) and Malphrus (2009).

A variety of modern cyberthreats and attacks, according to Muhammad and Sulaiman (2017), are crippling Nigeria's financial institutions and other agencies whose services rely on the internet platform. The claim is that Nigeria has entered one of its difficult periods as a result of the daily escalation of cybercrimes, which have cost the country billions of dollars.

In conclusion, we found that avalanche of literature already investigated the impacts of cybercrimes on the profitability of the sector and they majorly conclude that the impact has been detrimental to the sector. While the literature reviewed penetration of internet on the profitability of the banking sector seem to conclude that its effect has been positive. However, as far as we reviewed, no study is found to have examined the net effects of cybercrime and internet penetration combined on the performance of the sector. This leaves a void in the literature, which is what this study is trying to address.

3.0 Research methodology

3.1 Introduction

Three sub-sections make up this component of the study: the first covers the

descriptions of the data sources and variables; the second discusses the analytical framework; and the third describes the analytical methods to be used in the investigation.

3.2 Data sources and description

Two sources were used to gather the data: the Nigeria Deposit Insurance Corporation (NDIC) and the World Bank (WDI). The Deposit Insurance System (DIS) in Nigeria is managed by the NDIC, a statutory organization created by NDIC Act No. 16 of 2006. The NDIC as deposit-taking institution is given a license as part of the configuration of the country's need. One of such aspects is to safeguard customers and ensure the fulfillment of secured funds. The World Bank's data is used for the remaining variables, while the NDIC's data on cybercrime and banking industry performance serves as the study's main driving force (WDI, 2022).

The data are of quarterly frequency ranging from 2011Q1 to 2021Q1 as extracted from various sources. These series are originally annual data but due to limitations of some of the series, we quarterized them. Many benchmarking and quarterlization procedures are used, according to Bloem et al. (2001). In order to interpolate annual flow times data, the study uses the Proportional Denton Method, which is advised by the International Monetary Fund and the World Bank (IMF) (Osuji, 2013).

Table 1 presents information about the series employed in the paper for our estimation. These consist of our dependent variables of banking sector performance which will be measured by ROA and ROE, and the control variables that include internet penetration (Int_Pen), youth unemployment (Yun), real gross domestic product per capita (RGDPPC), and Human Capital Development (HCD).

Table 1: Data sources and description

Variables	Acronyms	Measurement	Source
Banking Sector Performance	Bank_Perf	Return on Asset (ROA) and Return on Equity (ROE)	Nigerian Deposits Insurance Corporation (NDIC)
Cybercrimes	CCRIME	Contribution of Cybercrime to Total Fraud in Nigeria (N' Billion)	Nigerian Deposits Insurance Corporation (NDIC)
Internet Penetration	ICTs	Fixed Internet suscriptions per 100 inhabitants	World Development Indicators (WDI) 2022
Economic	RGDP	Real Domestic Growth	World Development

Variables	Acronyms	Measurement	Source
Performance		Product	Indicators (WDI) 2022
Youth Unemployment Rate	Yun	% of total labor force ages 15-24	World Development Indicators (WDI) 2022
Poverty level	RGDPPC	Per capita income	World Development Indicators (WDI) 2022
Human Capital Development Index	HCD	Secondary School enrolment	Penn 10.1

Source: Compiled by the Author.

3.3 Analytical framework

A foundational model of a bank that maximizes performance (profit) is developed in this section. Without sacrificing generalization, no fixed costs is assumed at all, and that all inputs are perfect complements with no cross-price elasticity (in the long run).

Each bank i aims to maximize its earnings with the aid of the transformation function T and the pricing possibility set H . These factors show how a bank is perceived in terms of its level of competition and its clients' capacity to pay the rates it sets. The price opportunity setter includes Z , the level of equity.

For every output in the output vector Y_i (Y), each bank is required to estimate the price p using the opposite demand function. Then, bank i maximizes:

$$\Pi_i = \rho Y_i - w_i X_i$$

Subject to;

$$T(X_i, Y_i) = 0$$

$$H(\rho, Y_i, w_i, Z_i) = 0$$

$$\rho = f\left(\sum_{j=1}^N Y_j\right) = f(Y)$$

where N represents volume of firms in banking industry in the economy and f (Y) is inverse market demand. The related Lagrangian system is denoted by:

$$L = \rho Y_i - w_i X_i - \xi T(\bullet) - \theta H(\bullet) \quad - \quad - \quad - \quad - \quad - \quad - \quad - \quad (1)$$

The ideal input and output amounts and prices are obtained by simultaneously solving for p and X (denoted by asterisks):

$$\begin{aligned}\rho^* &= \rho(Y_i, w_i, Z_i) \\ X_i^* &= X_i^*(Y_i, w_i, Z_i)\end{aligned}$$

Profits are maximized if,

$$\frac{d\pi_i^*}{dY_i} = \rho^* + Y_i f'(Y) \frac{dY}{dY_i} - w_i \frac{dX_i^*}{dY_i} = 0 \quad (2)$$

where the demand for outputs Y_i determines the ideal number of inputs X_i^* . Y_i multiplied by results in:

$$\rho^* Y_i - w_i \frac{dX_i^*}{dY_i} Y_i = -(Y_i)^2 f'(Y) \left(\frac{dY}{dY_i} \right) \quad (3)$$

where " Y_i " stands for revenue. In this case, banks are supposed to conduct business in output markets characterized by price-segmentation, albeit facing perfectly competitive input markets. As a result, by altering the costs and prices of their output pricing schemes, banks can compete in response to changing characteristics of the market. Output numbers, input costs, and other factors that are all presented at the time of price setting decide how much of an impact they can have on pricing. We can overlook output prices in the empirical analysis even when not needed in the analytical framework, they are vulnerable to significant measurement mistakes.

We also alter and rearrange equation 3 to produce a model that is better consistent and in tandem with reality. We start by defining I as follows:

$$\frac{dY}{dY_i} = 1 + \frac{d \sum_{j \neq i} Y_j}{dY_i} = 1 + \lambda_i \quad (4)$$

where λ_i is the output-related hypothetical dispersion for firm i . Equation 3's result after substituting λ_i is:

$$\rho^* Y_i - w_i \frac{dX_i^*}{dY_i} Y_i = -(Y_i)^2 f'(Y) (1 + \lambda_i) \quad (5)$$

$$\frac{\rho^* Y_i - w_i \frac{dX_i^*}{dY_i} Y_i}{\rho^* Y_i} = - \frac{Y_i f'(Y) Y}{Y \rho^*} (1 + \lambda_i) \quad (6)$$

Even though the SCP hypothesis is devoid of a theoretical basis, nevertheless, we may be able to produce the SCP relationship using our basic profit model. We first get our core framework by multiplying Equation 4.5 by the number of enterprises:

$$\rho^* Y - \sum_{i=1}^N w_i \frac{dX_i^*}{dY_i} Y = - \sum_{i=1}^N \left(\frac{Y_i}{Y} \right)^2 (f'(Y) Y^2) (1 + \frac{\sum \lambda_i Y_i}{\sum Y_i^2}) \quad (7)$$

Dividing both sides by $\rho * Y$ gives:

$$\Pi^* = \rho^* Y_i - w_i \frac{dx_i^*}{dy_i} Y_i = - \left((HHI) \left(\frac{1}{\eta} \right) (1 + \mu) \right) \quad (8)$$

Where the Herfindahl–Hirschman Index, $HHI = \sum_{i=1}^N \left(\frac{Y_i}{Y} \right)^2$, $\left(\frac{1}{\eta} \right) = f'(Y) Y^2 / p * Y$ and $\mu = \frac{\sum \lambda_i Y_i}{\sum Y_i^2}$.

We need to make two more assumptions in order to arrive at the fundamental SCP relationship.

A constant degree of responsiveness of price to demand is the point number one to be noted. If not, the interpretation of an HHI coefficient may be skewed inversely (positively) by changes in the responsiveness of demand over a period of time in the absence of a proxy for.

The second premise focuses on how each firm conjectures that other firms will react to a change in production, or how much it conjectures that other firms will do so. Two alternatives exist. First, we assume it is constant and consistent across firms. Thus, it is taken out of the previous equation and a link between concentration and performance remains. The second option, assuming collusive action, is to openly acknowledge the collaboration between and HHI. We may demonstrate that growing market share MS_i or concentration HHI is projected to raise consciousness (μ) and, as a result, bolster collusive behavior, according to Stigler (1964).

Even though this still limits us from measuring precisely, It does enable us to assess its impact using HHI. After all, the performance improvement for the complicit oligopolist is greater than proportional to the rise in concentration. Instead, the rents lost by non-collusive behavior increase with market growth.

Altogether, we have demonstrated a fundamental connection between performance and structure that, if we assume and that it is a constant function of HHI, then the SCP relationship holds true. As a result, the fundamental equation becomes:

$$\Pi^* = ((HHI)(1 + \lambda)) \rho^* Y_i \quad (9)$$

Interpreting the joint effect of λ and HHI on banks' efficiency is what the model essentially does.

3.4 Analytical technique and empirical model

To achieve the objective of this study that examines the impact of internet penetration on banking sector performance which is indicated by two proxies: return on assets (RoA) and return on equity (ROE) (RoE) as defined in Table 1.,

was achieved employing the Two-stage Least Square (2SLS). The potential endogeneity/simultaneity that may stem from the reverse causality between CCRIME and IINT_PEN is addressed using instrumental variable procedure in the following reduced form equations:

$$\ln Bank\_Per_t = \alpha_0 + \alpha_1 \ln Int\_Pen_t + \alpha_2 \ln RGDP_t + \alpha_3 \ln HCD + \varepsilon_t \quad (10)$$

$$\ln Bank\_Per_t = \alpha_0 + \alpha_1 \overline{\ln Int\_Pen_t} + \alpha_2 \ln RGDP_t + \alpha_3 \alpha_2 \ln HCD + \varepsilon_t \quad (11)$$

Where, Bank_Per denotes bank performance capture with ROA and ROE, Int_Pen is the Fixed Internet suscriptions per 100 inhabitants, *RGDP_t* captures economic growth in Nigeria, *CD* is the Human Capital Development Index.

Finally, equation (12) incorporates the conditional implications and provides a sophisticated approach for modeling the moderating influence of cybercrimes on the link between internet usage and the performance of Nigeria's banking system. By include the multiplication of the cybercrime and internet penetration variables as one of the explanatory variables in the equation, the conditional effect is captured.

$$\ln Bank\_Per_t = \alpha_0 + \alpha_1 \overline{\ln Int\_Pen_t} + \alpha_2 \ln Ccrime_t + \alpha_3 \ln Int\_pen * Ccrime_t + \alpha_4 \ln RGDP_t + \alpha_5 \ln HCD + \varepsilon_t \quad (12)$$

where Cybercrime denotes CCRIME to total fraud in Nigeria; $\ln Int\_pen * Ccrime_t$ is the interactive term of internet penetration and cybercrime; and all other factors are still as they were before. By considering partial derivatives of equation (12), one may derive the total impact of internet penetration, which also accounts for the marginal impact of cybercrime on banking sector performance:

$$\frac{\partial Bank\_Per_t}{\partial Int\_Per_t} = \alpha_1 + \alpha_3 Ccrime_t \quad 13$$

This should be interpreted in terms of both sign and magnitude. Regarding sign, if $\alpha_1 > 0$ and $\alpha_3 < 0$, cybercrime dampens the impact of internet penetration on banking performance. However, if $\alpha_1 < 0$ and $\alpha_3 > 0$, it reflects that cybercrime ameliorates the impact of internet penetration on banking performance. Meanwhile, if $\alpha_1 > 0$ and $\alpha_3 > 0$, then cybercrime scale-up the impact of internet penetration on banking performance. Lastly, if $\alpha_1 < 0$ and $\alpha_3 < 0$, then cybercrime and internet penetration jointly scale-down banking performance in Nigeria. In terms of magnitude,

if $\frac{\partial Bank\_Per_t}{\partial Int\_Per_t} > 0$, internet penetration and cybercrime positive influence banking performance but
if $\frac{\partial Bank\_Per_t}{\partial Int\_Per_t} < 0$, both internet penetration and cybercrime negatively influence banking performance in Nigeria.

4.0 Results and discussions

There are three subsections within this section. The first portion goes into data concerns and quick analyses, the second section covers empirical analysis, and the third section examines the findings.

4.1 Data and preliminary analyses

4.1.1 Descriptive statistics

Table 2 displays the descriptive statistics for these variables. The table shows that CCRIME on average within this period is =N= 2.28 billion. The maximum CCRIME for the country was =N=7.2 billion and the minimum of =N= 0.004 billion which was recorded in 2018q3 and 2011q1. The average HCD in the period was 0.503 with the maximum HCD of 0.61 over this period recorded in 2015q1, and its minimum of 0.42 was recorded in 2014q2. The average INT_PEN rate over the period is 0.033 per 1000 head with a maximum of 0.075 and minimum of 0.002 per 1000 head in quqrtrs 2011q1 and 2012q4 respectively. Also, the average RGDP in the period under consideration was US\$472 billion while US\$398 billion and US\$517 billion were maximum and minimum values respectively. Furthermore, we observe the Nigerian banking performance using ROA and ROE. On the average, the ROA was 1.94 per cent with its largest value of 2.96 percent and minimum of 0.17 percent in 2012q4 and 2011q3 respectively. The ROE was on the average 18.09 per cent with its largest value of 26.80 percent and minimum of 0.82 percent in 2018q4 and 2011q3 respectively. Lastly, the YUN rate was an average of 13.35 percent at maximum value of 19.94 percent and minimum of 7.53 percent in 2020q4 and 2015q1 respectively.

Table 2: Descriptive statistics

Variables	Obs	Mean	Std. Dev.	Min	Max	Skew.	Kurt.	JB-Statistic
CCRIME	44	2.284	1.91	-.004	7.219	.969	3.11	6.640
HCD	44	.503	.041	.422	.608	.534	3.373	3.600
INT PEN	44	.033	.022	0.002	.075	.107	1.993	4.650
RGDP	44	4.720e+11	3.430e+10	3.980e+11	5.170e+11	-.962	2.692	6.320
ROA	44	1.943	.752	.167	2.958	-1.232	3.351	9.260
ROE	44	18.093	7.56	.817	26.804	-.99	2.677	6.570
YUN	44	13.348	4.286	7.529	19.939	.265	1.528	21.610

Source: Author's Computation (2023)

The skewness statistic shows that CCRIME, HCD, INT_PEN, and YUN are positively skewed, while RGDP, ROE and ROA are negatively skewed. The kurtosis statistic shows that the INT_PEN, RGDP, ROE and YUN are platykurtic; having kurtosis statistic of less than 3. While the kurtosis of CCRIME, HCD, and ROA is excessive ($k > 3$). The normality hypothesis can be

rejected for CCRIME and HCD at the 5% level of significance, indicating that they are not normally distributed. This is revealed by the Jarque-Bera statistic, which combines the skewness and kurtosis statistics to determine the normality of the variable. For the other model variables, the null is accepted at 1% and 10% level of significance, respectively.

By looking at the correlation between the relevant variables, we continue our preliminary investigation. Instead of using the modified variables in the study, the correlation was done on the primary variables. Table 3 displays the correlation's findings. The finding indicates that no two regressors have a perfect (one-to-one) correlation, which suggests that the multicollinearity issue is less likely than previously thought. ROA negatively correlates with YUN minimally other variables. This has no implication for multicollinearity as it is a correlation between a dependent and an independent variable.

Table 3: Matrix of correlations

Variables	CCRIME	HCD	INT_PEN	RGDP	ROA	ROE	YUN
CCRIME	1.000						
HCD	-0.348	1.000					
INT_PEN	-0.216	-0.304	1.000				
RGDP	0.473	0.002	0.033	1.000			
ROA	0.384	0.048	-0.694	0.277	1.000		
ROE	0.486	0.024	-0.647	0.435	0.887		
YUN	0.239	-0.177	0.403	0.670	-0.003	0.292	1.000

Source: Author's Computation (2023)

4.1.2 The unit root test results

Another preliminary analysis was done using the Unit root test. A test for the series' stationarity is the unit root test. This is essential because it is possible for a non-stationary variable to cause a misleading regression when it is regressed on another non-stationary variable. Three common unit root tests—the Augmented Dickey Fuller (ADF), Phillip-Perron (P-P), and Kwiatkowski-Phillips-Schmidt-Shin (KPSS) Tests—were used to run the test. In the absence of the unit root null hypothesis, a (level) series is stationary $I(0)$. When the series has been first-differenced and the null hypothesis is rejected, the series is $I(1)$. The results of the test are displayed in Table 4. The test shows that CCRIME, HCD, INT_PEN, RGDP, ROE and YUN are stationary after first difference, $I(1)$, while ROE is stationary at level, $I(0)$. The KPSS and PP values are very consistent with one another. The unit root results imply that a method for cointegrating these variables that is appropriate must take into consideration mixed orders of integration. Therefore, the ARDL bound testing cointegration approach is used in this work.

Table 4: Unit root results

	ADF Test			Phillip-Perron Test			KPSS Test		
	Level	First Diff	Remark	Level	First Diff	Remark	Level	First Diff	Remark
CCRIME	-6.12*	-5.17*	H(0)	-2.39	-2.84**	H(1)	0.06	0.042	H(0)
HCD	-9.82*	-5.95*	H(0)	-2.96**	-3.02***	H(1)	0.05	0.03	H(0)
INT_PEN	-5.23*	-4.35*	H(0)	-2.381	-2.35	H(2)	0.11	0.13	H(0)
RGDP	-2.67**	-3.64*	H(0)	-1.95	-2.28*	H(1)	0.23	0.09	H(0)
ROA	-6.94*	-5.08*	H(0)	2.61***	-3.05**	H(0)	0.08	0.06	H(0)
ROE	-4.93*	-4.27*	H(0)	-2.19	-2.95**	H(1)	0.09	0.07	H(0)
YUN	-1.46	-4.11*	H(1)	-0.13	-2.46	H(1)	0.20	.11	H(0)

Source: Author's Computation (2023)

Note: Following modeling with constant, constant & trend, and with none, the reported unit root results come from the best unit root models. Models having none, constant, and constant & trend are denoted, respectively, by superscripts a, b, and c. Additionally, the unit root null hypothesis is disproved at 1%, 5%, and 10% levels of statistical significance, respectively, as indicated by the asterisks *, **, and ***.

4.3.3: Examining the impact of cybercrime on banking performance in Nigeria

In making a choice between two estimators of Ordinary Least Square (OLS) and the 2_stage Least Square (2SLS), the Hausman test statistic developed by Hausman in 1978 was employed. Thus, the hypothesis tested is:

H_0 : The efficient estimator (OLS) is consistent Vs H_1 : The efficient estimator (OLS) is inconsistent

We estimated and analyzed the 2SLS estimator since the estimation of the Hausman was determined to be statistically significant at ($p < 0.05$ & $p < 0.1$) and rejected. As a result of its superior performance over the Dynamic OLS estimator, the analysis is based on the 2SLS estimates. The table 5 and 6 depict the findings for the 2SLS estimator as shown below. As was previously mentioned, two empirical models were estimated to proxy banking performance. The dependent variable in the first model is Return on Assets (ROA), while the regressor in the second model is Return on Equity (ROE).

The result depicts the observed relation between banking performance and internet penetration in Nigeria. In the first model, the findings show that there is a positive and statistically significant relationship ($P < 0.1$) between internet penetration and ROA. It shows that a 1% increase in internet penetration in Nigeria leads to approximately 0.42 percent increase in return on assets. In other words, this implies that as the internet penetrates in Nigeria, banking performance in terms of ROA rises. However, in model two, the result shows a non-positive and statistically insignificant relation between internet penetration

and banking performance in terms of ROE. This implies that a 1% increase in penetration results in 0.398% decrease in return on equity. We can thus infer that internet penetration has a positive and significant impact in banking sector in Nigeria. This finding corroborates with that of Funk (2019), Gonzalez & Loureiro (2014), and Oyewole, et al., (2013) in China, Spain and Nigeria respectively.

Accordingly, it is unexpectedly found that HCD, when all other factors are held constant, has a negative but statistically nonsignificant impact on ROA while it exhibits a negative and significant impact on ROE at ($p < 0.10$). This suggests that the performance of the banking industry is negatively impacted by HCD. That is human capital does not positively influence efficiency of banking sector in Nigeria. Last, the level of economic activity as measured by RGDP shows a statistically significant and positive impact on ROA and also, on ROAE. This implies that the level of economic activities positively influences banking performance in Nigeria, *ceteris paribus*.

Examining the goodness of fits of the models, the R-square values are found to be 40 percent and 44 percent respectively. This suggests that, after accounting for the amount of regressors, the models is able to explain around 40 and 44 percent variations in ROA and ROE, respectively. Meanwhile, using Chi-square statistics in the two models, at 1 percent, clearly the Chi-square values are found to be statistically significant. There is therefore no justification for not rejecting the null hypotheses. In conclusion, we may acknowledge that empirical models are suitable for assessing banking performance. As a result, the Chi-square test, which evaluates the estimated regression's overall significance, also evaluates the R-square's significance (Gujarati, 2004).

Table 5: Analysis of the effect of Internet Penetration on banking sector performance

	lnROA		lnROE	
	Coef.	t-value	Coef.	t-value
lnINT_PEN	0.421*	4.25	-0.398	-3.84
lnRGDP	4.38*	3.66	5,657*	4.53
lnHCD	-2.504	-2.18	-1.983***	-1.65
Constant	-120.57*	32.38	-152.23*	-4.51
	Diagnostics		Diagnostics	
R-squared	0.398		0.447	
Chi-square	24.457		27.371	
Prob > chi2	0.000		0.000	

Source: Author's Computation (2023)

Table 6 depicts the results of moderating role of cybercrimes on relation between banking performance and internet penetration in Nigeria. In the first model, the findings show that cybercrimes negatively and significantly ($P < 0.1$)

moderate the impact of internet penetration on ROA. It shows that a 1% increase in cybercrime damps the relations between internet penetration and ROA by 0.22%. In other words, this implies that cybercrime negatively affects the nexus between internet penetrates and ROA in the sector. Similarly, in model 2, the findings show that cybercrimes negatively and significantly ($P < 0.1$) moderate the impact of internet penetration on ROE. It shows that a 1% increase in cybercrime damps the relations between internet penetration and ROE by 0.44%. In other words, this implies that cybercrime negatively affects the nexus between internet penetrates and ROE in the sector. We can thus infer that cybercrimes negatively and significantly influence the impact of internet penetration on the performance of banking sector in Nigeria. This finding was supported by Funso & Olaniyan (2018), and Iyodo, Agbaji, & Abu (2016) in Nigeria.

Examining the goodness of fits of the models, using Chi-square statistics in the two models, at 1 percent, clearly the Chi-square values are found to be statistically significant. There is therefore no justification for not rejecting the null hypotheses. In conclusion, we may acknowledge that empirical models are suitable for assessing the moderating role of cybercrime on the nexus between internet penetration and banking performance in Nigeria. As a result, the Chi-square test, which evaluates the estimated regression's overall significance, also evaluates the R-square's significance (Gujarati, 2004).

Table 6: Analysis of the effect of Internet Penetration on banking sector performance

	lnROA		lnROE	
	Coef.	t-value	Coef.	t-value
lnINT_PEN	0.347	1.57	0.437	1.89
lnRGDP	-4.438	-1.51	-4.519	-1.46
lnHCD	4.508	1.9	5.919	2.39
LINT_PEN_CCRIME	-0.217	-3.25	-0.257	-3.67
Constant	123.87	1.53	129.52	1.53
	Diagnostics		Diagnostics	
R-squared	0.074		0.144	
Chi-square	20.89		28.215	
Prob > chi2	0.000		0.000	

Source: Author's Computation (2023)

5.0 Conclusion and policy recommendations

5.1 Conclusion

In this study, we investigated the cybercrime, internet penetration and performance of banking industry in Nigeria. This was done by examined the impact of cybercrimes on the nexus between internet penetration and banking industry performance in Nigeria.

We chose between using the Ordinary Least Square (OLS) and the Two-Stage Least Square (2SLS) estimators to analyze the effect of cybercrime on Nigerian banking performance. The Hausman test statistic, which Hausman invented in 1978, was used. We estimated and analyzed the 2SLS estimator since the estimation of the Hausman was determined to be statistically significant at ($p < 0.05$ & $p < 0.1$) and rejected. The findings from the analysis of the variables influencing banking performance in Nigeria revealed a positive and significant influence of internet penetration, and level of economic activities while human capital shows no significant impacts. In other words, this implies that as the internet penetration increases in Nigeria, banking sector performance increases, all things being equal. However, our findings reveal that cybercrime plays a negative moderating role on the nexus between internet penetration and banking performance.

5.2 Policy recommendations

In the light of the foregoing conclusion, the following consequences for policy are suggested:

- i. We recommend that the Central Bank of Nigeria (CBN) mandate the players in the industry particularly the deposit money banks (DMB) to dedicate about 0.1% of their net profit/gross profit towards the fight against cybercrime in Nigeria. This is to improve the shortage of finances already confronting the Nigeria's law enforcement agencies.
- ii. We recommend that appropriate approval protocols must be implemented throughout the entire banking system and operation in coordination with the Nigeria Deposit Insurance Commission (NDIC) and the Central Bank of Nigeria (CBN), with at least two approval requirements for wire transfers or clearance transfers and others.
- iii. There is also a need for the government at all levels in Nigeria, to deal ruthlessly with the socio-economic menace of youth unemployment and poverty which is the basic reason for the youth involvement in crimes in general, in Nigeria.

References

- Abdulkadir, D. S., Isiaka, S. B., & Adedoyin, S. I. (2012). Effects of strategic performance appraisal, career planning and employee participation on organizational commitment: An empirical study. *International Business Research*, 5(4), 124.
- Adegbola, I., & Fadara, O. (2022). Cyber Crime among Mathematical Science Students: Implications on their Academic Performance. *JOURNAL OF DIGITAL LEARNING AND DISTANCE EDUCATION*, 1(2), 47-54.
- Adesina, O. S. (2017). Cybercrime and poverty in Nigeria. *Canadian social science*, 13(4), 19-29. Adomi, E. E., & Igun, S. E. (2008). Combating cyber crime in Nigeria. *The Electronic Library*.
- Aduda, J., & Kingoo, N. (2012). The relationship between electronic banking and financial performance among commercial banks in Kenya. *Journal of finance and investment analysis*, 1(3), 99-118.
- Akinyomi, O. J. (2012). Examination of fraud in the Nigerian banking sector and its prevention. *Asian Journal of Management Research*, 3(1), 182-194.
- Andresen, M. A. (2013). Unemployment, business cycles, crime, and the Canadian provinces. *Journal of Criminal Justice*, 41(4), 220-227.
- Andrianaivo, M., & Kpodar, K. (2011). ICT, financial inclusion, and growth: Evidence from African countries.
- Andrianaivo, M., & Kpodar, K. (2012). Mobile phones, financial inclusion, and growth. *Review of Economics and Institutions*, 3(2), 30.
- Atrill, P., & McLaney, E. J. (2006). *Accounting and Finance for Non-specialists*. Pearson Education.
- Avizienis, A., Laprie, J. C., Randell, B., & Landwehr, C. (2004). Basic concepts and taxonomy of dependable and secure computing. *IEEE transactions on dependable and secure computing*, 1(1), 11-33.
- Becker, G. S. (1968). Crime and punishment: An economic approach. In *The economic dimensions of crime* (pp. 13-68). Palgrave Macmillan, London.
- Campbell, H. (2014). Religion and the Internet in the Israeli Orthodox context. In *New Media, Politics and Society in Israel* (pp. 58-77). Routledge.
- Cantor, D., & Land, K. C. (1985). Unemployment and crime rates in the post-World War II United States: A theoretical and empirical analysis. *American Sociological Review*, 317-332.
- Chittenden, F., Hall, G., & Hutchinson, P. (1996). Small firm growth, access to capital markets and financial structure: Review of issues and an empirical investigation. *Small business economics*, 59-67.
- Clough, J. (2012). The Council of Europe Convention on cybercrime: defining crime in a digital world. In *Criminal Law Forum* (Vol. 23, No. 4, pp. 363-391). Springer Netherlands.
- Cornwell, C., & Trumbull, W. N. (1994). Estimating the economic model of crime with panel data. *The Review of economics and Statistics*, 360-366.

- Dasgupta, S., Laplante, B., Wang, H., & Wheeler, D. (2002). Confronting the environmental Kuznets curve. *Journal of economic perspectives*, 16(1), 147-168.
- De Coning, E., & Stølsvik, G. (2013). United Nations Office on Drugs and Crime. *The International Journal of Marine and Coastal Law*, 28(1), 189-204.
- DeYoung, R., Lang, W. W., & Nolle, D. L. (2007). How the Internet affects output and performance at community banks. *Journal of Banking & Finance*, 31(4), 1033-1060.
- Dimitrov, V., & Jain, P. C. (2008). The value-relevance of changes in financial leverage beyond growth in assets and GAAP earnings. *Journal of Accounting, Auditing & Finance*, 23(2), 191-222.
- Drigă, I., & Socol, A. (2009). Liquidity risk management in banking. *Revista tinerilor economişti*, 13, 46-55.
- Drigă, I., & Socol, A. (2009). Liquidity risk management in banking. *Revista tinerilor economişti*, 13, 46-55.
- Ellison, L., & Akdeniz, Y. (1998). Cyber-stalking: the Regulation of Harassment on the Internet. *Criminal Law Review*, 29, 29-48.
- Fatoki, O. (2015). The causes and consequences of household over-indebtedness in South Africa. *Journal of social sciences*, 43(2), 97-103.
- Finklea, K. M., & Theohary, C. A. (2015, January). Cybercrime: conceptual issues for congress and US law enforcement. Washington: Congressional Research Service, Library of Congress.
- GAMBO, I., OLUWAGBEMI, O., & ACHIMUGU, P. (2011). Lack of interoperable health information systems in developing countries: an impact analysis. *Journal of Health Informatics in Developing Countries*, 5(1).
- Gonzalez, L., & Loureiro, Y. K. (2014). When can a photo increase credit? The impact of lender and borrower profiles on online peer-to-peer loans. *Journal of Behavioral and Experimental Finance*, 2, 44-58.
- Goulas, E., & Zervoyianni, A. (2012). *Growth, Deficits and Uncertainty in a Panel of 28 Countries* (No. 52_12).
- Hale, C., & Sabbagh, D. (1991). Testing the relationship between unemployment and crime: A methodological comment and empirical analysis using time series data from England and Wales. *Journal of Research in Crime and Delinquency*, 28(4), 400-417.
- Hamada, R. S. (1972). The effect of the firm's capital structure on the systematic risk of common stocks. *The journal of finance*, 27(2), 435-452.
- Ibikunle, F. and Eweniyi, O. 2013. Approach To Cyber Security Issues in Nigeria: Challenges and Solution. *International Journal of Cognitive Research in Science, Engineering and Education (IJCRSEE)*, 1(1), 100-110. Retrieved from <http://ijcrsee.com/index.php/ijcrsee/article/view/11/114>
- Iyodo, B. Y., Agbaji, J. S., & Abu, A. S. (2016). Consequences of bank frauds on the growth of Nigerian economy. *Global Journal of Commerce and Management Perspective*, 5(4), 19-28.

- Jaishankar, K. (Ed.). (2011). *Cyber criminology: Exploring Internet crimes and criminal behavior*. CRC Press.
- Janko, Z., & Popli, G. (2015). Examining the link between crime and unemployment: a time-series analysis for Canada. *Applied Economics*, 47(37), 4007-4019.
- Janko, Z., & Popli, G. (2015). Examining the link between crime and unemployment: A time-series analysis for Canada. *Applied Economics*, 47(37), 4007-4019.
- Kolapo, F. T., & Olaniyan, T. O. (2018). The impact of fraud on the performance of deposit money banks in Nigeria. *International Journal of Innovative Finance and Economics Research*, 1(6), 40-49.
- Koops, B.J., 2010. The Internet and its Opportunities for Crime. In: Herzog-Evans, M., (ed.) *Transnational Criminology Manual*. Nijmegen, Netherlands: WLP, pp.735-754.
- Lin, M. J. (2008). Does unemployment increase crime? Evidence from US data 1974– 2000. *Journal of Human resources*, 43(2), 413-436.
- Loader, B. D., & Thomas, D. (Eds.). (2013). *Cybercrime: Security and surveillance in the information age*. Routledge.
- Longe, O., Ngwa, O., Wada, F., Mbarika, V., & Kvasny, L. (2009). Criminal uses of information & communication technologies in sub-Saharan Africa: trends, concerns and perspectives. *Journal of Information Technology Impact*, 9(3), 155-172.
- Mago, S., & Chitokwindo, S. (2014). The impact of mobile banking on financial inclusion in Zimbabwe: A case for Masvingo Province. *Mediterranean Journal of Social Sciences*, 5(9), 221.
- Muradoglu, Y. G., & Baturevich, B. (2010). Would you follow MM or a Profitable Trading Strategy?. *Frontiers in Finance and Economics*, 7(2), 69-89.
- Nwaeze, C. (2008). Quality and internal control challenges in contemporary Nigeria banking. *Zenith Economic Quarterly, Zenith Bank PLC*, 3(2), 23-28.
- Ogbuji, C. N., Onuoha, C. B., & Izogo, E. E. (2012). Analysis of the negative effects of the automated teller machine (ATM) as a channel for delivering banking services in Nigeria. *International Journal of Business and Management*, 7(7), 180.
- Okpara, G. C. (2009). A synthesis of the critical factors affecting performance of the Nigerian banking system. *European Journal of Economics, Finance and Administrative Sciences*, 17(17), 34-44.
- Olubukola Stella Adesina (2017). Cybercrime and Poverty in Nigeria *Canadian Social Science* Vol. 13, No. 4, 2017, pp. 19-29.
- Olumide, O. O., & Victor, F. B. (2010). E-crime in Nigeria: Trends, tricks, and treatment. *The Pacific Journal of Science and Technology*, 11(1), 343-355.
- Olumide, O. O., Victor, F. B. (2010): *E-Crime in Nigeria: Trends, Tricks, and Treatment. The Pacific Journal of Science and Technology*, Volume 11. Number 1. May 2010 (Spring)

- Oluwatolani, O., Joshua, A., & Philip, A. (2011). The impact of Information Technology in Nigeria's banking industry. arXiv preprint arXiv:1108.1153.
- Oyewole, O. S., Gambo, J., Abba, M., & Onuh, M. E. (2013). Electronic payment system and economic growth: a review of transition to cashless economy in Nigeria. *International Journal of Scientific Engineering and Technology*, 2(9), 913-918.
- Ozeh, C. C., & Ohajionu, C. C. (2019). Unemployment, Migration and Cyber Criminality in Nigeria. *Internal Security Management in Nigeria: Perspectives, Challenges and Lessons*, 165-180.
- Pesaran, M. H., Shin, Y., & Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of applied econometrics*, 16(3), 289-326.
- Raphael, S., & Winter-Ebmer, R. (2001). Identifying the effect of unemployment on crime. *The journal of law and economics*, 44(1), 259-283.
- Ribadu, N. (2007). Cyber-crime and commercial fraud: A Nigerian perspective. Presented at the Congress Celebrating the Fortieth Annual Session of the UNCITRAL (United Nations Commission On International Trade Law), Vienna, Austria, 9-12 July. Retrieved from http://www.cnudmi.org/pdf/english/congress/Ribadu_Ibrahim.pdf
- Safeena, R., Date, H., & Kammani, A. (2011). Internet Banking Adoption in an Emerging Economy: Indian Consumer's Perspective. *Int. Arab. J. e Technol.*, 2(1), 56-64.
- Sreedevi, V., & Meena, K. (2011). ICT and financial inclusion. *International Journal of BusinessManagement, Economics and Information Technology*, 3(2), 331-334.
- Srivastava, S. (2014). Entangled urbanism: Slum, gated community, and shopping mall in Delhi and Gurgaon.
- Wada, F., & Odulaja, G. O. (2012). Electronic banking and cyber crime in Nigeria-a theoretical policy perspective on causation. *Afr J Comp ICTs*, 5(1), 69-82.
- Wall, D. (2001). *Crime and the internet: Cybercrimes and cyberfears*. London, New York: Routledge. [Crossref], [Google Scholar]
- Yar, M. (2005). The Novelty of 'Cybercrime' An Assessment in Light of Routine Activity Theory. *European Journal of Criminology*, 2(4), 407-427.

The financing dynamics of niche innovation in energy transitions: A South African perspective

Amanda-Leigh O'Connell

University of Johannesburg, South Africa

1. Introduction

"The global financial architecture does not need a simple evolution; it needs a radical transformation" (Guterres 2023). This powerful statement by the United Nations Secretary-General in his briefing to the General Assembly on Priorities for 2023 echoes a growing sentiment that finance mechanisms, institutions, and practices are ill-equipped to respond to the world's most pressing challenges, such as climate change (e.g., Loorbach, Schoenmaker and Schramade 2020; Naidoo 2020; Schot et al. 2022). This call for radical transformation is particularly relevant in the context of two urgent United Nations priorities for 2023: reducing emissions and achieving climate justice (Guterres 2023). Finance flows consistent with decarbonisation efforts are necessary to effectively respond to these priorities (IPCC 1990, 2022; United Nations 2015). Pertinently, the Secretary-General emphasises the need for fundamental reforms that prioritise and centre the needs and voices of vulnerable and developing countries (Guterres 2023). In such contexts, financing innovation is critical for both decarbonisation efforts and sustainable economic development to address persistent issues such as poverty and inequality (IEA 2021).

Addressing "grand societal challenges" like climate change is central to sustainability transitions literature, providing a valuable framework for analysing system change (Köhler et al. 2019, p.2). This field studies the complex interplay between the social and technological factors underpinning systems of production and consumption. Sustainability transitions theory conceptualises *transitions* as substantive and transformative changes in societal systems, like energy, healthcare, and mobility. These changes include "changing skills, infrastructures, industry structures, products, regulations, user preferences and cultural predilections" (Schot and Steinmueller 2018, p.1562). Consequently, system transitions entail shifts in underlying financing patterns, practices, and rules.

The sustainability transitions literature is increasingly focusing on the role of finance, as evidenced by a growing body of academic work (Egli et al. 2022; Köhler et al. 2019; Steffen and Schmidt 2021). However, significant knowledge gaps persist, prompting calls for further exploration and theorisation. This paper addresses an important gap in empirical and conceptual work in emerging and developing countries, where financing energy transitions is especially relevant (Steffen and Schmidt 2021). These countries are poised to be hardest hit by climate change while cumulatively contributing the least to its causes (IPCC 1990, 2022). Moreover, renewable energy innovation is critical for sustainable economic development. However, resource constraints and competing demands pose significant challenges in such contexts, underscoring the need to develop effective mechanisms for financing energy transitions (IEA 2021).

The primary aim of this research paper is to enhance the understanding of how financing rule¹ can shape how financial resources flow to, from and within an energy system and the implications thereof for actors seeking to introduce new renewable energy technologies in developing contexts. To analyse this dynamic, financing rules are framed as integral components of socio-technical *regimes* that consist of semi-coherent “formal and informal rules (e.g. shared beliefs and values, routines, regulations, institutionalized practices, capabilities, etc.) that mutually construct and are constructed by actors in a system” (Fuenfschilling and Truffer 2014, p.773).

The term *financing dynamics* is also introduced to assess the impact of these rules on the allocation and distribution of financial resources in a socio-technical system. The concept of market dynamics in economic theory informs this term, defined here as *the processes and factors involved in the allocation, provision, and management of funds, including forms of financing, use of funds, factors influencing financing decisions, and interactions between financing market actors and organisations that impact financing outcomes*. While loosely applied in research studies, a formal definition is lacking (see Dahlstrand and Cetindamar 2000; De la Torre, Feyen and Ize 2013; Hackbarth and Sun 2015). In this study, the term delimits the analysis to rules that may exert causal influence on financing dynamics. The term also embodies a central proposition of this study: the processes and factors involved in providing, allocating, and managing funds represent a critical nexus of actor interactions. Accordingly, the research question is formulated: *How do financing rules impact financing dynamics for decarbonised energy systems?*

This research adopts a single case-study approach for an in-depth analysis of the rules that underpin financing dynamics and their causal influence, drawing on theories and concepts from sustainability transitions and empirical work in innovation systems literature. The empirical domain is South Africa’s energy

system and the area of study, South Africa's energy transition and its emerging green hydrogen sector. As a developing country, South Africa continues to face economic, social, and environmental challenges despite significant progress post-Apartheid (National Planning Commission 2012; International Bank for Reconstruction and Development and The World Bank 2018; The World Bank 2022). With a high reliance on fossil fuels, significant renewable energy potential, and ambitious plans for a just energy transition, South Africa provides a relevant and informative context to investigate financing rules and dynamics in energy transitions (Department of Energy (DoE) 2019; Department of Mineral Resources and Energy (DMRE) 2021; IEA 2019; The Presidency 2022). The green hydrogen sector is earmarked as one of three priority investment areas for the just transition, providing a dynamic, evolving case to explore this relationship and its impact on innovation for system change (NBI (National Business Initiative) 2022; The Presidency 2022).

This study makes three main contributions to the literature. First, it introduces and operationalises the concept of financing dynamics, providing a novel framework for systematically analysing the impact of financing rules on the allocation and distribution of financial resources in a socio-technical system. Second, it proposes a set of finance rules that underpin socio-technical regimes in the South African context and presents a taxonomy for aligning financing instruments, types, and actors to these rules, bringing consistency and clarity to the literature. Lastly, it addresses the empirical gap in understanding financing dynamics for decarbonised energy in emerging and developing contexts. This case study expands current knowledge of financing innovation in a developmental context by exploring the role of financing rules and presenting a novel perspective on how actors navigate these rules to develop alternative practices within the South African context. While the focus is on a single case, the insights gained from the South African experience and the methodology employed can inform and contribute to the literature on financing innovation for energy transitions in the global South.

The remainder of this paper is structured as follows. Sections two and three present the literature reviewed on the financing dynamics of energy systems, and current knowledge of financing rules in transitions literature, respectively. Section four sets out the conceptual framework, and section five presents the research design and methodology. Section six presents the results and discusses and interprets the findings. The final section concludes by summarising the main findings and providing recommendations for future research.

2. Financing dynamics in energy transitions

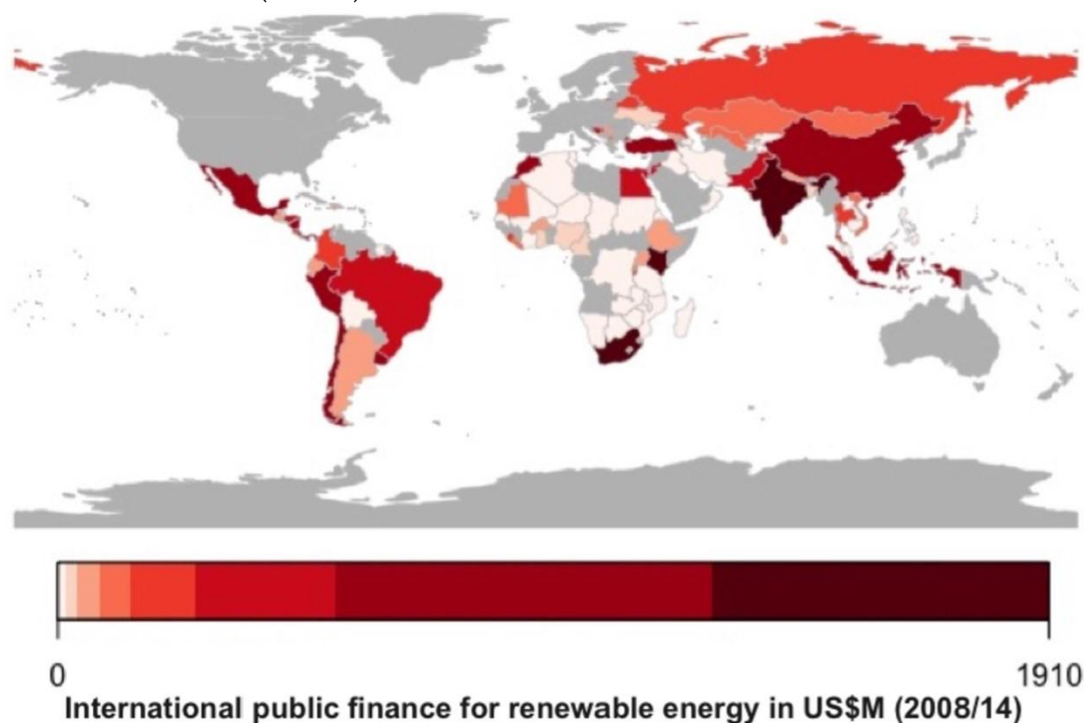
As introduced above, the term financing dynamics encompasses the processes and factors involved in the provision, allocation, and management of funds in a socio-technical system. Accordingly, this section examines the literature on financing dynamics in energy transitions, considering forms of financing, use of funds, factors influencing financing decisions, and interactions between financing and energy actors.

2.1 Forms of financing energy transitions

In this study, financing forms refer to the finance types, instruments, and mechanisms for mobilising financial resources for sustainable transitions. Various concepts have emerged to describe forms related to sustainable development, including transition finance (Caldecott 2022), climate finance (e.g., Anantharajah and Setyowati 2022; Hogarth 2012), green finance (e.g., Barua and Aziz 2021), sustainable finance (e.g., Fatemi and Fooladi 2013), and transformative investment (Schot et al. 2022). This diversity reflects an evolving field with ongoing debates surrounding appropriate financing classifications. The development of new terms suggests an underlying need to establish new financing rules – or at least the perceptions of new rules – to respond to the demands of sustainability transitions.

Public and private financing are primary delineations, with scholars exploring synergies, particularly the role of public financing in mobilising and de-risking private financing (e.g., Corrocher and Cappa 2020; Haščić et al. 2015; Isah et al. 2023; Owen, Brennan and Lyon 2018; Walwyn 2020; Zhang 2020). Sub-sets of private financing, such as investment finance, have also been examined, with comprehensive analyses of the literature on sustainability investing, including identifying underlying heuristics that inform practice (Penna, Schot and Steinmueller 2021). International and philanthropic financing types have received less attention in energy transitions, although both are significant in developing contexts. Ragosa and Warren (2019) illuminate varying patterns for international public finance for renewable energy (Fig. 1), while Isah et al. (2023) find Nigeria's reliance on bilateral agencies and multilateral development increased the risk profile of renewable technologies. Donor funding may also influence niche formation, increase energy stakeholders' transaction costs, and is strongly informed by international standards and norms (Isah et al. 2023; Jain, Hoppe and Bressers 2017; Kim 2019).

Figure 1. International public finance into renewable energy generation projects > 1 MW 2008–2014 (US\$M)



Source: Reproduced from Ragosa and Warren (2019, p.863)

Scholarly work on financing forms in transitions contexts also interrogates the instruments and mechanisms of financing, including equity investments, loans and bonds, grants, guarantees, and carbon credits, as well as newer instruments like crowdfunding, green bonds, and sustainability-linked bonds (e.g., Polzin and Sanders 2020; Sorenson et al. 2016; Owen, Brennan and Lyon 2018). For example, studies in developing contexts find that the number and type of Instruments and mechanisms employed shape the direction and pace of energy transitions by influencing the accessibility and distribution of funds while presenting tensions and trade-offs (Baker 2015; Isah et al. 2023; Ragosa and Warren 2019). For example, Ragosa and Warren (2019) find that providing international (public) finance is associated with a small increase in the likelihood of investment across all types of investors and significantly larger cumulative investment flows in developing countries.

The findings by Ragosa and Warren (2019) also reveal the heterogeneity of financing forms in developing countries, which suggests that case studies have an important role in understanding financing rules and their impact in specific contexts. However, the rules that underpin the development and adoption of financing instruments and mechanisms have largely not been addressed in the

academic literature in such contexts.

2.2 Use of funds in energy transitions

Achieving complex socio-technical change suggests shifts in allocating and managing funds among various actors and purposes. Sustainability transitions theory recognises that innovation often occurs in *niches*, defined as “protected spaces that allow nurturing and experimentation with the co-evolution of technology, user practices, and regulatory structures” (Schot and Geels 2008, p.538). Niche development involves a wide range of actors needing financial resources to engage in niche learning, networking and vision-building processes (Schot and Geels 2008). These processes are crucial for overcoming barriers and uncertainties associated with emerging technologies and fostering their successful diffusion and adoption.

However, the literature has yet to fully explore the use of funds beyond technology deployment and infrastructure investment in renewable energy, leaving a gap in understanding the financial needs of actors involved in socio-technical change. Polzin, Sanders and Serebriakova (2021) provide one such study, modelling the financing requirements of different energy transition scenarios to the necessary technology investments. Their model expands the scope of finance needs to include government and corporate research and development as critical mechanisms to support newer renewable energy technologies introduced into an energy system. They demonstrate that a deeper understanding of financing uses and their interplay with niche processes and actor interactions can contribute to more effective financing strategies for innovation and sustainable energy transitions.

2.3 Factors in decision-making

The factors informing financing decision-making in sustainable energy transitions speak to the financing rules that guide allocation processes. Empirical studies find that forms of finance, including public, private, and philanthropic financing, may have distinct institutional characteristics that shape the decision-making processes for actors. For instance, public financing may prioritise policy goals and social outcomes, while private financing would focus on profitability and risk management (Polzin and Sanders 2020; Walwyn 2020). Philanthropic financing, in turn, may be guided by the desire to create a long-lasting impact on sustainability and climate issues (see Jain, Hoppe and Bressers 2017; Kim 2019). One important implication of varying institutional characteristics is that actors may need to navigate different financing rules as niches develop – which supports the socio-technical system as an appropriate ontological level of analysis. However,

delineating these institutional characteristics and their implications for niche and regime interactions is underdeveloped.

While not necessarily addressed as a primary objective in scholarly work, many studies reveal key decision-making factors in their empirical studies on energy transitions, such as risk management, track records, and transaction size (e.g., Baker 2015; Geddes and Schmidt 2020). These decision-making factors may, in turn, influence regime resistance and directionality. For instance, the need to demonstrate a track record was suggested to limit niche innovation deployment and diffusion in developing contexts if not supported by public finance interventions (Geddes and Schmidt 2020). Baker (2015) also confirms the importance of track records in the power sector in South Africa's energy transition. The demand for policy certainty is another frequently raised decision-making factor – a good example of how financing rules interact with specific regime dimensions (Baker 2015; Isah et al. 2023; Walwyn 2020). Notably, factors may apply across socio-technical systems in both developing and developed contexts.

2.4 Actor interactions in energy transitions

The interactions between regime and niche processes have been the subject of theoretical and empirical energy transition studies (e.g., Elzen, Van Mierlo and Leeuwis 2012; Jain, Hoppe and Bressers 2017; Rosenbloom 2019; Ting and Byrne 2020). Understanding how actors collaborate, compete, and learn from one another can reveal mechanisms that drive the diffusion of innovations and identify resources, policies, and rules necessary for sustainable energy transitions. Yang, Schot and Truffer (2022) illuminate their importance by demonstrating that the nature of niche-regime interactions can influence the directionality of change in a socio-technical system. Using case studies, Sutherland, Peter and Zagata (2015) also show how niches can operate as loci for change in multiple systems, facilitating multi-regime interactions.

Despite its importance, elucidating financing-related interactions and their underlying rules is an underdeveloped study area in transition literature. Geddes and Schmidt (2020) interviewed experts in Australia, the United Kingdom, and Germany to identify critical determinants of interactions between the financial system's regime and energy niches. Notably, they find trust signalling and learning by co-investing to be important secondary effects of public finance interventions demonstrating the links between forms of financing, decision-making factors and actor interactions.

However, further research is warranted, and the empirical evidence suggests two key focus areas. First, the interrogation of power embedded in interactions and

how it is reproduced and reinforced through finance rules. According to Dutt (2022), power and politics are vital considerations, particularly in a global South context. For example, Anantharajah and Setyowati (2022) highlight the vulnerability of Asian and Pacific communities in donor engagements due to limited access to direct finance. Second, questions of inclusion and representation as characteristics of interactions should be considered (Hall et al. 2018; Naidoo 2020). These areas provide important avenues of inquiry for future research to enhance our understanding of the dynamics and complexities inherent in financing energy transitions.

3. Financing rules in socio-technical regimes

Sustainability transitions theory posits that the transformation of the global financial architecture, as envisioned by Guterres (2023), Schot et al. (2022) and others, would be predicated upon reconfiguring financing rules to support sustainable socio-technical systems. Identifying which financing rules need to change and what new configurations could entail is, therefore, a central question. However, few studies have attempted to operationalise and identify rules to construct the financing rule sets in socio-technical regimes. Seyfang and Gilbert-Squires (2019) offer a notable example of such an analysis; they map core financing rules to six regime dimensions for the United Kingdom's banking sector and a sustainable banking niche and demonstrate how such a comparative exercise can support recommendations for targeted policy interventions, albeit within a single system.

Through empirical work, some scholars have begun theorising alternative financing rules to respond to the demands of sustainability transitions. For instance, Loorbach, Schoenmaker and Schramade (2020) build on research in transition management and the multi-level perspective (MLP) to propose three principles to transform the financial system: integrated value, stewardship, and long-term societal value based capital allocation. Schot et al. (2022) draw on the MLP and deep transitions theory to propose 12 principles for transformative investment practice that address goal setting, investment strategy, investment process and sharing and learning. Importantly, introducing foundational principles such as these have implications for understanding the ontological hierarchy of financing rules in socio-technical regimes, their relative causal influence vis-à-vis incumbent rules, and the processes through which these new rules will be diffused to reconfigure incumbent regimes.

The extent to which principles such as those proposed by Loorbach, Schoenmaker and Schramade (2020) and Schot et al. (2022) can destabilise incumbent financing rules is closely tied to the question of structuration and the degree of

embeddedness. For financing rules, embeddedness has been predominantly understood through the lens of financialisation. Epstein (2005, p.3) defines *financialisation* as “the increasing role of financial motives, financial markets, financial actors and financial institutions in the operation of the domestic and international economies”. Critically, financialisation implies the wide diffusion of financing rules and their reproduction by actors in multiple socio-technical systems and across regime dimensions. For example, Walter and Wansleben (2020, p.36) use the concept to describe the “rise of finance as a new governmentality” for central banks, who have redefined their role relative to financial market expectations and orientated their macroeconomic policies accordingly.

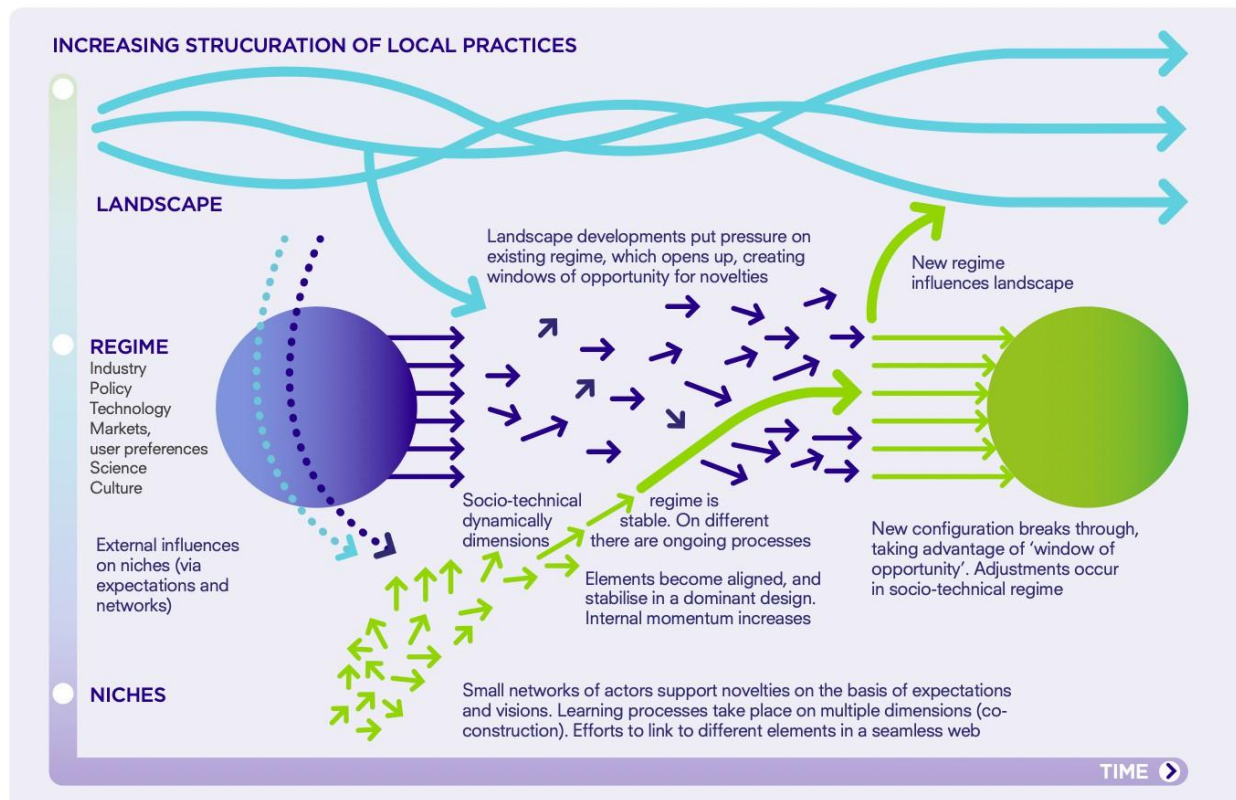
Several scholars have explored or considered the phenomenon in the context of energy transitions (e.g., Irfan et al. 2023; Klagge and Nweke-Eze 2020; Li, Cao and Xiang 2023). Sovacool et al. (2019) use case studies in Bangladesh, Netherlands, Malawi, Mexico and South Africa to demonstrate how financialisation (and experimentation and dispossession) can account for transition pathway selection. However, researchers have mainly focussed on the consequences of financialisation rather than exploring the underlying rules that support the phenomenon and the proliferation of those rules outside a financial system context. Studies by Klagge and Nweke-Eze (2020) in Kenya and Irfan et al. (2023) in G7 countries also suggest that financialisation may not be relevant in all contexts. Their studies operationalise private sector involvement and financial development, respectively – which does not necessarily preclude the diffusion of the financing rules associated with financialisation in those contexts. Indeed, financing rules may offer an alternative unit of analysis to deepen our understanding of the diffusion and impact of financialisation.

4. Conceptual framework

In sustainability transitions literature, the socio-technical system provides the point of departure for analysing system transitions, defined as “a network of the resources – material, human or institutional – that are required for fulfilling societal functions, such as transport, communication, nutrition, energy, or finance” (Penna, Schot and Steinmueller 2021, p.6). Understanding how and why change occurs within socio-technical systems is the central concern of sustainability transitions literature. Central to this inquiry is the socio-technical regime, which represents a socio-technical system’s structure, constituting deeply ingrained, semi-coherent rules that stabilise and guide system functions (Rip and Kemp 1998; Kemp, Schot and Hoogma 1998). A key proposition in sustainability transitions theory is that a system’s transformation is causally connected to changes in its regime (Geels 2002; Loorbach 2010; Schot and Geels 2008; Schot and Kanger 2018).

To analyse regimes and their dynamics, scholars have converged on a multi-level perspective (MLP) framework delineating six core regime dimensions: technology, science, industry, markets and user preferences, policy, and culture (Fig 2). Regime dimensions provide a framework to examine underlying rules, aiding scholars in understanding how changes within dimensions influence regime stability and transformation.

Figure 2. The Multi-level perspective of the socio-technical transition framework



Source: Reproduced from Boni, Giachi and Molas-Gallart (2019, p.19)

One important source of change in socio-technical stems from niches, which are emerging socio-technical configurations that enable the co-evolution of technology, user practices, and regulation in protected spaces (Geels 2002; Schot and Geels 2008). Change can also be triggered by *landscape* factors – exogenous social, economic, political, cultural, and environmental aspects that open windows of opportunity for niche innovations to mainstream and transform the system (Geels and Schot 2007). System change is also shaped by the interactions between regime, niche, and landscape processes (Schot and Kanger 2018).

In addition to niches and landscapes, the interconnections between various socio-technical systems may also play a significant role in driving change (see Elzen, Van

Mierlo and Leeuwis 2012; Konrad, Truffer and Voß 2008; Mah and Cheung 2020; Sutherland, Peter and Zagata 2015). An emerging interest is the study of *deep transitions*, defined as sustained, large-scale transformations of multiple socio-technical systems in a similar direction (Kanger 2022; Kemp et al. 2022; Schot and Kanger 2018). In this literature stream, deep transitions involve *meta-rules*, which are single rules that span multiple systems, and *meta-regimes*, which are sets of meta-rules. Meta-rules and meta-regimes coordinate the interconnections between technologies and industries across multiple systems (Schot and Kanger 2018).

In the context of socio-technical system change, several key concepts are relevant to this discussion. Transition pathways describe patterns of change in socio-technical systems, leading to new ways of achieving societal functions (Geels and Schot 2007). Directionality in sustainability transitions refers to socio-technical system outcomes that align with desired social and environmental objectives, shaped by various factors, such as strategic actions, technological advancements, and institutional changes (Ghosh and Schot 2019). Lastly, regime resistance acknowledges the propensity of incumbent regimes to resist transformative changes due to factors like vested interests, technology lock-in, and the entrenchment of institutionalised regime rules. Ghosh and Schot (2019) also draw attention to the degree of regime change, proposing a more refined framework that distinguishes between limited regime optimisation, substantive regime transformation, and the paradigmatic shift of a regime transition.

5. Research Design and methodology

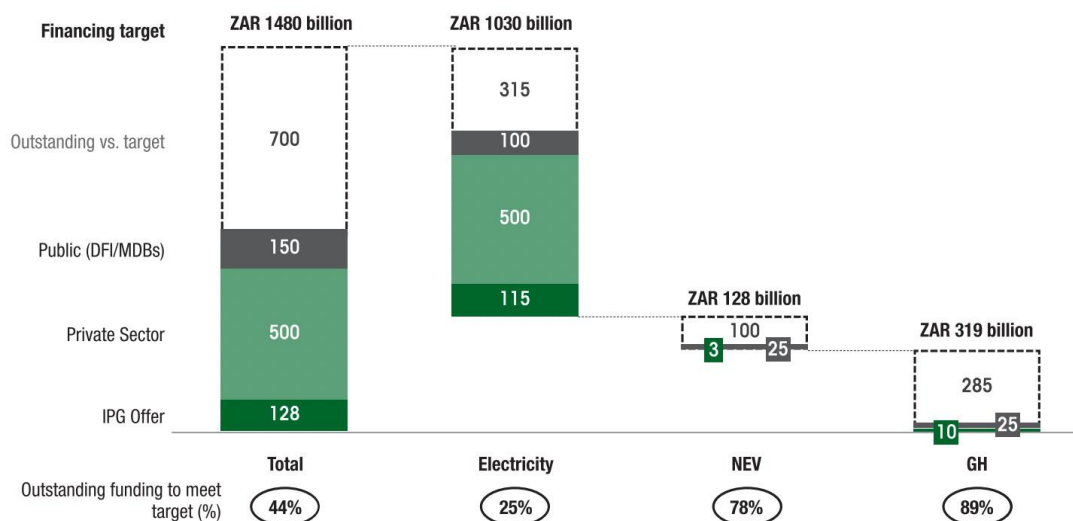
5.1 Case study selection

This study adopts a qualitative approach to investigate the impact of financing rules on financing dynamics for decarbonised energy systems. An instrumental single-case design was applied in the empirical domain of South Africa's energy system and emerging green hydrogen sector. South Africa was selected for its position as a developing country with high greenhouse gas (GHG) emissions, a predominantly fossil-based energy system, and a public commitment to pursuing a just energy transition (DOE 2019, 2020; DMRE 2021). Furthermore, the country's fiscal constraints and relatively weak structural power in the global economy necessitate the financial support of international private and public actors to realise this transition (NBI 2022; The Presidency 2022).

Among several emerging niches, the hydrogen sector, supported since circa 2003, stands as a critical investment area for innovation and development to support the country's just energy transition (Council for Scientific and Industrial Research (CSIR) 2006; DSI 2021; Department of Science and Technology 2007; The Presidency

2022). This prioritisation is evident in South Africa's Just Energy Transition Investment Plan², which projects the estimated financing needs by source for the priority areas of electricity, new energy vehicles, and green hydrogen (Figure 3). Consequently, this niche is characterised by extensive public discussion and documentation, considerable research interest, and good secondary data about financing needs, challenges and estimates to complement the primary data.

Figure 3. Projected funding needs and estimated availability by source for priority areas of electricity, new energy vehicles (NEV) and green hydrogen (GH).



Source: Reproduced from The Presidency (2022, p. 15)

The single-case study design allows for an in-depth and comprehensive investigation of financing rules and dynamics in the context of an important emerging niche technology in an unfolding energy transition in a developing country. This case study extends current knowledge of financing innovation in a development context by focusing on the role of financing rules. Furthermore, the study provides a novel perspective on how actors navigate these rules to develop alternative practices within the South African context. While this study focuses on a single case, the South African experience and the methodology adopted here can potentially inform and contribute to the literature on energy transitions in other developing countries.

5.2 Data collection and analysis

The case study design utilises a mix of data collection and analysis methods, and data was collected from both primary and secondary sources. The unit of analysis

is the financing structure (rules and rule sets) of South Africa's energy system. The units of observation are (i) actors providing financing and funding for renewable energy innovation development, adoption, or diffusion and (ii) actors seeking financing and funding for green hydrogen innovation development, adoption, or diffusion.

Subject selection and sampling were based on inclusion criteria developed for (i) financing and (ii) innovating actors, respectively. Financing actors had to participate in their organisation's financing approval, disbursement, or allocation processes. In addition, participants' organisations had to have energy-related assets, projects or investments in South Africa in their portfolios. Innovating actors were selected based on their involvement in green hydrogen innovation activities in South Africa, such as research and experimental development, IP-related activities, engineering and design, and skills development (see OECD/Eurostat 2018). These actors were either seeking funding for innovation activities or had acquired funding to finance their activities in the green hydrogen niche. A participant list detailing actors' roles and financing domains is included in Appendix A.

This study applied purposive sampling and snowball sampling techniques to identify participants who met the inclusion criteria, yielding a sample size of N=22 interviews. These techniques allowed for selection of participants with the necessary experience and knowledge to provide relevant insights into the research question. Snowball sampling facilitated multiple interviews within an organisation to interrogate and compare intra- and inter-organisational functions, interactions and perspectives. Interviews were conducted for approximately one hour each from July 2022 to November 2022. Secondary data, such as policy documents, Hansards, press releases, newspaper articles and workshop presentations, were used to source suitable interview participants, identify applicable regulative institutions, and elucidate contemporary patterns in financing flows. Data from secondary sources supplemented, complemented and triangulated interview data.

Critical realism informed the research design, with rules conceptualised as causal mechanisms or structures that engender observable phenomena (Mingers 2006, 2011). Initially, a sequential approach was planned for the data analysis. However, as the study progressed, it became clear that an iterative analysis process would better accommodate the complexity of the research. Three analytical processes were conducted to identify, map and analyse relationships between financing rules and financing dynamics and to understand their causal influence on system change (see Easton 2010; Eastwood, Jalaludin and Kemp 2014).

The first analytical process was concerned with the identification of the financing

rules that constitute financing structures. Using the three institutional pillars of Scott's (2014) analytical framework, rules were operationalised as regulative, normative and cultural-cognitive institutions pertaining to financing approval, disbursement, or allocation processes (Table 1). In addition to informing the theoretical development of the MLP, the three pillars have also been applied in empirical work as a taxonomy to analyse and understand the rules and institutions that constitute regimes (Seyfang and Gilbert-Squires 2019; Yang, Schot and Truffer 2022). Participant statements that depicted or described financing rules were classified accordingly.

Table 1. The three institutional pillars

	Regulative	Normative	Cultural-cognitive
Basis of compliance	Expedience	Social obligation	Taken-for-grantedness Shared understanding
Basis of order	Regulative rules	Binding expectations	Constitutive schema
Mechanisms	Coercive	Normative	Mimetic
Logic	Instrumentalism	Appropriateness	Orthodoxy
Indicators	Rules Laws Sanctions	Certification Accreditation	Common beliefs Shared logics of action Isomorphism
Affect	Fear Guilt/ Innocence	Shame/Honour	Certainty/Confusion
Bases for legitimacy	Legally sanctioned	Morally governed	Comprehensible Recognizable Culturally supported

Source: Reproduced from Scott (2014, p.60)

In a second process, the rules were analysed with participants' roles to identify rules specific to a particular form or use of financing. This analysis supported the development of a taxonomy of financing domains and their institutional characteristics as a basis for understanding how different domains overlap and interact with one another. In this process, innovating actors were key in supporting the development of philanthropic financing's attributes and characteristics. In the third process, the identified rules were mapped to the six dimensions of South Africa's energy regime to elucidate how financing rules may interact with the respective dimensions (see Ghosh and Schot 2019). Mapping was based on participants' perspectives of causal influence, and secondary data supported triangulation and corroboration in a retroductive analysis process. A key aspect of regimes is their semi-coherence, and this process also considers the relative coherence of rules and rule sets. While the study precludes statistical generalisation, inferences could be drawn on general tendencies and opportunities for alternative practice (see Welch et al. 2011).

6. Results and discussion

This section presents the analyses of South Africa's financing regimes and their impact on the energy regime's financing dynamics from the perspective of South Africa's emerging green hydrogen niche. Section 6.1 draws on the six dimensions of institutional pillars to develop a taxonomy of financing domains and proposes a set of attributes and institutional characteristics gleaned from the primary and secondary data (see Table 3). Section 6.2 presents a proposed set of financing meta-rules governing financing dynamics in South Africa. Section 6.3 discusses the causal influence of financing rules on financing dynamics and the implications for system change.

6.1 Institutional underpinnings of financing dynamics in South Africa

The first output of the analysis process is a summary table that proposes a set of institutional characteristics for different domains of financing that may underpin financing dynamics (Table 3). As Appendix A and B depict, all five domains were observed in the case study sample, with varying implications for financing dynamics and transition processes.

Table 2. Summary of institutional characteristics of financing domains in South Africa

Domain	Attributes (uses, actors, instruments)	Institutional characteristics emerging from the data
Public Financing	• Aims to promote economic growth, address social needs, and support government economic objectives. • Provided by governments, government agencies and bodies, development finance institutions, multilateral development banks, and export credit agencies. • Utilises development finance, grant funding, public-private partnerships, tax incentives, and special economic zones.	• Basis for compliance: Policy compliance, legal and regulatory requirements • Basis of order: Political system, public administration, national innovation system • Key mechanisms: Legislation, regulation, government mandates • Underlying logics: Public interest, social welfare • Key indicators: Public spending, economic growth, inequality, policy outcomes, commercialisation of IP, credit rating • Affects: Loyalty, trust/distrust, ideology • Basis for legitimacy: Social welfare, democratic processes, fiscal responsibility.

Corporate Financing	• Focuses on funding business operations, expansion and strategic initiatives, capital structure optimisation, and cash flow management. • Primarily funded by investment banks, commercial banks, development banks, private equity firms, venture capital firms, and hedge funds. • Instruments include debt finance, equity finance, asset-based finance, project finance, trade finance, and venture capital.	• Basis for compliance: Contractual obligations, legal requirements, public and private regulation • Basis of order: Market system, economic system • Key mechanisms: Contracts, negotiation, price • Underlying logics: Profitability, growth, shareholder value • Key indicators: Financial performance, market share, balance sheet size, credit rating • Affects: Trust, confidence, uncertainty, risk tolerance • Basis for legitimacy: Economic efficiency, economic function, market position, compliance
Investment Financing	• Targets wealth creation, portfolio diversification, and support for businesses at various development stages. • Sourced from institutional asset managers, investment banks, mutual funds, pension funds, sovereign wealth funds, private equity firms, venture capital firms, and hedge funds. • Instruments debt finance, equity finance, venture capital, and project finance.	• Basis for compliance: Fiduciary duties, legal requirements, public and private regulation • Basis of order: Financial markets, portfolio management • Key mechanisms: Investment, corporate governance, client mandates • Underlying logics: Portfolio theory, risk-return relationship • Key indicators: Return on Investment (ROI); portfolio performance, benchmarks, impact measurement • Affects: Confidence, optimism/pessimism, risk tolerance • Basis for legitimacy: Market performance, capital allocation, financial stability, compliance

Domain	Attributes (uses, actors, instruments)	Institutional characteristics emerging from the data
International Financing	• Aims to support global economic development, infrastructure projects, cross-border investments, poverty alleviation, international cooperation, and global challenges. • Provided by multilateral organisations, foundations, foreign governments, development finance institutions, bilateral aid agencies, and multilateral development banks. • Consists of development finance, foreign aid, project finance, trade finance, debt finance, equity finance, and asset-based finance.	• Basis for compliance: Bilateral and multilateral agreements, international law • Basis of order: Diplomacy, political economy, government mandates • Key mechanisms: Treaties, conventions, partnerships, declarations, memoranda of understanding • Underlying logics: Cooperation, sustainability, mutual beneficiation • Key indicators: Development indicators, cross-border collaboration, international trade • Affects: Nationalism, development, solidarity • Basis for legitimacy: Global governance, international development, structural power
Philanthropic Financing	• Positioned to address social, environmental, or cultural issues, promote social good, and drive positive change. • Offered by corporates (including financial actors), private foundations, high-net-worth individuals, and non-profit organisations • Encompasses grants, donations, endowments, impact investments, and corporate budgets.	• Basis for compliance: Donor intent, legal and regulatory requirements • Basis of order: Non-profit, non-government sector, discretion • Key mechanisms: Donations, grants, donor mandates • Underlying logics: Social impact, altruism • Key indicators: Impact measurement, programme effectiveness, spending • Affects: Social influence, duty • Basis for legitimacy: Social good, ethics

The financing rules associated with different domains can significantly impact system-level dynamics. For example, public financing has driven early technology and innovation through the Hydrogen South Africa research programme, laying the foundation for niche development and today's broader selection environment (I12, I13, I15, I16). In this domain, regulatory requirements and government mandate strongly influence decision-making: "So [the DFIs] are developing their

own kind of just transition criteria for lending. And they're saying it's not just the project...So if you've got a fossil fuel company that's wanting to invest in a renewable project...they look at the strategy of the organisation as a whole" (I15).

Corporate and investment financing become viable options as niche technologies commercialise and generate cash flows. One participant explained, "You still need the risk-adjusted return, and generally, the banking or finance industry is very good at pricing risk and understanding risk, right?... and therefore, going up the risk curve to try and finance new technologies is not something that come[s] natural" (I18). Although corporate and investment financing actors often participate in the same deals, their decisions are informed by different underlying logics and indicators (I1, I3, I5, I9, I10, I8, I21, I22).

Tensions rooted in underlying institutional characteristics can also shape dynamics (I1, I9, I10, I18). For example, one portfolio manager noted that "The banks are so aggressive now in the [renewable energy] market, that they are essentially excluding...trying to exclusively do these deals, whereas, in the past, there was a lot more collaboration" (I9). While the US\$ 8.5 billion international financing committed to South Africa for its Just Energy Transition Plan was much lauded in the press (see UK Cabinet Office 2022), a Head of Fixed Income Research expressed: "For example, CPO 26 funding. Who knows what is happening with that? I mean, I've chatted to everyone I know. There's [redacted] and his team sitting within the presidency. As far as I know, nobody's even seen the terms of reference for this...Everyone wants a slice of the pie; nobody's even sure of what the pie actually is" (I1).

Consequently, niche actors must navigate a diverse set of financing rules. For instance, one participant mentioned, "[Funding processes are] very different actually; every funder has its own mandates" (I11). Another participant exploring a green hydrogen project acknowledged the uncertainty this creates: "We have a very clear understanding of how that early reliance on donor funding is going to need to be coupled with quite soon or perhaps replaced with private funding, and we don't know what that's going to look like" (I2). Niche green hydrogen actors have collaborated with larger, more established regime actors to secure financing: "So our largest shareholder is a [private equity] institution. And the other shareholder is a renewable energy [company] that doesn't want to go into hydrogen themselves but wants a little bit of exposure" (I7).

The influence of underlying logics and affects can also be discerned in niche-regime interactions. For example, trust in corporate financing is critical: "Aspects of the project may stack up fine, but the people or the borrower itself or the sponsors to the project are not necessarily the right sponsors" (I22). The basis of

compliance can also be significant: “[Donor funding] is unaccountable, exclusionary. I also think that it’s, what would be the word? I think a lot of the time, it is controlling. You know, because there’s a strange competitiveness that seems to exist between these kinds of sources of donor funding...And also, often that there’s a certain desire to control what you do with that funding” (I14). In conclusion, the complex interplay between different types of financing and their associated rules impacts system-level dynamics. Understanding these nuances is critical for niche actors as they navigate the challenges of securing funding for innovative projects.

6.2 Financing meta-rules in financing dynamics

To address the research question of how financing rules impact financing dynamics, the second output of the analysis process focuses on the major financing rules that may govern financing dynamics in South Africa. In one sense, the institutional characteristics can be considered as nested rules that fit into an overarching financing architecture applicable across different socio-technical contexts. Drawing on the notion of meta-rules from Schot and Kanger (2018), Table 4 below maps overarching regulative, normative, and cognitive rules identified in the case study to the four dimensions of financing dynamics.

Table 3. Financing meta-rules underpinning financing dynamics

	Forms of Financing	Use of Funds	Decision-making Factors	Actor Interactions
Regulative Rules	• All forms of financing are subject to South African legislation and regulation, e.g., rules to prevent money laundering, consumer protection • Independent regulatory approval is a prerequisite for new instruments. • Non-compliance with regulatory requirements subject to fines, penalties, revocation of licenses, loss of reputation.	• Financing agreements specify the intended use of funds and restrict their use for other purposes. • Due diligence and risk assessments consider ESG implications and SDGs with exclusionary impacts.	• Financing actors participate in the market according to specific mandates or guidelines. • Actors must demonstrate compliance with mandates through regular reporting, audits, or other mechanisms.	• Formal and informal private regulations provide additional guidance or standards for industry actors, e.g., SDGs. • Perceived compliance with laws, regulations is imperative for all actors. • Collaboration among actors necessary for the development of new financial products and services.
Normative Rules	• Structural change should be pursued through available policy levers that have proven efficacy. • Fiscal support should be provided when necessary to expand market participation. • Private sector market dominance should not hinder policy transmission through public financing.	• Financing actors should adhere to accepted risk appetite profiles based on their domain and use of funds. • A strong economic rationale should support social development initiatives.	• Core finance principles are well- established and typically do not necessitate further theorisation. • Centralised decision-making processes and systems are preferred for optimal outcomes. • Intuition and instinct complement epistemological frameworks for decision-making. • Incorporating ESG considerations can reduce risk	• Demonstrating financial knowledge is important for non-finance actors to establish credibility. • Financing actors must fulfil their traditionally defined roles (and markets) to maintain market stability. • Specialist expertise is essential for the efficient allocation of resources.

Cultural-cognitive Rules	• Recognition of the de-risking role of the government to enable market growth and development. • Belief in the direct relationship between government incentives and tangible economic impact. • Perception of innovation as related to technology and intellectual property with limited opportunities for financing.	• Shared understanding that policy and regulation drive investment, innovation, and development. • Cognitive linkages between industry growth, market growth, and economic growth.	• Adherence to neo-classical economic principles, such as rational actors, portfolio theory, and risk-return. • Valuing financial mathematics and modelling as primary tools for understanding financial dynamics. • Emphasis on track records, proven success, and individual expertise in financial decision-making. • Scalability of investments/assets/projects is critical to recoup due diligence and transaction costs.	• Reliance on well-established channels for inter- and intra-organisational communication. • Shared belief in the importance of collaboration among various actors for knowledge sharing, risk management, and product development. • Importance of trust in actor interactions supporting relationship-based and network-based financing access.

The widespread diffusion of financing rules is an important aspect of financialisation, and this was evident in the case study sample for financing and innovating actors. For example, adherence to neo-classical principles and economic theories is a widely accepted and understood cultural-cognitive rule governing financing decision-making (I1, I3, I4, I10, I15, I19, I22). Processes of diffusion point to the significance of linkages with international actors across different contexts and geographies: “We do have a relatively close relationship with our offshore peers... there’s a lot of learnings that we can take from our international colleagues” (I22). The de-risking role assigned to the government is another (I2, I3, I10, I11, I15, I18): “What [de-risking] boils down to is that government will need to take some risk either directly financially or indirectly as in guarantees” (I15).

Similarly, environmental, social and governance [ESG] considerations have diffused through multilateral and international organisations and frameworks – although data collected suggested a good deal of variation in organizational practice (I4, I10, I11, I19). One participant stated: “I don’t think the importance of the green economy needs to be discussed or mentioned to funders. [There are] so many international initiatives, government initiatives, policies, by purpose that is driving that” (I11). This diffusion is important because relative performance is a key indicator (I3, I5, I9). As one portfolio manager reflected: “We were doing [ESG] 10 years ago, and no one else was, and we thought, are we doing something wrong? Because we’re missing out on deals because we had certain standards, and we weren’t comfortable” (I10).

Through the analysis process, one can also see how single rules are linked to form clusters that impact financing dynamics. For example, independent regulatory approval for new products (forms of financing: regulative), along with actor collaboration for new financial products (actor interactions: regulative), and financing actors must fulfil their traditionally defined roles (actor interactions: normative) may underpin this statement from an institutional asset manager: “We’ve explored so many options [for renewable energy assets]. You know, we’d asked [Development Bank of South Africa] at one point, why don’t they set up some kind of securitization vehicle in which they can package the assets and then we can buy listed notes into that? I mean, that never came through” (I1). This relative coherence of rules is a critical aspect underpinning stability and resistance to change (Schot and Kanger 2018).

The analysis also illustrates how landscape factors can shape the ontological hierarchy within rule clusters. Consider the rule sets associated with ESG: International organisations’ frameworks guide ESG implementation (actor interactions: regulative), coupled with due diligence and risk assessments that take ESG implications into account (use of funds: regulative), and ESG factors can

reduce risk (decision-making factors: normative). Here, the shared understanding that policy drives investment and innovation (use of funds: cultural-cognitive), along with the belief that social development initiatives should be backed by a robust economic rationale (use of funds: normative), can also contribute to how ESG principles manifest differently in various energy markets. As one Head of Sustainable Finance at a banking institution elucidated: “In South Africa, we will still finance thermal coal for Eskom, whereas in the UK, we’ve decided no more thermal no more coal at all, not even metallurgical...As a South African corporate, you have a responsibility to keep the economy going” (I19). From a critical realist perspective, this statement illustrates that the relative power of rules as causal mechanisms should be considered and that a rule may exert greater (or weaker) influence under specific contextual conditions.

6.3 Financing innovation: rules and energy system impacts

The third output summarises the possible causal influence of financing meta-rules and institutional characteristics on South Africa’s energy regime. The six regime dimensions of the multi-level perspective provided the framework for analysing and presenting the data (Table 5).

Table 4. Causal impact of financing rules on South Africa’s energy system’s financing dynamics

Policy	• Failure to retain institutional capacity in public finance institutions has limited policy transmission through government actors and constrained renewable energy markets (I1, I3, I9). • High degree of collaboration needed on new mechanisms has limited efficacy of policymakers, regulators and financing actors (I1, I3). • Established feedback routes to policymakers narrowed the set of possible interventions for alternative financing structures (I1, I3, I5, I9, I18). • Regulatory and policy gaps promoted self-regulation, gentleman’s agreements, and adherence ‘on paper’ (I4, I10, I21). • Bilateral conversations and collaborations with other governments were instrumental in shaping energy policy for public funding (I6, I12, I15, I16). • Public-private partnerships and interactions between the finance sector and policymakers influenced policy decisions (I14, I15, I17).

Culture	• Network-based and relationship-based financing narrows the pool of potential participants in the industry, with implications for inclusivity and economic redress (I2, I6, I7, I8, I13). • Comparative analysis and benchmarking constrained actors' ability to develop alternative rules that may reduce returns (I1, I3, I4, I5). • Complex intra-organisational and inter-organisational communication norms constrained opportunities for expanding networks and relationships to different types of actors (I1, I4, I5, I9, I17). • Data-driven decision-making has advanced in the industry, and using quantitative and qualitative data in decision-making across financing domains facilitated renewable energy investment (I4, I5, I9, I10, I13, I20). • A growth mindset permeated approaches to financing, shaping market perceptions and expectations (I1, I3, I7, I11, I17, I21). • The government's macroeconomic policies and socio-economic objectives shape financing strategies for renewable energy, permeating all funding types (I3, I6, I10, I11, I15, I17). • Philanthropic and international financiers were important funders of research into social and cultural dimensions of energy transitions (I2, I11, I13, I14).
Market, user preferences	• Appropriate policy incentives and interventions, such as tax incentives and carbon credit trading, encouraged private financing involvement and innovation (I7, I8, I15, I17, I18). • Financing business models dictate participation in energy markets (I1, I3, I9, I10, I18). • Investment financing actors in unlisted spaces have greater flexibility for engaging in niches (I1, I3, I7, I10, I18). • The pace of renewable energy capacity expansion positively impacts the rate of sell-downs (from corporate financing actors) and renewable energy market dynamics. (I9, I10, I18, I21). • Banking business models facilitated greater participation in energy markets vis-à-vis other financing domains (I1, I5, I9, I10, I17).

Industry	• Policy certainty is a key decision-making factor: explicit government support is needed for large-scale initiatives but can create opportunities for small businesses (niche actors) in the green hydrogen sector (I7, I9, I18, I25). • The efficacy of policy mechanisms to spur corporate and investment financing is limited in the current context of poor asset and economic growth (I1, I9, I10, I18). • Financing for new renewable energy technology deployment is primarily provided through loan instruments and led by the banking sector, which has determined the position of other financing actors on the yield curve (I1, I3, I9, I10, I22).
Science	• Funding allocations for research are tied to business plans, requiring innovators to align their projects with the financier's specific commercial and strategic goals (I2, I6, I8, I10, I13). • The growing interest in green hydrogen as a potential energy source created funding opportunities for research and development in this area (I6, I8, I11, I13). • Public, international funding played a significant role in shaping directionality within the energy sector through early investment and setting research agendas and mandates (I6, I7, I8, I14, I15). • Companies that provided funding for research through various mechanisms, such as grants, partnerships, and in-kind contributions, shaped the scientific landscape in the energy sector (I6, I8, I15)
Technology	• The conceptualisation of renewable energy innovation as intellectual property (amongst others) guided financing actors, focusing on securing and commercialising proprietary technologies (I1, I5, I9, I10, I16, I20). • The lack of robust venture capital capacity (a sub-optimal mix of financing actors) has hindered the development and diffusion of renewable energy technologies and limited the participation of new actors in the market (I8, I15, I17, I18). • Financing actors may not view innovation capacity as a critical factor in decision-making or asset valuation, shaping the corporate and investment financing available for niche innovation (I1, I4, I17, I18). • Firms incorporating innovative renewable energy technologies have increased the value of their assets and improved their value position (I3, I10).

While Table 5 describes a myriad of possible causal influences on the energy system, the discussion here explores how financing rules can influence system transformation, particularly concerning niche innovation. Specifically, three key implications for the energy transition are discussed in further detail.

Firstly, financing rules may interact to reinforce and reproduce exclusionary practices in the South African context, resulting in regime optimisation limited to technological change (see Ghosh and Schot 2019). Financing actors pay significant weight on track records while participants rely on networks and relationships to secure financing (I2, I3, I7, I8, I9, I10, I11, I14, I16, I17, I18, I22). For example, in public financing for research and development, one government funder explained: “If you look at the landscape of institutions of higher learning and suddenly you have those traditional universities that have been there for longer, that have a track record in research. It’s easy for them to access funding at an institution like CSIR, TIA because they have the record, they have the people, they have the expertise to research and make products” (I16). Consequently, the reliance on existing relationships and track records can inadvertently result in a quasi-closed system, limiting the pool of innovators.

Moreover, while quantitative methods play a crucial role in decision-making, several participants mentioned the importance of instinct and intuition (I3, I10, I9, I17). For example, one portfolio manager explained: “It comes down to judgements, judgement, and you can’t define judgement. It’s a blend of luck, experience and insight and as well, as you know, hard numbers which you can put on paper, maybe rules of thumb, basic rules of thumb, but that’s just one side of it” (I9). This reliance on judgment can prove challenging to new entrants: “On the finance side, what I’ve struggled with over the years is like, what are your criteria? What are those? What are the things we need to do for you to actually say, yes, we’ll pull the trigger on it? They tend to talk a little bit high level...it feels like you’re pulling teeth” (I14).

Secondly, the rules associated with specific financing types may also impact the directionality of change. The active role of international and philanthropic financing in the policy, science, and culture dimensions was evident in the study (Appendix B). Moreover, in the cases of green hydrogen, this influence has shaped (and continues to shape) both innovation activities and government policy (I2, I6, I7, I8, I11, I12). For example, a hydrogen research programme director indicated that “we basically align ourselves with what the expectations are, you know...when they issued the call for proposals, they will indicate in what areas they will really fund. So we submit accordingly” (I6). As corporate and investment financing only comes on board when technology can be commercialised, early public, international and philanthropic funders play a pivotal role in developing

transition pathways. In developing contexts, international and philanthropic financing are critical and, through financing terms and conditions, may serve to entrench global North interests.

Lastly, the conceptualisation of innovation may influence the perspectives on its responsibility, fostering, and financing in a system transition context. The remark “So why have [intellectual property] if you can’t commercialise it?” (I7) underscores a prevalent assumption that financing innovation predominantly lies in the domain of venture capitalists and government, and its primary purpose is to generate economic benefits (I1, I3, I4, I6, I7, I8, I9, I15, I16, I17, I18, I20). Furthermore, financing actors may not consider an organisation’s innovation capacity or propensity in decision-making processes (I3, I4). While participants recognise innovation’s crucial role, the data reveals multiple implicit conceptualisations. Innovation was characterised as intellectual property, a precursor to commercialisation, an inherent attribute of an asset, a method to augment asset and investment value, or, very broadly, as the introduction of technologic into a system. However, a broader understanding of innovation within the niche context acknowledges the numerous learning and niche-building processes necessary and the diverse financing needs required to sustain it.

7. Conclusion

In conclusion, this study has unpacked the complex ways financing rules may impact financing dynamics and shape decarbonisation processes in South Africa’s energy transition. By defining and operationalising *financing dynamics*, this study provides a method of systematically linking financing rules to financing outcomes for niche innovation. Additionally, the taxonomy created for this study brings consistency to how financing instruments, types, and actors align with financing rules to understand the implications for innovating actors.

Three crucial implications emerged from the discussion. Firstly, the potential for financing rules to inadvertently perpetuate exclusionary practices. The reliance on track records and established networks can narrow the diversity of innovative actors. Secondly, the rules tied to financing domains can influence the directionality of change. Notably, international and philanthropic funding can be pivotal in shaping the system’s policy, culture, and science dimensions through financing terms and conditions. Lastly, how innovation is conceptualised can influence actor participation in financing niche innovation in the context of system transitions.

These insights lead to several recommendations for policymakers. Firstly, it is essential to respond to the social aspects of financing. Acknowledging the role of

social networks and personal relationships in financial decision-making can help cultivate more inclusive financing rules, broadening the circle of potential participants in the energy sector and the distribution of economic benefits. Interventions that support the expansion of networks and relationships could assist in overcoming barriers to the collaboration that developing alternative financing rules demands. Secondly, global South actors should leverage the multilateral pathways that have successfully diffused ESG principles to push for more transformative changes. Finally, it is crucial to broaden the conceptualisation of innovation to include niche processes and systems perspectives and to explore and advocate for its explicit incorporation in financing decision-making. Such theorisation could foster a more conducive environment for diverse actors to contribute to niche innovation for sustainable solutions.

While this research offers valuable insights, it is important to acknowledge its limitations. The study focuses on a specific context – South Africa’s energy transition – and engages with a relatively narrow set of actors within the system, thus potentially limiting the scope of the perspectives and experiences captured. Furthermore, while informal financing was not examined in this study due to its limited relevance to the green hydrogen niche, exploring its attributes and institutional features would be important for developing a more comprehensive taxonomy of financing dynamics for developing contexts (see Kraemer-Mbula and Wamae 2010). As such, while illuminating, the findings may not be directly generalisable to other settings or energy systems. However, in line with the critical realist approach adopted in this study, the intention is not to generate universally applicable conclusions but to identify and understand the causal mechanisms (rules) and general tendencies that might be applicable in similar contexts.

Future research could employ statistical generalisation to support and expand these findings. Additionally, synthesising literature from innovation studies to construct rule sets using larger data samples could support comparative analyses through the operationalisation of financing dynamics. Here, the definition of financing dynamics and the proposed taxonomy can be helpful tools for researchers seeking to link financing rules to financing outcomes systematically. Scholarly work could also explore the further integration of innovation into decision-making frameworks and models within finance.

References

- Anantharajah, K. and Setyowati, A.B., 2022. Beyond promises: Realities of climate finance justice and energy transitions in Asia and the Pacific. *Energy Research and Social Science*, [online] 89, p.102550. Available at: <<https://doi.org/10.1016/j.erss.2022.102550>>.
- Baker, L., 2015. The evolving role of finance in South Africa's renewable energy sector. *Geoforum*, [online] 64, pp.146–156. Available at: <<https://doi.org/10.1016/j.geoforum.2015.06.017>>.
- Barua, S. and Aziz, S., 2021. *Making green finance work for the sustainable energy transition in emerging economies*. [online] *Energy-Growth Nexus in an Era of Globalization*. Elsevier Inc. Available at: <<https://doi.org/10.1016/B978-0-12-824440-1.00014-X>>.
- Boni, A., Giachi, S. and Molas-Gallart, J., 2019. Towards a Framework for Transformative Innovation Policy Evaluation. [online] (April). Available at: <<http://www.tipconsortium.net/publication/towards-a-framework-for-transformative-innovation-policy-evaluation/>>.
- Caldecott, B., 2022. Defining transition finance and embedding it in the post-Covid-19 recovery. *Journal of Sustainable Finance and Investment*, [online] 12(3), pp.934–938. Available at: <<https://doi.org/10.1080/20430795.2020.1813478>>.
- Corrocher, N. and Cappa, E., 2020. The Role of public interventions in inducing private climate finance: An empirical analysis of the solar energy sector. *Energy Policy*, 147(January). Available at: <<https://doi.org/10.1016/j.enpol.2020.111787>>.
- Council for Scientific and Industrial Research (CSIR), 2006. South African Hydrogen Economy Activities - National Hydrogen Energy & Fuel Cell R & D Strategy development. In: *The Hydrogen Economy for Sustainable Development*. [online] Reykjavic, Iceland. Available at: <https://www.un.org/esa/sustdev/sdissues/energy/op/hydrogen_seminar/presentations/15_mathe_csir.pdf>.
- Dahlstrand, Å.L. and Cetindamar, D., 2000. The dynamics of innovation financing in Sweden. *Venture Capital*, 2(3), pp.203–221. Available at: <<https://doi.org/10.1080/13691060050135082>>.
- Department of Energy, 2019. *The South African Energy Sector Report 2019*. [online] *Energy and Fuels*, Available at: <<http://www.energy.gov.za/files/media/explained/2019-South-African-Energy-Sector-Report.pdf>>.
- Department of Energy, 2020. *Department of Energy Annual Report 2019/2020*. [online] Department of Energy.pp.1–265. Available at: <<http://www.energy.gov.za/files/AnnualReports/DoE-Annual-Report2019-20.pdf>>.

- Department of Mineral Resources and Energy, 2021. *The South African Energy Sector Report 2021*. [online] Pretoria. Available at: <<http://www.energy.gov.za/files/media/explained/2021-South-African-Energy-Sector-Report.pdf>>.
- Department of Science and Technology, 2007. National Hydrogen and Fuel Cell Technologies Research, Development and Innovation Strategy. (May).
- DSI (Department of Science and Innovation), 2021. *Hydrogen Society Roadmap for South Africa 2021*. [online] Pretoria. Available at: <https://www.dst.gov.za/images/South_African_Hydrogen_Society_RoadmapV1.pdf>.
- Dutt, D., 2022. How power and politics shape niche-regime interactions: A view from the Global South. *Environmental Innovation and Societal Transitions*, [online] 43(April), pp.320–330. Available at:<<https://doi.org/10.1016/j.eist.2022.04.008>>.
- Easton, G., 2010. Critical realism in case study research. *Industrial Marketing Management*, [online] 39(1), pp.118–128. Available at: <<https://doi.org/10.1016/j.indmarman.2008.06.004>>.
- Eastwood, J.G., Jalaludin, B.B. and Kemp, L.A., 2014. Realist explanatory theory building method for social epidemiology: A protocol for a mixed method multilevel study of neighbourhood context and postnatal depression. *SpringerPlus*, 3(1), pp.1–12. Available at: <<https://doi.org/10.1186/2193-1801-3-12>>.
- Egli, F., Polzin, F., Sanders, M., Schmidt, T., Serebriakova, A. and Steffen, B., 2022. Financing the energy transition: four insights and avenues for future research. *Environmental Research Letters*, 17(5). Available at: <<https://doi.org/10.1088/1748-9326/ac6ada>>.
- Elzen, B., Van Mierlo, B. and Leeuwis, C., 2012. Anchoring of innovations: Assessing Dutch efforts to harvest energy from glasshouses. *Environmental Innovation and Societal Transitions*, [online] 5, pp.1–18. Available at: <<https://doi.org/10.1016/j.eist.2012.10.006>>.
- Epstein, G.A., 2005. Introduction: Financialization and the world economy. In: *Financialization and the World Economy*. [online] pp.3–16. Available at: <<http://search.lib.cam.ac.uk/?itemid=%7Cdepfacfmdb%7C434152>>.
- Fatemi, A.M. and Fooladi, I.J., 2013. Sustainable finance: A new paradigm. *Global Finance Journal*, [online] 24(2), pp.101–113. Available at: <<https://doi.org/10.1016/j.gfj.2013.07.006>>.
- Fuenfschilling, L. and Truffer, B., 2014. The structuration of socio-technical regimes - Conceptual foundations from institutional theory. *Research Policy*, [online] 43(4), pp.772–791. Available at: <<https://doi.org/10.1016/j.respol.2013.10.010>>.
- Geddes, A. and Schmidt, T.S., 2020. Integrating finance into the multi-level perspective: Technology niche-finance regime interactions and financial policy

- interventions. *Research Policy*, [online] 49(6), pp.1–13. Available at: <<https://doi.org/10.1016/j.respol.2020.103985>>.
- Geels, F.W., 2002. Technological transitions as evolutionary reconfiguration processes: A multi-level perspective and a case-study. *Research Policy*, [online] 31(8–9), pp.1257–1274. Available at: <[https://doi.org/10.1016/S0048-7333\(02\)00062-8](https://doi.org/10.1016/S0048-7333(02)00062-8)>.
- Geels, F.W. and Schot, J., 2007. Typology of sociotechnical transition pathways. *Research Policy*, 36(3), pp.399–417. Available at: <<https://doi.org/10.1016/j.respol.2007.01.003>>.
- Ghosh, B. and Schot, J., 2019. Towards a novel regime change framework: Studying mobility transitions in public transport regimes in an Indian megacity. *Energy Research and Social Science*, [online] 51(December 2018), pp.82–95. Available at: <<https://doi.org/10.1016/j.erss.2018.12.001>>.
- Guterres, A., 2023. *Secretary-General's briefing to the General Assembly on Priorities for 2023*. Available at: <<https://www.un.org/sg/en/content/sg/speeches/2023-02-06/secretary-generals-briefing-the-general-assembly-priorities-for-2023>>.
- Hackbarth, D. and Sun, D., 2015. Corporate Investment and Financing Dynamics. *SSRN Electronic Journal*. Available at: <<https://doi.org/10.2139/ssrn.2651313>>.
- Hall, S., Roelich, K.E., Davis, M.E. and Holstenkamp, L., 2018. Finance and justice in low-carbon energy transitions. *Applied Energy*, [online] 222, pp.772–780. Available at: <<https://doi.org/10.1016/j.apenergy.2018.04.007>>.
- Haščić, I., Cárdenas Rodríguez, M., Jachnik, R., Silva, J. and Johnstone, N., 2015. Public Interventions and Private Climate Finance Flows: Empirical Evidence from Renewable Energy Financing. *OECD Environment Working Papers*, [online] 80(80), pp.1–88. Available at: <[http://www.oecd-ilibrary.org/environment/public-interventions-and-private-climate-finance-flows-empirical-evidence-from-renewable-energy-financing_5js6b1r9lfd4-en%0Afile:///Users/g299/Documents/Library.papers3/Reports/2015/Haščić/Haščić_2015_Public Inter](http://www.oecd-ilibrary.org/environment/public-interventions-and-private-climate-finance-flows-empirical-evidence-from-renewable-energy-financing_5js6b1r9lfd4-en%0Afile:///Users/g299/Documents/Library.papers3/Reports/2015/Haščić/Haščić_2015_Public%20Inter)>.
- Hogarth, J.R., 2012. The role of climate finance in innovation systems. *Journal of Sustainable Finance and Investment*, 2(3–4), pp.257–274. Available at: <<https://doi.org/10.1080/20430795.2012.742637>>.
- IEA, 2019. *Africa Energy Outlook 2019, Overview: South Africa*. [online] *Energy Outlook*. Available at: <<https://webstore.iea.org/africa-energy-outlook-2019>>.
- IEA, 2021. *Financing Clean Energy Transitions in Emerging and Developing Economies*. *iea*.
- International Bank for Reconstruction and Development and The World Bank, 2018. *Overcoming Poverty and Inequality in South Africa: An assessment of drivers, constraints and opportunities*. Available at: <<https://doi.org/>>

- 10.1596/29614>.
- IPCC, 1990. *Climate Change: The IPCC Scientific assessment*. [online] *Climate Change*. Cambridge, NY: Cambridge University Press. Available at: <https://www.ipcc.ch/site/assets/uploads/2018/03/ipcc_far_wg_I_full_report.pdf>.
- IPCC, 2022. *Climate Change 2022: Impacts, Adaptation, and Vulnerability*. [online] *Climate Change 2021: The Physical Science Basis*. Cambridge, UK and New York, NY, USA: Cambridge University Press. Available at: <<https://doi.org/10.1017/9781009325844>>.
- Irfan, M., Rehman, M.A., Razzaq, A. and Hao, Y., 2023. What drives renewable energy transition in G-7 and E-7 countries? The role of financial development and mineral markets. *Energy Economics*, [online] 121(September 2022), p.106661. Available at: <<https://doi.org/10.1016/j.eneco.2023.106661>>.
- Isah, A., Dioha, M.O., Debnath, R., Abraham-Dukuma, M.C. and Butu, H.M., 2023. Financing renewable energy: policy insights from Brazil and Nigeria. *Energy, Sustainability and Society*, [online] 13(1), pp.1–16. Available at: <<https://doi.org/10.1186/s13705-022-00379-9>>.
- Jain, M., Hoppe, T. and Bressers, H., 2017. Analyzing sectoral niche formation: The case of net-zero energy buildings in India. *Environmental Innovation and Societal Transitions*, [online] 25, pp.47–63. Available at: <<https://doi.org/10.1016/j.eist.2016.11.004>>.
- Kanger, L., 2022. The spatial dynamics of deep transitions. *Environmental Innovation and Societal Transitions*, [online] 44(April), pp.145–162. Available at: <<https://doi.org/10.1016/j.eist.2022.06.005>>.
- Kemp, R., Pel, B., Scholl, C. and Boons, F., 2022. Diversifying deep transitions: Accounting for socio-economic directionality. *Environmental Innovation and Societal Transitions*, 44, pp.110–124. Available at: <<https://doi.org/10.1016/j.eist.2022.06.002>>.
- Kemp, R., Schot, J. and Hoogma, R., 1998. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technology Analysis and Strategic Management*, 10(2), pp.175–198. Available at: <<https://doi.org/10.1080/09537329808524310>>.
- Kim, J.E., 2019. Sustainable energy transition in developing countries: the role of energy aid donors. *Climate Policy*, [online] 19(1), pp.1–16. Available at: <<https://doi.org/10.1080/14693062.2018.1444576>>.
- Klagge, B. and Nweke-Eze, C., 2020. Financing large-scale renewable-energy projects in Kenya: investor types, international connections, and financialization. *Geografiska Annaler, Series B: Human Geography*, [online] 102(1), pp.61–83. Available at: <<https://doi.org/10.1080/04353684.2020.1729662>>.
- Köhler, J., Geels, F.W., Kern, F., Markard, J., Onsongo, E., Wieczorek, A.,

- Alkemade, F., Avelino, F., Bergek, A., Boons, F., Fünfschilling, L., Hess, D., Holtz, G., Hyysalo, S., Jenkins, K., Kivimaa, P., Martiskainen, M., McMeekin, A., Mühlemeier, M.S., Nykvist, B., Pel, B., Raven, R., Rohrer, H., Sandén, B., Schot, J., Sovacool, B., Turnheim, B., Welch, D. and Wells, P., 2019. An agenda for sustainability transitions research: State of the art and future directions. *Environmental Innovation and Societal Transitions*, [online] 31(December 2018), pp.1–32. Available at: <<https://doi.org/10.1016/j.eist.2019.01.004>>.
- Konrad, K., Truffer, B. and Voß, J.P., 2008. Multi-regime dynamics in the analysis of sectoral transformation potentials: evidence from German utility sectors. *Journal of Cleaner Production*, 16(11), pp.1190–1202. Available at: <<https://doi.org/10.1016/j.jclepro.2007.08.014>>.
- Kraemer-Mbula, E. and Wamae, W., 2010. Adapting the innovation systems framework to Sub-Saharan Africa. In: E. Kraemer-Mbula and W. Wamae, eds. *Innovation and the Development Agenda*, 1st ed. OECD and the International Development Research Centre. pp.65–90. Available at: <<https://doi.org/10.1787/9789264088924-7-en>>.
- De la Torre, A., Feyen, E. and Ize, A., 2013. Financial development: Structure and dynamics. *World Bank Economic Review*, 27(3), pp.514–541. Available at: <<https://doi.org/10.1093/wber/lht005>>.
- Li, W., Cao, N. and Xiang, Z., 2023. Drivers of renewable energy transition: The role of ICT, human development, financialization, and R&D investment in China. *Renewable Energy*, [online] 206(October 2022), pp.441–450. Available at: <<https://doi.org/10.1016/j.renene.2023.02.027>>.
- Loorbach, D., 2010. Transition management for sustainable development: A prescriptive, complexity-based governance framework. *Governance*, 23(1), pp.161–183. Available at: <<https://doi.org/10.1111/j.1468-0491.2009.01471.x>>.
- Loorbach, D., Schoenmaker, D. and Schramade, W., 2020. Finance in Transition: Principles for a Positive Finance Future. *RSM Series on Positive Change*. [online] Available at: <https://www.rsm.nl/fileadmin/Images_NEW/Positive_Change/2020_Finance_in_Transition.pdf>.
- Mah, D.N. and Cheung, D.M., 2020. Conceptualizing Niche-Regime Dynamics of Energy Transitions from a Political Economic Perspective: Insights from Community-Led Urban Solar in Seoul. *Sustainability (Switzerland)*, 12(12). Available at: <<https://doi.org/10.3390/SU12124818>>.
- Mingers, J., 2006. *Realising Systems Thinking: Knowledge and Action in Management Science*. Contemporary Systems Thinking. [online] Springer US. Available at: <<https://doi.org/10.1007/0-387-29841-X>>.
- Mingers, J., 2011. The Contribution of Systemic Thought to Critical Realism. *Journal*

- of *Critical Realism*, 10(3), pp.303–330. Available at: <<https://doi.org/10.1558/jcr.v10i3.303>>.
- Naidoo, C.P., 2020. Relating financial systems to sustainability transitions: Challenges, demands and design features. *Environmental Innovation and Societal Transitions*, [online] 36(November 2018), pp.270–290. Available at: <<https://doi.org/10.1016/j.eist.2019.10.004>>.
- National Business Initiative, 2022. Financing South Africa’s Just Transition. [online] Available at:<<https://www.nbi.org.za/wp-content/uploads/2021/08/NBI-Transition-Chapter-Decarbonising-SA-power-11-Aug-2021.pdf>>.
- National Planning Commission, 2012. *National Development Plan 2030: Our future - make it work*. [online] Pretoria. Available at: <<http://www.gov.za/issues/national-development-plan-2030>>.
- NBI (National Business Initiative), 2022. Decarbonising South Africa’s Power System: Just Transition and Climate Pathways Study for South Africa. [online] p.42. Available at: <<https://www.nbi.org.za/wp-content/uploads/2021/08/NBI-Transition-Chapter-Decarbonising-SA-power-11-Aug-2021.pdf>>.
- OECD/Eurostat, 2018. *Oslo Manual 2018: Guidelines for Collecting, Reporting and Using Data on Innovation*. 4th Editio ed. [online] *The Measurement of Scientific; Technological and Innovation Activities*. Luxembourg. Available at: <<https://doi.org/10.1787/9789264304604-en>>.
- Owen, R., Brennan, G. and Lyon, F., 2018. Enabling investment for the transition to a low carbon economy: government policy to finance early stage green innovation. *Current Opinion in Environmental Sustainability*, [online] 31, pp.137–145. Available at: <<https://doi.org/10.1016/j.cosust.2018.03.004>>.
- Penna, C.C.R., Schot, J. and Steinmueller, W.E., 2021. The Promise Of Transformative Investment: Mapping The Field Of Sustainability Investing. [online] Available at: <https://deeptransitions.net/wp-content/uploads/2021/11/DT-Working-Paper-Transformative-Investments_final.pdf>.
- Polzin, F. and Sanders, M., 2020. How to finance the transition to low-carbon energy in Europe? *Energy Policy*, [online] 147(December 2019), p.111863. Available at: <https://doi.org/10.1016/j.enpol.2020.111863>
- Polzin, F., Sanders, M. and Serebriakova, A., 2021. Finance in global transition scenarios: Mapping investments by technology into finance needs by source. *Energy Economics*, [online] 99, p.105281. Available at: <<https://doi.org/10.1016/j.eneco.2021.105281>>.
- Ragosa, G. and Warren, P., 2019. Unpacking the determinants of cross-border private investment in renewable energy in developing countries. *Journal of Cleaner Production*, [online] 235, pp.854–865. Available at: <<https://doi.org/10.1016/j.jclepro.2019.06.166>>.
- Rip, A. and Kemp, R., 1998. Technological change. In: S. Rayner and E.L. Malone,

- eds. *Human choice and climate change*. Columbus, OH: Battelle Press. pp.327–399. Available at: <<https://doi.org/10.1002/9781119721307>>.
- Rosenbloom, D., 2019. A clash of socio-technical systems: Exploring actor interactions around electrification and electricity trade in unfolding low-carbon pathways for Ontario. *Energy Research and Social Science*, [online] 49(October 2018), pp.219–232. Available at: <<https://doi.org/10.1016/j.erss.2018.10.015>>.
- Schot, J., Benedetti, R., Steinmueller, E., Keesman, S., Kanger, L., Davies, J., Ward, V. and Ghosh, B., 2022. Transformative Investment in Sustainability: An Investment Philosophy for the Second Deep Transition. (November).
- Schot, J. and Geels, F.W., 2008. Strategic niche management and sustainable innovation journeys: Theory, findings, research agenda, and policy. *Technology Analysis and Strategic Management*, [online] 20(5), pp.537–554. Available at: <<https://doi.org/10.1080/09537320802292651>>.
- Schot, J. and Kanger, L., 2018. Deep transitions: Emergence, acceleration, stabilization and directionality. *Research Policy*, [online] 47(6), pp.1045–1059. Available at: <<https://doi.org/10.1016/j.respol.2018.03.009>>.
- Schot, J. and Steinmueller, W.E., 2018. Three frames for innovation policy: R&D, systems of innovation and transformative change. *Research Policy*, 47(9), pp.1554–1567. Available at: <<https://doi.org/10.1016/j.respol.2018.08.011>>.
- Scott, W.R., 2014. Crafting an Analytic Framework I: Three Pillars of Institutions. In: *Institutions and Organizations: Ideas, Interests, and Identities*, Fourth. Thousand Oaks, California: SAGE Publications Inc. pp.473–474.
- Seyfang, G. and Gilbert-Squires, A., 2019. Move your money? Sustainability Transitions in Regimes and Practices in the UK Retail Banking Sector. *Ecological Economics*, [online] 156(February 2018), pp.224–235. Available at: <<https://doi.org/10.1016/j.ecolecon.2018.09.014>>.
- Sorenson, O., Assenova, V., Li, G.C., Boada, J. and Fleming, L., 2016. Expand innovation finance via crowdfunding: Crowdfunding attracts venture capital to new regions. *Science*, 354(6319), pp.1526–1528. Available at: <<https://doi.org/10.1126/science.aaf6989>>.
- Sovacool, B.K., Baker, L., Martiskainen, M. and Hook, A., 2019. Processes of elite power and low-carbon pathways: Experimentation, financialisation, and dispossession. *Global Environmental Change*, [online] 59(January), p.101985. Available at: <<https://doi.org/10.1016/j.gloenvcha.2019.101985>>.
- Steffen, B. and Schmidt, T.S., 2021. Strengthen finance in sustainability transitions research. *Environmental Innovation and Societal Transitions*, [online] pp.77–80. Available at: <<https://doi.org/10.1016/j.eist.2021.10.018>>.
- Sutherland, L.A., Peter, S. and Zagata, L., 2015. Conceptualising multi-regime interactions: The role of the agriculture sector in renewable energy transitions.

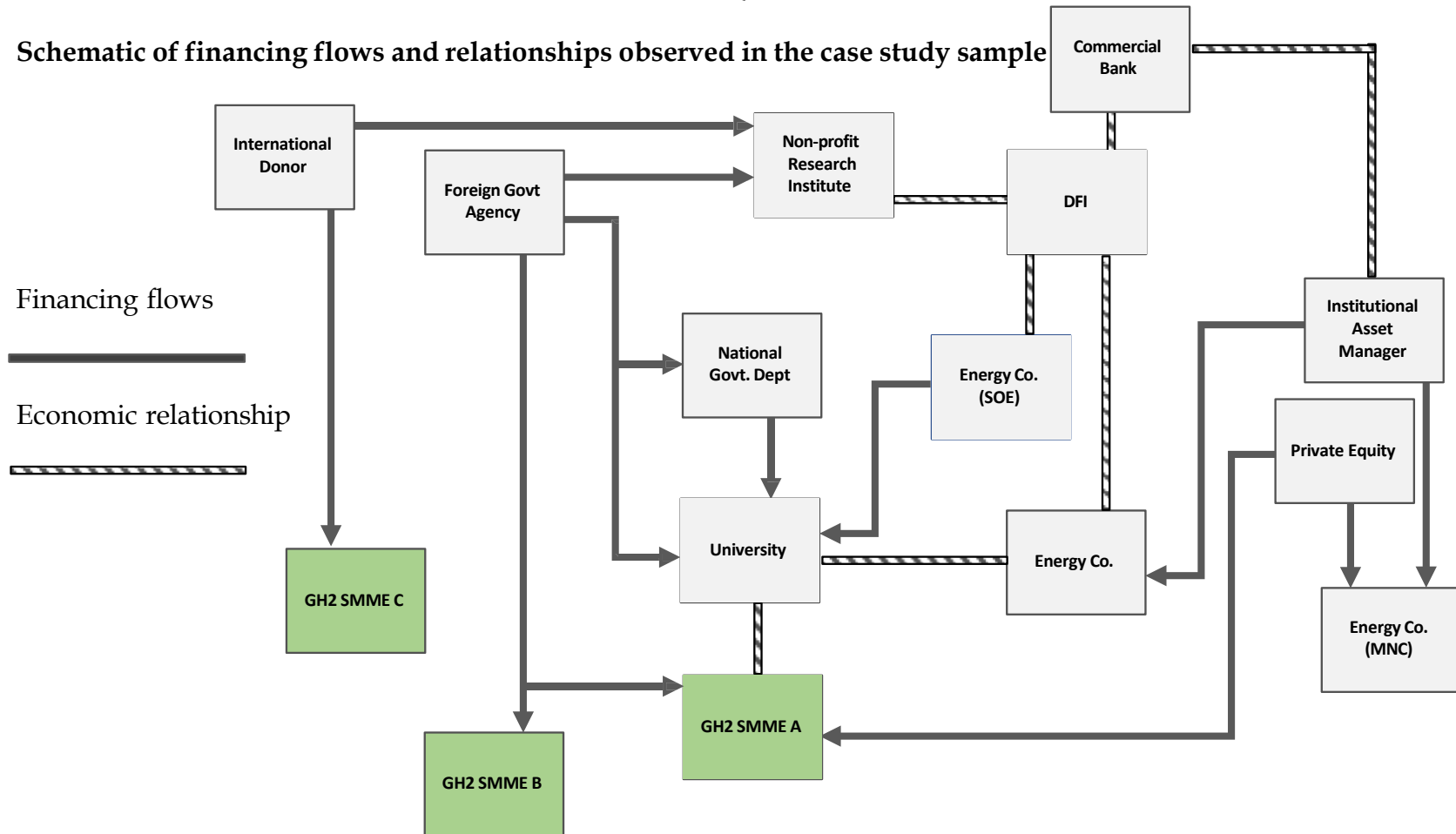
- Research Policy*, [online] 44(8), pp.1543–1554. Available at: <<https://doi.org/10.1016/j.respol.2015.05.013>>.
- The Presidency, 2022. *South Africa's Just Energy Transition Investment Plan (JET IP)*. [online] Pretoria: The Presidency, Republic of South Africa. Available at: <<https://www.thepresidency.gov.za/download/file/fid/2649>>.
- The World Bank, 2022. *Inequality in Southern Africa: An assessment of the Southern African customs union*. [online] Washington DC. Available at: <<https://documents1.worldbank.org/curated/en/099125303072236903/pdf/P1649270c02a1f06b0a3ae02e57eadd7a82.pdf>>.
- Ting, M.B. and Byrne, R., 2020. Eskom and the rise of renewables: Regime-resistance, crisis and the strategy of incumbency in South Africa's electricity system. *Energy Research and Social Science*, [online] 60(October 2019), p.101333. Available at: <<https://doi.org/10.1016/j.erss.2019.101333>>.
- United Kingdom Cabinet Office and Sharma, A., 2022. *Joint statement following the visit of COP26 President Alok Sharma to South Africa*. Available at: <<https://www.gov.uk/government/news/joint-statement-following-the-visit-of-cop26-president-alok-sharma-to-south-africa>>.
- United Nations, 2015. The Paris Agreement. *Towards a Climate-Neutral Europe: Curbing the Trend*, [online] pp.24–45. Available at: <<https://doi.org/10.4324/9789276082569-2>>.
- Walter, T. and Wansleben, L., 2020. The Assault of Financial Futures on the Rest of Time. In: I. Scoones and A. Stirling, eds. *The Policy of Uncertainty: Challenges of Transformation*, 1st ed. [online] London: Routledge. Available at: <<https://doi.org/10.4324/9781003023845>>.
- Walwyn, D.R., 2020. Turning points for sustainability transitions: Institutional destabilization, public finance and the techno-economic dynamics of decarbonization in South Africa. *Energy Research & Social Science*, [online] 70(January), p.101784. Available at: <<https://doi.org/10.1016/j.erss.2020.101784>>.
- Welch, C., Piekkari, R., Plakoyiannaki, E. and Paavilainen-Mäntymäki, E., 2011. Theorising from case studies: Towards a pluralist future for international business research. *Journal of International Business Studies*, 42(5), pp.740–762. Available at: <<https://doi.org/10.1057/jibs.2010.55>>.
- Yang, K., Schot, J. and Truffer, B., 2022. Shaping the directionality of sustainability transitions: the diverging development patterns of solar photovoltaics in two Chinese provinces. *Regional Studies*, [online] 56(5), pp.751–769. Available at: <<https://doi.org/10.1080/00343404.2021.1903412>>.
- Zhang, F., 2020. Leaders and followers in finance mobilization for renewable energy in Germany and China. *Environmental Innovation and Societal Transitions*, [online] 37(July), pp.203–224. Available at: <<https://doi.org/10.1016/j.eist.2020.08.005>>.

APPENDIX A

#	Role/Designation	Domain Experience
I1	Head of Fixed Income Research	Investment
I2	Director. Renewable Energy Entrepreneur	Public International Philanthropic
I3	Portfolio Manager: Equity	Investment
I4	ESG Analyst	Investment
I5	Multi-asset trader	Investment
I6	Hydrogen Research Programme Director	Public International Philanthropic
I7	Managing Director. Green Hydrogen SMME Portfolio Manager: Energy	Philanthropic Public International Investment
I8	Project Manager: Hydrogen Programme	Public
I9	Portfolio Manager: Equity	Investment
I10	Portfolio Manager: Debt/Equity, Renewable Energy	Investment
I11	Head: Renewable Energy Research	Public International Philanthropic
I12	Director: Hydrogen and Renewable Energy	Public
I13	Director: Energy Research Consultancy	Public International Philanthropic
I14	Renewable Energy Entrepreneur	Public International Philanthropic
I15	Government Industry Strategist. Green Hydrogen	Public International Investment
I16	Portfolio Manager: Energy Unit	Public
I17	Venture Capitalist. Angel Investor	Investment
I18	Head: Equities	Corporate
I19	Head: Sustainable Finance	Corporate Investment
I20	Chief Director. National government	Public
I21	Head: Sustainable Finance	Corporate Investment
I22	Head: Power and Infrastructure	Corporate Investment

APPENDIX B

Schematic of financing flows and relationships observed in the case study sample



Notes

- ¹ In transitions literature, the term “rules” replaces “institutions” as it would be referred to in institutional theory, given that the latter may be misconstrued as referring to public organisations (Geels and Schot 2007)
- ² The Just Energy Transition Partnership (JETP) was established at the United Nations Framework Convention on Climate Change 26th Conference of the Parties (COP26) between South Africa and the International Partners Group (IPG). The IPG constitutes the governments of France, Germany, the United Kingdom, the United States, and the European Union

Achieving sustainable industrial path development through territorial valuation dynamics – An analytical framework and case study of urban water management in Bengaluru

Johan Miörner

Lund University, Sweden

Christian Binz

Eawag: Swiss Federal Institute of Aquatic
Science and Technology, Switzerland;

Huiwen Gong

University of Stavanger, Norway

Abstract

This paper critically explores the link between innovation-driven industry dynamics and regional sustainability outcomes. While existing literature has largely examined the industrial dynamics associated with green industry emergence, there is a limited understanding of their impacts on social and environmental sustainability. In response, this paper emphasizes the role of 'valuation' processes in steering regional path development towards specific values, goals, and rationales. An analytical framework focusing on the interplay between industry dynamics, market valuation, and societal valuation is developed.

This framework is applied to an empirical case study in the Urban Water Management sector of Bengaluru, India, where shifts in societal and market valuation dynamics have led to a transition from 'pure' industrial path development focusing on creating economic value, to a more sustainable industrial path encompassing broader societal valuation dynamics.

Keywords: industrial path development, sustainable outcomes, valuation, urban water management, Bengaluru

1. Introduction

Scholarly work at the intersection of innovation studies and economic geography has directed increasing attention to exploring “green” or “sustainable” regional restructuring. There is extensive empirical interest in path development in renewable energy technologies and related industries (Dawley, 2014; Simmie, 2012; Steen, 2016; Mackinnon et al., 2021; Steen and Hansen, 2018; Binz et al., 2016; Cooke, 2012). Recently, there has also been an emerging conceptual interest, in terms of studies zooming in on 'green industrial path development', green growth, and green innovation-based renewal (Trippel et al., 2020; Capasso et al., 2019; Gibbs and Jensen, 2021; Nilsen and Njøs, 2021).

Recently, there have been calls to study the link between new path development and regional development from a broader economic and social perspective (Breul et al., 2021; Eadson and Van Veelen, 2023) and zoom in on the potentially adverse effects of industry dynamics on local developmental outcomes (Blažek et al., 2020). Nevertheless, in the new path development literature and evolutionary economic geography more broadly, there is still a limited understanding of the link between industry dynamics and regional development outcomes (Martin and Sunley, 2015; Coenen and Morgan, 2020; Boschma, 2017). This is particularly relevant when it comes to ‘green’ industry dynamics (Trippel et al., 2020; Eadson and Van Veelen, 2023). It is often implicitly assumed that local industry dynamics and innovation in, for example, clean tech, renewable energy or electric mobility, automatically lead to positive regional development outcomes.

Yet, studies have often focused rather narrowly on the industrial dynamics and regional system reconfiguration that are associated with green industry emergence, without exploring their impacts on social and environmental sustainability in much depth. In other words, few studies have interrogated the potential of green industry dynamics in terms of delivering on promises of sustainability in local contexts.

In this paper, we draw on insights from transition studies and understand sustainable path development outcomes as being mediated by the emergence of new, or transformation of existing, regional socio-technical configurations of technologies, actors and institutions ‘that works’ in terms of providing a societal function in a way that, in addition to economic value, provides environmental and societal benefits. Socio-technical configurations are often legitimized by pointing to one or a combination of several core values that provide a guiding rationale for actors and their value creating- and capturing activities.

Existing industrial path development literature has largely applied a technology-centred perspective on innovation, one which focuses on the development of new products and services, which are then valued in, and

diffused through, markets. In this paper, we argue in line with recent literature in regional studies and transition studies that if regional development outcomes are the focus, 'valuation' processes (Jeannerat and Kebir, 2016; Binz and Truffer, 2017; Carvalho and Van Winden, 2018; Yap et al., 2022; Huguenin and Jeannerat, 2017; Jeannerat and Crevoisier, 2022) that steer regional path development processes towards specific values, goals and rationales, have to be put centre stage.

Valuation refers to "the relational and institutional dynamics by which different objects and activities are socially valorized (i.e. transformed and commercialized) and evaluated (interpreted, recognized, legitimated and appraised) in the market" (Jeannerat and Kebir, 2016, p. 277). It involves the institutionalization of elements that shape recognition and engagement (Callon et al., 2002; Jansson and Waxell, 2011; Jeannerat, 2013). In addition to 'market valuation' as the primary mechanism for providing incentives to industry actors, it is crucial to consider 'societal valuation' (Huguenin and Jeannerat, 2017) if we want to capture how industry dynamics lead to sustainable outcomes.

Inspired by previous literature (Jeannerat and Kebir, 2016; Huguenin and Jeannerat, 2017; Jeannerat and Crevoisier, 2022), we connect industrial path development literature with the concept of territorial valuation (Jeannerat and Crevoisier, 2022) in regional contexts. This refers to mutually reinforcing processes of market valuation and societal valuation that are shaped by regional relational and institutional set-ups that reflect (a set of) core values. By studying valuation contexts, we open the 'black box' of how industry dynamics lead to certain sustainable outcomes (or not) in a particular local/regional context.

Middle-income cities (MICs) provide a potentially unique context to study valuation processes and their effect on regional sustainability outcomes. They have particularly strong environmental and social pressures, coupled with low path dependencies in established infrastructure systems and dynamic industrial change processes. Their firms are often mostly supplying to international global production networks, or because market interests and economic prosperity are valued higher than protecting the environment or achieving social justice. MICs are often characterised by limited organizational capabilities and depending on more dynamic, distributed and/or market-oriented transition trajectories.

We develop an analytical framework that aims to provide a comprehensive understanding of valuation contexts, particularly focusing on the interplay between industry dynamics, market valuation, and societal valuation. This framework illustrates the impact of these elements on the socio-technical configurations that emerge as a result of industrial activities and valuation processes. Furthermore, it highlights how the establishment of these configurations is inherently connected to specific outcomes.

Our analytical framework is illustrated with an empirical case study in the Urban Water Management (UWM) sector of Bengaluru, India, which provides an emblematic example of a 'green' path development trajectory in a changing valuation context. Industry formation around small-scale wastewater treatment solutions in this city showed a gradual reorientation from 'pure' industrial path development' to more and more 'sustainable path development'. Tracing shifts in the relevant societal and market valuation dynamics is a key factor explaining this outcome.

2. Analytical framework

2.1. Regional industrial path development and regional outcomes

We define regional industrial path development in line with previous studies as "the emergence and growth of new industries and economic activities in regions" (MacKinnon et al., 2019, p. 114). It refers to the emergence of a critical mass of functionally related actors (firms and other supporting actors) and institutions that form an industrial development trajectory which can retain a degree of persistence over time (Hassink et al., 2019; Steen and Hansen, 2018). Over the last two decades, the path development literature has flourished with studies highlighting different explanatory factors, sources and mechanisms of path development (see Hassink et al., 2019; Miörner, 2019).

Scholars have recently started to question the taken-for-granted benefits of regional industrial path development, and the lack of normative directionality (Marques and Morgan, 2021).

Calls have been made to better understand the link between industrial change processes and regional development outcomes. For example, Blažek et al. (2020) investigated evolutionary pathways of decline and the 'dark side' of path development and outlined industrial development trajectories that are detrimental for the regional economy. Breul et al. (2021) developed a framework for understanding synergetic relationships and sources of economic inequality associated with regional industrial evolution and showed how the development of a tourism path in Namibia had both positive and negative impacts on the agricultural path in that country. They conceptualised the effect on existing regional industries by the emergence of a new one as 'reformation processes' shaped by inter-industry relationships and demonstrate how the creation of a tourism path had profound consequences for regional agriculture activities, triggered by resource reformation processes.

When it comes to the link between regional industrial evolution and sustainability, there is an emerging body of literature concerned with analysing industrial path development which combines economic growth with normative directionality (Uyarra et al., 2019; Tödtling et al., 2021), most prominently in the form of 'green' industrial path development (Trippel et al., 2020; Sotarauta et al.,

2021; Nilsen and Njøs, 2021) or development opportunities in the ‘foundational economy’ (Coenen and Morgan, 2020; Hansen, 2022).

Green industrial path development refers to the rise of new industries with a green profile, like renewable energies, electric mobility or organic agriculture, and the ‘greening’ of existing industries through innovation-based renewal that leads to more environmentally friendly products and services or production processes (Trippel et al., 2020, p. 189). In other words, the concept captures industrial change processes that seek to generate *economic value* by developing more environmentally friendly products or production processes.

The allure of generating growth via industrial change while at the same time achieving sustainability goals is highly attractive to policy makers and regional planners (Capasso et al., 2019; Tödtling et al., 2021). However, it is well established that innovation processes around environmental innovations and their benefits may be unevenly and asymmetrically distributed across space (Guðmundsdóttir et al., 2018; Mccauley and Stephens, 2012; Gibbs and Jensen, 2021). In other words, it is not clear how, where and under what conditions such ‘green’ industry dynamics are linked to environmental outcomes such as a reduction in pollution and greenhouse gas emissions, or to the substantial transformation in regional infrastructure systems such as housing, energy, water and transport (Andersen et al., 2020; Binz and Truffer, 2017; Van Welie et al., 2018; Huenteler et al., 2016; Klerkx et al., 2012; Elzen et al., 2012).

In this paper, we seek to capture how creating new or reorienting existing industrial activities in a region, can lead not only to economic prosperity but may also contribute to solving key environmental and/or social challenges in the same region. We are thus looking to understand the broader processes of generating *societal value*, that forge links between industry dynamics and (positive) environmental and social outcomes in the same region, which are realized by reorienting industrial activities towards finding solutions for local problems.

2.2. Socio-technical configurations and outcomes

We start by defining ‘sustainable outcomes’ as inherently connected to transformation processes in fundamental infrastructure sectors (e.g. housing, energy, water, transport, etc.). Key societal functions such as energy provision, water supply, sanitation, or mobility are provided by ‘socio-technical configurations’ of actors, technologies and institutions that co- evolve and become aligned in stable patterns of coordination (Kemp et al., 1998; Geels and Turnheim, 2022; Markard et al., 2012). Socio-technical configurations encompass both the supply and demand side of a sector and are characterized by a distinct way of providing a societal function and are motivated by, and reflected in, a set of economic, environmental and societal values (Fuenfschilling and Truffer, 2014).

The term ‘configurations that work’ (Markard, 2011) is often used to refer to configurations that have been successfully aligned and institutionalized to a degree so that they have become a legitimate and efficient way of providing a societal function in contexts ranging from the local to the global (Miörner and Binz, 2021). E.g. car-based transport, industrial agriculture, centralized urban water management. In most sectors, one or several alternative, more emergent, configurations exist that allow fulfilling the same function, which are however less deeply standardized, more expensive, etc., but also promising substantial gains in terms of social/environmental sustainability.

For example, in the energy sector there is a well-established regime configuration of actors, institutions and technologies that generates ‘bulk energy’ in centralized (coal, gas, nuclear, wind power, etc.) plants and distributes it in expansive grid-based infrastructure systems (Geels and Turnheim, 2022). With this ‘bulk energy’ logic, some regions are hosting energy production plants and may capture some economic rents through increased tax returns from the power producers. The transformative capacity of this logic for regional economies remains rather limited, though. In parallel, an alternative configuration increasingly emerges around decentralized energy generation using solar photovoltaic technology, local storage and power management systems and distribution in small-scale energy grids, which are managed by local micro-utilities (*ibid.*). This configuration is associated with a lower environmental footprint, a valuation frame that emphasizes energy self-sufficiency of local communities with variegated economic and social co-benefits in regions, especially in low- and middle-income contexts.

Shifts in valuation will also imply shifts in technologies, business models, policy frameworks, financing mechanisms, and consumer behaviour. Those values in turn relate directly to the regional valuation context, which has a strong impact on regional industrial dynamics.

The key question is thus through what mechanisms, and under what conditions, regional industry dynamics can be re-oriented in ways that create more substantive social, economic and environmental local co-benefits. To answer this question, we need to understand the interplay between industry dynamics, the valuation processes that guide firm-level innovation activities and assign values to solutions and practices, and how valuation contexts change over time.

2.3 Valuation dynamics

Recent studies in economic geography and transition studies have started to pay attention to valuation dynamics and their territorial dimension. Jeannerat and Kebir (2016) elaborate on how knowledge is valued in a market through the co-evolution of a production and consumption system – extending the production-focused or technology-driven perspective in territorial innovation

models. Binz and Truffer (2017) consider how the interplay of knowledge-based innovation and valuation processes jointly define the conditions of innovation success. Carvalho and Van Winden (2018) explore how ‘intangibles’ inherent in products and services are central to how they are valued, how the economic significance of innovations derive from other factors than from superior “functionality”. Yap et al. (2022) show that latecomer countries can proactively drive local valuation processes to achieve transformative leapfrogging. Finally, Huguenin and Jeannerat (2017) argue that value creation should not be considered the result, or being a ‘by-product’, of innovation, but that it is inherently and deeply linked to societal and market valuation processes.

Valuation processes are often territorially defined and guide the innovation activities of actors in non-trivial ways. Valuation dynamics refers to “the relational and institutional dynamics by which different objects and activities are socially valorized (i.e. transformed and commercialized) and evaluated (interpreted, recognized, legitimated and appraised) in the market” (Jeannerat and Kebir, 2016, p. 277). They include a process of “qualification” governed by values and social conventions of “quality” which actors refer to in their manifold activities (Beckert, 2009; Carvalho and van Winden, 2018)

In other words, demand for innovations is not exogenously given but socially constructed through dynamic dialogue between supply and demand in two dimensions of value creation: Markets (producers-consumers) and Society (industry-society).

Market valuation refers to the ways through which innovations generate economic value for the region and its actors, realized through the production of new goods or services that are sold in local or external markets. Market valuation is a key driver of the diffusion and upscaling of new innovations under conditions of uncertainty. In the absence of clear routines and pricing mechanisms, valuation challenges have to be overcome in order to legitimize new actors and/or products and services (Huguenin and Jeannerat, 2017).

Societal valuation on the other hand, refers to the ways through which innovations generate value that has broader societal utility (Huguenin and Jeannerat, 2017). Societal valuation is closely associated with a continuous process of problem-setting and -solving, often reflecting context-specific challenges. Societal values create meaning by changing the ways through which actors see and engage with the social world. They provide rational justification for activities and goals, and are thus reflected in guiding visions and directionality (Huguenin and Jeannerat, 2017; Weber and Rohrer, 2012).

Industry dynamics can be considered the fundamental inception ground for developing technological innovation and new services. Their value on markets and in society are however not predetermined. Instead, economic and societal values related to the locally developed innovation have to be created/actuated

through valuation dynamics in markets and in society (Huguenin and Jeannerat, 2017). Thus, achieving sustainable development outcomes accordingly depends not only on the functional dimension of innovation and the generation of economic values, but on the ability to develop aligned technological and institutional structures that reflect and non-economic values and guide regional actors (ideally) towards more sustainable activities and goals.

2.4 Analytical framework: regional valuation contexts and industry dynamics

Valuation dynamics are reflected in the way through which actors, technologies and institutions are organized in configurations that provide essential societal services such as water, energy and transportation, and thus shape the outcomes in terms of economic, societal and environmental impacts.

To conceptualize the link between regional industry dynamics and sustainable outcomes, we expand on the notion of 'territorial valuation' introduced by Jeannerat and Crevoisier (2022). More specifically, we define the regional 'valuation context' for industry dynamics as territorially defined institutional set-ups that guide valuation processes in the market- and societal dimensions respectively.

Values are socially constructed and actors draw on them to mobilize support for their activities, to rationalize and legitimize, their actions. Actors can subscribe only to a finite number of alternative combinations of value positions (Truffer, 2021). The valuation context typically favours specific technological solutions, business models, forms of governance, etc. that become reflected in the resulting socio-technical configurations that are developed in the region. I.e. if the guiding values shift, technologies and infrastructure solutions usually have to adapt. In other words, valuation alters the directionality / degrees of institutionalization of different socio-technical configurations, which has direct impacts on the types of innovation pursued by local industrial actors. Core valuation dynamics (market- and societal valuation) become reflected in regional structures through institutionalization processes. For example through quality conventions, product categories, modes of cooperation and competition, technology standards, labels, routines, regulations, norms, visions, strategies, actor roles, accountability principles, market segments, etc.

The analytical framework is illustrated in Figure 1. The fundamental proposition of our framework is that sustainable regional development outcomes of industrial path development hinge on the alignment of industrial activities with the region's most pressing economic, societal, and environmental challenges. Central to achieving this alignment are value creation and negotiation processes within the regional valuation context.

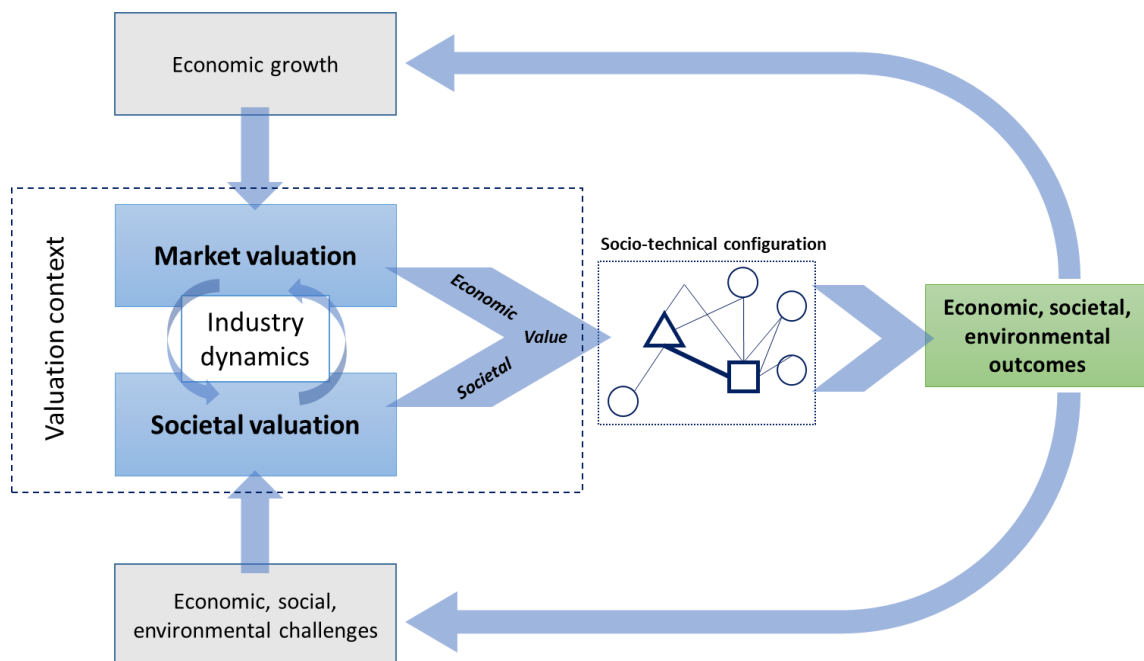
The valuation context shapes the incentives for industrial actors to align their

activities with key local challenges. It cannot be expected that this happens only through market valuation, but requires structures that assign societal values to sustainable practices. If the industrial path is largely detached from societal valuation dynamics, such as catering primarily to international export markets or being driven solely by economic growth, positive spillovers for regional economic growth may occur, but critical challenges may remain unaddressed.

The most potent driver of regional sustainable outcomes is thus the alignment and reinforcement of market and societal valuation dynamics. This occurs when the market valuation process takes into account and creates economic incentives to develop societal and environmental values.

Such alignment can be brought about through intentional interventions by key actors or endogenous or exogenous shocks. Actors coordinate their actions to influence and express attachment or detachment to different value frames (Callon et al., 2002; Jeannerat and Kebir, 2016). Activities undertaken by end-users, state actors, and intermediaries can lead these actors to create, enable, motivate, mediate, or legitimize attachment or detachment from market and societal valuation processes. Additionally, key inflection points, events, or crises, such as financial crises, environmental disasters, or rapidly changing landscape conditions, can lead to shifts in core valuation dynamics.

Figure 1: Analytical framework (own elaboration)



3. Decentralized wastewater treatment and reuse systems in Bengaluru

Bengaluru is the capital of Karnataka and is often referred to as the "Indian Silicon Valley" due to its significant ICT sector. The city is situated on the Deccan plateau, 900 meters above sea level, and is renowned for its enjoyable

climate throughout the year. As with many other South Asian cities, Bengaluru has witnessed tremendous population growth in recent decades. In 2015, the population of Bengaluru was estimated to be around 10.6 million (the third largest city in India), a rise of about 4.5 million residents compared to the year 2000 (Statista, 2021).

The rapid urban growth of Bengaluru has created and continues to create significant infrastructure challenges, leading to inefficiencies in delivering essential services such as urban mobility, electricity, water, and sanitation. Staggering population growth and rainfall seasonality have led to the continuous depletion of local surface water sources and groundwater reserves (Britannica, 2022; Drangert and Sharatchandra, 2017; Kuttuva et al., 2018). Furthermore, the city has grown without a comprehensive long-term development plan for wastewater disposal and treatment.

The Urban Water Management (UWM) sector in Bengaluru presents a unique case to examine the connection between industry dynamics and regional development outcomes. Over the last 20 years, the sector has undergone a significant transition, resulting in a vigorous UWM industry and in a new socio-technical configuration centred around decentralized wastewater treatment and reuse systems. Policies have been adopted that promote decentralized water and wastewater technologies in new building projects, essentially delegating responsibility for wastewater treatment and disposal to private actors and residents instead of expanding the centralized sewage network.

Consequently, Bengaluru is an 'extreme' case (Flyvbjerg, 2006) by being among the few cities worldwide that have successfully shifted from centralized to decentralized wastewater treatment systems on a large scale. It is a rare example of a booming megacity where 100% on-site wastewater treatment and reuse is officially mandated for residential complexes and commercial buildings over a certain size. It is estimated that over 3000 small-scale wastewater treatment plants (SSTPs) cumulatively treat 10-20% of the city's total generated wastewater (Klinger et al., 2020; Nath et al., 2022). This transition has occurred largely without strong steering from the government, relying instead on a series of regulations and mandates that served as triggers for market valuation dynamics. As a result, many transition challenges have been addressed by the industry, civil society, NGOs, and other organizations, with varying degrees of success and efficacy over time.

The empirical analysis presented in this study is grounded in a comprehensive investigation based on different sources of data. An extensive review of secondary sources was conducted, which included academic papers, grey literature, newspaper articles, and official documents such as mandates, government memos, and minutes. In order to gather first-hand information, interviews were conducted with 38 stakeholders in the UWM field in

Bengaluru, encompassing firms, policymakers, intermediaries, experts, and residents (see Appendix 1 for a list of interviews). The duration of each interview ranged from 20 to 120 min (average 55 min). Moreover, the research was enriched by 19 site visits to de-centralised wastewater treatment and reuse systems in Bengaluru. To further validate the data collected through interviews and secondary sources, the authors also participated in two workshops organised by ATREE, a local research institute. These workshops provided opportunities to engage in discussions on relevant issues with a broad range of stakeholders. The entire empirical data collection process spanned a total of 10 weeks in Bengaluru, taking place during March-April and September 2022.

In the following sub-sections, we analyse the UWM transition in Bengaluru during three phases of development, focusing on shifts in the valuation context and changing incentives for industrial actors due to (the interplay of) market- and societal valuation dynamics. For each phase, we describe the main activities in the UWM sector, analyse valuation context and how this shapes the particular set-up of actors, institutions and technologies that provide UWM services to residents, through market- and societal valuation dynamics that steer and incentivise activities among industry actors. Furthermore, we analyse the outcomes that this configurational set-up give rise to, and how these relate to broader goals of industrial growth vis-à-vis regional sustainability challenges.

3.1. Phase 1: Addressing wastewater treatment during rapid urban expansion

In 2004, the Karnataka government initiated a novel mandate requiring the installation of small-scale wastewater treatment plants (SSTPs) in all new residential and commercial buildings above a specific size. This marked a significant shift in urban sanitation governance, moving from state-led expansion of centralized sewage networks to a hands-off approach that made private firms and residents accountable for building, operating, and maintaining on-site SSTPs and ensuring satisfactory treatment quality.

To facilitate this new approach, the government introduced a permitting pathway for real estate developers planning and constructing new buildings. This process outlined the necessity of an SSTP before beginning construction, known as "Consent to Establish," and incorporated an inspection of the plant before use, referred to as "Consent to Operate." Additionally, once operations began, state authorities mandated regular water sampling and laboratory testing by the SSTP operator. Upon completion of a new building, the real estate developer typically transfers the responsibility for SSTP operation and maintenance (O&M) to a Resident Welfare Association (RWA). RWAs assume legal ownership of residential properties after construction and sale, and are self-governed by residents through voluntary efforts.

The need for policy intervention and a departure from the status quo in UWM

emerged from two directions. Firstly, the rapid rate of urban expansion exerted immense pressure on the existing UWM approach. Expanding the centralized system was deemed infeasible, considering its current condition and the swift growth of new apartment complexes on the city's outskirts. Several interviewees emphasized the issues with the conventional UWM approach as a significant catalyst for change, driven by the desire to overcome constraints imposed by the need for proper wastewater treatment in newly developed areas.

Secondly, an ongoing policy reorientation at the national level, beginning in the early 2000s, served as a stimulus for exploring decentralized UWM at the state and city levels. This shift was motivated by the negative environmental impacts of inadequate wastewater treatment. National policies, such as the Environmental Impact Assessment Act, began to explicitly recognize SSTPs as a potential means to reduce pollution and enhance environmental protection.

3.1.1.Valuation context and industry dynamics

If zooming in on how the initial mandate and permitting pathway shaped the valuation context for industrial actors and their activities in Bengaluru during the first phase, the case analysis provides several relevant insights.

First and foremost, the analysis reveals that the rationale for 'urban expansion' was driving the societal valuation of SSTPs in the first phase. Policies were enacted to facilitate Bengaluru's rapid growth by ensuring wastewater treatment was provided by private firms and residents at the building-level. However, we could not identify any institutional structures that addressed other societal challenges, even though the prevailing public discourse emphasized environmental benefits.

If turning to market valuation, the mandate and 'hands-off' approach spurred the rapid growth of a SSTP market and a new industrial path in Bengaluru. The market was fuelled by economic incentives and a local 'water hype,' attracting firms with diverse backgrounds. The demand side consisted of real estate developers who were legally obligated to install SSTPs, and RWAs, who were responsible for operation and maintenance after handover. Our analysis shows that an industrial path emerged and was quickly inundated with technology vendors, SSTP designers, consultants, engineering firms, O&M service providers, and laboratories.

However, the early SSTP market lacked institutional structures that incentivized firms to provide well-functioning plants. Real estate developers were more focused on fulfilling the mandate than providing proper wastewater treatment in new buildings. There were no standards, certification schemes, labels, or categories to guide technology choices or identify high-quality products and services. Furthermore, few enforcement mechanisms existed to

ensure long-term functionality of the SSTPs.

Firms mainly competed on price, with little emphasis on quality. The regional valuation context discouraged innovation and favoured rent-seeking behaviour. A lack of accountability along the value chain led to issues, such as the existence of 'nightflyers' (Mulder, 2020), who exploited the absence of long-term accountability and, in collaboration with questionable consultants, provided poorly designed or unsuitable SSTPs. Additionally, the market structures did not foster knowledge accumulation or skill enhancement. Most firms in the industry were either unskilled or relied on technologies developed for large-scale systems, which were often unsuitable for smaller plants. This resulted in serious consequences for RWAs' ability to operate and maintain plants after handover.

3.1.2. Socio-technical configuration and outcomes

Over the 10 years that followed after the introduction of the mandate, more than 2500 SSTPs were installed in Bengaluru, showing the power of the market when it comes to diffusing a divergent solution when market- and societal valuation dynamics are aligned. However, studies have shown that a vast majority (up to 85%) of the plants did not work properly (Klinger et al., 2020; Nath et al., 2022). Due to poor technology choice and design, many plants were immensely difficult and expensive to operate and maintain, and RWAs often lacked any incentive to operate the plants properly due to a lack of enforcement. For example, there are several examples of how RWAs simply have bypassed or turned off their plants to save energy and operation costs (Klinger et al., 2020). Any actor could call themselves an SSTP vendor and since contracts for new projects were awarded through commission rackets involving consultants, or through personal relationships, and not by degree of technical expertise, there were few incentives for vendors to develop new systems that were easier to operate and maintain.

The observed industry dynamics in the first phase clearly thus shows the potential of well- aligned societal- and market valuation dynamics. The mandate was introduced with the underlying rationale of enabling urban expansion, by reducing the pressure on the state to expand the centralized sewerage network. As such, it targeted a narrowly defined local challenge. Market valuation dynamics subsequently incentivised the creation of an SSTP industry in Bengaluru which quickly came up with solutions to make new buildings fulfil the mandate and continue expanding the city.

However, it also shows the limitations of a market-focused, and laissez faire, approach in relation to fostering broad sustainable regional development outcomes. The socio-technical configuration that emerged as a result from the industry dynamics in the first phase was one of cheap and poorly designed SSTPs, and a huge variety of technologies and designs. Actor roles and

responsibilities were not institutionalized, and no structures were in place to support actors at different stages of the value chain to make informed decisions.

Overall, this led to a socio-technical configuration optimized primarily to fulfil the mandate and enable a continued urban expansion. In other words, the valuation context heavily shaped the socio-technical configuration through the value incentives provided to industry actors. Over the decade that followed, vast amounts of partly- or untreated wastewater were released into storm drains, ending up in streams and water bodies, as a result of poorly functioning SSTPs. This had tremendous environmental impacts, which will be discussed in the next section.

3.2.Phase 2 – Addressing environmental pollution

Around 2012, the UWM sector in Bengaluru reached a turning point due to escalating environmental issues and a rising sense of urgency among citizens to address the situation. There was a growing realization among various actors that the current largely unregulated approach failed to provide effective solutions to local challenges.

The media and discussions within and between RWAs started highlighting issues with SSTPs. Many of our interviewees noted that the failures of these systems raised awareness among residents, who were previously largely unaware of their SSTP's existence. They began to recognize that a malfunctioning plant could lead to problems such as unpleasant odours, noise pollution, in addition to fines for non-compliance with regulations. Consequently, the plants started to become a high-priority agenda item and a source of annoyance for residents.

Moreover, the health of Bengaluru's renowned lakes gained attention, with problems manifesting as visually striking and disastrous events, such as local lakes foaming and catching fire (see National Geographic, 2018). These incidents received extensive coverage in local, national, and international media, prompting increased efforts from civil society, industry, and government to improve existing water and sanitation solutions or explore alternative options. Additionally, as Bengaluru faced increasing water scarcity, the city was pressured to find alternatives to relying on the Cauvery River, its main freshwater source, and private bore wells for water supply. Reports revealed that the city had reached its maximum withdrawal limit from the Cauvery River (Kuttuva et al., 2018). This sense of urgency led to the introduction of new mandates, enforcement routines, and guidelines by city and state authorities during the second phase.

First, more stringent wastewater treatment standards were introduced at the state level. Although criticized as potentially too high for SSTPs, making treatment difficult and costly (Evans et al., 2014; Schellenberg et al., 2020), these

new standards were an important signal of a commitment to address environmental pollution. Furthermore, in 2014, the state introduced the Zero Liquid Discharge (ZLD) mandate (KSPCB, 2014), prohibiting the discharge of wastewater (treated or untreated) into lakes and rivers, in response to the severe pollution of Bengaluru's lakes.

Second, at the national level, a directive was issued requiring all states to reuse treated wastewater for non-potable purposes, such as toilet flushing, cleaning, gardening, and industrial processes. New buildings were to be equipped with dual piping to facilitate the use of treated wastewater for toilet flushing. This directive was echoed in a state-level requirement, mandating the installation and use of SSTPs in both new and existing buildings above a certain size. This marked the first time the government called for the installation of SSTPs in already-constructed buildings.

3.2.1 Valuation context and industry dynamics

During the second phase, societal valuation dynamics expanded to address a wider range of local challenges. This shift led to the emergence of new core values, which were incorporated into the policies and requirements introduced. The primary focus was on "environmental sustainability," while "water security" also gained importance, albeit to a lesser degree.

The UWM sector increasingly came to be seen as a major contributor to the city's poor environmental conditions, and it became widely accepted that the sector should be part of the solution. Although framed as water-saving efforts, the ZLD and retrofit mandates primarily targeted the release of untreated wastewater into urban water bodies. They recognized that even treated water could be contaminated, and thus, the existence of an SSTP did not guarantee proper wastewater treatment.

This broadening of societal valuation dynamics also solidified the UWM sector's future direction in the city, in terms of a combination of centralized and decentralized approaches. By acknowledging the potential environmental benefits of decentralized systems, this was no longer seen as mere quick fixes for urban expansion. Instead, they became integral components of a more sustainable and secure urban water management strategy.

However, despite the reflection of a broader set of local challenges in new mandates and requirements, Bengaluru's institutional framework failed to offer detailed guidance on fulfilling these requirements, such as specifying approved solutions or assigning responsibilities beyond those of RWAs and real estate developers. Our interviews indicate that the regulatory environment for water reuse remained somewhat of a vacuum, with the absence of substantial policies beyond the ZLD mandate. In essence, while the valuation context shifted towards a more comprehensive set of societal valuation dynamics, it remained

institutionalized only to a low degree.

When it comes to market valuation dynamics, they shifted in the second phase as new demand-driven incentives emerged, influencing the behaviour of actors in the industrial path. On one hand, new sophisticated demand extended beyond merely fulfilling the original mandate, and was partly driven by environmental concerns, such as public outcry over lakes catching fire, and new legal requirements. Additionally, the inconvenience of a malfunctioning SSTP for residents who could not simply release water into storm drains contributed to this demand. As treated wastewater was mandated for toilet flushing, residents became more psychologically invested in ensuring proper treatment and O&M of SSTPs. Consequently, the more sophisticated demand from residents led high-end real estate developers to increasingly focus on SSTP issues.

On the other hand, this emerging demand did not result in the institutionalization of support mechanisms like labels, training programs, or certification schemes to help customers evaluate and choose SSTP vendors or technologies. As a result, both real estate developers and residents were left with insufficient information to address the issue with malfunctioning plants. In addition, many RWAs continued to face high operating expenditures for wastewater treatment, along with the investment required to retrofit buildings with SSTPs and dual piping for flushing to meet legal requirements. The need for new capital investments and increasing operating expenses took a long time to recover through water savings, leaving many RWAs without financial incentives to properly operate their SSTPs. In particular, RWAs with access to cheap municipal water had fewer incentives to spend additional money to meet the new requirements. In contrast, apartment buildings without piped water connections and unreliable private borewells, which relied on expensive tanker water during summer months, had greater incentives to make the necessary investments.

Our analysis shows that due to the low degree of institutionalization of societal- and market valuation dynamics, only a limited number of industrial actors were incentivised to innovate. These were primarily actors that were able to identify market potentials emerging from the emerging sophisticated demand by niche customers, primarily high-end real estate developers and particularly knowledgeable or environmentally aware RWAs. In that sense, their innovative activities do reflect the interrelatedness of societal- and market valuation dynamics, even though the effect was mostly seen at the level of individual firms and their consumers, rather than at the level of the industrial path.

First, a limited number of companies developed innovative business models to capitalize on market opportunities arising from the intersection of malfunctioning plants and increasing sophisticated demand. These firms

concentrated their efforts on SSTP rehabilitation, upgrading, and retrofitting. In addition, they began providing O&M services through qualified personnel and offering more advanced monitoring services. This approach addresses a fundamental gap in the SSTP market inherited from the first phase, where companies exclusively focused on the design and construction of the plants, leaving the operation to RWAs.

Second, we can identify some technological innovation that focuses on market opportunities resulting from the failure of existing plants and the dissatisfaction of real estate developers with conventional technological solutions. Examples include 'nature-based' or 'hybrid' system designs, which incorporate anaerobic or combinations of aerobic and anaerobic technologies in mid- to top-tier residential segments. Marketed as low-maintenance systems, these designs promise to reduce maintenance costs by over 90% compared to conventional SSTPs. Furthermore, some firms are targeting the demand for high-quality water for reuse among residents. They install a combination of micro- and ultrafiltration, or even reverse osmosis (RO) membranes, which in theory treat water to near-potable quality.

Third, we nevertheless observed some broader changes in the industry dynamics concerning the profile of SSTP vendors. These changes were the result of a market selection and retention process that gradually unfolded since the initial mandate, rather than deliberate institutionalization efforts by key stakeholders. More specifically, there was a growth in the number of high-end providers, stemming from both accumulated experience and the entry of new, highly skilled participants. This included entrepreneurs who left positions at multinational companies, such as Thermax, ION Exchange, and VA Tech Wabag, to establish new firms in the SSTP sector. Additionally, acquiring technology through a memorandum of understanding and then selling it under a company's own brand became a common practice. This contributed to the elimination of some bottom-end 'nightflyers' from the market.

3.2.2 Socio-technical configuration and outcomes

The evolution of the socio-technical configuration surrounding decentralized UWM in Bengaluru was influenced, at least temporarily, by a greater emphasis on societal values. Specifically, the configuration aimed to address not only the challenges of urban expansion but also environmental sustainability and water scarcity. During the second phase, the health of Bengaluru's lakes improved, partially due to the implementation of more stringent treatment requirements and new policies reflecting local challenges. A subset of the SSTP industry contributed to meeting these challenges by experimenting with solutions aligned with new societal value propositions, despite unclear market valuation dynamics.

In other words, the development of market structures providing value for new

solutions lagged behind. There were limited financial incentives for most RWAs to ensure proper SSTP operation and maintenance, and even fewer for efficient water reuse. Although there were calls for water reuse in the industry, no mechanisms existed to match the supply of treated wastewater with demand. Consequently, the ZLD mandate led to a paradoxical situation where RWAs were incentivized to ensure proper treatment but were uncertain about how to utilize all the treated wastewater. Even with maximum on-site reuse, only 30-70% of the treated wastewater could be utilized, leading to overwatering of gardens or illegal release of water into storm drains in some cases. When the water is properly treated, this situation results in a massive waste of water. If treatment is not up to standard, it contributes to environmental pollution.

These outcomes can be traced to the fact that a *laissez faire* approach still dominated in terms of market valuation dynamics, and had been extended to the societal valuation dimension in the second phase. On one hand, new policies were enacted to ensure that the socio-technical configuration was oriented towards a broader set of societal challenges. For industrial actors that could find ways of capitalising on this change in the valuation context, this provided a strong impetus to engage in more experimental activities and develop new technological or business model innovations. On the other hand, how to reach this goal was still very much up to the industry steered by demand-side dynamics. The observed potentials in the market originated primarily from a demand-impulse for better quality SSTP solutions and services. The valuation context, as defined in the analytical framework, was still under-developed, in terms of providing very little guidance in terms of *how* to reach the societal goals.

Overall, this led to a limited alignment of industrial activities with the new societal challenges of environmental sustainability and water security. Even though the situation improved considerably during the second phase, due to the aforementioned activities at the higher end of the SSTP industry spectrum, and the crowding out of 'nightflyers' at the lower end, much of the mid-tier of the SSTP industry went on with 'business as usual'.

3.3 Phase 3: Addressing urban water resilience

In the most recent phase, which is arguably still ongoing, a striking feature is the widespread acceptance and integration of SSTPs as a vital component of Bengaluru's water and sanitation system. This shift is particularly evident at the government level, where senior officials are increasingly recognizing the potential of on-site reuse in discussions with stakeholders and in public discourse. The public discourse has also reoriented, shifting focus from the problems and challenges associated with SSTPs to the opportunities and potential benefits of decentralized sanitation and self-sufficiency at the apartment complex level.

In terms of the regulatory environment, KSPCB started to explicitly target the lacking formalization of the UWM sector and issued a memorandum specifying a list of approved SSTP technologies, details on plant operation, specifications for mechanical equipment, advice on the location of SSTP within apartment complexes, standards of the treated wastewater, and prescribes the installation of sensors for measuring various indicators. However, the list of approved technologies would, according to several of our interviewees, have limited innovation in the sector, leading to push-back from some industry actors. Consequently, the list was abandoned in favour of a technology-neutral formulation.

At the national level, the extension of two major policy programs (AMRUT and the Swachh Bharat mission) explicitly envisioned treated wastewater to be an important water source for Indian cities in the coming decade. This permeated through the state- and city levels, with authorities having publicly sanctioned a broader set of reuse applications, for example using treated wastewater in the construction sector. Although no formal regulations govern this reuse application, the authorities have clearly communicated their intent to promote increased reuse in the industry and further explore safety aspects. With regards to the latter, a report from the Indian Institute of Sciences confirmed the safety of using treated wastewater for mixing concrete.

The most profound changes in the third phase however relates to the emergent activities by a broader set of stakeholders than government actors, targeting institutional change in a broader sense than during previous phases. Our findings support the notion of an emerging self-organization of industry actors, new and existing intermediaries and community actors, around issues and value propositions in the UWM sector. Even to some effects are yet to be seen, several interviewees highlighted how these activities started to have an impact on the valuation context in the third phase.

3.3.1 Valuation context and industry dynamics

From being characterised by a layering of societal challenges that the UWM sector should seek to align to, the third phase included a stronger integration of different societal challenges on one hand, and between societal and market valuation dynamics on the other. This was largely driven by an emerging self-organisation among industry actors, new and existing intermediaries, and community actors. Their activities targeted failures in the alignment between industry and societal goals. As such, they focus directly on market valuation dynamics (by targeting market failures) and on societal valuation dynamics (by forging links between societal goals), and on the interplay between market- and societal valuation.

In terms of societal valuation dynamics, actors started to realise that the broader challenge of “water security” encompasses both previous value orientations

already existing in the UWM sector (environmental sustainability and urban expansion). In order to increase the reuse of treated wastewater, it is adamant that the other two societal challenges are also targeted (the installation of SSTPs in apartment complexes, which align with the challenge of urban expansion, and to ensure proper treatment and fulfil the ZLD mandate, which align with the challenge of environmental sustainability). As such, we observe a strong reorientation of societal valuation dynamics towards promoting urban water resilience and water security as the overall objective or actors' activities.

Similarly, the market valuation dynamics reflected in the activities by public actors show signs of formalization, more clearly pointing towards 'quality' as an important factor that determines the relationship between supply and demand in the sector, rather than only price. This is reflected by the KSPCB guidelines which provides a more detailed specification of the technological requirements of SSTPs and their monitoring. It is also reflected in a further sophistication of the demand-side, with RWAs pushing towards, and experimenting with, expanded reuse within apartment complexes, sometimes up to potable quality. Taken together, to be competitive on the SSTP market, firms can no longer compete by providing the cheapest possible way of fulfilling the mandate, but need to take into account the demand (from residents through real estate developers) for well-functioning, affordable and easy-to-maintain systems.

A distinct difference from previous phases is that the reorientation was supported by concrete activities that targeted the institutionalization of supporting mechanisms in the valuation context, beyond open-ended changes to regulatory framework conditions. For example, RWAs started to self-organise platforms to exchange knowledge and experiences on SSTP operation and maintenance. This took place through informal exchanges and personal contacts, but also in using the Bangalore Apartment Federation (BAF) as a platform for interaction. From the industry-side, firms initiated training programs for operators of their plants and the technologies they employ, and arranged courses directed to RWAs and real estate developers. There have also been discussions about O&M firms about voluntary certification schemes, and several firms have started to get engaged in public outreach activities and education efforts. What these activities have in common is that they seek to effectively link societal valuation- with market valuation dynamics, by highlighting how properly designed, operated and maintained SSTPs may provide financial benefits that extend far beyond avoiding fines for non-compliance.

This constitutes a major change to the valuation context which can be seen to have an impact on industry dynamics. The incremental improvements to technology and business models that begun in the second phase continued, with SSTP vendors increasingly considering treatment quality as an important

factor, a continued upgrading of skills, and an improvement in O&M capabilities. However, the changed valuation context also triggered some more radical novel solutions that was made available on the market by local firms.

For example, one model that emerged was for SSTP vendors to offer RWAs to take over existing malfunctioning treatment plants, rehabilitate them and operate them for a fee, under the condition that they own the produced water which could then be sold back to the RWAs. This model creates societal value in terms of creating incentives for reuse, while at the same time exploiting market opportunities that emerge from water scarcity.

Another prominent innovation in the same domain has been a high-quality container-sized add-on that is attached to an existing SSTP in order to further treat the wastewater to a quality that corresponds to the requirements for potable reuse. This add-on enables a considerably broader range of reuse applications, even though it requires a well-maintained SSTP as a starting point. The firm that has developed this solution argues that it allows for a quick return on investment if the client does not have municipal water supply, indicating a strong alignment between the societal value of water security and market valuation dynamics.

Finally, a third example is a platform (mobile app) for matching supply of treated wastewater with demand in, for example, the construction industry. The idea is to identify well-maintained SSTPs and let water users order tankers that source the water from a nearby treatment plant. To date, the entrepreneur is supplying water mainly from plants that they have designed themselves, to large construction projects nearby, but in the future the idea is to open up the platform to a broader range of self-managed SSTPs.

3.3.2. Socio-technical configuration and outlook

One main difference compared to the first two phases and the third one is the development of feedback mechanisms between the end-users of the emerging socio-technical configuration (residents) and the technology providers (SSTP vendors). Previously, vendors had continued over decades to sell plants that often failed and were not affected by the problems faces by RWAs in the O&M of their plants. In the third phase, many builders started to put pressure on SSTP vendors to take responsibility for their systems and, maybe more importantly, themselves took a larger interest in the question of SSTP design in their construction projects.

In combination with the layering of additional regulations, behaviour change on the demand- side, and the entrance of new actors with innovative technologies and business models, this holds the promise of giving rise to a contextually adaptable socio-technical configuration that reflect both market- and societal values. The key for long term success, however, is not technological

innovation, but lies in the development of a valuation context that reinforces the relationship between market- and societal valuation dynamics. If projecting the current tangent into the future, one scenario is that treated wastewater becomes a valued resource also in market terms, and that this would lead RWAs to 1) ensure proper O&M so that the inherent value in the wastewater can be captured by residents, and 2) resist centralized sewerage connections even if provided, since this would take away an income stream for the RWA. Both of these however rely on a well-functioning on- and off-site reuse market. An inherent risk in the current scenario is that the market for reused water contracts when construction moves further to the outskirts of the city and leaves well-functioning SSTPs underutilized in terms of reuse capacity.

Nevertheless, if RWAs and industries can directly harness the value of controlling a stable water source by reducing reliance on tanker water or ever-deeper bore wells, this could serve as a powerful trigger for exploiting the potentially unique situation with a large number of SSTPs already constructed in Bengaluru.

4. Discussion/ Conclusion

Scholarly research in innovation studies and economic geography has been increasingly focusing on "green" or "sustainable" regional restructuring. Despite this, there has been a limited understanding of the link between industry dynamics and regional development outcomes. Studies have generally focused on the industrial dynamics and regional system reconfiguration associated with green industry emergence, but often neglect in-depth exploration of their impacts on social and environmental sustainability.

In response to this, this paper highlights the importance of 'valuation' processes in steering regional industrial path development towards specific values, goals, and rationales. Valuation involves the institutionalization of elements that shape recognition and engagement, and it includes both 'market valuation' and 'societal valuation.' By studying valuation contexts, the paper aims to shed light on how industry dynamics lead to certain sustainable outcomes.

An analytical framework has been developed to understand such valuation contexts, focusing on the interplay between industry dynamics, market valuation, and societal valuation. This framework was used to analyze an empirical case study of the Urban Water Management (UWM) sector in Bengaluru, India, showcasing a 'green' path development trajectory within a changing valuation context. The study identifies shifts in societal and market valuation dynamics as key factors in driving more sustainable path development.

The case analysis shows how the UWM transition and associated industry dynamics in Bengaluru progressed from a 'pure' form of industrial path

development, focusing on mainly the expansion of the industry based on the generation of economic values through market transactions, to a regional industry that is more aligned with local challenges and the generation of societal values. This evolution has been largely driven by changes in the valuation dynamics and context.

In the initial phase, the focus was predominantly on market forces and market valuation, adhering to a *laissez-faire* approach. The primary role of policymakers was to establish an overall framework, leaving the details and execution to the industry. The second phase saw a shift towards societal valuation, while maintaining the *laissez-faire* approach. The overall framework was still set by policymakers, but the industry was left to determine the specifics. Here, the valuation context defined the goals, although the means to reach them were not provided. This phase was characterized by a lack of efficient links between societal and market valuation.

In the third and most recent phase, we observe a more mature valuation context. Societal values have become well-established and are increasingly backed by institutional setups within the valuation context. This phase highlights an explicit focus on bridging societal and market valuation dynamics.

To be developed.

Appendix 1: List of interviews and site visits

Type of actor	No.of interviews
Firms focusing on water reuse	5
Industrial water users	2
Policy actors and experts	3
Real estate developers and architects	7
Resident Welfare Association representatives and staff	7
SSTP vendors and operators (alternative)	7
SSTP vendors and operators (conventional)	4
UWM experts	3
	38
Site visits (conventional systems)	9
Site visits (alternative systems)	7
Site visits (reuse systems)	3
	19

References

- Andersen, A. D., Steen, M., Mäkitie, T., Hanson, J., Thune, T. M. and Soppe, B. (2020) 'The role of inter-sectoral dynamics in sustainability transitions: A comment on the transitions research agenda', *Environmental Innovation and Societal Transitions*, 34, pp. 348-351.
- Binz, C. and Truffer, B. (2017) 'Global Innovation Systems—A conceptual framework for innovation dynamics in transnational contexts', *Research Policy*, 46(7), pp. 1284- 1298.
- Binz, C., Truffer, B. and Coenen, L. (2016) 'Path Creation as a Process of Resource Alignment and Anchoring: Industry Formation for On-Site Water Recycling in Beijing', *Economic Geography*, 92(2), pp. 172-200.
- Blažek, J., Květoň, V., Baumgartinger-Seiringer, S. and Trippel, M. (2020) 'The dark side of regional industrial path development: towards a typology of trajectories of decline', *European Planning Studies*, 28(8), pp. 1455-1473.
- Boschma, R. (2017) 'Relatedness as driver of regional diversification: a research agenda', *Regional Studies*, 51(3), pp. 351-364.
- Breul, M., Hulke, C. and Kalvelage, L. (2021) 'Path Formation and Reformation: Studying the Variegated Consequences of Path Creation for Regional Development', *Economic Geography*, 97(3), pp. 213-234.
- Britannica (2022) *Bengaluru India*. Available at: <https://www.britannica.com/place/Bangalore-India> (Accessed: 22 December 2022).
- Callon, M., Méadel, C. and Rabearisoa, V. (2002) 'The economy of qualities', *Economy and society*, 31(2), pp. 194-217.
- Capasso, M., Hansen, T., Heiberg, J., Klitkou, A. and Steen, M. (2019) 'Green growth – A synthesis of scientific findings', *Technological Forecasting and Social Change*, 146, pp. 390-402.
- Carvalho, L. and Van Winden, W. (2018) 'Making waves: the valuation of innovations in San Sebastian's surf economy', *European Planning Studies*, 26(1), pp. 75-93.
- Coenen, L. and Morgan, K. (2020) 'Evolving geographies of innovation: existing paradigms, critiques and possible alternatives', *Norsk Geografisk Tidsskrift-Norwegian Journal of Geography*, 74(1), pp. 13-24.
- Cooke, P. (2012) 'Transversality and Transition: Green Innovation and New Regional Path Creation', *European Planning Studies*, 20(5), pp. 817-834.
- Dawley, S. (2014) 'Creating new paths? Offshore wind, policy activism, and peripheral region development', *Economic Geography*, 90(1), pp. 91-112.
- Drangert, J.-O. and Sharatchandra, H. (2017) 'Addressing urban water scarcity: reduce, treat and reuse—the third generation of management to avoid local resources boundaries', *Water Policy*, 19(5), pp. 978-996.
- Eadson, W. and Van Veelen, B. (2023) 'Green and just regional path development', *Regional Studies, Regional Science*, 10(1), pp. 218-233.
- Elzen, B., Barbier, M., Cerf, M. and Grin, J. (2012) 'Stimulating transitions towards sustainable farming systems': Springer Netherlands, pp. 431-455.

- Evans, A. E., Varma, S. and Krishnamurthy, A. (2014) 'Formal approaches to wastewater reuse in Bangalore, India'.
- Flyvbjerg, B. (2006) 'Five misunderstandings about case-study research', *Qualitative Inquiry*, 12(2), pp. 219-245.
- Fuenfschilling, L. and Truffer, B. (2014) 'The structuration of socio-technical regimes— Conceptual foundations from institutional theory', *Research Policy*, 43(4), pp. 772- 791.
- Geels, F. W. and Turnheim, B. (2022) *The Great Reconfiguration: A Socio-Technical Analysis of Low-Carbon Transitions in UK Electricity, Heat, and Mobility Systems*. Cambridge: Cambridge University Press.
- Gibbs, D. and Jensen, P. D. (2021) 'Chasing after the wind? Green economy strategies, path creation and transitions in the offshore wind industry', *Regional Studies*, pp. 1-12.
- Guðmundsdóttir, H., Carton, W., Busch, H. and Ramasar, V. (2018) 'Modernist dreams and green sagas: The neoliberal politics of Iceland's renewable energy economy', *Environment and Planning E: Nature and Space*, 1(4), pp. 579-601.
- Hansen, T. (2022) 'The foundational economy and regional development', *Regional Studies*, 56(6), pp. 1033-1042.
- Hassink, R., Isaksen, A. and Trippl, M. (2019) 'Towards a comprehensive understanding of new regional industrial path development', *Regional Studies*, 53(11), pp. 1636-1645.
- Huenteler, J., Schmidt, T. S., Ossenbrink, J. and Hoffmann, V. H. (2016) 'Technology life- cycles in the energy sector – Technological characteristics and the role of deployment for innovation', *Technological Forecasting and Social Change*, 104, pp. 102-121.
- Huguenin, A. and Jeannerat, H. (2017) 'Creating change through pilot and demonstration projects: Towards a valuation policy approach', *Research Policy*, 46(3), pp. 624-635.
- Jansson, J. and Waxell, A. (2011) 'Quality and regional competitiveness', *Environment and planning A*, 43(9), pp. 2237-2252.
- Jeannerat, H. (2013) 'Staging experience, valuing authenticity: Towards a market perspective on territorial development', *European Urban and Regional Studies*, 20(4), pp. 370- 384.
- Jeannerat, H. and Crevoisier, O. (2022) 'From competitiveness to territorial value: transformative territorial innovation policies and anchoring milieus', *European Planning Studies*, pp. 1-21.
- Jeannerat, H. and Kebir, L. (2016) 'Knowledge, Resources and Markets: What Economic System of Valuation?', *Regional Studies*, 50(2), pp. 274-288.
- Kemp, R., Schot, J. and Hoogma, R. (1998) 'Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management', *Technology Analysis & Strategic Management*, 10(2), pp. 175-195.
- Klerkx, L., Van Mierlo, B. and Leeuwis, C. (2012) 'Evolution of systems approaches to agricultural innovation: concepts, analysis and interventions': Springer Netherlands, pp. 457-483.

- Klinger, M., Ulrich, L., Ramprasad, C., Wolf, A. T., Reynaud, N., Narayan, A. S., Siemsen, P., Lüthi, C. and Philip, L. (2020) 'Technology, Implementation and Operation of Small-Scale Sanitation in India – Performance Analysis and Policy Recommendations.', *4S Project Report Vol. I*.
- KSPCB (2014) *Karnataka State Pollution Control Board - Notice of Directions under Section 24/25 of the Water (Prevention and Control of Pollution) Act, 1974*. Bengaluru.
- Kuttuva, P., Lele, S. and Mendez, G. V. (2018) 'Decentralized wastewater systems in Bengaluru, India: success or failure?', *Water Economics and Policy*, 4(02), pp. 1650043.
- MacKinnon, D., Dawley, S., Pike, A. and Cumbers, A. (2019) 'Rethinking Path Creation: A Geographical Political Economy Approach', *Economic Geography*, 95(2), pp. 113- 135.
- Mackinnon, D., Karlsen, A., Dawley, S., Steen, M., Afewerki, S. and Kenzhegaliyeva, A. (2021) 'Legitimation, institutions and regional path creation: a cross-national study of offshore wind', *Regional Studies*, pp. 1-12.
- Markard, J. (2011) 'Transformation of Infrastructures: Sector Characteristics and Implications for Fundamental Change', *Journal of Infrastructure Systems*, 17(3), pp. 107-117.
- Markard, J., Raven, R. and Truffer, B. (2012) 'Sustainability transitions: An emerging field of research and its prospects', *Research Policy*, 41(6), pp. 955-967.
- Marques, P. and Morgan, K. (2021) 'Innovation without Regional Development? The Complex Interplay of Innovation, Institutions, and Development', *Economic Geography*, 97(5), pp. 475-496.
- Martin, R. and Sunley, P. (2015) 'Towards a Developmental Turn in Evolutionary Economic Geography?', *Regional Studies*, 49(5).
- Mccauley, S. M. and Stephens, J. C. (2012) 'Green energy clusters and socio-technical transitions: analysis of a sustainable energy cluster for regional economic development in Central Massachusetts, USA', *Sustainability Science*, 7(2), pp. 213-225.
- Miörner, J. (2019) *(Re-)shaping regional economies: Regional innovation system dynamics and new industrial path development*. Lund University, Lund.
- Miörner, J. and Binz, C. (2021) 'Towards a multi-scalar perspective on transition trajectories', *Environmental Innovation and Societal Transitions*, 40, pp. 72-188.
- Mulder, M. R. (2020) *Shaping the emerging alternative service regime from a Technological Innovation System perspective: the case of decentralized water and sanitation*. Utrecht University.
- Nath, S., Singh, S. and Balasubramanian, S. (2022) *Challenges and opportunities in the residential decentralised wastewater treatment and reuse sector in Bengaluru*, Bengaluru: ATREE.
- National Geographic (2018) *Why This Lake Keeps Catching on Fire*. Available at: <https://www.nationalgeographic.com/science/article/bangalore-india-lake-bellandur-catches-fire-pollution> (Accessed: 2023-04-13).
- Nilsen, T. and Njøs, R. (2021) 'Greening of regional industrial paths and the role of sectoral characteristics: A study of the maritime and petroleum sectors in

- an Arctic region', *European Urban and Regional Studies*, pp. 096977642110384.
- Schellenberg, T., Subramanian, V., Ganeshan, G., Tompkins, D. and Pradeep, R. (2020) 'Wastewater Discharge Standards in the Evolving Context of Urban Sustainability- The Case of India', *Frontiers in Environmental Science*, 8.
- Simmie, J. (2012) 'Path dependence and new technological path creation in the Danish wind power industry', *European Planning Studies*, 20(5), pp. 753-772.
- Sotarauta, M., Suvinen, N., Jolly, S. and Hansen, T. (2021) 'The many roles of change agency in the game of green path development in the North', *European Urban and Regional Studies*, 28(2), pp. 92-110.
- Statista (2021) *Number of inhabitants in Bengaluru in India from 1975 to 2015*. Available at: <https://www.statista.com/statistics/911018/india-population-in-bangalore/> (Accessed: 22 December 2022).
- Steen, M. (2016) *BECOMING THE NEXT ADVENTURE?: Exploring the complexities of path creation: The case of offshore wind power in Norway*. Dissertation: Norwegian University of Science and Technology, Faculty of Social Sciences and Technology Management, Department of Geography.
- Steen, M. and Hansen, G. H. (2018) 'Barriers to Path Creation: The Case of Offshore Wind Power in Norway', *Economic Geography*, 94(2), pp. 188-210.
- Trippl, M., Baumgartinger-Seiringer, S., Frangenheim, A., Isaksen, A. and Rypestøl, J. O. (2020) 'Unravelling green regional industrial path development: Regional preconditions, asset modification and agency', *Geoforum*, 111, pp. 189-197.
- Tödting, F., Trippl, M. and Desch, V. (2021) 'New directions for RIS studies and policies in the face of grand societal challenges', *European Planning Studies*, pp. 1-18.
- Uyarra, E., Ribeiro, B. and Dale-Clough, L. (2019) 'Exploring the normative turn in regional innovation policy: responsibility and the quest for public value', *European Planning Studies*, 27(12), pp. 2359-2375.
- Van Welie, M. J., Cherunya, P. C., Truffer, B. and Murphy, J. T. (2018) 'Analysing transition pathways in developing cities: The case of Nairobi's splintered sanitation regime', *Technological Forecasting and Social Change*, 137, pp. 259-271.
- Weber, K. M. and Rohrer, H. (2012) 'Legitimizing research, technology and innovation policies for transformative change: Combining insights from innovation systems and multi-level perspective in a comprehensive 'failures' framework', *Research Policy*, 41(6), pp. 1037-1047.
- Yap, X.-S., Truffer, B., Li, D. and Heimeriks, G. (2022) 'Towards transformative leapfrogging', *Environmental Innovation and Societal Transitions*, 44, pp. 226-244.

Trade liberalisation and productivity growth: role of resource allocation and catching up - Evidence from Indian manufacturing industry

Parameswaran Mavannoor

Centre for Development Studies

Abstract

This study examines the impact of trade liberalisation introduced in India since 1991 on the productivity growth of Indian manufacturing industry. The study focuses on the role of resource allocation from less productive to more productive firms as well as from nonexporting to exporting firms in industry level productivity growth. The recent theoretical literature analysing the impact of trade liberalisation on productivity uses heterogeneous firms framework. In this, firms operate at varying levels of productivity and trade liberalisation in these models increases aggregate productivity through resource allocation from less productive firms to more productive firms. Increased competition also forces firms to catch up with the best productivity firms; this within firm productivity growth also improves industry level productivity. However, a close examination of the theoretical literature modelling the impact of increased competition through trade liberalisation on firms' incentive to invest in productivity enhancing activities shows that it depends on many firm-specific and industry specific features. Thus the theoretical literature is inconclusive on whether increased competition incentivise firms to invest more in productivity enhancing activities. However, the literature elicit a number of firm-specific and industry specific factors shaping the final outcome, thus providing some guidance to the empirical analysis.

The present study used firm level panel data for the period 1992-93 to 2005-06 and firm level productivity is estimated using a gross output production function. Unbiased estimation of the contribution of resource allocation to aggregate productivity requires addressing *omitted price bias* while estimating productivity. We use the recent methodology presented in Gandhi, Navarro, and Rivers (2020a) to estimate a gross output production function, which addresses both omitted price bias and simultaneity bias.

The results of the study suggest that in a number of industries resource allocation to more productive firms accounts for a significant share of productivity growth. The contribution of resource allocation to exporters was also found to be significant in driving aggregate productivity growth. The study also finds evidence for productivity catch up among firms. The results also show that the impact of import competition on productivity growth depends on the extent of concentration in the domestic market. Similarly, import competition also increases productivity growth of initially low productive firms if they are operating in a concentrated market.

Keywords: Economic reforms, Trade liberalisation, productivity, India, productivity catch up

1. Introduction

In 1991 Indian economy witnessed beginning of major changes in its economic policy regime. Economy started moving to a liberalised trade and industrial policy regime. Initially, most of these policy reforms were targeted manufacturing sector with the objective of improving its growth and international competitiveness. Later, a large number of studies examined various dimensions of performance of the manufacturing sector during the post reform period.¹ One of the aspects received lot of attention is productivity growth of the manufacturing sector during the post reform period. The studies analysing the productivity impact of liberalisation either compared the productivity growth between pre and post reform periods or examined the impact of various dimensions of liberalisation on manufacturing productivity. As we shall see below, majority of the studies that compared productivity growth between pre and post reform period could not find a significant improvement in productivity during the post reform period.²

The two major aspects of economic liberalisation in India are industrial deregulation and trade liberalisation. The policy of industrial deregulation removed restrictions on entry, capacity expansion and product diversification; allowing firms to exploit economies of scale and scope as well as to invest in activities having highest rate of return. Similarly, trade liberalisation, which reduced the tariff and non-tariff barriers on import, must have increased the competitive pressure on firms. Therefore, it is quite surprising that manufacturing productivity has not responded in a significant way to liberalisation policy. Against this background, the present study focuses on the firm level dynamics that shape the productivity response of the manufacturing sector to trade and industrial deregulation. The study also examines firm and industry specific features that shape firms' response to trade and industrial policy reform.

Recent developments in the theoretical literature analysing the effect of liberalisation, particularly of trade liberalisation, considers heterogeneous firm framework, in which firms are assumed to have different levels of productivity. In this framework, trade liberalisation enhances aggregate industry level productivity through two channels. First, resource reallocation from less productive to more productive firms. Second, by incentivising firms to catch up with the best productivity firms. The present study empirically examines the role of resource allocation and catching up process in driving productivity growth at industry. We analyse the productivity effect of resource allocation in general, i.e., allocation among firms irrespective of their export status and of allocation between exporting and non-exporting firms. Further, the study also examines the firm specific and industry specific features that shape the impact of import competition on firms' productivity. The present study improves upon

existing studies by using an estimation framework that allows unbiased estimation of the contribution of resource allocation to productivity. Further, the study also correct for the simultaneity problem in the estimation of gross production function using a recent methodology that can also accommodate imperfect competition.

This study is organised in ten sections, including this introduction. Section 2 reviews the liberalisation policy introduced in India since 1991. Section 3 discusses the theoretical literature analysing various channels through which liberal trade and industrial policy regimes can affect firms' productivity. Section 4 reviews the empirical literature examining the productivity growth of Indian manufacturing industry in the context of economic policy reform. Section 5 presents research questions and objectives of the present study. Sections 6 and 7 respectively explains methodology and data of the study. Section 8 describes estimation procedure and also discusses the results. Section 9 concludes the study. An Appendix presents the data editing rules, definition and construction of variables, and further details on data and estimates.

2. Economic liberalisation in India

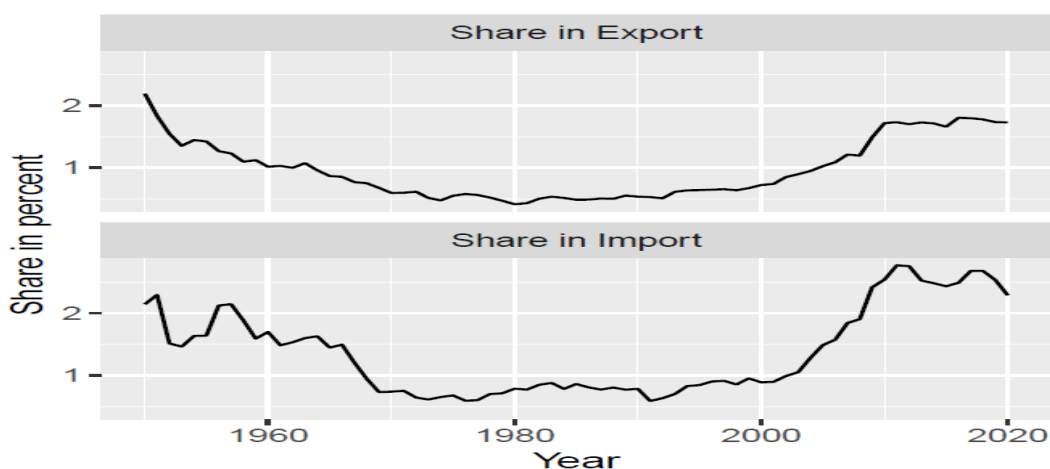
This section briefly reviews the evolution of industrial trade policies in India.³ Immediately after independence, India adopted an import-substituting industrialisation strategy. The important elements of this strategy were (1) planning, (2) protection of the domestic industry from import using tariff and non-tariff barriers, (3) prominent role for the public sector by reserving core sectors exclusively to public sector investment, and (4) state control on entry and expansion of existing production capacity of plants using industrial licensing. Since both import and domestic production were subjected to licensing restrictions, supply could not respond to demand, resulting in shortages of many commodities and consequent rise in their prices. To tackle this, government introduced price controls on those items which were deemed to be "essential items". Thus price controls were also evolved to become the integral part of the import substituting industrialisation strategy of India. The restrictions on the private sector further tightened after 1969. This include enacting of Monopolises and Restrictive Trade Practices (MRTP) Act and introduction of Small Scale Industries (SSI) reservation policy. The MRTP act curbed the expansion of big firms and restricted the large firms from producing goods that were reserved for production in small scale sector. The SSI reservation policy reserved a number of products exclusively for the production in small scale units. The number of commodities reserved for small scale production stood at 800 in 1980 (see: Chatterjee, 2012, p.19).

The economy performed reasonably well in terms of growth rate during the first fifteen years of planning era. However, growth rate during the subsequent fifteen years was quite disappointing. The growth rate of manufacturing halved

to 3.5 percent from 7.0 percent (see: Balakrishnan (2010)). Further, it was also argued that industrialisation behind tariff and non-tariff walls led to the emergence of an highly inefficient manufacturing sector. The import substituting industrialisation also resulted in allocation of resources away from exporting sectors to import competing sectors, defying India's comparative advantage. The severe balance of payment crises that India faced during this period were addressed mainly by further tightening the import controls, rather than devaluing the rupee. This led to an overvalued exchange rate and this further eroded the competitiveness of Indian exports, resulting in continuous decline in the share India's export in total world export (see Figure 1). Ultimately, this resulted in the severe balance of payment crisis in 1991, forcing the government to liberalise its industrial and trade policy regimes.

Though the 1991 reforms were a drastic change in the economic policy regime of the country, it was not something unexpected. Support for a liberal economic policy environment was slowly emerging in India since around 1980. The prolonged phase of industrial stagnation since the mid- 1960s led to a re-look into the role of trade and industrial policies in India. The Government of India appointed a number of committees to examine the various aspects of economic policies. These committees include Alexander Committee and Abid Hussain Committee to examine the import and export policies and Narasimham Committee to look into the financial sector. All these committees recommended for a more liberal policy regimes in their respective fields. Further, academic research also brought out the ill effects of import substituting industrialisation strategy and this also provided the required push for an alternative liberal economic policy regime (See: Bhagwati and Desai (1970), and Bhagwati and Srinivasan (1975)). Consequently, 1980s witnessed a number of moves towards a liberal trade and industry policy regime.

Figure 1: Trends in India's share in World Trade



The liberalisation measures initiated in 1991 focused on reforming the industrial and trade policy regimes of the country in order to stimulate the its industrial sector. The major policy reforms reforms introduced include the following. Industrial licensing requirement was abolished for majority of the industries. Reservation of industries for public sector production and investment was also discarded in the case of majority of industries. Monopoly Restrictive Trade Practices (MRTP) Act was amended in 1991 removing completely the restrictions on expansion and mergers by large firms. From 1997 onwards, products that had been reserved for small scale production were progressively dereserved. From, over a thousand products in 1996, only 22 remained reserved by 2008 and none by 2015 (see: Martin, Nataraj, and Harrison, 2017).

The reforms in the external sector focused on merchandise trade, ex- change rate regimes and foreign direct investment. Though exchange rate reform is a part of the general macroeconomic policy, it has implications for trade and industrial policies of the country (see: Rodrik, 1992b). Following Bhagwati (1988) trade liberalisation can be defined in terms of removal of discriminatory quantitative restrictions and reductions in tariff rate and the exchange rate regime being consistent with outward orientation with equality of effective exchange rates for importables and exportables.

The post 1991 period witnessed significant reduction in the tariff and non-tariff barriers to import. All most all the segments of the merchandise trade witnessed reduction in import tariff (see Table 1). Import licensing requirement was abolished in 1991 for the import of capital and intermediate goods. It remained in place for consumer goods for another decade and was phased out by 1 April 2002 under WTO obligations, except for a small number of items that remain restricted on grounds of health, safety and so on. Even after significant reduction in tariff and non-tariff barriers, industries experienced procedural complexities in case of import and export and hence high transaction cost. In this context, Sengupta and Roy (2018) observes that in the later part of the 1990s, Indian trade policy reform focused on simplifying these procedures in order to reduce the transaction costs.

Table 1: Trends in average tariff rate

Product	1990	1996	2000	2005	2010
Chemical and related products	77.93	40.74	35.32	15.79	8.22
Leather and related products	78.72	27.01	24.23	12.28	5.99
Wood, cork product and furniture	61.09	27.07	29.04	13.04	7.39
Textiles and garments	92.26	48.79	31.57	15.11	7.91
Non-metallic mineral products	74.22	47.7	36.93	14.76	7.8
Metal and manufactures	72.77	31.88	34.15	15.23	5.75
Machinery and transport equipment	57.79	33.17	25.34	15.49	7.98
Miscellaneous manufacturing	101.67	42.63	35.64	14.47	8.01

Source: Sengupta and Roy (2018), calculated using data from TRAINS database.

In order to deal with the immediate balance of payment crisis in 1991, the rupee was devalued by about 19% in two stages. Thus trade liberalisation is accompanied with compensating depreciation. In the following years, India moved to an entirely market-determined exchange rate regime, with the Reserve Bank of India (RBI) intervening occasionally to ensure exchange rate stability. The real effective exchange rate fell sharply after the initial devaluation, but then remained fairly stable with recurring episodes of real appreciation, caused by high domestic inflation relative to India's trading partners and inflows of foreign portfolio capital.

During the post 1991 period, significant changes were also introduced in the foreign direct investment (FDI) regime. Till 1979, FDI inflows were largely restricted in terms of ceiling on foreign equity participation, dividend balancing, local content requirement and export obligations (Kumar, 1994). The 1980s witnessed gradual easing of regulations on foreign capital and import of capital goods and technology were allowed. Liberalisation of FDI policies since 1991 was done in phases, with focus on manufacturing industries during the 1990s. Initially, automatic approval of FDI was sanctioned up to 51% of equity in a list of 34 industries specified as high priority, capital intensive and high technology industries. This was also subjected to the condition that foreign equity would cover foreign exchange requirement for import of capital goods. This condition was later waived in 1996. Later the number of industries eligible for approval under the automatic route increased from 35 to 111. Following the extensive liberalisation of FDI in manufacturing industries, later the focus on FDI policies shifted to infrastructure, insurance and services sectors and by 2000, 100% foreign equity participation was allowed in most of the manufacturing sectors (see: Sengupta and Roy, 2018).

One area that eluded reforms in the early and later phase of the policy reform was the labour market. It has been argued that labour market regulations in India are not industry friendly and posed it as one of the major reasons for the disappointing performance of the manufacturing industry during the post reform period (see: Panagariya, 2007). However, the recent years witnessed a number of significant changes in the labour market having implications for the performance of industry.

3. Liberalisation and productivity: Theory

The industrial and trade policy reforms introduced in India since 1991 are expected to impact the long run growth rate of its manufacturing sector in many ways. The policy reforms allowed greater role for the market forces in the allocation of resources across sectors, firms and over time. The removal of industrial licensing allowed free entry of firms. Free entry can increase the actual and potential competition in an industry. Further, this would also

increase the efficiency of resource allocation as more productive firms can freely enter into the production. Similarly, removal restrictions on the expansion of production capacity and diversification of products by firms and phasing out of small scale industry (SSI) reservation are all expected to allow firms to exploit economies of scale and scope, which in turn can increase their productivity.

Among various liberalisation measures, trade liberalisation attracted lot of scholarly attention for its potential impact on manufacturing industry. Earlier theoretical models examined the impact of trade liberalisation on domestic industry using competitive market models. These models predicted an increase aggregate productivity after opening to international trade through inter-industry specialisation according to the comparative advantage of the country. Import substituting policy regime has an inherent domestic market bias, as sectors protected by import restrictions were far more attractive to invest, hence resources will move away from exporting sectors to import substituting sectors, defying the country's comparative advantage. This implies that trade liberalisation can generate aggregate productivity gains through resource allocation to exporting sectors, which makes best use of them. However, this source of gain is argued to be of short run nature, as it get exhausted once economy specialise according to its comparative advantage (see: Rodrik, 1992b).

Later theoretical literature used models in which domestic industry is characterised by imperfect competition to examine the impact of trade liberalisation. Earlier models in this genre used representative firm framework. In these models trade liberalisation can affect the level of output of the firm and thereby its ability to exploit economies of scale and its incentive to invest in technological progress. If the domestic marginal costs are above border price, import competition can result in a reduction in the domestic output. However, it is also possible for trade liberalisation to result in an expansion of output of the import competing sectors. This happens when the perceived demand elasticity faced by the import-competing firms is likely to increase as result of trade liberalisation. Consequently their market power will erode, reducing the incentive to prop up domestic price by reducing the level of output. It is possible that the inward shift in the demand curve faced by these firms to be more than offset by the change in its slope, prompting an increase in domestic sales, enabling the firms to realise economies of scale(see: Rodrik, 1992b)⁴.

Further developments in the theoretical literature based on imperfect competition discarded representative firm approach in favour of heterogeneous firm framework, in which firms are assumed to have different productivity realisations. The literature based on firms' productivity heterogeneity identifies additional sources of aggregate productivity growth after

trade liberalisation. This literature is based on the realistic feature of industries that all industries are characterised by firms having different levels of productivity (see: Melitz, 2003; Melitz and Ottaviano, 2008). Melitz (2003) uses the monopolistic competition model of Dixit and Stiglitz (1977) to analyse the effect of trade liberalisation on aggregate industry level productivity through resource allocation among firms. In the model, products are symmetrically differentiated and firms are heterogeneous in terms of their productivity. The model also allowed endogenous differences in the “toughness” of competition across industries. Mark-ups are endogenous across firms that responds to the toughness of competition in a market. The toughness of competition is reflected in the number and average productivity of competing firms in that market. The analysis focuses on the long run effects of trade on the relative behaviour and performance of firms with different productivity levels. In this model, a more productive firm will be bigger (larger output and revenue), charges a lower price and earn higher profits than a less productive firm. Exposure to international trade in the model necessitate an increase in the cut-off productivity level, below which firms cannot survive. The least productive firms with productivity levels below the new cut-off will no longer earn a positive profit and therefore has to exit. This also results in reallocation of market shares towards more productive firms and hence higher aggregate productivity. The model also postulates that exporting requires a fixed entry cost and therefore only high productive firms can enter export market. Hence, increase in the scale of exporting firms after trade liberalisation would result in further gain in aggregate productivity (See also: Clerides, Lach, and Tybout, 1998; Bernard and Jensen, 1999; Aw, Chung, and Roberts, 2003; Pavcnik, 2002; Bernard, Jensen, and K., 2006). The exposure to trade thus generates a type of Darwinian evolution within an industry: the most efficient firms thrive and grow - they export and increase their market share and profits. Some less efficient firms still export and increase their market share but incur losses in profit. Some even less efficient firms remain in the industry but do not export and incur losses of both market share and profit. Finally, the least efficient firms are driven out of the industry.

The above theoretical predictions regarding the increase in aggregate productivity after trade liberalisation is through resource reallocation from less productive to more productive firms and exit of lesser productive firms. There are also theoretical arguments predicting increase in within firm productivity after trade liberalisation because of heightened competition. One early idea this respect is the so called *X-efficiency* argument. The *X-efficiency* argument is based on the presence of entrepreneurial slack in the import competing sectors. With trade liberalisation and consequent increased competition from import, entrepreneurs are forced to increase their effort. Similar predictions can be obtained from optimising models where entrepreneurs undertake a labour-leisure choice and their labour

supply curve is backward bending.⁵ So a reduction profit in import competing sectors after trade liberalisation forces entrepreneurs to increase their effort level. But these models have an unappealing and frequently overlooked implication of increased managerial slack in export oriented industries precisely for the same reason - improved profitability in these sectors should encourage entrepreneurs to have more leisure.

The very recent theoretical literature examining the productivity effect of trade liberalisation models its impact on firms' incentive to improve their technology. Perla, Tonetti, and Waugh (2021) uses the standard heterogeneous firm model in differentiated product market as in Melitz (2003). Firms are monopolistic competitors, who differ in their productivity and have opportunity to export. Firms upgrade their technologies to the one already in use by the other more productive firms. The paper analyses how trade liberalisation affects firms' incentive to upgrade their technology and the implications of this for the aggregate productivity growth. Firms' incentive to adopt better technology depends on two competing forces, the expected benefit of new technology in terms of profit and the opportunity cost of adoption. The opportunity cost of adopting a new technology is the forgone profits from producing with the current technology. This formulation implies that firm's incentive to adopt new technology depends upon its position in the productivity distribution. Hence, impact of trade liberalisation on firm's incentive to adopt new technology and production practices depends upon its location in the productivity distribution. In the model, with trade liberalisation, firms occupying lower part of the productivity distribution contract as competition from foreign firms reduce their profit and those occupying higher side of the productivity distribution increase their profit by expanding their output and export. For low productivity firms, this reduces the opportunity cost of adopting new technology and increases the benefit of adopting new technology, resulting in more frequent technology adoption and consequent higher productivity growth. Firms occupying upper part of the productivity distribution continue to operate with their existing technology. As the aggregate growth rate depends upon the frequency at which firms upgrade their technology and production practices, the model shows that growth rate is higher in open economies. The model also implies a decline in productivity dispersion after trade liberalisation as low productive firms continuously upgrade their technology.⁶ Though, a number of empirical studies show a positive association between firms' productivity and extent of trade openness, the mechanism generating this association is not clearly explored. Perla, Tonetti, and Waugh (2021) provides a firm level mechanism having implications for aggregate productivity growth.

Another set of theoretical models focused on the impact of import competition on firms' incentive to invest in innovation and consequent growth. However,

these models show that the direction of the effect is ambiguous and sensitive to the modelling assumptions about domestic market structure, cost structure of firms and so on. Rodrik (1992a) analysed the effect of import competition under two market structures; one in which the domestic market structure is monopoly and in the other an oligopoly market. Under monopoly, the incentive to invest in R&D is greater larger the scale of output. Therefore, in the first model, import competition, which shrinks the market share of the domestic producer, reduces the incentive to invest in innovation. In the second model, where the domestic industry is an oligopoly behaving in Bertrand fashion, import competition stimulates R&D effort. Thus, these results, as in any oligopoly model, depends on the particular behavioural assumptions about the conduct of firms in the market.

Smulders and Klundert (1995) examined the effect of market concentration on innovation effort using an endogenous growth model under a monopolistic competition in the domestic market. This model predicts that firms' incentive to invest in R&D is greater when the domestic market is highly concentrated and market power of firms is lower. Import competition, in this model encourages R&D by reducing the mark-up as well as by increasing the level of domestic concentration through the exit of the inefficient firms and absorption of their market share by extant ones.

One important assumption of Smulders and Klundert (1995) is that the cost structure of the domestic and foreign firms is symmetric. Traca (2002) relaxed this assumption and examined the impact of import competition on domestic monopolist's effort on innovation. The effect of import competition on domestic firm's output is divided into two parts, namely *market share effect* and *pro-competitive effect*. The first one captures the reduction in domestic firm's output due to a fall in its market share caused by the substitution to import. The second one refers to the expansion in output induced by the lower price due to the reduction in mark-up brought out by foreign competition. How import competition affects R&D depends on the relative strengths of these two effects. When the pro-competitive effect dominates the market share effect, the domestic firm would face a net expansion in the output and this along with lower mark-up could provide greater incentive to invest in R&D and productivity enhancing activities. This model shows that in mature industries, where productivity gap between domestic and foreign industry is low, pro-competitive effect dominates and firm would enjoy an expansion in the market. Whereas, in infant industries, where productivity gap is high, market share effect dominates and import competition would reduce R&D or productivity enhancing investments.

Another theoretical literature relevant in the present context examined the effect of product market competition on firms' innovation effort and these include Aghion, Harris, Howitt, and Vickers (2001) and Aghion (2003). It is Schumpeter

(1942) who suggested that market structure and innovation effort of firms were related and a competitive one was not conducive for innovation. Here, the argument is that reward for a successful innovator comes in the form of monopoly rents. Intense competition, by lowering the monopoly rents, reduces the incentive to innovate. Aghion, Harris, Howitt, and Vickers (2001) and Aghion (2003) argue that the incentive to perform R&D depends not on the rents of a successful innovator per se, but rather on the innovator's incremental rents, i.e. the difference between rents of a successful innovator and that of an unsuccessful one. These models consider a duopoly with Bertrand price competition and predict that product market competition and innovation effort have an inverted U shape relationship. In the extreme case of monopoly, profit is independent of monopolist's technological leadership and therefore has no incentive to invest in technological progress. On the other hand, at high levels of competition firms find it difficult to appropriate rents from innovation and has lower incentive to become innovative. Thus, in these models injecting greater competition into an already concentrated industry can generate greater incentive to invest in productivity enhancing activities. The theoretical literature reviewed above suggests that impact of increased import competition on aggregate productivity growth of the manufacturing sector depends upon many factors specific to the industry and firm like firms' position in the productivity distribution and the extent of domestic concentration.

An open trade policy regime also gives firms access to variety of intermediate inputs, including capital goods, raw-materials and technology. It has been shown that increase in the variety of intermediate inputs in production increases productivity (see: Grossman and Helpman, 1991). However, as we have discussed above, whether firms will adopt these new technologies and intermediate inputs depends upon how trade liberalisation affects the incentive of the firm to adopt new technology and production practices.

4. Liberalisation and productivity: Review of empirical literature

A number of studies examined the impact of liberalisation on productivity of manufacturing industry using various data sets and different methodologies. Majority of the studies focused on the organised sector of the manufacturing industry and used either industry or plant level data obtained from Annual Survey of Industries (ASI) or firm level data obtained from sources like PROWESS database of Centre for Monitoring Indian Economy (CMIE). This section provides a very selective review of these studies, as recent exhaustive review of them are given in Goldar (2014) and Krishna, Goldar, Das, Aggarwal, Erumban, and Das (2022). The set of studies examining the impact of liberalisation on manufacturing productivity can be grouped into two. First group consists of studies comparing the productivity growth between pre and post liberalisation periods. The sec-

ond group examined the impact of variables related to liberalisation on the observed productivity along with other control variables using econometric methods.

Two earlier studies examined the impact of liberalisation, particularly trade liberalisation on manufacturing productivity growth are Krishna and Mitra (1998) and Balakrishnan, Parameswaran, Pushpangadan, and Suresh Babu (2006). These studies used firm level data to estimate and test for a change in productivity growth after 1991. The studies used dummy variable to capture the change in productivity growth in the post reform period and also controlled for any change in markup. The first study found an increase in productivity growth and the second one a decline. As the dummy variable can capture all the developments after 1991, the observed change in productivity growth cannot be attributed to liberalisation alone.

Studies that estimated total factor productivity growth of Indian manufacturing industry for the pre and post reform period include Trivedi, Prakash, and Sinate (2000), Unni, Lalitha, and Rani (2001), Goldar and Kumari (2003), Goldar (2004), Trivedi (2004), Das (2004), Raj and Mahapatra (2009), Banga and Goldar (2007), Virmani and Hashim (2009), Trivedi, L., Jain, and Gupta (2011), and Das and Kalita (2011). These studies estimated the total factor productivity of organised manufacturing industry using growth accounting method based either on value added production function or gross output production function. These studies used data from Annual Survey of Industries (ASI). Some of these studies also estimated TFP growth at the two-digit level of National Industrial Classification (NIC). The overall conclusion of these studies is that growth rate of TFP during the post reform period is lower than that in prereform period. On the other hand, Unel (2003) and Bollard, Klenow, and Sharma (2013), who also estimated TFP growth for the pre and post reform periods, found an increase in TFP growth after the liberalisation. These two studies estimated productivity growth using growth accounting method based on value added production function.⁷ Srivastava, Gupta, and Datt (2001), estimate productivity growth using firm level data, also shows a slower productivity growth during the post reform period.⁸

The slower productivity growth during the post reform period should not led one to conclude that liberalisation has no productivity enhancing effect. As we have seen above the process of initial tariff reduction may not have much effect on the extent of import competition as it was offset by the devaluation of the rupee. Also, the import of consumer goods were freed from import restrictions only from 2000 onwards. Further, the change in policy regime may have first shaken the system and hence firms might need some time to adjust. This implies that liberalisation can show its effect on productivity only after some time. Hashim, Kumar, and Virmani (2011) and Virmani and Hashim

(2011) tested the hypothesis of J curve relationship between productivity and liberalisation. The hypothesis implies that initially liberalisation can have a negative effect on productivity as it takes time for firms to respond to the structural transformation implied by the trade liberalisation. Later, when firms adjusted for the new policy regime, liberalisation can accelerate the growth rate of productivity. These studies find that later part of the post liberalisation period witnessed higher productivity growth, evidence supporting their hypothesis. Goldar (2014) also shows that the TFP growth rate during the later part of the post reform period (2001-02 to 2007-08) is significantly higher than that in the previous periods.

Another set of studies explicitly tested the impact of various dimensions of liberalisation on the productivity of the manufacturing industries. Chand and Sen (2002) investigated the impact of trade liberalization on manufacturing productivity growth. It finds that variables related to trade liberalisation do have a positive effect on productivity. It also shows that liberalization of the intermediate good sectors has greater positive impact on TFP growth than that of the final good sectors. Goldar and Kumari (2003) finds positive effect of tariff reduction on productivity. Similarly, Das (2006) finds positive effect of removal of non-tariff barriers on productivity in the capital goods and intermediate goods industries. Mitra and Ural (2008) investigates the determinants of productivity of Indian manufacturing industry using two digit industry level data for Indian states, find that trade liberalization increases productivity in all industries across all states. Productivity is also found to be higher in the less protected industries. Sivadasan (2009), using plant level data, examined the impact of tariff liberalisation and FDI liberalisation on mean plant level productivity. The study finds significant increases in productivity in the FDI and tariff liberalised industries. The study also examined the role of intensive (within-plant productivity growth) and extensive (reallocation from less to more productive plants) margins in the post-reform productivity improvement, and find that within plant productivity has a predominant role. Topalova and Khandelwal (2011) finds that heightened competition, resulting from lower tariffs on final goods, as well as access to better inputs, owing to lower input tariffs, appear to have increased firm-level productivity, with reduction in input tariffs having a larger impact. Goldberg, Khandelwal, Pavcnik, and Topalova (2010) shows that trade liberalisation allowed the firms to access a variety of inputs enabling them to exploit economies of scope in their product mix by introducing new products. This also contributed to higher TFP and R&D investment by firms. Das (2016) presents industry level evidence on the connection between trade policy reforms and manufacturing productivity. The study covers the period 1990-91 to 2009-10. It shows a sharp fall in tariff levels as well as quantitative restrictions in the 1990s as compared to the 2000s. It also notes that the periods of 2000s were more aimed at simplifying the trade procedures as well as boosting exports. The analysis of trade openness

on productivity finds that the reduction in effective rate of protection has lagged impact on productivity and removal of import coverage rate has relatively smaller impact on TFP. The study observes that the biggest impact of rise in TFP has actually comes from improvement in domestic industrial policy.

Two important studies that are relevant for the current study are Harrison, Martin, and Nataraj (2013) and Sivadasan (2009). Using plant level data from ASI, these two studies examined the role of market share reallocation from less productive to more productive plants and plants' own productivity growth in driving the aggregate productivity growth. Harrison, Martin, and Nataraj (2013) covered the period from 1985 to 2004 and Sivadasan (2009) from 1986-87 to 1994-95. Both studies found that, compared to plant's own productivity growth, the contribution of market share reallocation to aggregate productivity growth is not important.

5. Research problem and objectives of the study

The main elements of the liberalisation measures that can affect the manufacturing sector's productivity growth are as follows. (1) Increased competition arising from trade liberalisation and removal of entry barriers. Trade liberalisation in terms of reduction in tariff and non-tariff barriers can increase the availability of closely substitutes products in Indian market, which in turn can increase the competitive pressure on firms. (2) Removal of entry barriers can also increase competition from potential entrants, which can force the extant firms to perform well. (3) The abolition of restrictions on the expansion of size, product diversification and SSI reservation enable firms to exploit economies of scale and scope. Further, as the firms move from the regime of "licence raj" which controlled firms' ability to use economies of scale and scope to a regime in which market forces can decide the allocation resources, resources reallocation from less productive to more productive firms can increase the productivity growth at the aggregate level. (4) As we have mentioned above, liberalisation can increase the firms' access to variety of intermediate inputs and technologies from abroad and utilisation of these inputs can increase firms productivity. As we have seen above, the impact of increased competition on productivity is not clear. The theoretical literature gives the impression that it depends on features specific to firms and industry, leaving it to the empirical analysis to decide which features matter.

Majority of the empirical studies reviewed in Section 4 either compared the growth rate of productivity between pre and post reform periods or explicitly correlated different indicators of liberalisation with the estimated productivity growth. Only Harrison, Martin, and Nataraj (2013) and Sivadasan (2009) examined the role of resource re-allocation in the productivity growth at the aggregate industry level. However, these studies

could not find a significant role for resource re-allocation in driving productivity growth. None of the studies examined the role of resource allocation to exporting firms in the aggregate productivity growth, though it has been presented as an important driver of productivity growth. In Indian context, trade liberalisation was accompanied with industrial policy reform involving removal of restrictions on entry, expansion of scale of operation and diversification of product mix. Removal of these restrictions have the potential to greatly contribute to aggregate productivity growth through resource reallocation from less productive to more productive firms. Therefore, it is quite surprising that studies that examined the role of resource allocation in driving the aggregate productivity growth could not find a major role for it. These studies suffer from two weaknesses that can be the reason for this. First, both these studies uses plant level data from ASI, not firm level data. As plants have maximum technical capacity set at the time of their construction, more productive firms expand their production by setting up more plants rather than over utilising or building over their existing plants.⁹ This implies that reallocation effect on productivity cannot be captured using plant level data. Second, these studies are not correcting for the bias induced by omitting the firm level price. As we shall see in the next section, productivity estimation without correcting for the bias due to the omission of firm level price underestimate the contribution of reallocation to aggregate productivity. Third, as we have seen above, trade liberalisation and industrial policy reforms, like tariff reduction and abolition of the SSI reservation, introduced in India gradually, and hence longer time data is required to assess their impact on productivity through resource allocation. Against this background the present study focuses on the following objectives.

1. Assess the contribution of resource allocation at aggregate productivity growth at the industry level. For this study uses firm level panel data for a longer time period and also correct for the bias induced by the omission of firm level prices.
2. To assess the increase in industry level productivity growth due to expansion of exporting firms.
3. To test whether the impact of import competition on firms' productivity growth depends upon its location on the productivity distribution.
4. To test whether domestic market structure has any role in shaping the impact of import competition on firm level productivity growth.

6. Methodology

6.1 Issues in estimation

We use production function approach to estimate productivity at the firm

level. Production function of j^{th} firm in t^{th} year is specified as follows,

$$Y_{jt} = F(C_{jt}, L_{jt}, M_{jt})e^{v_{jt}} \quad (1)$$

Where C , L , and M respectively denote capital, labour, and material including energy. The Hicks neutral productivity shock v_{jt} is decomposed as $\omega_{jt} + \epsilon_{jt}$, where ω_{jt} is the firm-specific productivity and η_{jt} is a random error term capturing measurement error and idiosyncratic shocks to production. The production function in log form is given below, where the lower case letters denote values in log form.

$$y_{jt} = f(c_{jt}, l_{jt}, m_{jt}) + \omega_{jt} + \epsilon_{jt} \quad (2)$$

Standard OLS estimation of the production function (2) runs into two major problems.¹⁰ The first problem is often referred to as “simultaneity” and stems from the fact that ω_{jt} is observable to the firm and so enters into its input demand function. However, ω_{jt} is unobservable to the econometrician and hence ends up in error term. The presence of ω_{jt} in the error term makes OLS estimates of the production function inconsistent (see: Griliches and Mairesse, 1998).

The second problem originates from the unavailability of information on firms’ output Y_{jt} in our dataset. This forces us to use a proxy, namely firm’s revenue R_{jt} deflated by an industry-level price deflator, as prices at firm level are also not available. This is not a problem in a perfectly competitive market in which all firms charge same price. When the markets are imperfectly competitive, firm-level estimated productivity is likely to be biased as in these markets revenue differences between firms underestimate the output differences.¹¹ For instance consider two firms with one percent relative output differences. In a perfectly competitive homogeneous good industry two firms would also have a one percent relative revenue differences. In a differentiated product industry with elasticity of demand σ the firm with higher output must have reduced its price relative to the other by¹² per cent in order to increase its relative output by one per cent. Then its relative revenue will be only $\frac{\sigma-1}{\sigma}$ per cent higher, i.e. less than one percent. Thus in differentiated product industry, revenue differences across firms understate the output differences. Since the definition of productivity is based on output, the true productivity differences gets underestimated. The amount of this measurement bias depends on the value of elasticity of substitution, greater the level of product differentiation (lower the σ) greater the underestimation of the true productivity. One can show that this would lead to underestimation of the productivity of firms having sales above industry average sales ($\bar{R}_t$) and overestimation of productivity of firms having sales lower than industry average sales ($\bar{R}_t$), as shown below.¹³

$$\hat{\omega}_{jt} = \omega_{jt} - \frac{1}{2(\sigma - 1)}(R_{jt} - \bar{R}_t)$$

It is obvious that it would lead to underestimation of the aggregate productivity growth through resource allocation from less productive to more productive firms. Because the problem is caused by omitting individual prices from estimation process, this problem is usually referred to as “omitted price bias”. Writing the production function in terms of deflated sales revenue, $\tilde{r}$:

$$\tilde{r}_{jt} = f(c_{jt}, l_{jt}, m_{jt}) + \omega_{jt}^* + \eta_{jt}^*$$

where $\tilde{r}_{jt} = r_{jt} - p_t = y_{jt} + p_{jt} - p_t$. Whenever $p_{jt} \neq p_t$, the OLS estimator is affected by the omitted price bias.

6.1 .1 Dealing with biases

We address the omitted price bias by adopting an approach originally developed by Klette and Griliches (1996) and latter further improved by Melitz (2000). The resulting revenue production function is estimated using the methodology suggested by Gandhi, Navarro, and Rivers (2020a), which provides procedure for getting consistent estimates of the production function under monopolistic competition. The procedure can be explained as follows:

Consider an industry producing N symmetrically differentiated varieties with a common elasticity of substitution σ between any two varieties. The demand for each firm’s output Y_j is generated by a representative consumer with utility:

$$U = \left[\sum_{j=1}^N Y_j^\rho \right]^{\frac{1}{\rho}} ; \quad 0 < \rho < 1 \quad (3)$$

Where Y_j is the quantity of j^{th} good consumed. Demand function for the above utility function (3) is given below:

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\sigma} \frac{1}{N} \left(\frac{R_t}{P_t} \right) \quad (4)$$

Where $R_t = \sum_{j=1}^N P_{jt} Y_{jt}$ is the total industry sales, $\frac{R_t}{P_t}$ is the total industry’s deflated sales and $\sigma = \frac{1}{1-\rho}$. The demand function summarises how in a differentiated product industry, a firms’ output is jointly determined by its price relative to the industry price index P , the number of firms in the industry and the aggregate industry sales. The exact price index P underlying the demand function (4) for the consumer utility specified in

(3) is:

$$P_t = \left[\frac{1}{N} \sum_{j=1}^N P_{jt}^{1-\sigma} \right]^{\frac{1}{1-\sigma}} \quad (5)$$

Taking log of the demand function:

$$y_{jt} = -\sigma(p_{jt} - p_t) + \bar{r}_t$$

where $\bar{r}_t$ is the average industry deflated sales.

$$\sigma(p_{jt} - p_t) = \bar{r}_t - y_{jt}$$

$$(p_{jt} - p_t) = \frac{1}{\sigma} \bar{r}_t - \frac{1}{\sigma} y_{jt}$$

Using the revenue production function relation, we get,

$$\bar{r}_{jt} = y_{jt} - \frac{1}{\sigma} y_{jt} + \frac{1}{\sigma} \bar{r}_t$$

$$\bar{r}_{jt} = y_{jt} \left(\frac{\sigma-1}{\sigma} \right) + \frac{1}{\sigma} \bar{r}_t$$

$$\bar{r}_{jt} = \frac{\sigma-1}{\sigma} (\beta_c c_{jt} + \beta_l l_{jt} + \beta_m m_{jt} + \omega_{jt} + \eta_{jt}) + \frac{1}{\sigma} \bar{r}_t$$

$$\bar{r}_{jt} = \gamma_c c_{jt} + \gamma_l l_{jt} + \gamma_m m_{jt} + \gamma_r \bar{r}_t + \omega_{jt}^* + \eta_{jt}^* \quad (6)$$

where $\gamma_i = \frac{(\sigma-1)}{\sigma} \beta_i$ and $\gamma_r = \frac{1}{\sigma}$, $\omega_{it}^* = \frac{(\sigma-1)}{\sigma} \omega_{it}$. The revenue production function (6) is estimated using the production function estimation procedure developed by Gandhi, Navarro, and Rivers (2020b). This methodology estimates the revenue production function consistently and the details of which are given below.

6.2 Estimation of revenue production function

The issue of identification of the production function arises when firms optimally choose their inputs as a function of their productivity, but productivity is unobservable to the econometrician, hence ending up in the error term. This give rise to the problem of simultaneity or transmission bias, which makes OLS estimates of the production function inconsistent. In gross output production function, consisting of capital, labour and material inputs, simultaneity leads to underestimation of the coefficient of the

relatively fixed input, namely capital and overestimation of the co-efficient of relatively flexible inputs, namely raw materials or energy. A number of methods have been developed to estimate the production function consistently in the face of simultaneity problem. The recent methods include the semi-parametric methodology of Olley and Pakes (1996) and further improvement of the semi-parametric methodology by Levinsohn and Petrin (2003). The semi-parametric methodology obtains consistent estimates of the production function using firms' raw material consumption to control for the variation in inputs that is correlated with the unobserved productivity. However, Akerberg, Caves, and Frazer (2015) shows that semi parametric methodology of Levinsohn and Petrin (2003) can identify value added production functions only, not the gross output production function. Later, Gandhi, Navarro, and Rivers (2020a) (hence forth GNR) developed a methodology that non-parametrically identifies the gross output production function. The identification strategy is based on the structural link between production function and firms' first order conditions for a flexible input. The paper shows that a non-parametric regression of the flexible input's revenue share on all inputs identifies flexible input elasticity. We use GNR methodology to estimate the above revenue production function, and the details are explained below.

We start with specifying a CES utility function as given below

$$U = \left[\sum_{j=1}^N Y_j^\rho \right]^{\frac{1}{\rho}} \quad 0 < \rho < 1$$

where Y_j is the quantity consumed of variety j . Note that in the above utility function, consumption of all varieties enters symmetrically. The parameter ρ represents the love-of-variety effect of consumers. $0 < \rho < 1$ to make sure that products are imperfect substitutes to each other. If $\rho = 1$ commodities are perfect substitutes.

Consumers' budget constraint is given by

$$E = \sum_{j=1}^N P_j Y_j$$

From the utility function and the budget constraint, following demand function for the j^{th} variety can be derived as in Klette and Griliches (1996), Melitz (2000) and De Loecker (2011b). From here onwards we use both individual and time index in equations.

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\sigma} \frac{1}{N} \frac{Y_t}{P_t} \quad (7)$$

where P_t is the price index, Y_t is the total demand in the industry and N indexes the number of firms in the industry. The above demand function can also be written as follows

$$Y_{jt} = \left(\frac{P_{jt}}{P_t} \right)^{-\sigma} \bar{Y}_t \quad (8)$$

where $\bar{Y}_t$ is the average real demand in the industry (average sales deflated by the price index). The inverse demand function with demand shock X_{jt}

$$P_{jt} = P_t \left(\frac{Y_{jt}}{\bar{Y}_t} \right)^{-\frac{1}{\sigma}} e^{X_{jt}} \quad (9)$$

where P_{jt} is the output price of firm j . P_t is the industry price index, $\bar{Y}_t$ is the average output (in real terms) in the industry, X_{jt} is an observable (to the firm) demand shock, and σ is elasticity of demand.

Let

$$Y = F(k_{jt}, l_{jt}, m_{jt}) e^{\nu}$$

where $\nu_{jt} = \omega_{jt} + \epsilon_{jt}$ Using production function and demand function we can write as follows:

$$P_{jt} Y_{jt} = P_t \frac{Y_{jt}^{1-\frac{1}{\sigma}}}{\bar{Y}_t^{-\frac{1}{\sigma}}} \cdot e^{X_{jt}} \quad (10)$$

Differentiating with respect to M_{jt} to maximise the expected profit, the first order condition is

$$\frac{\partial P_{jt} Y_{jt}}{\partial M_{jt}} = \left(1 - \frac{1}{\sigma} \right) P_t \frac{Y_{jt}^{-\frac{1}{\sigma}}}{\bar{Y}_t^{-\frac{1}{\sigma}}} \frac{\partial}{\partial M_{jt}} F(c_{jt}, l_{jt}, m_{jt}) e^{\omega_{jt} + \epsilon_{jt}} e^{X_{jt}} = \rho_t \quad (11)$$

where ρ is the unit price of M .

From (11), we can derive the following share equation.

$$\ln \left[1 - \frac{1}{\sigma} \right] + \ln \left[D(c_{jt}, l_{jt}, m_{jt}) E \left(e^{(1-\frac{1}{\sigma})\epsilon_{jt}} \right) \right] - \left(1 - \frac{1}{\sigma} \right) \epsilon_{jt} = s_{jt}$$

where

$$s_{jt} \equiv \ln \left[\frac{\rho_t M_{jt}}{P_{jt} Y_{jt}} \right]$$

In the above expression, the expected markup, i.e., the ratio of price to marginal cost is

$$\frac{1}{1 - \frac{1}{\sigma}}$$

$D(\cdot)$ is the output elasticity of intermediate input and ϵ_{jt} is the ex post shock. The above equation nests the one that can be derived for the competitive case, the only difference being the addition of the expected markup, which is equal to one in the competitive case.

Let $\tilde{E}_{jt} = 1 - \frac{1}{\sigma} \epsilon_{jt}$ and taking expectations inside, then the above equation becomes,

$$s_{jt} = Y + \ln D(k_{jt}, l_{jt}, m_{jt}) + \ln \tilde{E}_{jt} - \tilde{\epsilon}_{jt} \quad (12)$$

where $\tilde{E} = E[e^{\epsilon_{jt}}]$ and $Y = \ln 1 - \frac{1}{\sigma}$. The intermediate input elasticity can be rewritten so that we can break it into two parts: a component that varies with inputs and a constant μ , i.e., $\ln D(c_{jt}, l_{jt}, m_{jt}) = \ln D^\mu(c_{jt}, l_{jt}, m_{jt})$, then the above equation become

$$\begin{aligned} s_{jt} &= (Y + \mu) + \ln \tilde{E} + \ln D^\mu(c_{jt}, l_{jt}, m_{jt}) - \tilde{\epsilon}_{jt} \\ &= \varphi + \ln \tilde{E} + \ln D^\mu(c_{jt}, l_{jt}, m_{jt}) - \tilde{\epsilon}_{jt} \end{aligned} \quad (13)$$

As equation (13) makes clear, without observing prices, we can non-parametrically recover the scaled ex-post shock $\tilde{\epsilon}_{jt}$ (and hence ϵ_{jt}), the output elasticity of intermediate input up to a constant $\ln D^\mu(c_{jt}, l_{jt}, m_{jt}) = \ln D(c_{jt}, l_{jt}, m_{jt}) - \mu$; and the mark up up to a constant, $\varphi = Y + \mu$.

Now we can correct our estimates for $\tilde{E}$ and solve the differential equation that arises from equation (13). Since, we can still identify the elasticity only up to a constant μ we have to be careful about keeping track of it as we can only calculate $D^\mu(c_{jt}, l_{jt}, m_{jt}) dm_{jt} = e^{-\mu} D(c_{jt}, l_{jt}, m_{jt}) dm_{jt}$. It follows that

$$f(c_{jt}, l_{jt}, m_{jt}) e^{-\mu} + C(c_{jt}, l_{jt}, m_{jt}) e^{-\mu} = \int D^\mu(c_{jt}, l_{jt}, m_{jt}) dm_{jt} dm_{jt} \quad (14)$$

From this equation it is clear that we will not be able to separate the integration constant $C(c_{jt}, l_{jt})$ from the unknown constant μ . Now we recover the both the constant μ and the constant of integration as follows.

Note that what we observe in the data is firm's real revenue, which in logs is given by $r_{jt} = (p_{jt} - p_t) + y_{jt}$.

$$p_{jt} - p_t = -\frac{1}{\sigma} - (y_{jt} - \bar{r}_t) + X_{jt}$$

Using the above relation we can re-write the revenue function

$$r_{jt} = \left(1 - \frac{1}{\sigma}\right) y_{jt} + \frac{1}{\sigma} \bar{r}_t + X_{jt}$$

$$r_{jt} = \left(1 - \frac{1}{\sigma}\right) f(c_{jt}, l_{jt}, m_{jt}) + \frac{1}{\sigma} \bar{r}_t + X_{jt} + \left(1 - \frac{1}{\sigma}\right) \omega_{jt} + \tilde{\epsilon}_{jt}$$

We know that

$$\Upsilon = \ln \left(1 - \frac{1}{\sigma}\right) \Leftrightarrow e^\Upsilon = \left(1 - \frac{1}{\sigma}\right)$$

$$\varphi = \Upsilon + \mu$$

$$\Rightarrow \varphi - \mu = \Upsilon$$

$$e^\varphi e^{-\mu} = e^\Upsilon = \left(1 - \frac{1}{\sigma}\right)$$

Now rewriting the above equation

$$r_{jt} = e^{\varphi t} e^{-\mu} f(c_{jt}, l_{jt}, m_{jt}) - (e^{\varphi t} e^{-\mu} - 1) \bar{r}_t + [e^{\varphi t} e^{-\mu} \omega_{jt} + X_{jt}] + \tilde{\epsilon}_{jt}$$

We now follow a similar strategy as before. Multiplying both sides of equation (14) by e^φ , and subtract it from the revenue function we get.

$$r_{jt} - e^{\varphi t} \int D^\mu(c_{jt}, l_{jt}, m_{jt}) dm_{jt} = e^{\varphi t} e^{-\mu} f(c_{jt}, l_{jt}, m_{jt}) - (e^{\varphi t} e^{-\mu} - 1) \bar{r}_t + [(e^{\varphi t} e^{-\mu}) \omega_{jt} + X_{jt}] + \tilde{\epsilon}_{jt} - e^{\varphi t} e^{-\mu} C(c_{jt}, l_{jt})$$

Let us define R_{jt} as follows,

$$R_{jt} \equiv \ln \left(\frac{\frac{P_{jt} Y_{jt}}{\Pi_t}}{e^{\tilde{\epsilon}_{jt}} e^{\varphi t} \int D^\mu(c_{jt}, l_{jt}, m_{jt}) dm_{jt}} \right)$$

Now

$$R_{jt} = -e^{\varphi t} e^{-\mu} C(c_{jt}, l_{jt}) - (e^{\varphi t} e^{-\mu} - 1) \bar{r}_t + [e^{\varphi t} e^{-\mu} \omega_{jt} + X_{jt}]$$

Note that R_{jt} is an observable variable that can be derived from the data. Also without having recovered φ_t from the share regression first, it would not be possible to identify time varying mark-ups. Also note that in the above equation both σ_t and y_t change with time, and hence y_t cannot be used to identify σ_t unless we restrict $\sigma_t = \sigma$, as in Klette and Griliches (1996) and De Loecker (2011a).

We can recover only a linear combination of productivity and demand

shock,

$$\left(\left[1 - \frac{1}{\sigma} \right] \omega_{jt} + \mathcal{X}_{jt} \right)$$

The reason is that since we do not observe prices, we have no way of disentangling whether, after controlling for inputs, a firm has higher revenue because it is more productive (ω_{jt}) or because it can sell at a higher price $\mathcal{X}_{jt}$.

We can write

$$\omega_{jt}^{\mu} = \left(\left[1 - \frac{1}{\sigma} \right] \omega_{jt} + \mathcal{X}_{jt} \right)$$

as a function of the parts that remains to be recovered,

$$\omega_{jt}^{\mu} = \mathcal{R}_{jt} + e^{\varphi_t - \mu} \mathcal{C}(c_{jt}, l_{jt}) + (e^{\varphi_t} e^{-\mu} - 1) \bar{r}_t$$

and impose the Markovian assumption on this combination,

$$\omega_{jt}^{\mu} = h(\omega_{jt-1}^{\mu}) + \eta_{jt}^{\mu}$$

We can use following moment restrictions,

$$E(\eta_{jt}^{\mu} \mid c_{jt}, l_{jt}, \bar{r}_t) = 0$$

to identify the constant of integration $\mathcal{C}(c_{jt}, l_{jt})$ as well as μ , and hence the level of mark up.

6.3 Estimation of Productivity

The revenue production function after correcting for deviation between firm's own price and industry price deflator is

$$\left(1 - \frac{1}{\sigma}\right) f(c_{jt}, l_{jt}, m_{jt}) + \frac{1}{\sigma} \bar{r}_t + \mathcal{X}_{jt} + \left(1 - \frac{1}{\sigma}\right) \omega_{jt} + \tilde{\epsilon}_{jt}$$

Write $f(c_{jt}, l_{jt}, m_{jt})$ as $x_{jt}\beta$,

$$\frac{\sigma - 1}{\sigma} \omega_{jt} = r_{jt} - \left(\frac{\sigma - 1}{\sigma}\right) x_{jt}\beta - \frac{1}{\sigma} \bar{r}_t$$

Here we ignore the error term and demand shock, both of them are merged with the productivity term.

$$\omega_{jt} = \left(\frac{\sigma}{\sigma - 1}\right) r_{jt} - x_{jt}\beta - \frac{1}{\sigma - 1} \bar{r}_t$$

6.4 Estimating the contribution of resource allocation to productivity

This subsection explains the methodology used to estimate the contribution of resource allocation to more productive firms in aggregate productivity growth. We estimate the contribution of resource allocation to aggregate productivity growth in general and the contribution of resource allocation to exporting firms. The methodology used for these estimations are explained below.

We construct an index of productivity P by subtracting productivity of the average firm in the first year from ω_{jt} as follows.

$$P_{jt} = \omega_{jt} - \bar{\omega}_1$$

Where $\bar{\omega}_1 = \frac{1}{N} \sum_{j=1}^N \omega_{j1}$ and N is the number of observation in an industry in 1992-93.¹⁵ This methodology has been used in several studies such as Pavcnik (2002). It ensures that productivity index has desired properties such as transitivity and insensitivity to the units of measurement. We estimate industry level productivity (WP) as weighted average of firms' productivity (P_{jt}), as follows.

$$WP_t = \sum_{j=1}^N S_{jt} P_{jt}$$

Where S_{jt} is the share of j^{th} firm in the total output of the industry in year t . There is a useful relationship between weighted average productivity, WP_t , and simple average productivity or unweighted productivity, UP_t . This relationship is used to estimate the contribution of resource allocation to aggregate productivity growth. As shown below, the difference between the two is the *total covariance* between productivity and output share across

firms, which is a measure of aggregate productivity accounted by the greater allocation of resources to more productive firms. This relationship has been used by Pavcnik (2002) to estimate aggregate productivity growth due to resource allocation.

$$WP_t = UP_t + \sum_{j=1}^N (S_{jt} - S_t)(P_{jt} - P_t)$$

where

$$UP_t = \frac{1}{N} \sum_{j=1}^N P_{jt} = P_t$$

$$S_t = \frac{1}{N} \sum_{j=1}^N S_{jt}$$

If more productive firms get large share of the total output of the industry, as one expects in a market economy, the covariance term would get a positive value. Given the increasing role assigned to market forces in resources allocation over time owing to the liberalisation process, we expect that the covariance term would not only be positive, but also increases over time.

6.4.1 Estimating productivity contribution of resource allocation to exporters

For estimating the contribution of resource allocation to exporters, we adopted the same methodology described above. The details are given below. Let the number of exporting and non-exporting firms in the sample in the year t respectively are N_1 and N_2 , then the weighted productivity of the exporters (WP^X) and non-exporters (WP^{NX}) are defined as follows.

$$WP_t^X = \sum_{j=1}^{N_1} S_{jt}^X P_{jt}^X$$

$$WP_t^{NX} = \sum_{j=1}^{N_2} S_{jt}^{NX} P_{jt}^{NX}$$

where S^X is the share of the j^{th} exporting firm in the total output of all the exporting firms in the year t . S^{NX} is defined similarly. Now defined the simple average productivity of exporting and non-exporting firms in the year t as follows

$$UP_t^{X,NX} = \frac{WP_t^X + WP_t^{NX}}{2}$$

The contribution of resource allocation to exporters to aggregate productivity (COV^X) is defined as follows,

$$COV_t^X = WP_t - UP_t^{X,NX}$$

COV^X gives a measure of the change in the industry level productivity due to resource allocation to exporting firms.

6.5 Testing for catching up and productivity convergence

We test for the presence of catching up and consequent convergence in productivity across firms and the role of trade liberalisation and industry and firm specific factors in this process. For this we use the framework developed by Bernard and Jones (1996), which is explained below.

Let productivity of firm j is evolving according to

$$\omega_{jt} = \gamma_j + \lambda[\omega_{1t-1} - \omega_{jt-1}] + \omega_{jt-1} + \epsilon_{jt} \quad (15)$$

Where γ is asymptotic rate of productivity growth and λ parametrising the speed of catch up denoted by $(\omega_{1t} - \omega_{jt})$, where ω_{1t} is the highest productivity performance in the industry. Here, catch up is parametrised as a function of one year lagged gap between firm's productivity and the best productivity in the industry. Let $(\omega_{1t} - \omega_{jt}) = \theta_{jt}$, we can re-write the difference equation (15) to yield

$$\hat{\omega}_{j,t,T} = -\frac{1 - (1 - \lambda)^T}{T} \hat{\omega}_0 + \frac{1}{T} \sum_{s=0}^{T-1} (1 - \lambda)^{T-s} (\omega_t - \gamma_1 + \hat{\epsilon}_{js}) \quad (16)$$

Where $\hat{\omega}$ is the average productivity growth rate of firm j relative to that of the best productivity firm between time zero and T and $\hat{\epsilon}_{jt} = (\epsilon_{jt}/\epsilon_{1t})$. From model (16) we derive the following estimable model.

$$g_{jt,t+s} = \beta_0 + \beta_1 \omega_{jt} + u_i \quad (17)$$

Where $g_{jt,t+s} = \frac{\omega_{j,t+s} - \omega_{jt}}{s}$ is the average annual growth rate of productivity between t and $t + s$ years. If the catching up hypothesis is valid, β_1 gets a negative significant value. This implies that firms having lower productivity in the initial year record higher productivity growth in the subsequent years and thereby catch up with the best productivity level. The speed of convergence λ is estimated using the following relationship.

$$\theta_1 = -\frac{1 - (1 - \lambda)^T}{T} \quad (18)$$

The speed of convergence in this framework can be interpreted as the rate at which productivity level is converging to the best productivity level, which may itself be growing over time. As we have discussed above, the productivity catching up process among firms can also be influenced by the industry specific and firm specific variables and variables related to trade liberalisation. Therefore, the model (17) is augmented by including variables that capture competition from import, domestic market structure, and the position of the firm in the initial productivity distribution. The augmented model is given below and this is what we estimate in the next section.

$$\dot{\omega}_{jlt} = \alpha_0 + \alpha_1 \omega_{j0} + \alpha_2 \overline{IPR}_{it} + \alpha_3 \overline{HI}_{it} + \alpha_4 \overline{IPR}_{it} \times D_{25} + \alpha_5 \overline{IPR} \times \overline{HI} + \alpha_6 \overline{IPR} \times \overline{HI} \times D_{25} + \alpha_7 T + u_{jlt} \quad (19)$$

where $\dot{\omega}_{jlt}$ is the average productivity growth of the j^{th} firm in i^{th} industry during t and $t - 5$ years, IPR is the average import penetration rate during t and $t - 5$ years. HI is the average of the Herfindhal Index of market concentration in the domestic market during t and $t - 5$ years. The variable D_{25} is a dummy variable capturing the position of the firm in the distribution of initial productivity (ω_{j0}). It takes value one if the initial productivity of the firm belongs to the first quartile of the distribution, otherwise zero. T is the time trend to capture the trend in the growth rate of productivity.

7. Data

The study uses firm level panel data obtained from Centre for Monitoring Indian Economy (CMIE) for the period 1992-93 to 2005-06. We would like to make following points on the choice of the period of analysis. As the study is focusing on the firm dynamics after the liberalisation, the period of analysis should start immediately after the beginning of the liberal policy regime. As shown by Sengupta and Roy (2018) and Das (2016), by the year 2000 liberalisation of trade in terms of reduction or removal of tariff and non-tariff barriers were almost complete. These authors argue that from the year 2000 onwards, India focused on export promotion and simplification of the trade procedures. In this sense the study covers a period of time beyond the phase of significant policy reforms as far as manufacturing sector is concerned. Therefore, it is presumed that study period up to 2005-06 is sufficient to elicit the manufacturing firms' response to liberalisation.

After editing the data for possible erroneous and outlier observations, the

sample consists of 31129 observations on 3065 firms. The firms are grouped according to the two-digit classification of NIC 1998. Appendix A.1 provides the details on the data cleaning procedures adopted. Tables A.2 and A.3 in the Appendix respectively report industry-wise and year-wise distribution of the sample. The industries, (1) Tobacco Products (NIC-16), (2) Leather and Leather Products (NIC-19), (3) Wood and Wood Products (NIC-20), (4) Publishing and Printing (NIC-22), and (5) Coke and Petroleum Products (NIC-23) were clubbed together under the title “Other Industries” due to insufficient number of observations in each industry. In each year, the sample firms cover around 60 per cent of the total output reported in Annual Survey of Industries (ASI). The panel is unbalanced. Since it is not mandatory for firms to report their balance sheet to the data collecting agency, firms’ entry and exit from the sample is primarily related to reporting rather than their actual entry or exit from the industry. Obviously, some of the entry and exit can be of this type; however, with the available information we are unable to identify this type of entry and exit separately. Apart from firm level data, data obtained from Annual Survey of Industries (ASI) are also used. The study also uses trade data obtained from COMTRADE through World Integrated Trade Solutions Software (WITS) for estimating the import penetration ratio at three digit industry level of NIC 1998. The subsection A.2 presents the details of construction of variables used in this study.

8. Estimation and results

We use Gandhi, Navarro, and Rivers (2020a) method to estimate the production function. The estimation procedure consists of two steps. First share regression is estimated and then the constant of integration C and the Markov process $h()$. For estimating the share regression, a complete polynomial of degree two in c , l and m is specified. We use the sum of the squared residuals as $\sum_{jt} \epsilon_{jt}^2$ the objective function. The estimator would solve the following,

$$\min_{\gamma'} \sum_{jt} \left\{ s_{jt} - \ln \left(\frac{\gamma'_0 + \gamma'_c c_{jt} + \gamma'_l l_{jt} + \gamma'_m m_{jt} + \gamma'_{cc} c_{jt}^2 + \gamma'_{ll} l_{jt}^2 + \gamma'_{mm} m_{jt}^2 + \gamma'_{cl} c_{jt} l_{jt} + \gamma'_{cm} c_{jt} m_{jt} + \gamma'_{lm} l_{jt} m_{jt}}{\gamma'_{mm} m_{jt}^2 + \gamma'_{cl} c_{jt} l_{jt} + \gamma'_{cm} c_{jt} m_{jt} + \gamma'_{lm} l_{jt} m_{jt}} \right) \right\}^2 \quad (20)$$

The constant of integration is also approximated by a polynomial of degree two in c and l and we used the appropriate moment conditions to estimate the parameters. The estimated parameters of the production function and the parameter¹ is reported in Table A.4. The z values of the estimated coefficients are based on bootstrap standard errors and the number of replications is 250.

8.1 Results: Resource Allocation and Productivity

We estimate firm level productivity using the procedure explained in sub-section 6.3 and then estimate the contribution of resource allocation to industry level productivity growth using the procedure described in sub-section 6.4. For reporting the productivity growth during the period of analysis, the entire period has been divided into three sub-periods, based on the growth rate of organised manufacturing sector (see Table A.1 in the Appendix for the growth rate of manufacturing sector in terms of value added.). The three sub-periods are (I) 1993-94 to 1995-96, (II) 1996-97 to 2000-01, and (III) 2001-02 to 2005-06. Further, as described above, estimated productivity values are normalised using the average industry productivity in the year 1992-93. Therefore, the average productivity numbers in the subsequent years indicate the growth rate of average industry productivity over the year 1992-23. Results on productivity and resource allocation are reported in Tables 2, 3 and 4.

In Table 2, WP is the weighted average productivity growth of the each industry and COV is the contribution of the share reallocation to WP. For instance in Food Products industry, the weighted average productivity (WP) in period III is 17.9 percent higher than that in 1992-93. And this entire growth is contributed by resource allocation to more productive firms. In this industry in all sub-periods, the unweighted average productivity (UP) declined, suggesting fall in within firm productivity. This fall in UP is offset by efficient resource allocation in all the sub-periods, allowing the WP productivity to record positive growth rates.

Table 2: Productivity Growth and Resource Allocation						
Period	WP	UP	COV	WP	UP	COV
	Food Products			Textiles		
I	0.088	-0.073	0.161	0.070	-0.048	0.117
II	0.183	-0.168	0.350	0.221	-0.021	0.242
III	0.179	-0.161	0.341	0.523	0.177	0.346
	Paper & Paper Products			Chemicals		
I	0.124	-0.045	0.169	0.061	-0.136	0.197
II	0.130	-0.004	0.133	0.275	-0.235	0.511
III	0.235	0.052	0.183	0.552	-0.142	0.695
	Rubber & Plastic Products			Non-metallic Mineral Products		
I	0.063	0.014	0.050	-0.011	-0.026	0.016
II	0.238	0.203	0.035	0.129	0.073	0.056
III	0.550	0.407	0.142	0.164	0.076	0.088
	Basic Metals			Metal Products		
I	0.075	0.069	0.006	0.180	0.066	0.114
II	0.096	0.053	0.043	0.315	0.113	0.202
III	0.159	0.091	0.068	0.595	0.213	0.382
	Non-electrical Machinery			Computing Machinery, TV & Radio		
I	0.125	0.155	-0.030	0.184	0.178	0.006
II	0.160	0.098	0.062	0.464	0.336	0.128
III	0.262	0.159	0.103	0.537	0.320	0.217
	Electrical Machinery			Medical Precision & Optical		
	Instruments					
I	0.096	0.075	0.020	0.233	0.325	-0.092
II	0.211	0.161	0.051	0.383	0.440	-0.058
III	0.293	0.197	0.096	0.458	0.521	-0.062
	Transport Equipment & Parts			Other Industries		
I	0.160	0.147	0.013	0.039	0.161	-0.122
II	0.240	0.198	0.042	-0.054	0.033	-0.087
III	0.469	0.364	0.105	-0.236	-0.036	-0.200

Note: Period:(I)1993-4 to 1995-96, (II)1996-97 to 2000-01 and (III)2001-02 to 2005-06

Table 2 shows that all the industries except “Other Industries” recorded positive productivity growth in all the sub-periods. Second, in all the industries, except Non-electrical Machinery, Medical Precision and Optical

Instruments and Other Industries, COV is positive in all the sub-periods. In Non-electrical Machinery, COV is negative only in period one. This suggests that in large majority of the industries in all the sub-periods, resource allocation to more productive firms contributed to higher productivity.

Also note that in all the industries, except Medical Precision and Optical Instruments and Other Industries, the value of COV is increasing overtime. This may be suggesting the increasing importance of resource allocation overtime to productivity growth. Table 4 presents the contribution of COV to WP in percentage. Find that in all the industries, except Medical Precision and Optical Instruments and Other Industries, the relative contribution of COV to WP is significant. For instance in Food Products 190 percent of average productivity in period III is contributed by resource allocation to more productive firms, thus offsetting the decline in within firm productivity. Similarly, in Non-electrical Machinery, 39 per cent of growth in productivity in period III is contributed by improved resource allocation. In majority of the industries, at least in one period, COV accounts for a larger share of growth in productivity. Further, in five industries the relative contribution of resources allocation to aggregate productivity growth is increasing over time. Overall, the results suggest that improved resource allocation has contributed to the productivity growth of Indian manufacturing industry during the period of analysis.

Table 3 presents the productivity growth of exporters (WP^X) and non-exporters (WP^{NX}) and the contribution of resource allocation to exporting firms to aggregate productivity growth (WP). The Tables shows that in six industries exporting firms recorded higher productivity growth than non-exporting firms in all the sub-periods.¹⁶ In nine industries exporters recorded higher productivity level during the third sup-period than non-exporters. The higher productivity growth of the exporters suggests that resource allocation to exporting firms can also contribute to aggregate productivity growth (WP). COV^X in Table 3 reports the contribution of resource allocation to exporting firms to WP. The Table shows that in ten industries COV^X is positive in sub-period III. The column COV^X in Table 4 reports the contribution of COV^X to WP in percent. It shows that in Paper and Paper Products, Non-electrical Machinery, Transport Equipment and Parts, Chemicals, Non-metallic Mineral Products the contribution of resource allocation to exporting firms is significant.

Table 3: Productivity Growth: Exporters Vs Non-exporters

Period	WP ^x	WP ^{Nx}	COV ^x	WP ^x	WP ^{Nx}	COV ^x
	Food Products			Textiles		
I	0.082	0.117	-0.012	0.056	0.010	0.037
II	0.259	0.112	-0.003	0.174	0.438	-0.086
III	0.202	0.126	0.015	0.515	0.354	0.088
	Paper & Paper Products			Chemicals		
I	0.158	-0.010	0.050	0.088	-0.032	0.033
II	0.213	-0.045	0.045	0.317	0.116	0.059
III	0.268	0.147	0.027	0.610	0.307	0.094
	Rubber & Plastic Products			Non-metallic Mineral Products		
I	0.068	0.171	-0.056	-0.004	-0.038	0.011
II	0.237	0.365	-0.063	0.149	-0.003	0.056
III	0.556	0.373	0.085	0.184	0.033	0.056
	Basic Metals			Metal Products		
I	0.074	0.086	-0.005	0.184	0.371	-0.097
II	0.101	0.077	0.007	0.325	0.506	-0.100
III	0.160	0.106	0.026	0.619	0.605	-0.017
	Non-electrical Machinery			Computing Machinery, TV, Radio		
I	0.133	0.050	0.034	0.187	0.182	0.000
II	0.163	0.085	0.036	0.474	0.480	-0.013
III	0.263	0.195	0.033	0.536	0.669	-0.066
	Electrical Machinery			Medical Precision & Optical		
	Instruments					
I	0.099	0.081	0.005	0.172	0.524	-0.115
II	0.197	0.258	-0.016	0.322	0.615	-0.086
III	0.280	0.349	-0.022	0.389	0.857	-0.164
	Transport Equipment & Parts			Other Industries		
I	0.168	0.007	0.073	0.039	-0.009	0.024
II	0.241	0.113	0.063	-0.062	-0.049	0.002
III	0.478	0.278	0.091	-0.237	-0.250	0.007

Note: Period:(i)1993-4 to 1995-96, (ii)1996-97 to 2000-01 and (iii)2001-02 to 2005-06.

8.2 Results on productivity catch up

The regression model (19) is estimated using fixed effect and random effect methods, and Hausman test is used to select the valid estimates. Table 5 reports the results. Hausman tests shows that fixed effect estimates are valid and hence we use these estimates for discussion in the following paragraphs.

The results shows that initial productivity has significant negative effect on subsequent productivity growth, providing strong evidence on productivity convergence among firms. The coefficient of the initial productivity implies a convergence rate 0.504; suggesting that on average 50 percent of the difference between firm's productivity and best productivity level is covered every year. Import penetration rate and Herfindhal Index of market concentration have significant negative effect on productivity growth. The negative effect of IPR supporting the theoretical result that import liberalisation is actually reducing firms' scale of operation and thereby dampening their incentive to invest in productivity enhancing activities.

Herfindhal index of market concentration has negative impact on productivity suggesting that market concentration is actually reducing firms' incentive to undertake productivity enhancing investments as argued in some the theoretical literature. However, it is interesting to note that the interaction between Herfindhal index and Import Penetration Rate is positive and significant. It suggests that firms operating in highly concentrated industries when subjected to competition from import have greater incentive to improve their productivity. This results supports the general theoretical results of Smulders and Klundert (1995), Traca (2002), Aghion, Harris, Howitt, and Vickers (2001) and Aghion (2003) that firms having larger market share, when subjected to heightened competition, have greater incentive to enhance their productivity. Another interesting results is related to the role of firms' position in the productivity distribution. The coefficient of the variable $IPR \times D_{25}$ is negative and significant, suggesting when firms having relatively low initial productivity perform badly when subjected to import competition. However, this results get reversed when the firms having relatively lower productivity also operate in a concentrated market, as shown by the significant positive coefficient of the interaction variable $IPR \times HI \times D_{25}$. This results also modifies the theoretical results of Perla, Tonetti, and Waugh (2021) that firms occupying lower part of the productivity distribution have greater incentive to invest in productivity enhancing activities, when they are subjected import competition. The results of this study implies that firms having relatively lower productivity should also have a larger market share to invest in productivity enhancing activities when they are subjected to competition from import.

Table 4: Contribution of Resource Allocation to Productivity
(in %)

Period	COV	COV $\times$	COV	COV $\times$
	Food Products		Textiles	
I	182.813	-13.055	168.09	52.85
II	191.782	-1.638	109.58	-38.76
III	190.131	8.496	66.19	16.83
	Paper & Paper Products		Chemicals	
I	136.6267	40.50	323.79	54.30
II	102.7267	35.05	185.49	21.42
III	78.0038	11.65	125.76	16.95
	Rubber & Plastic Products		Non-metallic Mineral Products	
I	78.688	-88.686	146.40	99.22
II	14.853	-26.542	43.20	43.53
III	25.925	15.473	53.94	33.89
	Basic Metals		Metal Products	
I	7.352	-7.230	63.41	-54.19
II	45.001	7.616	64.03	-31.82
III	43.041	16.470	64.24	-2.81
	Non-electrical Machinery		Computing Machinery, TV, Radio	
I	-23.9547	26.84	3.48	0.21
II	38.8933	22.68	27.54	-2.81
III	39.2688	12.69	40.39	-12.30
	Electrical Machinery		Medical Precision & Optical Instruments	
I	21.2479	5.43	-39.64	-49.55
II	23.953	-	-15.04	-22.52
	7.710			
III	32.707	-	-13.61	-35.85
	7.411			
	Transport Equipment & Parts		Other Industries	
I	8.0052	45.40	-309.68	61.62
II	17.3616	26.32	161.51	3.60
III	22.4343	19.43	84.70	2.88

Notes: (1) The figures are in per cent. (2) The column COV reports the contribution of COV to industry productivity. (3) The column COV_x reports the contribution of COV_x to industry productivity.

The trend variable is significant and positive suggesting that productivity growth in general has an upward trend during the period of analysis. The variable Period Dummy, which take value one for the period 2 and zero in

period 3 is significant and positive, suggesting favourable impact of period two on productivity, though growth rate of the manufacturing sector is lower during this period.

Table 5: Productivity Catch-up: Determinants

Variables	Fixed Effect	Random Effect
Initial Productivity (ω_{j0})	-0.194** (-64.33)	-0.0878** (-32.43)
$\overline{IPR}$	-0.0263** (-4.08)	-0.00668 (-0.90)
$\overline{HI}$	-0.110* (-2.15)	0.0847** (4.38)
$\overline{IPR} \times \overline{HI}$	0.106** (4.39)	0.0857** (2.94)
$\overline{IPR} \times D_{25}$	-0.268** (-12.25)	-0.349** (-13.42)
$\overline{IPR} \times \overline{HI} \times D_{25}$	0.798** (7.33)	1.013** (8.87)
Time trend	0.00317** (7.84)	-0.000482 (-1.17)
Period Dummy	0.0124** (8.68)	0.00902** (5.47)
Constant	0.935** (64.42)	0.434** (34.68)
Hausman Test (χ^2)	12075.90 (0.00)	
No. Observations	15465	

Notes: (1) Dependent variable is the average productivity growth between t and $t - 5$ time points. (2) t values, based on standard errors robust to heteroscedasticity and autocorrelation, are given in parentheses below the coefficients; ** and

* indicate significant respectively at 1 percent and 5 percent. For Hausman test, P -value is given in parenthesis below the χ^2 value.

9. Concluding remarks

Against the background of industrial and trade policy reforms introduced in India since 1991, this study examined the role of resource allocation and catching up in growth in productivity. The study also examined the importance of industry specific and firm specific factors in shaping firms' productivity response to increased competition from import. The study used an estimation framework that corrects for the omitted price bias while estimating firm level productivity. This correction is required to get unbiased estimates of aggregate productivity impact of resource allocation on productivity growth. The study also examined the importance of industry specific and firm specific factors in deciding the productivity impact of trade liberalisation. The study used firm level panel data from CMIE for the period 1992-93 to 2005-06. It used production function approach to estimate productivity at the firm level. For estimating the production function, the study used a recently developed methodology to estimate it consistently and the methodology also accommodates imperfect competition in the product market.

The study estimated the contribution of resource allocation in general and resource allocation between exporting and non-exporting firms to aggregate productivity. The analysis is done by dividing the entire period of study into three sub-periods. Productivity growth rates show that all the industries except one recorded positive productivity growth during all the sub-periods. It also shows that resource allocation from less productive to more productive firms contributed to significantly to the growth in productivity during the post reform period. Further, in all the industries, except in one, the contribution of resource allocation to productivity growth is increasing over time. The study also examined the contribution of resource allocation to exporters. The findings suggest that compared to the non-exporting firms, in six industries exporting firms recorded higher productivity growth in all the sub-periods and in sub-period III nine industries recorded higher productivity. This suggests that allocation of resources to exporting firms increases aggregate productivity. The results of the study shows that same.

The study also examined the role factors specific to industry and firm in deciding firms' response to import competition. The results shows that the positive impact of import competition on productivity depends on the market structure of the industry. If the market structure is concentrated, import competition increases productivity. Findings also suggests that the productivity enhancing effect of import competition on firms occupying lower part of the initial productivity distribution depends upon if they are operating in a concentrated market or not. Overall, these results give greater insights into the domestic factors that decide the productivity effect of import competition.

A Appendix

A.1 Data Editing Rules

We adopted following set of thumb rules to identify and remove possible erroneous and outlier observations.

1. First, we dropped all those observations having zero or negative values for value added.
2. Second, we adopted a resistant rule to identify outlier observations. Let $\ln(Q/C)$, $\ln(Q/L)$, $\ln(Q/M)$, and $\ln(Q/E)$ respectively denote the log of ratio of output (Q) to capital(C), labour (L), raw-material ex- pense (M) and energy expenditure (E)(all variables are in constant price). We removed all those observations having values for these variables above or below two times inter-quartile range from median at three digit level of NIC.

A.2 Construction of Variables

Output (Q): The value of output is in 1993-1994 prices obtained by deflat- ing reported nominal value. In order to mitigate the discrepancy between unobserved firm level price deflator and industry level price deflator, price indices used are at a more disaggregate industry level than that reported in Table 1.

Capital (C): We use Gross Fixed Asset (GFA) expressed in 1993-94 prices. The data base reports GFA in historical costs. We use perpetual inventory method to estimate GFA at constant prices for each year as shown below.

$$P_t C_t = \frac{P_t}{P_{t-1}} P_{t-1} C_{t-1} + P_t I_t$$

Where P , C and I respectively refer to price of capital, gross physical capital and gross investment obtained from the relation, $I_t = (GFA_t - GFA_{t-1})$. This method to estimate capital stock requires the base year capital stock C_{t-1} valued at replacement cost. We choose 1999-2000 as the base year as it has the largest number of observations and revalued GFA of this year into replacement cost using a revaluation factor. As Srivastava (1996) gives details on the estimation of revaluation factor, the discussion is kept brief here. It is based on the following assumptions:

1. Based on CSO (1989), we assume that life of a plant and machinery is 20 years. This implies that no firm has any capital stock in 1999-2000 of vintage earlier than 1980-81. For firms incorporated after 1980-81, we assume that the earliest vintage of capital in their capital mix dates back to the year of

incorporation. Indeed, year of incorporation and the vintage of oldest capital in the firm's asset mix may not coincide for some firms, but the assumption is made for want of a better alternative.

2. The price of capital changed at a constant rate from 1980-81 or from the date of incorporation of the firm (which ever is later) up to 1999- 2000. Growth rates of price of capital were obtained from the price indices for fixed capital formation in manufacturing collected from various issues of the *National Account Statistics* of India. The growth rate of capital price is not firm specific but it varies with the year of incorporation, provided the firm was incorporated after 1980-81.
3. The growth rate of investment is same for all firms. Here growth rate of gross fixed capital formation in manufacturing in 1980-81 prices is assumed to apply to all firms. Again, different average annual growth rates were applicable for firms established after 1980-81.

GFA at replacement cost is obtained by multiplying the balance sheet value of GFA in 1999-2000 by the revaluation factor. This revalued GFA was deflated to obtain the real capital stock for that year. We used price index for Machinery and Machine Tools for deflation, as plant and machinery account for 71.5 per cent of GFA (Reserve Bank of India, 1990:103).

Note that we use gross value of capital. Some of the recent studies in Indian context used gross value rather than net value of capital for purposes similar to the present study and Balakrishnan, Parameswaran, Pushpangadan, and Suresh Babu (2006) provides rationale of this.

Materials (M): The input materials include both raw materials consumed and the energy consumed during the process of production. First, cost of raw materials in 1993-94 prices is obtained by deflating the nominal cost using raw material price indices. Raw material price indices were constructed at more disaggregated level than that reported in Table 1 using weights obtained from the *Input-Output Transaction Table of India for 1999-2000* published by the Central Statistical Organisation (CSO) and appropriate price indices taken from the *Index Numbers of Wholesale Prices in India, base 1993-94= 100*. Energy costs also prepared in 1993-94 prices, obtained by deflating the reported nominal cost using an energy price index. The energy price index was constructed using weights obtained from the *Input-Output Transaction Table of India for 1993-94* and appropriate price indices

Labour(L): It is the number of workers employed, obtained by dividing firms' reported total wage bill by the average annual wage for a worker in the three digit industry of ASI. If the variation in the wage bill across firms reflects the

variation in quality of labour employed, our measure of labour input would capture labour quality also.

Import Penetration Rate (IPR): . IPR is the ratio of import to total domestic production at the three digit level of NIC 1998. It is defined as follows.

$$IPR = \frac{\text{Import}}{\text{Domestic Production} + \text{Import-Export}}$$

Value of import and export at three digit level of NIC 1998 were obtained from COMTRADE through World Integrated Trade Solutions System (WITS). Total domestic production at three digit level of NIC 1998 were obtained from Annual Survey of Industries.

Herfindhal Index (HI): Herfindhal Index of domestic concentration at three digit level of NIC-1998 is computed as follows

where S_j is the share of the of the j^{th} firm in the total domestic sales of all firms in the industry. Domestic sales is arrived at by subtracting exports from the sales of each firm.

Table A.1: Growth rate of Value Added in Manufacturing

NIC Code	Industry	1991-92 to 1995-96	1996-97 to 2000-01	2001-02 to 2005-06
15	Food and Food Products	10.0	2.1	-0.9
17 + 18	Textiles	7.4	5.6	4.6
21	Paper and Paper Products	8.4	9.6	6.2
23	Coke and Petroleum Products	21.4	-2.8	21.4
24	Chemical and Chemical Products	17.1	2.7	8.6
25	Rubber and Plastic Products	9.8	6.8	6.4
26	Non-metallic Mineral Products	3.2	11.6	10.2
27	Basic Metals	21.2	-0.2	28.1
28	Metal Products	14.4	3.2	10.6
29	Non-Electrical Machinery	14.9	-0.5	6.4
30+32	Computing Machinery, TV, Radio and Electronic equipment	11.9	0.1	8.2
31	Electrical Machinery	13.2	2.0	8.5
33	Medical precision and Optical Instruments	22.7	0.1	11.1
34+35	Transport Equipment and Pars	20.6	-7.1	25.9
16+19+20+22	Other Industries	5.6	-2.3	4.0
	All Manufacturing	13.8	1.6	11.7

Notes: Annual average Growth rate in %

Table A.2: Industry-wise distribution of the Sample

	Industry	No.
1	Food Products (15)	3377
2	Textiles (17+18)	3818
3	Paper and Paper Products (21)	1046
4	Chemicals (24)	7235
5	Rubber and Plastic Products (25)	2170
6	Non-metallic MineralProducts (26)	1565
7	Basic Metals (27)	2733
8	Metal Products (28)	855
9	Non-electrical Machinery (29)	2424
10	Computing Machinery, TV, Radio (30+32)	1065
11	Electrical Machinery (31)	1171
12	Transport Equipment and Parts (34+35)	2280
13	Medical Precision and Optical Instruments(33)	298
14	Other Industries(16,19,20,22 and 23)	1092

Notes: NIC-1998 2 digit numbers are given in parenthesis.

Table A.3: Year-wise Distribution of the Sample

Year	Observations	Share of ASI	Exporters
1992-93	1335	59.4	825
1993-94	1669	60.5	1023
1994-95	1965	66.2	1218
1995-96	2149	65.2	1367
1996-97	2298	70.3	1437
1997-98	2458	61.4	1529
1998-99	2663	62.6	1606
1999-00	2836	65.1	1635
2000-01	2676	65.6	1606
2001-02	2508	62.8	1543
2002-03	2369	58.2	1491
2003-04	2252	58.1	1467
2004-05	2116	66.5	1410
2005-06	1835	66.1	1278

Table A.4: Production Function Estimates

Variable	Coef.	z value	Coef.	z value
Food Products			Textiles	
c	0.099	7.39	0.15	12.07
l	0.282	11.48	0.183	7.85
m	0.683	62.21	0.673	101.41
$\frac{1}{\sigma}$	0.342	13.53	0.311	17.86
Paper & Paper Products			Chemicals	
c	0.100	4.68	0.123	12.68
l	0.222	4.56	0.287	14.26
m	0.712	67.3	0.656	113.37
$\frac{1}{\sigma}$	0.273	10.1	0.379	16.83
Rubber & Plastic Products			Non-metallic Mineral Products	
c	0.177	10.03	0.234	11.71
l	0.041	1.18	0.252	4.99
m	0.673	80.88	0.556	34.83
$\frac{1}{\sigma}$	0.147	6.95	0.074	3.65
Basic Metals			Metal Products	
c	0.093	8.74	0.137	2.87
l	0.153	6.2	0.313	3.41
m	0.751	85.9	0.575	20.06
$\frac{1}{\sigma}$	0.032	3.74	0.253	8.55
Non-electrical Machinery			Computing Machinery, TV, Radio	
c	0.155	11.11	0.091	3.21
l	0.352	15.65	0.279	8.18
m	0.575	68.34	0.636	44
$\frac{1}{\sigma}$	0.106	4.81	0.139	4.89
Electrical Machinery			Medical Precision and Optical Instruments	
c	0.158	10.31	0.069	1.21
l	0.261	7.42	0.375	3.62
m	0.664	60.72	0.558	24.16
$\frac{1}{\sigma}$	0.113	6.34	0.073	1.07
Transport Equipment & Parts			Other Industries	
c	0.187	10.54	0.199	6.51
l	0.256	9.32	0.357	6.84
m	0.604	63.29	0.573	29.06
$\frac{1}{\sigma}$	0.074	6.38	0.064	4.94

Note: (1). z values based on bootstrap standard errors are reported; number of replications is 250.

References

- ACKERBERG, D. A., K. CAVES, AND G. FRAZER (2015): "Identification Properties of Recent Production Function Estimators," *Econometrica*, 83(6), 2411–2451.
- AGHION, P. (2003): "Empirical estimates of the relationship between product market competition and Innovation," in *Institutions, Innovations and Growth: Selected Economic Papers.*, ed. by J. Touffut, pp. 142–169. Edward Elgar, UK.
- AGHION, P., C. HARRIS, P. HOWITT, AND J. VICKERS (2001): "Competition, Imitation and Growth with Step-by-Step Innovation," *Review of Economic Studies*, 68, 467–492.
- AW, B. Y., S. CHUNG, AND M. J. ROBERTS (2003): "Productivity, Output, and Failure: A Comparison of Taiwanese and Korean Manufacturers," *The Economic Journal*, 113(491), F485–F510.
- BALAKRISHNAN, P. (2010): *Economic Growth in India: History and Prospect*. Oxford University Press, New Delhi.
- BALAKRISHNAN, P., M. PARAMESWARAN, K. PUSHPANGADAN, AND M. SURESH BABU (2006): "Liberalization, Market Power, and Productivity Growth in Indian Industry," *Journal of Policy Reform*, 9(1), 55–73.
- BANGA, R., AND B. GOLDAR (2007): "Contribution of Services to Output Growth and Productivity in Indian Manufacturing: Pre- and Post-Reforms," *Economic and Political Weekly*, 42(26), 2769–2777.
- BERNARD, A. B., AND J. B. JENSEN (1999): "Exceptional exporter performance: cause, effect, or both?," *Journal of International Economics*, 47(1), 1–25.
- BERNARD, A. B., J. B. JENSEN, AND S. P. K. (2006): "Trade costs, firms and productivity," *Journal of Monetary Economics*, 53, 917–937.
- BERNARD, A. B., AND C. I. JONES (1996): "Comparing Apples to Oranges: Productivity Convergence and Measurement Across Industries and Countries," *American Economic Review*, 86(5), 1216–1238.
- BHAGWATI, J. N. (1988): "Export-Promoting Trade Strategy: Issues and Evidence," *The World Bank Research Observer*, 3(1), 27–57.
- BHAGWATI, J. N., AND P. DESAI (1970): *India: planning for industrialization*. Oxford University Press.
- BHAGWATI, J. N., AND T. N. SRINIVASAN (1975): *Foreign trade regimes and economic development: India*. National Bureau of Economic Research : Distributed by Columbia University Press.
- BHATTACHARJEA, A. (2023): "Industrial policy in India since independence," *Indian Economic Review*, pp. 1–34.
- BOLLARD, A., P. J. KLENOW, AND G. SHARMA (2013): "India's mysterious manufacturing miracle," *Review of Economic Dynamics*, 16(1), 59–85.
- CHAND, S., AND K. SEN (2002): "Trade Liberalization and Productivity

- Growth: Evidence from Indian Manufacturing," *Review of Development Economics*, 6(1), 120–132.
- CHATTERJEE, A. (2012): *Industrial Policy and Economic Development in India 1947-2012*. New Century Publications.
- CLERIDES, S. K., S. LACH, AND J. R. TYBOUT (1998): "Is Learning by Exporting Important? Micro-Dynamic Evidence from Colombia, Mexico, and Morocco," *The Quarterly Journal of Economics*, 113(3), 903–947.
- CSO (1989): *National Income: Sources and Methods* Central Statistical Organisation, New Delhi.
- DAS, D. K. (2004): "Manufacturing Productivity under Varying Trade Regimes, 1980-2000," *Economic and Political Weekly*, 59, 423–433.
- (2006): "Improving Industrial Productivity: Does Trade Liberalisation Matter? Evidence from Indian Capital and Intermediate Goods sectors," in *India: Industrialisation in a Reforming Economy: Essays for K L Krishna*, ed. by S. Tendulkar, A. Mitra, K. Narayanan, and D. K. Das, pp. 321–348. Academic Foundation, New Delhi.
- (2016): "Trade Policy and Manufacturing Performance: Exploring the Level of Trade Openness in India's Organized Manufacturing in the Period 1990 – 2010," DRG Study Series 41, Department Research Group, Reserve Bank of India.
- DAS, D. K., AND G. KALITA (2011): "Aggregate Productivity Growth in Indian Manufacturing: An Application of Domar Aggregation," *Indian Economic Review*, 46(2), 275–302.
- DE LOECKER, J. (2011a): "Product Differentiation, Multiproduct Firms, and Estimating the Impact of Trade Liberalization on Productivity," *Econometrica*, 79, 1407–1451.
- (2011b): "Recovering markups from production data," *International Journal of Industrial Organization*, 29(3), 350–355.
- DEL GATTO, M., G. I. P. OTTAVIANO, AND M. PAGNINI (2008): "Openness to Trade and Industry Cost Dispersion: Evidence from a Panel of Italian Firms," *Journal of Regional Science*, 48(1), 97–129.
- DIEWERT, W. E. (1976): "Exact and Superlative Index Numbers," *Journal of Econometrics*, 4(2), 115–145.
- DIXIT, A. K., AND J. E. STIGLITZ (1977): "Monopolistic Competition and Optimum Product Diversity," *The American Economic Review*, 67(3), 297–308.
- FEENSTRA, R. C. (1994): "New Product Varieties and the Measurement of International Prices," *The American Economic Review*, 84(1), 157–177.
- GANDHI, A., S. NAVARRO, AND D. A. RIVERS (2020a): "On the Identification of Gross Output Production Functions: Article and online appendix," *Journal of Political Economy*, 128(8), 2973–3016.
- (2020b): "On the Identification of Gross Output Production Functions," *Journal of Political Economy*, 128(8), 2973–3016.
- GOLDAR, B. (2004): "Indian Manufacturing: Productivity Trends in Pre- and

- Post-Reform Periods," *Economic and Political Weekly*, 39, 5033–43.
- (2014): "Productivity in Indian Manufacturing in the Post-reform Period: A Review of Studies," in *Productivity in Indian Manufacturing: Measurements, Methods and Analysis*, ed. by V. Kathuria, R. Raj S N, and K. Sen, chap. 4, pp. 75–105. Routledge, India.
- GOLDAR, B., AND A. KUMARI (2003): "Import Liberalisation and Productivity Growth in Indian Manufacturing Industries," *The Developing Economies*, 41(4), 436–60.
- GOLDBERG, P. K., A. K. KHANDELWAL, N. PAVCNIK, AND P. TOPALOVA (2010): "Imported intermediate inputs and domestic product growth: evidence from india," *Quarterly Journal of Economics*, 125(4), 1727–1767.
- GOOD, D. H., I. M. NADIRI, AND R. C. SICKLES (1996): "Index Number and Factor Demand Approaches to Estimation of Productivity," Working Paper 5790, National Bureau of Economic Research, 1050 Massachusetts Avenue Cambridge, MA.
- GRILICHES, Z., AND J. MAIRESSE (1998): "Production Functions: The Search for Identification," in *Practicing Econometrics, Essays in Methods and Applications*, ed. by Z. Griliches, pp. 383–411. Edward Elgar, Cheltenham.
- GROSSMAN, G. M., AND E. HELPMAN (1991): *Innovation and Growth in the Global Economy*. MIT Press, Cambridge MA.
- HARRISON, A. E., L. A. MARTIN, AND S. NATARAJ (2013): "Learning versus Stealing: How Important Are Market-Share Reallocations to India's Productivity Growth?," *World Bank Economic Review*, 27(2), 202 – 228.
- HASHIM, D. A., A. KUMAR, AND A. VIRMANI (2011): "J-Curve Hypothesis of Productivity and Output Growth: A Case Study of Indian Manufacturing in the Post Reforms Period," *Indian Economic Review*, 46(1), 1–21.
- HORN, H., H. LANG, AND S. LUNDGREN (1995): "Managerial effort incentives, X-inefficiency and international trade," *European Economic Review*, 39(1), 117–138.
- KATHURIA, V., R. R. S. N, AND K. SEN (2010): "Organised versus Unorganised Manufacturing Performance in the Post-Reform Period," *Economic and Political Weekly*, 45(24), 55–64.
- KLETTE, T. J., AND Z. GRILICHES (1996): "Inconsistency of common scale estimators when output prices are unobserved and endogenous," *Journal of Applied Econometrics*, 11, 343–361.
- KRISHNA, K., B. GOLDAR, D. K. DAS, S. C. AGGARWAL, A. A. ERUMBAN, AND P. C. DAS (2022): "India Productivity Report," .
- KRISHNA, P., AND D. MITRA (1998): "Trade liberalization, market discipline and productivity growth: new evidence from India," *Journal of Development Economics*, 56(2), 447–462.
- KUMAR, N. (1994): *Multinational enterprises and industrial organization: The Case of India*. Sage Publications.
- LEVINSOHN, J., AND A. PETRIN (2003): "Estimating Production Functions

- Using Inputs to Control for Unobservables," *Review of Economic Studies*, 70(2), 317–341.
- MARTIN, L. A., S. NATARAJ, AND A. E. HARRISON (2017): "In with the Big, Out with the Small: Removing Small-Scale Reservations in India," *American Economic Review*, 107(2), 354–386.
- MELITZ, M. J. (2000): "Estimating Firm-Level Productivity in Differentiated Product Industries," Unpublished document, Department of Economics, Harvard University.
- (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1675–172
- MELITZ, M. J., AND G. I. P. OTTAVIANO (2008): "Market Size, Trade, and Productivity," *Review of Economic Studies*, 75(1), 295–316.
- MITRA, D., AND B. P. URAL (2008): "Indian manufacturing: A slow sector in a rapidly growing economy," *The Journal of International Trade and Economic Development*, 17(4), 525–559.
- OLLEY, G., AND A. PAKES (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica*, 64(6), 1263–1297.
- PANAGARIYA, A. (2007): "Why India Lags Behind China and How It Can Bridge the Gap," *World Economy*, 30(2), 229 – 248.
- PAVCNIK, N. (2002): "Trade Liberalization, Exit, and Productivity Improvements: Evidence from Chilean Plants," *Review of Economic Studies*, 69(1), 245–276.
- PERLA, J., C. TONETTI, AND M. E. WAUGH (2021): "Equilibrium Technology Diffusion, Trade and Growth," *American Economic Review*.
- RAITH, M. (2003): "Competition, Risk, and Managerial Incentives," *The American Economic Review*, 93(4), 1425–1436.
- RAJ, S. N. R., AND M. K. MAHAPATRA (2009): *On Measuring Productivity Growth in Indian Industry: Analysis of Organized and Unorganized Sector in Selected Major States* pp. 81–104. Physica-Verlag HD, Heidelberg.
- RODRIK, D. (1992a): "Closing the Productivity Gap: Does Trade Liberalisation Really Help?," in *Trade Policy, Industrialisation and Development: New Perspectives*, ed. by G. K. Helleiner. Clarendon Press, Oxford, UK.
- (1992b): "The Limits of Trade Policy Reform in Developing Countries," *Journal of Economic Perspectives*, 6(1), 87–105.
- SCHUMPETER, J. A. (1942): *Capitalism, Socialism and Democracy*. Harper & Row, New York.
- SENGUPTA, A., AND S. S. ROY (2018): "India's trade policy: Which way now?," *DECISION*, 45(2), 129–145.
- SIVADASAN, J. (2009): "Barriers to Competition and Productivity: Evidence from India," *The B.E. Journal of Economic Analysis & Policy*, 9(1), Article 42.
- SMULDERS, S., AND T. V. D. KLUNDERT (1995): "Imperfect Competition, Concentration and Growth with firm specific R&D," *European Economic Review*, 39(1), 139–160.

- SRIVASTAVA, V. (1996): *Liberalisation, Productivity and Competition: A Panel Study of Indian Manufacturing*. Oxford University Press, New Delhi.
- SRIVASTAVA, V., P. GUPTA, AND A. DATT (2001): "The Impact of India's Economic Reforms on Industrial Productivity, Efficiency and Competitiveness: A Panel study of Indian Companies - 1980-1997," Research Report, National Council of Applied Economic Research.
- TOPALOVA, P., AND A. KHANDELWAL (2011): "Trade Liberalization and Firm Productivity: The Case of India," *Review of Economics and Statistics*, 93(3), 995-1009.
- TRACA, D. A. (2002): "Import as Competitive Discipline: The role of productivity gap," *Journal of Development Economics*, 69, 1-21.
- TRIVEDI, P. (2004): "An Inter-state Perspective on Manufacturing Productivity in India 1980-81 to 2000-01," *Indian Economic Review*, 39, 203-37.
- TRIVEDI, P., L. L., R. JAIN, AND Y. K. GUPTA (2011): "Productivity, Efficiency and Competitiveness of the Manufacturing sector," DRG Study 37, DEPR, Reserve Bank of India.
- TRIVEDI, P., A. PRAKASH, AND D. SINATE (2000): "Productivity in Major Manufacturing Industries in India: 1973-74 to 1997-98," DRG Study 20, Reserve Bank of India.
- UNEL, B. (2003): "Productivity Trends in India's Manufacturing Sectors in the Last Two Decades," Working Paper WP/03/22, International Monetary Fund (IMF).
- UNNI, J., N. LALITHA, AND U. RANI (2001): "Economic Reforms and Productivity Trends in Indian Manufacturing," *Economic and Political Weekly*, 36(41), 3914-3922.
- VIRMANI, A., AND D. A. HASHIM (2009): "Factor Employment, Sources and Sustainability of Output Growth: Analysis of Indian Manufacturing," Working Paper 3/2009-DEA, Ministry of Finance, Government of India.
- VIRMANI, A., AND D. A. HASHIM (2011): "J-Curve of Productivity and Growth: Indian Manufacturing Post-Liberalization," Working Paper WP/11/163, International Monetary Fund.

Notes

- ¹ For a review of these studies see Bhattacharjea (2023).
- ² Goldar (2014) provides a detailed review of these studies
- ³ For a recent detailed reviews of evolution of industrial and trade policies of India see Bhattacharjea (2023); Sengupta and Roy (2018) respectively
- ⁴ Raith (2003) also predicts that increased competition can dampen firms' incentive to improve its performance when it results in a reduction in its scale of output.
- ⁵ This argument is also modelled without assuming backward bending supply curve, in a principal-agent framework by Horn, Lang, and Lundgren (1995).
- ⁶ Note that the underlying mechanism in this model is distinct from the standard "market size" effect, i.e., opening to trade increases the size of the market and, hence, raises the value of adoption.
- ⁷ For more details on these studies see Goldar (2014).
- ⁸ There are studies examining the productivity growth of unorganised manufacturing also. These studies include Kathuria, N, and Sen (2010), Unni, Lalitha, and Rani (2001), Raj and Mahapatra (2009) and Raj and Mahapatra (2009). The first two studies show a decline in TFP during the post reform period, while the latter two studies suggest a decline during the pre-reform period and a reversal during the post reform period. Therefore, it is very difficult to reach a definite conclusion.
- ⁹ For instance, Maruti Suzuki added two plants, one in 2007 and another in 2017 and in 2021 it announced opening up of another plant. Similarly, ACC Cements also added a number of plants to their capacity.
- ¹⁰ Estimating of production function involves many other problems that we do not consider here.
- ¹¹ In monopolistic competition, a market structure relevant for many of the industrial products, productivity differences are translated into differences in prices, as prices are mark-ups over marginal cost. So higher productivity leads to lower price and higher market share.
- ¹² Derivation of this result is available from the author upon request. See also Del Gatto, Ottaviano, and Pagnini (2008).
- ¹³ Exact price index in this framework is nothing but price per unit of utility. So when we divide expenditure by P we get U (see: Feenstra, 1994; Diewert, 1976)
- ¹⁴ 1992-93 is the first year of the period of analysis.
- ¹⁵ For a review see Good, Nadiri, and Sickles (1996).
- ¹⁶ The industries are NIC 21, 24, 26, 29, and 34+35.

Digital trade and wealth inequality: Evidence from Asia Pacific region

Sanjeev Vasudevan

Madras School of Economics, India;

Mini Thomas

BITS Pilani Hyderabad Campus, India

Abstract

This paper examines the impact of digital trade on wealth inequality for 40 developed and developing member countries of the UNESCAP Asia Pacific region, for the time period from 2005-2021. Data on digital trade, measured using two indicators, namely, trade in digitally deliverable services and trade in ICT goods, is extracted from UNCTAD. Data on within- country wealth inequality, measured using two indicators, namely, wealth share of top 1 percentile and top 10 percentile of adult population, is sourced from World Inequality Database. The overall trend in digital services trade restrictiveness and wealth inequality within the Asia Pacific region is initially analysed. Based on a rich panel dataset inclusive of standard control variables, this study then estimates the digital trade-wealth inequality nexus with the help of panel fixed effects and instrumental variable estimation techniques. This study puts forth the following key findings. First, we find empirical evidence in support of the positive and significant impact of international trade, in both digitally deliverable services and ICT goods, on within-country wealth inequality in the Asia Pacific region. Second, we observe marked heterogeneity between developed and developing member countries of ESCAP, with the effect of digital trade on wealth inequality turning out to be significant for developed countries and insignificant for developing countries. This study finally puts forth crucial policy implications for streamlining digital trade to achieve the Sustainable Development Goal of reduced wealth inequality.

Keywords: Digital trade; digitally deliverable services; wealth inequality; Asia Pacific; panel fixed effects; instrumental variable estimation

JEL: F10; F13; F14; O15

1. Introduction

The rise of digital economy in last few decades has significantly altered the international trade landscape globally. Digital trade is defined as digitally enabled transactions in trade in goods and services which can be delivered either digitally or physically, involving consumers, firms and governments (Lopez-Gonzalez and Jouanjean, 2017). Growing popularity of online platforms have resulted in an increasing number of small goods packages crossing international borders, whereas new digital technologies have changed how services are produced and delivered across borders. Digital trade is based on movement of data, with data acting both as a means of production as well as an asset which can be traded. Ensuring that digital trade is inclusive, both within and across countries, poses a formidable challenge to policymakers worldwide. Small and medium sized firms, which were earlier supplying only in domestic markets due to cost constraints, can now cater to a large and growing global customer base, by taking advantage of access to online market places offered through digital technologies. Online sellers are able to offer products at a cheaper price due to lower capital requirements. However, there exists a wide disparity between the developed and developing world when it comes to uptake of information and communication technology (ICT). The proportion of individuals using the internet, an ICT Sustainable Development Goal (SDG) indicator, stood at a towering figure of 87.26 percent for North America and Europe in 2020. In contrast, only 40.53 percent of individuals in South Asia and 57.83 percent of individuals in Southeast Asia used the internet even during the pandemic year of 2020 (ITU, 2021).

Globalisation and technological change have resulted in uneven spatial distribution of wealth and rising within-country inequality, especially in advanced economies (Iammarino, 2018). UNCTAD (2021) highlights the increasing tendency of the world's largest digital platforms, such as Apple, Microsoft, Amazon, Alphabet (Google), Facebook, Tencent and Alibaba to invest in all parts of the global data value chain. Interestingly, all of these companies have their headquarters in the Economic and Social Commission for Asia and the Pacific (ESCAP) member countries. These companies are keen on upgrading their data collection, data transmission, data storage, data analysis, data processing and final use by heavily investing in user-facing platform services, submarine cables and satellites, data centers and Artificial Intelligence. Reaping the benefits of economies of scale, network effects and privileged access to data, these corporations have come to acquire massive financial, market and technological power and global reach in recent years. Accelerated digitalisation in the wake of COVID19 pandemic have further enhanced their size, profits, market value and market dominance. Four of these digital platforms accounted for 67 percent of the global revenue from cloud infrastructure services in the last quarter of 2020 (UNCTAD, 2021). In this new data driven world order, developing countries risk being mere providers of raw

data to global digital platforms, while also having to pay for the digital intelligence obtained from their data.

Net Wealth of a household is defined as all the non-financial assets as well as financial assets over which the household can enforce ownership rights and which also provides economic benefits to their owners, net of any debt (Zucman, 2019). Non-financial assets such as real estate, land and buildings and financial assets such as bank deposits, equities, bonds, life insurance and pension funds are included in this definition of wealth. Wealth inequality is defined as the unequal distribution of wealth and assets among the countries of the world and within countries (Njangang et al., 2022). According to the World Inequality Database (2022), the top 1% of the world's population possessed nearly 39 times the wealth of the lowest 50% in 2021. However, a cross-country scrutiny indicates that wealth distribution among the rich varies depending on the per capita income levels. In 2021, the top 1% in the developed countries possessed 24% of national wealth, whereas the comparable statistics in the developing countries was 42%. Besides, Piketty and Zucman (2014) highlighted that wealth inequality as a share of total national wealth is on the rise in the last 40 years. The uneven patterns of wealth distribution and their macroeconomic consequences warrant an investigation of the major determinants of wealth inequality within and across countries.

International trade, among others, is identified as a critical determinant of wealth inequality (Foellmi & Oechslin, 2010; Chang & Wu, 2016). However, effects of trade on wealth inequality remain ambiguous. Tadadjeu et al. (2021) showed that trade openness increases wealth inequality, for a sample of 45 developed and developing countries. On the contrary, Islam (2018) argued that international trade and wealth inequality are inversely related in countries with lower levels of democracy. Technological progress can drive wealth inequality because of its in-built skills-bias. In contrast, digitalization provides more opportunities for the lower income groups to get market access by lowering transaction costs (Zhu et al., 2022). The proliferation of digitalization in the post-Covid period further amplifies the importance of digital trade in shaping inequality levels within and across countries (Wang & Xu, 2023). Apart from the direct effect, digitalization affects wealth inequality through several indirect channels, such as entrepreneurship and managerial skills, financial development and innovation (Njangang et al., 2022).

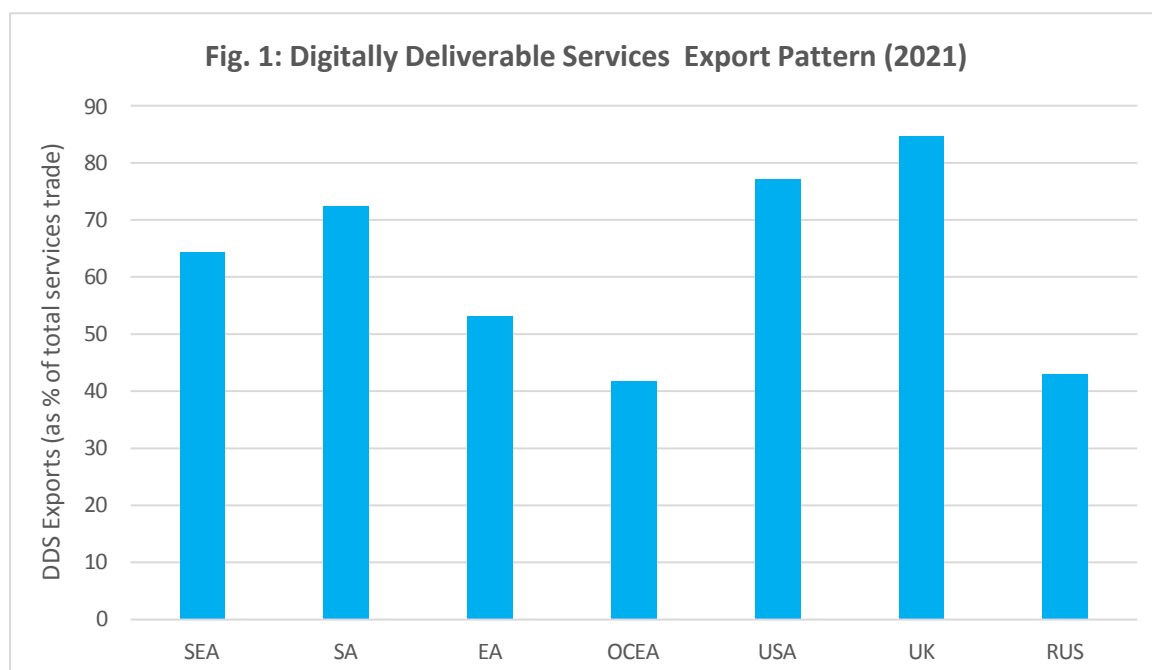
Zhu *et al.* (2022) studied the impact of digital service trade on within-country income inequality, for a panel of 100 countries. They defined income inequality in terms of the Gini index and digital service trade in terms of digitally deliverable services trade as a share of total services trade of a country. Using two stage least squares-instrument variable estimation, they found that digital service trade had a negative and significant impact on income inequality, in case of high-income and middle-income countries. However, the negative

impact was found to be insignificant for low-income countries. Njangang et al. (2022) studied the impact of information and communication technology on wealth inequality for a panel of 45 countries. They used different measures of ICT such as internet penetration, mobile penetration, ICT service exports, ICT index, ICT quality and quantity. They measured wealth inequality using Credit Suisse data for three proxies, namely, wealth share of top one percent, wealth share of top ten percent and billionaire wealth to GDP ratio. They found that ICT increases wealth inequality, with democracy playing an important role in mitigating this impact.

Baldwin (2005) adapted the heterogeneous firm model of Meltiz and proved that online markets helped in decreasing income inequality by reducing the fixed cost of exporting. Lendle and Olarreaga (2017) studied the role of online markets in making international trade more inclusive, by bringing down income inequality. Information frictions are lesser in online markets since there is a much smaller need to search for clients or to establish a distribution channel, compared to offline markets. In online markets, the cost incurred by the seller in finding the right customer is quite negligible. It is also much cheaper for a seller to build reputation online due to the inbuilt mechanisms offered by e-commerce websites, wherein customers give ratings of various online vendors based on their earlier purchases.

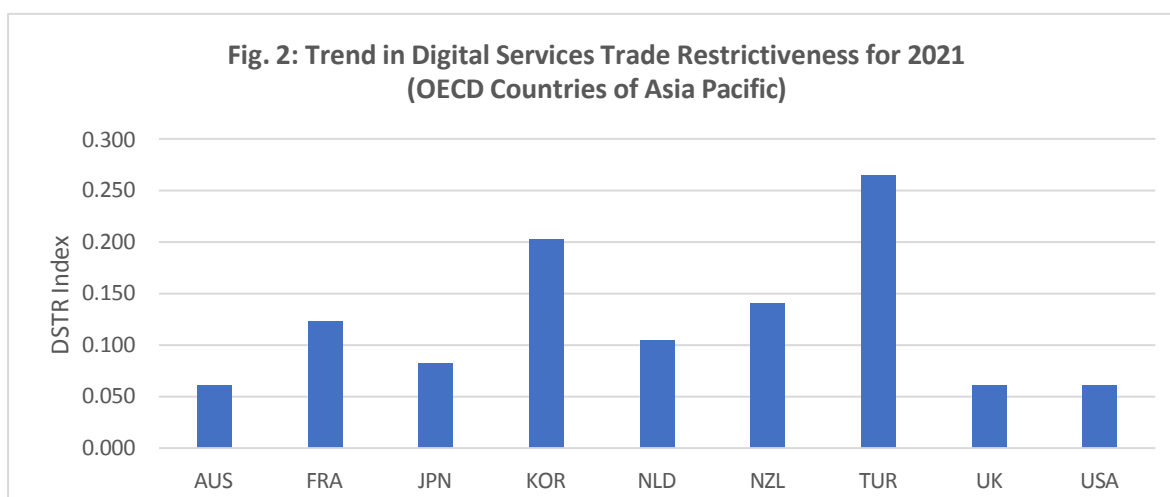
While the effects of international trade and digitalization have been examined independently, whether and how digital trade affects wealth inequality remains unexplored. Our study proposes to fill this gap. The study gains more relevance in the prospect of achieving the SDG of reduced inequality. The Asia-Pacific region provides a suitable testing ground for our study, since heightened income and wealth inequalities have accompanied rapid economic growth in this region (Zhuang, 2023). Also, growth in digital services trade experienced by this region is the highest in the world (ADB, 2022).

2. Overview of Digital Trade and Wealth Inequality within Asia Pacific region



Source: Authors' compilation from UNCTADStat

Figure 1 throws light on the digital trade pattern unfolding within the Asia Pacific region, and specifically among the ESCAP member states, for the latest year for which data is available, namely, 2021. Developed countries such as United Kingdom and United States exhibited the highest share of digitally deliverable service (DDS) exports in total services trade, at 85 percent and 77 percent respectively. Among the other regions within Asia Pacific, South Asia and Southeast Asia emerge as leading players in the digital services export market. In South Asia, comprising countries such as India which is a world leader in Information Technology exports, DDS exports as a share of total services trade, stood at 72 percent. DDS exports for East Asia, in spite of having economic powerhouses such as China, Japan and South Korea, formed only 53 percent of total international trade in services of East Asian region. Russia as well as the Oceania region, which includes many island nations as well as developed countries such as Australia and New Zealand, were found to be lagging behind in the digital services export market.

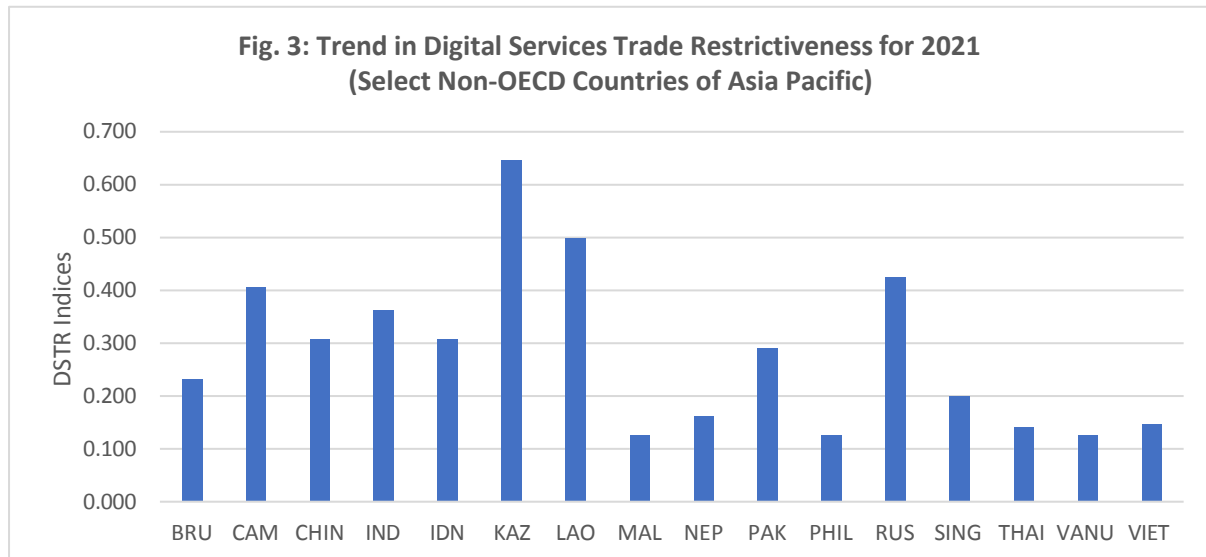


Source: Authors' compilation from OECDStat

The Digital Services Trade Restrictiveness Index published by OECD, quantifies the cross-cutting barriers that inhibit or completely prohibit the ability of firms to supply services which are traded digitally (Ferencz, 2019). This is a composite index comprising five major dimensions, namely, infrastructure and connectivity, electronic transactions, e-payment systems, intellectual property rights and other barriers to trade in digitally enabled services. The value of the index varies from 0 to 1, with 0 indicating an open regulatory environment and 1 indicating a completely closed regulatory environment. Among the ESCAP member states, data pertaining to this index is available for 25 countries. Among the OECD Countries which belong to the Asia Pacific region, it is found that Turkey has the maximum amount of trade restrictions, with a DSTRI of 0.264 in the year 2021, when it comes to engaging in digital trade (Figure 2). Majority of the barriers pertaining to digital services trade for Turkey, South Korea and New Zealand can be attributed to the infrastructure and connectivity dimension of the index. UK, USA and USA have the most conducive regulatory environment for firms engaging in digital services trade, with a DSTRI of 0.061. All of these nine OECD members are also developed countries.

Figure 3 reveals the trends in digital services trade restrictiveness for select non-OECD countries in the Asia Pacific region, for the year 2021. Kazakhstan, Lao PDR, Russia and Cambodia are found to have a highly unfavourable regulatory regime for engaging in digital trade. Kazakhstan has a DSTRI of 0.647 for the year 2021. Philippines, Malaysia and Vanatu are found to have the least amount of trade barriers pertaining to digital services trade, with a DSTRI of 0.127. Emerging market economies such as India, China and Indonesia also exhibit relatively high DSTRI values of more than 0.3, majority of which can be attributed to trade barriers in infrastructure and connectivity. India also has

Significant trade barriers relating to e-payment systems, which contributes to its high levels of digital services trade restrictiveness. Singapore, despite being a developed country, has a moderately high DSTRI of 0.2. Except for Singapore, all other ESCAP member states included in Figure 3 are developing countries.



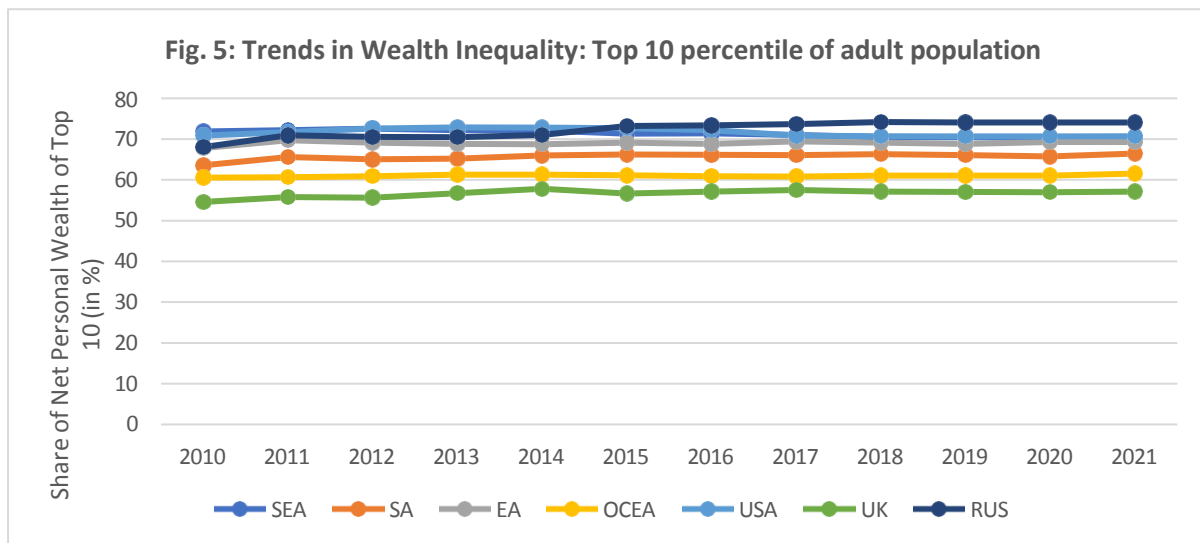
Source: Authors' compilation from OECDStat

Figures 4 and 5 showcase the wealth inequality trends within the Asia Pacific region during the past one decade. While examining the share of net personal wealth held by top 1 percentile of the adult population of a country or region with the help of Figure 4, it is found that Russia has witnessed sharp increase in wealth inequality in the last decade. As of 2021, close to 50 percent of the Russia's wealth is concentrated in the hands of the top 1 percentile of their adult population. Southeast Asia, on the other hand, has been registering a slightly declining trend in wealth inequality in the past decade, with close to 40 percent of wealth of Southeast Asian region held by the top 1 percentile of their adult population in 2021. Other sub-regions such as South Asia, Oceania, USA and UK have also witnessed a rising trend in wealth inequality during this time-span. Wealth inequality in UK is found to be the lowest, with about 22 percent of the wealth being concentrated in the hands of the top 1 percentile, during the year 2021. The corresponding share for United States for the year 2021 stood at 35 percent. Wealth inequality in East Asia also remained more or less the same during the past decade, with occasional ups and downs.

Figure 5 reveals the trends in wealth inequality within Asia Pacific region using another indicator, namely, the share of net personal wealth held by top 10 percentile of the adult population of a country or region. This indicator re-confirms the trend of Russia having witnessed significant increase in wealth inequality in the past one decade. Close to 75 percent of the Russia's wealth is

concentrated in the hands of top 10 percentile of the adult population in 2021. East Asia, Southeast Asia and USA also have about 70 percent of the wealth being held by the top 10 percentile of adults. The whole of Asia Pacific region possesses a conducive environment for sustained growth in wealth inequality, with a minimum of 60 percent of wealth being held by top 10 percentile of adult population in the year 2021, with no drastic changes in this wealth distribution pattern over the past one decade.

Source: Authors' compilation from World Inequality Database



Source: Authors' compilation from World Inequality Database

3. Data, variable description and estimation technique

3.1 Data and variable description

3.1.1 Data sources

Our sample covers 40 member countries of the United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP) over the time period 2005–2021. Data has been obtained from various sources: Data on wealth inequality is obtained from the World Inequality Database (WID), published by Paris School of Economics. Indicators of digital trade are obtained from the United Nations Conference of Trade and Development (UNCTAD) database. We sourced the control variables of the econometric analysis from World Development Indicators (WDI) database, and Global Data Lab. Table 1 presents the descriptive statistics. The full description of the data is as follows.

3.1.2 Measures of wealth inequality

Following the literature, we consider two measures of wealth inequality, namely the *top 1% wealth share* and *top 10% wealth share*. Njangang et al. (2022) put forth three arguments in support of the use of these variables. First, the top wealth shares are not sensitive to wealth changes at the bottom of the wealth distribution. Second, the probability that the wealth of individuals with the highest wealth share will increase is greater than the probability that the wealth of less wealthy individuals will increase. In other words, the wealth concentration in the hands of top 1-10% in an economy is expected to increase further. Finally, these measures of wealth inequality are highly associated with the most commonly used measures of income inequality. Based on the existing literature, both of these measures are obtained as percentages from the World Inequality Database.

3.1.3 Measures of digital trade

The existing literature focuses largely on the production and consumption of ICT goods and services as proxies of digitalization. In the current study, we employ two measures of digitalization of international trade in goods and services. The first measure is the trade in digitally deliverable services as a share of total services trade (Zhu et al., 2022). We label this measure in our analysis as '*digital services trade*'. The second measure is trade in information and communication technology goods as a share of total merchandise trade, labelled as '*ICT goods trade*'. Both measures indicate the magnitude of digital trade in goods and services, of a country. Data on both these measures are obtained from UNCTAD database.

3.1.4 Control variables

We include the *per capita real GDP* to control for economic development, which is one of the major determinants of wealth inequality. However, the effect of economic development on inequality (or wealth inequality in particular) remains ambiguous. One stream of literature finds that economic development significantly reduces wealth inequality, indicating a statistically significant negative relationship (Savvides, 1998; Berisha and Meszaros, 2019; Njangang et al., 2022). Contrary to this, an alternate strand finds a significant positive or mixed effect of economic development on inequality (Nguyen, 2022; Nchofoung and Asongu, 2022; Ndoya and Asongu, 2022). These studies found that countries with higher levels of economic development have higher levels of wealth inequality.

To account for the effect of foreign flow of capital, we define *foreign direct investment* as the net inflow of foreign direct investment as a share of GDP. The existing literature postulates that the effects of foreign direct investment on wealth inequality can be negative (Tadadjeu et al., 2021; Njangang et al., 2022;

Zhu et al., 2022) or positive (Herzer and Nunnenkamp, 2013; Rose and Cuaresma, 2016; Huynh, 2021; Yin and Choi, 2023).

Studies which associate education with greater wealth and faster wealth accumulation indicate a positive relationship between the two variables (Conley and Ryvicker, 2004; Zhu et al., 2022). However, another stream of literature states that education increases the income of the bottom stratum of the society and thereby reduces wealth inequality (Abdullah et al., 2015). We employ average years of schooling as the measure of education (Ndoya and Asongu, 2022), obtained from Global Data Lab. We label the variable as *education*.

Based on the resource curse hypothesis that higher possession of natural resources leads to increased wealth inequality, we postulate a positive relationship for this variable (Tadadjeu et al., 2021). Further, existing literature on income inequality also finds a significant positive effect of natural resources. We label the variable as *natural resources* and proxy it with the rent paid on natural resources as a percentage of GDP, obtained from WDI database.

We measure *urbanization* as the proportion of population living in urban areas. The literature finds a positive effect of urbanization on inequality in general and wealth inequality in particular (Yin and Choi, 2023). Higher concentration of people in urban centres seeking livelihood is likely to result in increased unemployment and disparities in income and wealth holdings. Hence, we expect a positive relationship.

Table 1. Descriptive statistics

Variables	Obs	Mean	SD	Min	Max
Wealth inequality					
Top 1% wealth share	680	27.30	5.43	15.48	47.86
Top 10% wealth share	680	60.47	4.68	47.69	75.78
Digital trade					
Digital services trade	680	58.03	36.50	0.00	157.17
ICT goods trade	680	9.63	8.35	0.00	54.30
Control variables					
Per capita real GDP	680	8.56	1.41	5.94	11.10
Foreign direct investment	680	4.34	7.49	-37.17	86.48
Education	675	2.10	0.45	0.53	2.62
Natural resources	680	6.36	10.00	0.00	81.91
Urbanization	680	53.59	24.19	12.98	100.00

The descriptive statistics show sufficient variability within the data. Variables

such as per capita real GDP and education are reported in the natural log form. All other variables are in percentages.

3.2 Model specification and empirical approach

We employ a panel data econometric model of wealth inequality for our empirical analysis. The model addresses wealth inequality as a function of various macroeconomic and country-specific demographic variables. Our study aims to investigate the effect of digital trade on wealth inequality. We hypothesize that digital trade is positively associated with wealth inequality. Therefore, we estimate the model as specified below.

$$Wealth\ Inequality_{it} = \beta_0 + \beta_1 Digital\ trade_{it} + \beta_2 X_{it} + \mu_i + \pi_t + \varepsilon_{it}, \quad (1)$$

Where the *Wealth Inequality_{it}* is the country specific measure of wealth inequality of country *i* for year *t*. *Digital trade_{it}* stands for country specific indicators of digital trade, as mentioned in the previous section. *X_{it}* is the vector of the baseline controls, including per capita real GDP, foreign direct investment, education, natural resources, and the extent of urbanization. Subscripts *i* refers to the countries with *i*= 1,2, 3,...40. Subscript *t* refers to the time in years where *t*= 1,2, 3. 17. μ_i captures the country fixed effects and π_t captures the yearfixed effects.

We estimate various combinations of equations in the analysis, using two measures of wealth inequality and two measures of digital trade, along with different combinations of fixed effects. The 40 countries across a period of 17 years make the total number of observations 680. We estimate the equation above using a panel data approach. The Breusch Pagan Lagrangian Multiplier test favours the panel data framework over the ordinary least-squares counterpart, confirming the presence of unobserved heterogeneity across cross-sectional units. Given the selective nature of the sample countries chosen, we employ the panel fixed effect estimator for the benchmark analysis.

One of the potential econometric issues is the endogeneity arising from reverse causality running from wealth inequality to ICT (Njangang et al., 2022). Further, Zhu et al (2022) and Yin and Choi (2023) stated that endogeneity issues arise because of the reverse causality between digital technology and income. Specifically, a growth in wealth inequality may have driven the increased use of digital technology. Further, skewed distribution of wealth may also lead to unequal access to digital infrastructure. Considering the above possibilities, the model treats indicators of digital trade as endogenous in the analysis. We correct the endogeneity issues using the 1-period lagged values of endogenous variables as instruments for estimation (Li et al., 2020). Given the longitudinal nature of the data, we estimate the regression equation using panel instrumental variable fixed effects (IV-FE) method. Since we employ only one instrument for an endogenous variable, the estimation is exactly identified.

4. Econometric results and discussion

4.1 Baseline results

Tables 2 and 3 report the estimation results of equation (1) with digital services trade and ICT goods trade respectively, as proxies for digital trade. Both tables report top 1% wealth share and top 10% wealth share as proxy variables for wealth inequality. In both tables, columns (1) and (4) estimate the equation without any control variables. We add the control variables subsequently along with year fixed effects.

The results given Table 2 provide empirical evidence in support of a positive and highly significant impact of digital trade on wealth inequality. Specifically, the coefficient of digital services trade has a magnitude suggesting that 10 units of increase in digital services trade increases wealth share of the ultra rich, on an average, by 0.06 - 0.08 units. As digitally deliverable services are more readily accessible and affordable for the wealthy, it provides them with more opportunities to enhance wealth accumulation. Similar results are obtained for top 10% wealth share. The empirical results indicate that 10 units increase in digital services trade enhances wealth share, on an average, by 0.07-0.09 units for the top decile.

Table 2. Effects of digitally deliverable services trade on wealth inequality						
Dependent variables Share	Top 1 % Wealth Share			Top 10 % Wealth		
	(1)	(2)	(3)	(4)	(5)	(6)
Digital service trade	0.00855*** (0.00302)	0.00821*** (0.00314)	0.00615* (0.00333)	0.00946*** (0.00293)	0.00750** (0.00300)	0.00736** (0.00323)
Per capita real GDP		3.471*** (0.435)	2.703*** (0.531)		3.176*** (0.417)	2.933*** (0.514)
Foreign direct investment		0.0106 (0.00933)	0.0132 (0.00940)		0.0221** (0.00893)	0.0230** (0.00910)
Education		-4.287*** (0.709)	-4.909*** (0.731)		-4.776*** (0.679)	-4.971*** (0.708)
Natural resources		0.0177** (0.00888)	0.0223** (0.00925)		0.0177** (0.00849)	0.0186** (0.00895)
Urbanization		-0.0181 (0.0380)	-0.0478 (0.0392)		0.0683* (0.0364)	0.0628* (0.0380)
Constant	26.80*** (0.183)	6.874** (2.698)	16.41*** (4.318)	59.92*** (0.178)	38.94*** (2.582)	41.87*** (4.179)
Country fixed effects	No	No	Yes	No	No	Yes
Year fixed effects	No	No	Yes	No	No	Yes
Observations	680	675	675	680	675	675
R-squared	0.012	0.127	0.162	0.016	0.155	0.171
Number of cross-sections	40	40	40	40	40	40

All estimations are carried out using panel fixed effects model. Per capita real GDP and education are in natural logs. Other variables are in percentages. Wealth inequality is measured using the share of top 1% and top 10%. Digital services trade is measured as a share of total services trade. Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5% and 10% respectively.

Similar to Table 2, the results of Table 3 provide evidence of a positive and significant effect of digital trade on wealth inequality. Specifically, the coefficients of ICT goods trade suggest that that 10 units of increase in ICT digital trade increases wealth share of the ultra-rich, on an average, by 0.22 - 0.42 units. Similar results are obtained for top 10% wealth share. The empirical results indicate that 10 units increase in ICT goods trade enhances wealth share, on an average, by 0.23-0.41 units for the top decile. Our findings indicate a strong positive effect of digital trade through services and goods, on wealth inequality. Comparing digital services trade with ICT goods trade, we observe that the effects of digital trade in the form of ICT goods have a stronger impact on wealth inequality.

Table 3. Effects of ICT goods trade on wealth inequality

Dependent variables	Top 1 % Wealth Share			Top 10 % Wealth		
Share	(1)	(2)	(3)	(4)	(5)	(6)
ICT goods trade	0.0421*** (0.00995)	0.0292*** (0.0102)	0.0221** (0.0109)	0.0416*** (0.00968)	0.0239** (0.00977)	0.0231** (0.0105)
Per capita real GDP		3.257*** (0.440)	2.562*** (0.526)		2.999*** (0.422)	2.766*** (0.510)
Foreign direct investment		0.00754 (0.00919)	0.0107 (0.00930)		0.0191** (0.00881)	0.0201** (0.00901)
Education		-4.142*** (0.707)	-4.779*** (0.733)		-4.644*** (0.677)	-4.834*** (0.710)
Natural resources		0.0204** (0.00891)	0.0241*** (0.00922)		0.0199** (0.00854)	0.0207** (0.00893)
Urbanization		-0.00674 (0.0373)	-0.0388 (0.0390)		0.0796** (0.0357)	0.0734* (0.0377)
Constant	26.90*** (0.110)	7.989*** (2.759)	17.02*** (4.289)	60.07*** (0.107)	39.78*** (2.643)	42.65*** (4.154)
Country fixed effects	No	No	Yes	No	No	Yes
Year fixed effects	No	No	Yes	No	No	Yes
Observations	680	675	675	680	675	675
R-squared	0.027	0.129	0.163	0.028	0.155	0.170
Number of cross-sections	40	40	40	40	40	40

All estimations are carried out using panel fixed effects model. Per capita real GDP and education are in natural logs. Other variables are in percentages. Wealth inequality is measured using the share of top 1% and top 10%. ICT

goods trade is measured as a share of total merchandise trade. Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5% and 10% respectively.

Most of the control variables exhibit statistically significant results. We find that per capita real GDP has a positive and significant effect on wealth inequality, in both tables. It indicates that unequal distribution of wealth is likely to worsen with enhanced economic development. The results are in line with the findings of Yin and Choi (2023). The coefficient of foreign direct investment indicates a significant positive effect, although it is the case only for the equations with top 10% wealth share as dependent variable. Our findings are similar to the works of Rose and Cuaresma (2016) and Yin and Choi (2023). Education has a highly significant, negative effect in all specifications suggesting that increased education leads to reduced wealth inequality. Abdullah et al (2015) and Njangang et al (2022) obtained similar findings. Further, natural resources indicate a significant positive effect across all specifications indicating that the possession of natural resources widens wealth inequality. Ndoya and Asongu (2022) reported similar findings for income inequality. Besides, Tadadjeu et al (2021) had reported a positive effect of natural resources specifically on wealth inequality. The urbanization variable is insignificant in the equation for top 1% wealth share, and weakly significant with a positive sign for top 10% wealth share. The results indicate that increased urbanization leads to widening of wealth inequality. These results are in tandem with the findings of Yin and Choi (2023).

4.2 Developed versus developing countries

We split the sample into developed and developing countries based on the World Bank's income classification. There are 10 developed and 30 developing countries in the sample. All high- income countries are classified as developed countries, and all other countries are classified as developing countries. It is worth noting that our sample includes only one 'low income' country, Afghanistan. Other 29 countries belong to the middle income (upper or lower) category. Table 4 reports the results of panel fixed effects estimation for developed and developing countries. Columns 1-4 represent developed countries, and columns 5-8 represent developing countries respectively.

The results of Table 4 indicate a positive and highly significant effect of digital trade on wealth inequality in developed countries. The coefficient of digital services trade indicates that 10 units of increase in digital services trade increases wealth inequality, on an average, by 0.13 -

0.16 units. Similar results are obtained in ICT goods trade, wealth inequality increases, on an average, by 0.36 units for the top decile. However, the results for developing countries are contrary to the baseline findings. We find that digital services trade has no significant effect on wealth inequality in developing

countries. However, we find a positive effect of ICT goods trade on the wealth concentration of top 10%. The variations in the effects of digital trade, in goods and services, across different income groups are not surprising. For instance, Zhu et al (2022) find that the effect of digital trade on income inequality is insignificant for low income countries. The results also indicate that the countries at higher levels of economic development are likely to engage more in digital trade, and thereby experience increased wealth inequality.

Regarding the control variables, we find that foreign direct investment and natural resources are highly significant for developed countries, and insignificant or weakly significant in the context of developing countries. We observe that the effects of education are high in the context of developing economies. A deviant result from the baseline findings is the switching sign of urbanization in developed countries. The urbanization variable suggests a negative and significant effect on wealth inequality. These results are in line with the findings of Ndoyu and Asongu (2022) who observed a negative effect of urbanization for high income countries within their sample.

Table 4. Developed versus developing countries

Dependent variables	Developed countries				Developing countries			
	Top 1% wealth share		Top 10% wealth share		Top 1% wealth share		Top 10% wealth share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Digital service trade	0.0138*** (0.00364)		0.0161*** (0.00391)		0.00185 (0.00543)		0.000327 (0.00519)	
ICT goods trade		0.0362** (0.0176)		0.0150 (0.0193)		0.0210 (0.0138)		0.0259** (0.0131)
Foreign direct investment	0.0137 (0.00832)	0.0123 (0.00859)	0.0349*** (0.00893)	0.0328*** (0.00942)	-0.00602 (0.0179)	-0.00747 (0.0179)	-0.0120 (0.0171)	-0.0139 (0.0170)
Education	3.639 (3.770)	4.305 (3.927)	-3.481 (4.045)	-3.424 (4.306)	-4.714*** (0.859)	-4.637*** (0.848)	-4.360*** (0.821)	-4.311*** (0.809)
Natural resources	0.0922*** (0.0277)	0.101*** (0.0291)	0.0932*** (0.0297)	0.117*** (0.0319)	0.0171 (0.0106)	0.0186* (0.0106)	0.0138 (0.0101)	0.0156 (0.0101)
Urbanization	-0.350*** (0.0871)	-0.221*** (0.0818)	-0.227** (0.0935)	-0.0682 (0.0897)	0.0412 (0.0458)	0.0410 (0.0456)	0.162*** (0.0438)	0.160*** (0.0435)
Constant	42.81*** (12.19)	31.10** (12.32)	83.17*** (13.08)	70.60*** (13.50)	35.07*** (2.103)	34.87*** (2.095)	62.47*** (2.008)	62.28*** (1.998)
Country fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	170	170	170	170	505	505	505	505
R-squared	0.480	0.444	0.469	0.406	0.109	0.114	0.113	0.121
Number of cross-sections	10	10	10	10	30	30	30	30

All estimations are carried out using panel fixed effects model. Countries are classified as developed and developing based on the World Bank classification of income level. Per capita real GDP has been excluded from the estimation. Education is in natural logs. Other variables are in percentages. Wealth inequality is measured using the share of top 1% and top 10%. Digital services trade is measured as a share of total services trade. ICT goods trade is measured as a share of total merchandise trade. Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5% and 10% respectively.

4.3 Endogeneity concerns

In this section, we re-estimate equation (1) using panel IV-FE method by incorporating 1-period lagged values of endogenous variables as instruments. Since we employ only one instrument for an endogenous variable, the equation is exactly identified. We do not report the results of the first stage regression. However, to confirm the significance, we show the F-statistics and corresponding probability values from the first stage.

Table 5. Endogeneity correction – IV FE estimates

Dependent variables Share	Top 1 % Wealth Share		Top 10 % Wealth	
	(1)	(2)	(3)	(4)
Digital service trade	0.0114** (0.00447)		0.0105** (0.00419)	
ICT goods trade		0.0346** (0.0147)		0.0287** (0.0138)
Per capita real GDP	2.665*** (0.570)	2.455*** (0.563)	2.863*** (0.533)	2.672*** (0.527)
Foreign direct investment	0.0139 (0.00938)	0.00922 (0.00918)	0.0246*** (0.00879)	0.0204** (0.00860)
Education	-4.201*** (0.760)	-3.924*** (0.766)	-4.153*** (0.711)	-3.924*** (0.717)
Natural resources	0.0198** (0.00915)	0.0226** (0.00909)	0.0164* (0.00857)	0.0189** (0.00852)
Urbanization	-0.0687 (0.0424)	-0.0554 (0.0419)	0.0381 (0.0397)	0.0503 (0.0393)
Constant	15.93*** (4.635)	16.77*** (4.589)	41.75*** (4.340)	42.57*** (4.298)
Year fixed effects	Yes	Yes	Yes	Yes
Observations	636	636	636	636
First stage F – stat (prob)	63.06 (0.00)	48.64 (0.00)	63.06 (0.00)	48.64 (0.00)
R-squared (within)	0.1501	0.1571	0.1582	0.1647
Number of cross-sections	40	40	40	40

All estimations are carried out using instrumental variable panel fixed effects model. Per capita real GDP and education are in natural logs. Other variables are in percentages. Wealth inequality is measured using the share of top 1% and top 10%. ICT goods trade is measured as a share of total merchandise trade. Digital service trade and ICT goods trade are treated as endogenous variables. They are instrumented by their 1 period lag values, respectively. Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5% and 10% respectively.

The endogeneity corrected estimates confirm the baseline results from Tables 2 and 3, suggesting that digital trade in services and goods, has a positive and significant effect on wealth inequality. The coefficients of digital trade from Table 5 are larger than the ones obtained from Table 2, indicating that endogeneity leads to a downward bias. Specifically, the coefficients indicate that 10 units increase in digital trade, in services and goods, increases wealth inequality, on an average, by 0.10 – 0.34 units. We also find that per capita real GDP, foreign direct investment, education and education continue to be important determinants of wealth inequality, even after taking care of endogeneity.

Table 6. Developed versus developing countries – IV FE estimates

Dependent variables	Developed countries				Developing countries			
	Top 1% wealth share		Top 10% wealth share		Top 1% wealth share		Top 10% wealth share	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Digital service trade	0.0250*** (0.00556)		0.0222*** (0.00580)		0.00510 (0.00756)		0.00329 (0.00696)	
ICT goods trade		0.0519** (0.0256)		0.000424 (0.0275)		0.0353* (0.0189)		0.0384** (0.0173)
Foreign direct investment	0.0167* (0.00877)	0.0145 (0.00889)	0.0372*** (0.00914)	0.0346*** (0.00954)	-0.00593 (0.0182)	-0.00907 (0.0182)	-0.0113 (0.0167)	-0.0149 (0.0167)
Education	3.159 (4.117)	4.180 (4.244)	-5.448 (4.291)	-5.863 (4.556)	-4.123*** (0.890)	-3.878*** (0.874)	-3.700*** (0.818)	-3.482*** (0.803)
Natural resources	0.0791** (0.0308)	0.0928*** (0.0327)	0.0824** (0.0321)	0.118*** (0.0351)	0.0149 (0.0104)	0.0176* (0.0104)	0.0115 (0.00955)	0.0145 (0.00960)
Urbanization	-0.412*** (0.111)	-0.170* (0.0970)	-0.283** (0.116)	-0.0551 (0.104)	0.0154 (0.0490)	0.0141 (0.0486)	0.133*** (0.0451)	0.130*** (0.0447)
Constant	48.68*** (14.33)	27.25* (13.96)	92.53*** (14.94)	75.66*** (14.99)	34.76*** (2.214)	34.31*** (2.207)	62.17*** (2.037)	61.73*** (2.027)
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	160	160	160	160	476	476	476	476
First stage F – stat (prob)	23.12 (0.00)	23.65 (0.00)	23.12 (0.00)	23.65 (0.00)	35.38 (0.00)	32.42 (0.00)	35.38 (0.00)	32.42 (0.00)
R-squared	0.430	0.413	0.436	0.383	0.102	0.109	0.103	0.113
Number of cross-sections	10	10	10	10	30	30	30	30

All estimations are carried out using instrumental variable panel fixed effects model. Countries are classified as developed and developing based on the World Bank classification of income level. Per capita real GDP has been excluded from the estimation. Education is in natural logs. Other variables are in percentages. Wealth inequality is measured using the share of top 1% and top 10%. Digital services trade is measured as a share of total services trade. ICT goods trade is measured as a share of total merchandise trade. Standard errors in parentheses. ***, **, and * indicate significance at 1%, 5% and 10% respectively.

Table 6 reports the endogeneity corrected estimates for developed and developing countries. Similar to Table 4, columns 1-4 of Table 6 represent developed countries, and columns 5-8 represent developing countries respectively. The results of Table 6 strengthen the findings from Table 4. We observe that digital trade has a positive and highly significant effect on wealth inequality in developed countries. The coefficient of digital services trade indicates that 10 units of increase in digital services trade increases wealth inequality, on an average, by 0.22 -

0.25 units. Similar results are obtained for ICT goods trade, as 10 units increase in ICT goods trade leads to an increase in wealth inequality, on an average, by 0.52 units for the top decile. The results for developing countries, reported in columns 5-8, are consistent with the findings of Table 4. We find that digital services trade has no significant effect on wealth inequality in developing countries. Nevertheless, we find a positive and significant effect of ICT goods trade on the wealth concentration of top 10%, for developing countries.

5. Conclusions and Policy Implications

This study examined the effect of digital trade on wealth inequality for 40 developed and developing member countries of the UNESCAP, for the period 2005-2021, using a rich panel dataset, estimated using panel fixed effects and instrumental variable estimation techniques. The study puts forth the following important findings. First, we find empirical evidence in support of the positive and significant impact of international trade, in both digitally deliverable services and ICT goods, on wealth inequality among ESCAP countries. Second, we observe marked heterogeneity between developed and developing member countries of ESCAP, with the effect of digital trade on wealth inequality turning out to be significant for developed countries and insignificant for developing countries. Third, education is found to have an inverse relationship with wealth inequality, with greater educational attainment leading to lower wealth inequality in both developed and developing countries. Fourth, we find that the effect of urbanization on wealth inequality varies depending on the income level of the member country. Our estimates are robust to potential endogeneity bias and subsample analysis.

Wealth inequality has garnered quite a bit of attention in recent years, especially after the publication of the book “Capital in the Twenty First Century” by Thomas Piketty. However, there have been only very few studies which have explored the relationship between international trade and wealth inequality, partly due to paucity of reliable data on wealth distribution. In contrast, there have been numerous studies which have examined the nexus between international trade and income inequality. There is now a general consensus that trade reforms have resulted in a worsening of within-country income inequality on the one hand, but a reduction in income inequality

between countries on the other (Bourguignon, 2016). The impact of international trade on wealth inequality remains an ambiguous question. Our study contributes to the existing literature by delving into how the emergence of digital trade in recent years have influenced within-country wealth inequality in the Asia Pacific region. Our empirical results indicate that only the wealthier sections of a country have been able to reap the benefits of digital trade in the Asia Pacific region. The positive relationship between digital trade and wealth inequality should be a cause of concern for governments and policymakers in the region, and targeted steps should be taken to resolve the skewed wealth distribution in this region and achieve goal 10 of UN SDGs of reduced inequality.

The future of world economy and international trade lies in digitalization. The regulatory framework of digital trade should be strengthened to make it more inclusive and ensure that the benefits are not cornered by a select few within each country. Human capital accumulation of the poorer sections of society through education and skills training can play an important role in enhancing the uptake of digitalization and accelerate digital trade to bring about material upliftment of the people at the bottom of the economic ladder. Since movement of data is the underlying phenomenon behind digital trade, strict regulations should be in place with regard to data security, data storage, data usage and selling and purchasing of data. Monopolistic and restrictive trade practices of huge multi-national companies engaging in digital trade as well as their top executives who have amassed huge wealth from it need to be regulated by national governments as well as through inter-governmental cooperation. Possibility of imposing a wealth tax on individuals or corporates who have disproportionately gained from the growth in digital trade can also be explored, especially for developed countries within the Asia Pacific region. Obstacles faced by small and medium scale enterprises, youth as well as women entrepreneurs participating in digital trade, such as financing or trade credit constraints, information asymmetry, regulatory compliance and barriers to entry should be resolved at the earliest, by designing targeted policy interventions. The policy approach to regulating digital trade in developing countries and developed countries should be different, due to their differential impact on wealth inequality.

Based on the initial descriptive analysis carried out using OECD digital services trade restrictiveness indices, it was found that barriers to digital services trade are quite low in developed countries, but quite high in developing countries. At the same time, we also found that the smooth and seamless flow of digital trade has also resulted in worsening of wealth inequality in the developed countries. Hence, developing countries should follow a guarded approach while loosening their regulations with regard to digital trade, and ensure that social security mechanisms are in place for the people who are left behind in the emerging international trade regime, thus making the country's digital trade policy inclusive.

References

- Baldwin, R. (2005). Heterogeneous firms and trade: testable and untestable properties of the Melitz model. *NBER Working Paper* . <https://doi.org/10.3386/w11471>
- Bourguignon, F. (2016). The globalization of inequality. Princeton University Press.
- Chang, E. C., & Wu, W.-C. (2016). Preferential trade agreements, income inequality, and authoritarian survival. *Political Research Quarterly*, 69(2), 281–294. <https://doi.org/10.1177/1065912916636688>
- Ferencz, J. (2019). The OECD Digital Services Trade restrictiveness index. *OECD Trade Policy Papers*, (221). <https://doi.org/10.1787/16ed2d78-en>
- Foellmi, R., & Oechslin, M. (2010). Market imperfections, wealth inequality, and the distribution of trade gains. *Journal of International Economics*, 81(1), 15–25. <https://doi.org/10.1016/j.jinteco.2010.03.001>
- Iammarino, S. (2018). FDI and Regional Development Policy. *Journal of International Business Policy*, 1(3–4), 157–183. <https://doi.org/10.1057/s42214-018-0012-1>
- Islam, Md. R. (2018). Wealth inequality, democracy and economic freedom. *Journal of Comparative Economics*, 46(4), 920–935. <https://doi.org/10.1016/j.jce.2018.01.002>
- The ITU ICT SDG indicators. ITU. (n.d.). <https://www.itu.int/en/ITU-D/Statistics/Pages/SDGs-ITU-ICT-indicators.aspx>
- Lendle, A., & Olarreaga, M. (2017). Can online markets make trade more inclusive. *ADB Working Paper Series*.
- López González, J., & Jouanjean, M.-A. (2017). Digital trade: developing a framework for analysis. *OECD Trade Policy Papers*. <https://doi.org/10.1787/524c8c83-en>
- Njangang, H., Beleck, A., Tadadjeu, S., & Kamguia, B. (2022). Do icts drive wealth inequality? evidence from a dynamic panel analysis. *Telecommunications Policy*, 46(2), 102246. <https://doi.org/10.1016/j.telpol.2021.102246>
- Piketty, T., & Zucman, G. (2014). Capital is back: Wealth-income ratios in rich countries 1700–2010. *The Quarterly Journal of Economics*, 129(3), 1255–1310. <https://doi.org/10.1093/qje/qju018>
- Tadadjeu, S., Njangang, H., Asongu, S., & Nounamo, Y. (2021). Natural resources and wealth inequality: a cross-country analysis. *Journal of Economic and Administrative Sciences*. <https://doi.org/10.1108/jeas-05-2021-0099>
- UNCTAD. (2021). (rep.). *Digital economy report 2021: Cross-border data flows and development: For whom the data flow*.
- Wang, J., & Xu, Y. (2023). Digitalization, income inequality, and public health: evidence from developing countries. *Technology in Society*, 73, 102210. <https://doi.org/10.1016/j.techsoc.2023.102210>
- Zhu, W., Li, X., Wang, H., & Sun, B. (2022). Digital service trade and income inequality: A panel data analysis for 100 countries. *Applied Economics Letters*, 30(11), 1457–1461. <https://doi.org/10.1080/13504851.2022.2061895>

- Zhuang, J. (2022). Income and wealth inequality in Asia and the Pacific: trends, causes, and policy remedies. *Asian Economic Policy Review*, 18(1), 15–41. <https://doi.org/10.1111/aepr.12399>
- Zucman, G. (2019). Global wealth inequality. *Annual Review of Economics*, 11, 109–138. <https://doi.org/10.1146/annurev-economics-080218-025852>.

On the resiliency of incubator supported start-ups: Case studies from a developing country context

Phumuza Langa

Technology Transfer Office,
University of Johannesburg, South Africa;

Joe Amadi-Echendu

Dept of Eng & Technology Management,
University of Pretoria, South Africa

Abstract

To become established as a going concern beyond incubation support, a start-up company must enter the market and sustain its business. Invariably, the real business environment is characterized by volatility, uncertainty, complexity and ambiguity (VUCA) which a start-up company must confront. VUCA inherent in business and the real world of markets often manifest as transient, intransigent, and evolutionary stressors arising from supervening change events. This implies that, for a start-up to survive, it must possess some dynamic capability to deal with transient, intransigent, and evolutionary stressors imposed by supervening VUCA-laden changes that occur in the real world of business and markets. By means of multiple case studies, the discourse in this paper explores the resiliency of incubator supported start-ups within a developing country context. The case studies provide some insight with regard to the dynamic and transformative capabilities developed by start-ups that receive formalized incubation support within the ambit of a country's national systems of innovation. The discourse posits that dynamic capabilities and resiliency contribute to strategic management of innovation particularly through new business ventures, noting that start-up companies are crucial and indispensable

for innovation-driven transformation of the economies of countries in the Global South.

Keywords: Dynamic capability, Resiliency, Start-up incubation, Innovation-driven transformation, Knowledge economy

1. Introduction

The support of start-ups through business incubation has gained popularity globally with many nations introducing incubation policies and practices within their countries (Dubihlela & Van Schaikwyk, 2014). Although formalized incubation was largely initiated, particularly in the Global North, by developed nations in Europe and North America, developing nations in the Global South have also been implementing formalised incubation support since the 1990s. For instance, countries like South Africa introduced incubation programs such as *Godisa* (re: Buys & Mbewana (2007)) which has now evolved to small enterprise development agency (SEDA). Thus, there is a general understanding within science, technology and innovation policy circles that formal enterprise support instruments such as incubators are relatively useful in facilitating the development of start-ups and small firms (see, for example, Akçomak & Taymaz (2004); Tavoletti (2012); Schutte and Chauke (2021); Theodoraki, Messeghem and Audretsch (2022)).

A matter of interest is to learn how incubators assist start-ups to overcome the 'liability of newness' phenomenon (re: Stinchcombe's (1965)). The study by Aspelund, Berg-Utby and Skjevdal (2005) "suggests that initial resources controlled by technology-based new ventures are important antecedents for their survival in later stages." Aspelund *et al* (2005) surmise that a heterogeneous entrepreneurial team coupled with a reasonable degree of technological radicalness increase the likelihood of survival of a start-up firm. The ongoing challenge remains the transformation of a start-up into a sustainable business organisation in the short term, subsequently to become established as a going concern in the medium term with the capability of continued existence into the long term.

Thus, a more pertinent question is whether incubation as formalized within the ambit of national systems of innovation facilitates the resiliency of start-ups, given that, in the real world of business and markets, start-ups are actually confronted by transient, intransigent, and evolutionary stressors (i.e., strains and tensions) emanating from supervening volatile, uncertain, complex, and ambiguous (VUCA)

change events. This is an important issue to examine, especially because of the acknowledged importance of start-ups and small firms, coupled with the pressing need to adopt innovation-driven approaches, not only to transform the economies of developing countries in the Global South, but also to make such economies more resilient to supervening local and global VUCA events inherent in the real world of business and markets.

In this regard, this article attempts to answer the question, "*what dynamic capabilities do start-ups gain during formalized incubation to make them more resilient ex post facto?*" The proposition here is that dynamic capabilities are indispensable for the resiliency of start-ups, hence, the main argument proffered in this paper is based on case studies of some start-ups in a developing country of the Global South. Section 2 of this discourse provides an overview on the concepts of dynamic capability and resiliency. Case studies of three incubated start-ups are briefly described in section 3, and a discussion regarding the dynamic capabilities gained through incubation is presented in section 4. Section 5 includes some concluding remarks on the dynamic capabilities of start-ups during the respective phases/stages of resiliency.

2. Resiliency and dynamic capability

2.1. Resiliency

The concept of *resiliency* has taken on renewed importance in the current era characterized by VUCA stressors that are embedded in *casus fortuitus* and supervening *force majeure* events. Increasing concerns regarding the effects of climate change plus local, regional and international geopolitical tensions in highly interconnected and globalized systems of business and markets have re-invigorated studies on resiliency. Intriguingly, many studies tend to be based on a negativistic perception of uncertainty, thus leading to a narrow view of resiliency in terms of reaction to adversity. It is important to recognize that opportunity represents a positivistic perception of uncertainty. Thus, the view of resiliency also extends to pro-activeness to exploit opportunities that arise due to supervening VUCA events (re: Amadi-Echendu (2023)).

Adversity and opportunity affect established firms as well as start-ups, even though intuition suggests that, by their very nature (e.g., small size and low inertia), start-ups should exhibit higher levels of agility and adaptability than established
going
concerns

¹ especially when confronted by VUCA stressors. Extant literature (e.g., Sheffi and Rice (2005), Linnenluecke (2015), Annarelli and Nonino (2016), Xiao and Cao (2017)) more or less argue that resiliency is the ability that an organization

possesses in order to recover from the effects of adverse disruptions to 'normal' business circumstances and/or threats to known organizational vulnerabilities. Foris *et al* (2022) discuss how some start-ups react to disruptions in various key business elements such as supply chains, transportation and freight service network. As a matter of course, a start-up will confront adversities as well as opportunities, hence a more appropriate view of resiliency is the capacity to accommodate stressors imposed and/or induced by supervening change events, taking into consideration that negativistic and positivistic VUCA effects can emanate from change events. In general, the stressors may be a conflation of transient, intransigent, or evolutionary strains and tensions. This view leads to the decomposition of resiliency into phases or stages (see, for example, Duchek (2020); Amadi-Echendu (2023)), with the implication that start-ups require dynamic capabilities so as to transcend the phases/stages of resiliency in order to transform to going concerns.

2.2. Dynamic capabilities

According to Teece *et al* (1997), dynamic capabilities may be defined as *"the firm's ability to integrate, build, and reconfigure internal and external competences to address rapidly changing environments."* This definition extends to include sensing, seizing, and transforming parameters that can enable an enterprise to upgrade its ordinary capabilities towards high-order capabilities (Teece, 2007). Winter (2003) emphasizes the need for having this distinction between ordinary or static capabilities during steady-state operational conditions and dynamic capabilities, that is, capabilities that are more sensitive to change events. This point is that emergence from incubation will require a start-up firm to shift from ordinary to dynamic capabilities in order to transition to a going concern in the VUCA laden realities of business and markets.

Extant literature suggests that studies on dynamic capabilities have tended to focus on established large companies (Rindova and Kotha, 2001; Zahra *et al* (2006)). However, some studies (see, for example, Helfat and Winter (2011), Teece (2012), Farkas (2022)) have examined dynamic capabilities within start-ups. There is acknowledgement that because of their very nature, established large companies tend to be more resistive to change, in a sense that they are less agile than smaller companies. The indication is that well established companies tend to rely on their stock of internal resources (i.e., steady-state capabilities) to respond to change.

In contrast, start-ups generally have meagre internal resources and therefore tend to source external resources to improve their ability to deal with steady-state circumstances and change imposing events. Furthermore, entrepreneurial flair tends to more pronounced in start-ups than in large companies, that is,

entrepreneurial aptitude and behaviour or entrepreneurship engenders more dynamic capabilities within start-ups. Farkas (2022) points out that start-ups are more adept at gaining dynamic capabilities when changes in strategic direction occur more rapidly. Established and larger entities have policies manifesting into numerous fixed routines that lock-in a substantial part of their resources, causing such firms to be less adept to track and deal with rapid changes in strategy. Hence, it can be argued that dynamic capabilities tend to be better aligned towards adaptability, entrepreneurial ambidexterity, flair, and innovativeness. That is, dynamic capabilities facilitate the balance between explorative and exploitative activities.

Ma *et al* (2015) examined what constitutes dynamic capabilities for technology start-ups and found that they are largely associated with management's entrepreneurial spirit and leadership skills when start-ups are concerned. Similarly, Park *et al* (2022) reflects on the importance of entrepreneurial ambidexterity as was seemingly apparent in the rapid development of Covid-19 vaccines. Zhang *et al* (2022) indicate that Chinese SMEs deploy extant knowledge to renew, adjust or update resources and operational capabilities to cope with changing circumstances. Ochie *et al* (2020) examined the strategies that firms in developing countries deploy to scale up organizational ambidexterity under conditions of environmental uncertainty. Hanchi and Kerzazi (2020) discuss innovation capacity as a construct that start-ups can deploy through sensing, seizing, and transforming.

2.3. Start-ups and incubators

This section provides a brief review on start-ups and incubators as a prelude to the case studies discussed later in section 3. For brevity, start-ups are generally viewed as small entities that focus on commercializing innovative ideas (De', Pandey & Pal (2020)). Some start-ups are categorized as technology-based, implying that the technology to be commercialized arises from formal (and sometimes informal) research and development (R&D) activities (Ahn *et al* (2022)). A technology-based start-up (TBS) is synonymous with a new technology-based firm (NTBF); that is, a business based on the exploitation of inventions and technologically intensive innovations associated with high risk (Story and Tether (1998), Martin (2004)) The prevailing view is that start-ups tend to involve high levels of R&D and intensive scientific and technical expertise. Spin-off firms and small-to-medium sized enterprises also tend to be viewed as start-ups in terms of technology intensity and the 'age' of the business organisation (Camisón-Haba *et al* (2019)).

Incubation more or less fits within the survival stage of the small business growth model posited by Lewis and Churchill (1983). That is, incubation should enable a

start-up to demonstrate that it is a workable business entity, capable of breaking even and becoming profitable with staying power to grow. Hackett and Diltz (2004) indicate that “a business incubator is a shared office-space facility that seeks to provide its incubatees (i.e., portfolio – or client or tenant companies) with a strategic value-adding intervention system.” Extrapolating from Pettersen *et al* (2015), Al-Mubarak and Busler (2017), Shih and Aaboen (2019), and Leitão, Pereira and Gonçalves (2022), it is reasonable to surmise that incubators generally provide administrative assistance whilst facilitating access to local and regional (and sometimes global) (i) networks of agents and institutions for economic and financial, governance, knowledge, and socio-political resources, and (ii) markets. It is in this regard that incubators are considered to be useful instruments in the development of start-ups.

3. Case studies

This section deals with the case studies under analysis. It takes an approach of briefly introducing the company by means of profiling, this is then followed by discussing information about the entrepreneur composition, the technology and the market that the business conducts in. The cases were sourced as secondary data from an annual report of a South African enterprise development agency. The case study incubatees are drawn from industries such as manufacturing, agriculture and waste management. The incubation support that was provided is discussed together with the VUCA stressors that it meant to resolve and the subsequent dynamic capabilities that were built in the process to render the TBS resilient.

3.1. Case 1: Fabricated filters

The company is in the business of manufacturing fabricated filters that are utilised by heating, ventilation and air conditioning companies to name a few. It is in the business of designing and fabricating filtration systems for installation and maintenance that have applications ranging from dust, oil, water, fuel and air to medical applications and odour control. Sectors serviced include manufacturing, mining, power generation, automotive shipping and consumer markets.

- The entrepreneur: Owns 100% of the business. He has a mechanical engineering qualification with over 15 years of experience in the company. This reflects a strong involvement in the business and most likely has extensive influence in the strategic direction, particularly in instances of business environmental changes which require dynamic capabilities.
- The technology: In the business of hydraulic technologies. In combination with ancillary services such as maintenance training and baghouse design
- The market: This industry is characterised by fluctuations in raw material prices which can directly affect the manufacturing costs and impact growth. These raw materials are mostly commodities such as rubber, plastics, fibers and chemicals which are generally highly sensitive to changes in the business environment, therefore it's a very dynamic industry which constantly needs adaptation for survival. Furthermore, it is an industry that could experience market growth which will be driven by the rise in the global automotive industry on account of the growing population and increasing disposable income levels, especially in the developing regions, such as Asia Pacific, Middle East & Africa, and Central & South America.

A statement is made by the entrepreneur to reflect the interplay between support received and the capability obtained:

*"We have realised a vast **improvement in efficiency**, thanks to the **streamlined processes** that emerged from interventions and recommendations. I am confident things will run smoothly as the company continues to **grow**."*

From the statement above it becomes evident that some of the stressors that the business could have faced relate to limited growth and limited efficiencies due to processes not being optimally streamlined. Teixeira *et al* (2019) for instance finds a relationship between startup growth and the dynamic capability of the enterprise to understand market signals, learning from customers and designing scalable, repetitive and profitable business models. This is possibly the case in terms of the entity under observation. The incubation support received as mentioned by the annual report relates to services in the form of business planning, strategy and succession planning. A case study by the Jordan Enterprise Development Corporation corroborates this relationship by reflecting how supporting firms

through incubation services such as business plan preparations and business development advice can lead to revenue growth (Al-Damen (2021)). Furthermore, the entity was assisted with training on implementing key quality management systems (QMS). This could be the explanation why the entity gained the capability to improve its efficiencies from an operations perspective and have better streamlined processes. Neo *et al* (2020) also reveals that successful implementation of QMS in small and micro enterprises ensures improved control of processes. Another form of support that the entity received was to attend a trade mission and it was given skills on electronic e-marketing which led to upgrading their website to suit international markets. This is in line with Mu (2017) who found that in developing a marketing capability, a firm can improve its performance.

3.2. Case 2: Agricultural Primary Cooperative

This agricultural primary cooperative is mainly in the business of crop production (vegetables). These range from cabbages, spinach and beetroot to name a few. They supply mainly to the local market and local vegetable traders.

- The entrepreneur: Owned by a group of young gricultural enthusiasts.
- The technology: Agricultural ploughing equipment such as tractors, irrigation systems and pack houses.
- The market: According to (www.visualcapitalis.com), Agriculture makes up nearly 20% of Sub-Saharan Africa's economy which is a higher percentage than any other region worldwide. Growth and improvement in this market will be influences by technological advances, such as precision technology that can measure rainfall, soil information, and soil productivity. Also, technologies such as remote sensing which can provide information on weather and climate.

The cooperative aimed to increase its plant production (by volume and by variety) and intended to improve its fresh produce that could be available for supply to local markets. The constant pursuit of utililising existing agricultural lands to intensify production by farming business is viewed as critical for productivity in agriculture ecosystem and this can assist in increasing supply of crops to the market (Strassburg *et al* (2014)). The entity under observation had this stressor of land productivity and required assistance. The enterprise support received relates to assistance with the purchase of materials (such as seedlings) and equipment (such as wheelbarrows and a gasoline pump engine). Additionally, provision of a grant for production input under Covid-19 intervention was received. The interventions above resulted in an increase in turnover for the business and subsequent additional employment opportunities for locals. Ruete (2015) highlights the importance of supporting small scale farmers in developing countries with mechanisms such as finance as being vital.

3.3. Case 3: Waste Management

The company's core business is to provide green economy, waste management and recycling services to industries such as the mining and engineering sectors.

- The entrepreneur: Founder run entity. The owner has some level of training on waste management as such has a capability to run the business within this market.
- The technology: In the business of providing green economy, waste management and recycling services to industries such as the mining and engineering sectors.
- The market: Major activities performed on waste are its collection and disposal. Collection services refer to collection & transportation, storage & handling, and sorting of waste. Whereas disposable services refer to segregation of waste into landfills, recycling, compositing, and anaerobic digestion (www.alliedmarketresearch.com).

The identified challenge for this particular entity was the lack of business skills and limited knowledge in supply chain management in respect of procurement and tender processes that relate to getting more business. This limitation resulted in the firm not accessing some business opportunities. The incubation support provided relates to mentoring and coaching on procurement and tendering processes. Furthermore, assistance was received in the form of machinery and equipment. This resulted in the company getting more contracts with some key clients and also an increase in revenue. Illustrating how resources can be deployed to fill in gaps that disadvantaged entrepreneurs have as a result of a lack of formal education and managerial experience that could hinder business growth. Some of the interventions being for instance personalized mentoring. This is consistent with Assenova (2020) who posits that incubation support can result in scaling, and profitability among disadvantaged entrepreneurs.

Lastly, Masteika and Cepinskis (2015) review the relationship between dynamic capabilities and supply chain management. Indicating that it makes the organisation more flexible and can assist it in being quick to adapt to new market trends when market turbulences arise. This case reveals the supply chain capabilities built by the company through incubation, these could assist in *resiliency* for the entity.

4. Incubator-driven dynamic capabilities

From the case studies above and the dynamic capabilities identified in section 2; viz – entrepreneurial adaptability, ambidexterity, flair and innovativeness, a table is created below to reflect the incubation support that assists in developing the dynamic capabilities in question. The Table below reflects the dynamic capabilities and the associated Incubation support.

The table below summarises the discussions above and highlights the connection made between the

Table1: Dynamic capability developed from incubator support

Dynamic Capability	Incubator support
Entrepreneur flair	There have been examples from the case studies above that reflect for instance personalize mentoring assisted entrepreneurs to have more capacity to improve their business. This mentoring
	Filled the gap where the entrepreneur had a lack of formal education and managerial experience for that was hindering business growth. Another example of how incubation support assisted the entrepreneur, relates to mentoring and coaching on procurement and tendering processes.
Adaptability	It has been shown that the incubator support above assisted in helping one of the companies to understand supply chain process. Also preparing the entity to be more flexible and quicker to adapt to new market trends when market turbulences arise. Another support mechanism provided related to E- marketing which led to improved performance, a reflection of the business to adapt in times of change.
Ambidexterity	In the case of the Filter business case for example, it is revealed that incubator support in the form of business planning, strategy and succession planning was provided. An observation is made that this assisted the business in having an ambidextrous capability of constantly being in tune with market signals and in the same time implementing this the day to day running of the business by designing scalable, repetitive and profitable business models.

Innovativeness	An example is made on the case studies above that an entity was assisted with training on implementing key quality management systems (QMS). This could be the explanation why the entity gained the capability to improve its efficiencies from an operations perspective and have better streamlined processes. This could lead to having an innovation capability in the market when there are business environment changes and be in a position to be more resilient.

5. Concluding remarks

The case studies support the view that entrepreneurial adaptability, ambidexterity, flair and innovativeness are dynamic capabilities that can prepare the incubated entity to be resilient *ex post facto*. In particular, the case studies point to ambidexterity as a dynamic capability that assists the incubated entities to be constantly in tune with market signals whilst implementing their daily operations. Furthermore, innovativeness is a dynamic capability that can assist the incubated entity to improve its efficiencies with prompt transformation of strategy to align with rapid changes in business and market conditions (re: Hanchi and Kerzazi, 2020). A key observation here is that the dynamic capabilities are not mutually exclusive, rather, they form a set of simultaneous capabilities that a start-up firm can utilise to overcome the 'liability of newness' post incubation support in the VUCA characterized real world of business and markets.

The VUCA stressors that confront start-ups are nothing short of remarkable given the increasing concerns regarding the effects of climate change events, the remnant ripple effects of Covid-19 pandemic, the ongoing geopolitical tensions exacerbated by Russia/Ukraine conflict, increasingly significant cyberattacks, technological disruptions, shifting demographics, energy crises, etc. The fact remains that the phenomenon of 'polycrisis', i.e., the convergence, confluence and fusing of different VUCA events (local, regional and global) impose transient, intransigent and evolutionary stressors on all categories of firms ranging from start-ups to well established companies. In the case study developing country, the increasing shortage of electricity supply resulting in loadshedding and its secondary effects (e.g., on water supply and transportation) exacerbate and intensify the strains and stresses on start-ups.

Interestingly, many start-ups in developing countries tend to be motivated by *necessity* to a greater extent, rather than being motivated by *opportunity*. It is arguable that opportunity-driven entrepreneurship occurs more in developed

economies, and such entrepreneurs tend to start-up new businesses to pursue autonomy, freedom, and flexibility. Consequently, start-ups in developed economies may readily take advantage of highly formalized and structured incubation services in order to develop necessary dynamic capabilities. However, start-ups motivated by necessity (as prevalent in developing economies of the Global South) obviously require extensive formal and informal support in order to develop necessary dynamic capabilities for resiliency *ex post facto* incubation. It is no wonder that Contreras and Dornberger (2023) suggest that sustainable entrepreneurship should be regarded as a field of knowledge in the developing countries of the Global South.

References

- Ahn, S., Kim, K.S., Lee, K. (2022). Technological Capabilities, Entrepreneurship and Innovation of Technology-Based Start-Ups: The Resource-Based View. *Journal of Open Innovation*, 8, 156. <https://doi.org/10.3390/joitmc8030156>
- Al-Damen, R.A. (2021). Business Incubator and Its Impact on Business Success: A Case Study of Jordan Enterprise Development Corporation (JEDCO). *Journal of Management and Sustainability*, 11 (1), 35.
- Al-Mubarak, H.M., Busler, M. Challenges and opportunities of innovation and incubators as a tool for knowledge-based economy. *J Innov Entrep* 6, 15 (2017). <https://doi.org/10.1186/s13731-017-0075-y>
- Akçomak, I.S, and Taymaz, E., (2004). Assessing the effectiveness of incubators: The case of Türkiye. *ERC Working Papers in Economics* 04/12, Middle East Technical University, Türkiye.
- Amadi-Echendu, J.E., (2023). An Approach to Operationalise Resiliency towards Sustainable Futures. *PICMET '23 Conference. "Managing Technology, Engineering and Manufacturing for a Sustainable World"*. July 23 - 27, 2023. Crowne Plaza - Monterrey, Mexico PICMET23R0015
- Annarelli, Alessandro & Nonino, Fabio. (2015). Strategic and operational management of organizational resilience: Current state of research and future directions. *Omega*. 10.1016/j.omega.2015.08.004.
- Assenova, V.A. (2020). Early-Stage Venture Incubation and Mentoring Promote Learning, Scaling, and Profitability Among Disadvantaged Entrepreneurs. *Organization Science*, 31(6), 1560-1578. <https://doi.org/10.1287/orsc.2020.1367>
- Aspelund A., Berg-Utby, T., and Skjevda, R. (2005). Initial resources' influence on new venture survival: a longitudinal study of new technology-based firms, *Technovation*, Volume 25 (11), pp 1337-1347, ISSN 0166-4972, <https://doi.org/10.1016/j.technovation.2004.06.004>
- Buys, A.J., and Mbewana, P.N. (2007). Key success factors for business incubation in South Africa: The Godisa case study. *South African Journal of Science* 103, September/October 2007, pp356-358.
- Camisón-Haba S., Clemente-Almendros J., Gonzalez-Cruz T. (2019). How technology-based firms become also highly innovative firms? The role of knowledge, technological and managerial capabilities, and entrepreneurs' background. *Journal of Innovation and Knowledge*, 4(3), 162-170. DOI: 10.1016/j.jik.2018.12.001
- Contreras, F., and Dornberger, U., (2023). Sustainable Entrepreneurship as a Field of Knowledge: Analyzing the Global South. *Sustainability* 2023, 15(1), 31; <https://doi.org/10.3390/su15010031>
- Curtis, A., McVay, S. & Toynbee, S. (2020). The changing implications of research

- and development expenditures for future profitability. *Rev Account Stud* 25, 405–437. <https://doi.org/10.1007/s11142-019-09528-6>
- De' R.,1, Pandey N., Pal A. (2020). Impact of digital surge during Covid-19 pandemic: A viewpoint on research and practice. *International Journal of Information Management*,
- Duchek, S., (2020). Organizational resilience: a capability-based conceptualization. *Bus Res* 13, 215–246 (2020). <https://doi.org/10.1007/s40685-019-0085-7>
- Hackett S., M., David M. Dilts. (2004). A Systematic Review of Business Incubation Research. *The Journal of Technology Transfer*, 29(1), 55-82. DOI:10.1023/B:JOTT.0000011181.11952.0f
- Hanchi, S., Kerzazi, L. (2022) Startup innovation capability from a dynamic capability-based view: A literature review and conceptual framework. *Journal of Small Business Strategy* 2020, Vol. 30, No. 02, 72-92 ISSN: 1081-8510 (Print) 2380-1751 (Online)
- Helfat C.E., Winter S.G. (2011). Untangling Dynamic and Operational Capabilities: Strategy for the (N)ever-Changing World. *Strategic Management Journal*, November, 32 Issue (11), pp 1243–1250
- Farkas, T. (2022). Dynamic Capabilities of Start-Ups. *Review of Economic Theory and Policy*, 17(3), 84–94.
- Leitão, J., Pereira, D., Gonçalves, Â. (2022). Business Incubators, Accelerators, and Performance of Technology-Based Ventures: A Systematic Literature Review. *J. Open Innov. Technol. Mark. Complex.* 2022, 8, 46. <https://doi.org/10.3390/joitmc8010046>
- Lewis, Virginia L. and Churchill, Neil C., *The Five Stages of Small Business Growth* (1983). University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship, Available at SSRN: <https://ssrn.com/abstract=1504517>
- Linnenluecke, M.K., (2015). Resilience in Business and Management Research: A Review of Influential Publications and a Research Agenda. *International Journal of Management Reviews*. <https://doi.org/10.1111/ijmr.12076>
- Mai, Y., Wu, Y.J., Wang, Y.M. (2022,). How Does Entrepreneurial Team Relational Governance Promote Social Start-Ups' Organizational Resilience? *International. Journal of Environmental Research and Public Health*.19, 6677. <https://doi.org/10.3390/ijerph19116677>
- Ma, X., Zhou, Z., and Fan, X. (2015) The process of dynamic capability emergence in technology start-ups – an exploratory longitudinal study in China, *Technology Analysis & Strategic Management*, 27:6, 675–692, DOI:10.1080/09537325.2015.1034266
- Mu, J. (2017). Dynamic Capability and Firm Performance: The Role of Marketing Capability and Operations Capability. *IEEE Transactions on Engineering Management*, pp (99), 1-12. <http://dx.doi.org/10.1109/TEM.2017.2712099>

- Neo, S., Mukwakungu, S., Lumbwe, A., Sukdeo, N. (2020). Quality Management System Implementation. The South African Small, Medium and Micro Enterprises. 1-6. 10.1109/ICTMOD49425.2020.9380607.
- Ochie, C., Nyuur1, R., Ludwig, G., Cunningham, J. (2022). Dynamic capabilities and organizational ambidexterity: New strategies from emerging market multinational enterprises in Nigeria. Thunderbird International Business Review published by Wiley Periodicals LLC.
- Park, A., Goudarzi, A., Yaghmaie, P. (2022). Rapid response through the entrepreneurial capabilities of academic scientists. Nature Nanotechnology. 17, 802–807. <https://doi.org/10.1038/s41565-022-01103-6>
- Pettersen, I.B., Aarstad, J., Høvig, Ø.S. et al. Business incubation and the network resources of start-ups. J Innov Entrep 5, 7 (2015). <https://doi.org/10.1186/s13731-016-0038-8>
- Ruete, M. (2015). Financing for Agriculture: How to boost opportunities in developing countries. International Institute for Sustainable Development (IISD).
- Schutte, F. and Chauke, T. (2021). The role and effectiveness of business incubators in growing small businesses: A focus on the manufacturing industry. Academy of Entrepreneurship Journal Volume 27, Special Issue 4, 2021. ISSN: 1087-9595
- Shih, T., Aaboen, L. (2019). The network mediation of an incubator: How does it enable or constrain the development of incubator firms' business networks?, Industrial Marketing Management, 80: 126–138.
- Stinchcombe, A.L. (1965). Social Structure and Organizations. In: March, J.P., Ed., Handbook of Organizations, Rand McNally, Chicago, 142-193.
- Strassburg, B. B. N., Latawiec, A. E., Barioni, L. G., Nobre, C. A., Da Silva, V. P., Valentim, J. F., Vianna, M., & Assad, E. D. (2014). When enough should be enough: Improving the use of current agricultural lands could meet production demands and spare natural habitats in Brazil. Global Environmental Change, 28, 84-97. <https://doi.org/10.1016/j.gloenvcha.2014.06.001>
- Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic Capabilities and Strategic Management. Strategic Management Journal (Vol. 18). 175
- Teece, D.J. (2007) Explicating Dynamic Capabilities: The Nature and Micro Foundations of (Sustainable) Enterprise Performance. Strategic Management Journal, 28, 1319-1350.
- Teece, D.J. (2007). Dynamic Capabilities: Routines versus Entrepreneurial Action. Journal of Management Studies, 49(8), 1395-1401. <https://doi.org/10.1111/j.1467-6486.2012.01080.x>
- Tavoletti, E. (2012). Business incubators: Effective infrastructures or waste of

- public money? Looking for a Theoretical Framework, Guidelines and Criteria. *Journal of the Knowledge Economy* 4, 423–443 DOI 10.1007/s13132-012-0090-y
- Theodoraki, C., Messeghem, K., and Audretsch, D.B. (2022). The effectiveness of incubators' co-opetition strategy in the entrepreneurial ecosystem: Empirical evidence from France," in *IEEE Transactions on Engineering Management*, vol. 69, no. 4, pp. 1781-1794, Aug. 2022, doi: 10.1109/TEM.2020.3034476.
- Rindova, V.P., Kotha, S. (2001). Continuous "Morphing": Competing through Dynamic Capabilities, Form, and Function. *The Academy of Management Journal*, 44(6), 1263-1280. <https://doi.org/10.2307/3069400>
- Small Enterprise Development Agency (SEDA) Annual report 2021-2022. South Africa.
- Sheffi, Yossi & Rice, Jr, James. (2005). A Supply Chain View of the Resilient Enterprise. *MIT Sloan Management Review*. 47.
- Storey, D.J., Tether, B.S. (1998). New technology-based firms in the European union: an introduction. *Research Policy*
- Teixeira, E.G., de Moura, G.L., Lopes, L.F.D., Marconatto D.A..B, Fischmann, A.A. (2021). The influence of dynamic capabilities on startup growth. *RAUSP Management Journal*, 56(1).
- Winter, S.G. (2003). Understanding dynamic capabilities. *Strategic Management Journal*, 29(10), pp 991- 995
- Xiao, Lei & Cao, Huan. (2017). Organizational Resilience: The Theoretical Model and Research Implication. *ITM Web of Conferences*. 12. 04021. 10.1051/itmconf/20171204021.
- Zahra, S.A., Sapienza, H.J, Davidsson, P. (2006). Entrepreneurship and Dynamic Capabilities: A Review, Model and Research Agenda. *Journal of Management Studies*, 43(4), 917-955. <https://doi.org/10.1111/j.1467-6486.2006.00616.x>
- Zhang, Y., Long, J., Zhao, W. (2022). Building dynamic capabilities of small and medium-sized enterprises through relational embeddedness: evidence from China. *Electronic Commerce Research*. <https://doi.org/10.1007/s10660-022-09579-z>

Note

¹ Going concerns refers to companies that are financially stable enough to meet their obligations and continue their businesses for the foreseeable future.

Transcending the formal- informal divide in gig work on digital platforms: The evidence from India

Balaji Parthasarathy, Janaki Srinivasan,

Mounika Neerukonda, Bilahari M,

Pradyumna Taduri, Amruta Mahuli

International Institute of Information Technology, Bangalore

Funda Ustek-Spilda, Kavita Dattani, Mark Graham

Oxford University

Abstract

Digital platforms have grown thanks to institutional changes that have emphasised flexible labour markets, aided by technological advances. As platforms mediate an increasing number of services, the dominant work opportunity they offer is gig work to independent contractors. Thanks to a power asymmetry between gig workers and platforms, and little regulatory control over gig work, the conditions of work are largely determined by platforms. Nevertheless, there is hope vested in the developmental possibilities of the platform economy and the gradual formalization of gig work. This paper examines the possibilities and the tensions inherent to any formalization by drawing on evidence from India, which has a large and rapidly growing platform economy.

1. Digital platforms and why study them

The evolution and spread of information and communication technologies (ICTs or IT for short), has given rise to a digital economy comprising “the digital sector [including hardware and software production, telecommunications and IT consulting] plus emerging digital and platform

services”.¹ Platforms provide “the basic infrastructure to mediate between different groups”,² and extract value from the interactions between “those who offer services and others who are looking for them”.³ In the digital economy, platforms have emerged as a new business model, in the form of “online marketplaces that involve at least three parties where the platform provider serves as an intermediary coordinating supply and demand sides of the other two parties”.⁴ By lowering transactions costs, and with what geographers refer to as space-time compression,⁵ platforms have created new work possibilities. The work opportunities offered by digital platforms fall into two categories.⁶ One, is employment i.e, when a person hired directly by a platform with an employment relationship. Those hired under this category are typically responsible for the functioning of the platform and constitute a tiny fraction of those hired. By contrast, the work of the vast majority of hires perform platform-mediated individual ‘gigs’ or tasks, and are categorized as self-employed or independent contractors. Gig work is a “non-standard form of employment”, that designates all employment that is definite, and neither full time nor part of a subordinate and bilateral employment relationship.⁷

The gig work that platforms mediate can be placed into two broad categories.⁸ The first is physical (or geographically tethered) gig work: location-based digital labor or “work on- demand via apps” including “activities such as transport, cleaning and running errands” and “forms of clerical work ...channelled through apps managed by firms”.⁹ Examples include platforms such as Uber or Airbnb, which are also referred to as lean platforms, as they provide services with a minimal ownership of assets.¹⁰

The second is digital gig work: “contingent (task- or project-based) intangible work delivered digitally and done for money, organised via online outsourcing platforms that are marketplaces bringing together buyers and sellers”.¹¹ The tasks may vary from “menial and monotonous parceled activities requiring some sort of judgment beyond the understanding of artificial intelligence (e.g. tagging photos, valuing emotions or the appropriateness of a site or text, completing surveys) to bigger and more meaningful works such as the creation of a logo, the development of a site or the initial project of a marketing campaign.”¹²

It was estimated the revenues of the platform economy to be worth US\$4.8 billion in 2016, and that it could grow up to US\$25 billion in 2020.¹³ By 2020, however, the revenues of the global platforms were more than twice that estimate - US\$52 billion, and the number of digital platforms had also grown to 777 (from 142 in 2010).¹⁴ Much less is known about the size of the workforce in the platform economy “owing to non-disclosure of data on the part of the platforms.”¹⁵

Irrespective of the absolute number, there is much debate on the quality and

experience of gig work in the platform economy. Gig work is not what the International Labour Organization (ILO) terms decent work: “work that is productive; ensures equality of opportunity and treatment for all women and men; delivers a fair income, security in the workplace and social protection for families; provides prospects for personal development; and gives workers the freedom to express their concerns, organize and participate in decisions that affect their working lives.”¹⁶

In contrast to the ILO’s argument about gig work, platforms have projected themselves as “a revolution in labour markets, suggesting that they can help lift people out of poverty.”¹⁷ Development agencies, such as the United Nations Development Program (UNDP) and the World Bank, have also promoted digital gig work as a vehicle for development. Youth unemployment is a particularly pressing concern in the global South. Of the world’s 1.2 billion people between the ages of 15 and 24, 87% are to be found in the global South, two thirds of whom are either unemployed or “trapped in low quality jobs”.¹⁸ Digital gig work is viewed as having the potential to alleviate un(der)employment, when augmented with broader national employment strategies.

To gauge the extent to which gig work offers developmental possibilities, as opposed to confining people in jobs of low quality that do not qualify as decent, this paper will examine gig work on digital platforms in India which, in 2020, was home to 8% of all platforms, second only to the US (29%).¹⁹ Specifically, the paper will examine the conditions of geographically tethered gig workers on digital platforms. The leading think tank of the Government of India, estimated that, in 2020-2021, the number of geographically tethered gig workers was 7.7 million (2.6% of non-agricultural work-force), a figure expected to treble by 2029-2030 to 23.5 million (6.7% of non-agricultural work force).²⁰

Before turning directly to India, the following section will briefly explore the broader social and economic context for the emergence of the platform economy globally. Specific reference will be made to the growing emphasis on market-driven development as opposed to state-led models, along with rising call for flexibility and the deployment for ICTs. Section 3 will examine the socio-economic backdrop for the emergence of the platform economy in India, while Section 4 will describe the key challenges facing gig work in the country. Section 5 will take examples of the hope vested in the platform economy and its developmental potential by formalising gig work. The paper will conclude with a brief discussion in Section 6 about the lessons from India, and highlight the institutional limits of the developmental possibilities of gig work.

2. The socio-economic context for the emergence of the digital platform economy

This section will situate the emergence of the digital platform economy within

the socio- economic reconfiguration that took place as ICTs became the handmaiden for broad institutional changes that were the leitmotif of the late twentieth century. The 1980s witnessed the ideological ascendancy of the “Washington consensus”, so called because of its promotion by the World Bank, the International Monetary Fund and the US Treasury.²¹ Driven in part by the disappointing record of autarkic, statist development models adopted by many countries of the Global South, the Washington Consensus sought ‘public sector reform’ by ‘reinventing governments’ to eliminate inefficiency and corrupt processes in the public sector, and by limiting the role of the state to supporting the smooth functioning of markets.²² A ‘new institutional economics’ emerged to make the case for such support in the form of minimizing transactions costs or the effects of imperfect information, and to enforce contracts and property rights.²³

ICTs also facilitated another aspect of the Washington Consensus i.e., free capital movement transcended state boundaries, and an open international trading regime, institutionalised in the form of the World Trade Organisation, to lay the basis for an informational capitalism organised in a global economy which had “the capacity to work as a unit in real time, or chosen time, on a planetary scale”.²⁴ The effort to free economic activities from the fetters of the state, and to incentivise the private sector and entrepreneurs, were complemented by studies which highlighted the importance of deploying of ICTs, to move away from mass production to flexible specialisation, as a means of nimbly adapting to shifting market demands.²⁵

Flexibility also became evident in the treatment of labour. With flexibility and the diffusion of ICTs in an increasingly global economy, the “traditional form of work, based on full-time employment, clear-cut occupational assignments, and a career pattern over the life-cycle is being slowly but surely eroded away” in all OECD countries.²⁶ The rise in flexible and non- standard forms of work arrangements to support the mobility, profitability and productivity of capital has led to de-institutionalization and the “individualization of labor in the labor process”.²⁷ Digital gig work is thus a continuation of the earlier trends of flexible and non- standard forms of employment.

In the Global South, where a career-long full-employment was, of course, a privilege not available much of the workforce, the term “informal sector” was used to describe alternate work options.²⁸ Scholars then conceptualised informal economic activities as the negation of *formal* economic activities, and the informal economy as what the formal economy was not: invisible, uncountable/unaccountable, temporary, unstable and lacking protection against work-related risks.²⁹ Conceptualising the informal economy as the opposite of the formal economy led to teleological assumptions that the former would develop into the latter. if and when the necessary economic and infrastructural inputs and policies were in place.³⁰ To the extent that the informal economy was seen to have potential the qualities of informal economies, it focused only on petty trade and small-scale establishments,³¹ for

instance, emphasising the wealth of opportunities to test entrepreneurial ideas and the elasticity of possible returns (e.g. there being no cap on how much a person could make, in contrast to defined job roles and remuneration in the formal economy).

If the academic literature has found it difficult to conceptually characterise formality and informality, their definition and measurement in practice has been no less fraught. In India, for instance, the terms “informal” and “unorganised” are used interchangeably. “The unorganised sector [economy] consists of all unincorporated private enterprises owned by individuals or households engaged in the sale and production of goods and services operated on a proprietary or partnership basis and with less than ten total workers.”³² Thus, the organised economy is anything that the unorganised economy is not. Similarly, informal workers are “those working in the unorganised sector or households, excluding regular workers with social security benefits provided by the employers and the workers in the formal sector without any employment and social security benefits provided by the employers’ terms.”³³

By the 1980s, the emergence of a global economy, tied together by global commodity chains (GCCs),³⁴ not only challenged the distinction between the economies of the Global South and the Global North,³⁵ but also the formal-informal dichotomy. Thus, researchers argued that the growth of the informal economy is inherent to the capitalist economy, and that the formal/informal economies must be understood as “a series of transactions that connect different economies and spaces to one another”, rather than in dichotomous terms.³⁶ Specifically, the argument was that reliance on economic enterprises and workers in the informal economy represented attempts by the formal economy to reduce input labour and input costs and thereby increase competitiveness; it also represented the reaction to the power of organized labor, to state regulation of the economy (notably, taxes and social legislation), and to global competition.³⁷ Indeed, in India, even within the formal or the organized sector, which accounted for only 13.1 percent of employment in 2017-18, 40 percent of the workforce was informal.³⁸

As already mentioned, most platforms do not recognize gig workers as employees; instead, workers are classified as self-employed persons or independent contractors, which potentially allows employers to minimize their legal and social responsibilities, such as providing minimum wages or other employment benefits. Playing the role of an intermediary allows a platform provider to shift most of the costs, risks and liabilities to the other two parties. with no guarantee that a worker will not be terminated on the completion of a gig or task

In the context of the platform economy, there are broader concerns about the

relationship between gig workers and the operations of the platform. A primary concern has to do with the mechanism of control over workers. Platforms embody “an *architecture* - a design for products, services, and infrastructure facilitating network users’ interactions - plus set of *rules*; that is, the protocols, rights, and pricing terms that govern transactions” (original emphasis).³⁹ As the implementation of the architecture and rules of digital platforms increasingly relies on advances in artificial intelligence and machine learning, there are concerns about “algorithmic management”, defined as work settings where “human jobs are assigned, optimized, and evaluated through algorithms and tracked data”.⁴⁰ Algorithmic management continuously tracks workers’ behavior; constantly evaluates the performance of workers from client reviews, and the client’s acceptance or rejection of their work; automatically implements decisions, without human intervention; and, requires workers to interact with a “system” rather than humans, thus depriving them of opportunities for feedback or discussion and negotiation with their supervisor, as would be typical in offline jobs.⁴¹ Worse, there is low transparency.

The low transparency, or opacity, renders algorithmic management particularly problematic as it implies that workers have no access to the logic or criteria behind key decisions being made about their performance evaluation and, consequently, future opportunities.⁴² Opacity may result from one of three reasons. First, it could be the result of intentional institutional action or “willful self-protection by corporations in the name of competitive advantage, but this could also be a cover for a new form of concealing sidestepped regulations, the manipulation of consumers, and/or patterns of discrimination.”⁴³ Second is opacity arising from technical illiteracy, or not possessing the specialized skills required to read, write, and comprehend code. The third, which is understanding “the algorithm in action, operating on data”,⁴⁴ is the most insidious because of the complexity that arises when algorithms, however clearly written, *learn* from interacting with the heterogeneous properties of vast quantities of data. The broader argument is about how behavior in cyberspace is governed by code, and that the architectures of some spaces are more regulable than others, i.e., the architecture of cyberspace is its politics.⁴⁵

Notwithstanding these broad arguments, the politics, the regulation and the manifestation of the platform economy is unlikely to be the same the world over. Various studies emphatically confirm the lack of “systematic links” between ICTs and employment, as specific outcomes depend largely on “macro-economic factors, economic strategies, and sociopolitical context”.⁴⁶ The broad relationship between technology and work can only be understood as a “complex interaction within a social system” comprising “management decisions, systems of industrial relations, cultural and institutional environments and government policies.”⁴⁷ The next section will therefore turn to understanding the platform economy in the Indian social system.

3. The emergence of the platform economy in India

Since the 2007 launch of the e-commerce platform Flipkart (which was subsequently acquired by Walmart in 2018),⁴⁸ India has witnessed the emergence of several platforms that offer “work on-demand via apps”. To appreciate why and how they have attracted workers, it is crucial to note that a feature of the Indian economy is its decreasing dependence on

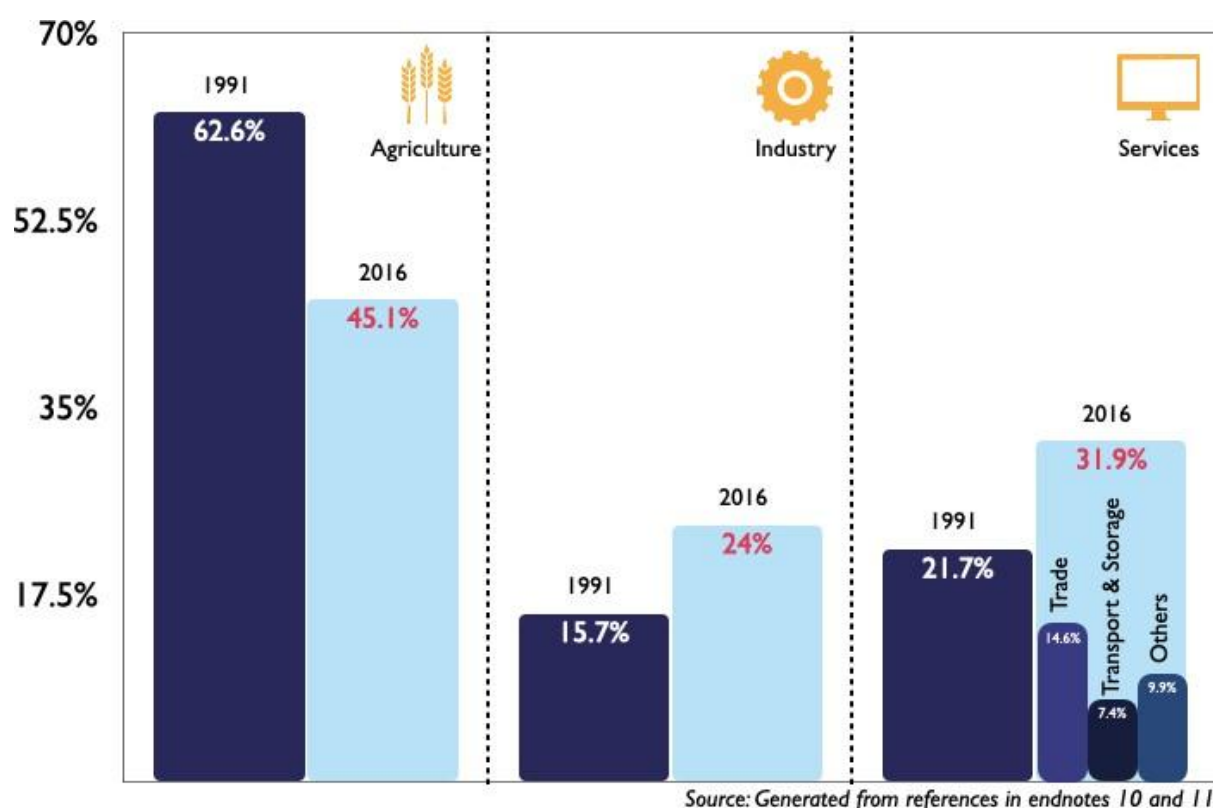


Figure 1: Changing proportion of employment by sector, 1991-2016

agriculture for employment, as Figure 1 shows.⁴⁹ The figure also shows that trade, and transport and storage (service sub-sectors in which many platforms operate), have led the way in offering jobs over the last two decades.⁵⁰

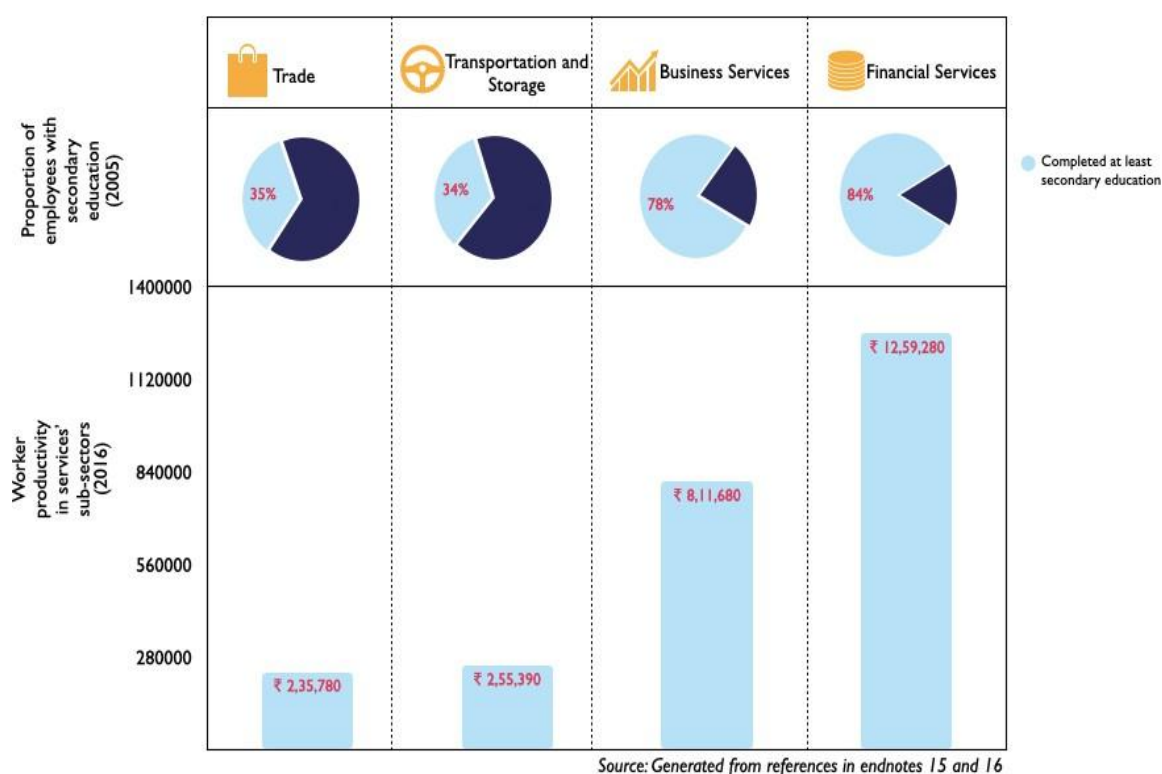


Figure 2: Comparing worker productivity in current Rupees (2016), and levels of secondary education (2005), by services sub-sectors

Despite remaining the largest source of employment, agriculture's contribution to gross value-added in 2016 was only 15.23 percent.⁵¹ The biggest contribution (51.13 percent) came from services, with the rest coming from industry. Likewise, between 1984 and 2010, the annual growth in productivity in services (4.9 percent) was ahead of the national figure (3.7 percent).⁵² But services include a range of activities which vary in the productivity and the skills they employ. In 2016, for instance, as Figure 2 shows, the productivity per worker in trade, and in transport, storage and communications was barely a third of what it was for business services, and less than a fourth of financial services.⁵³ Figure 2 also shows that the differences in productivity correspond to differences in education. In 2005, the share of employees with a secondary education in sub-sectors such as transport and communication, and wholesale and retail trade, was less than half of what it was in financial services and in business services.⁵⁴

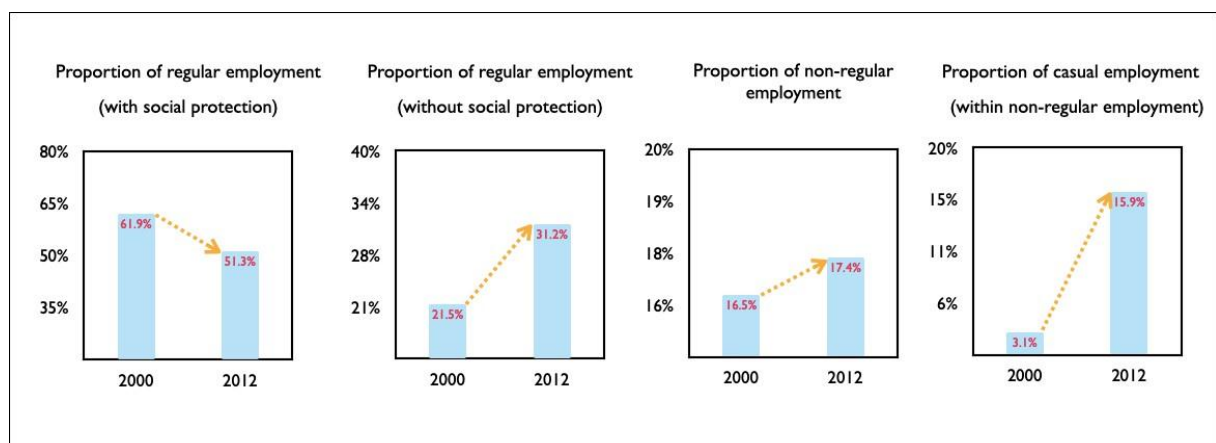
As the employment profile of the country has changed, those leaving agriculture have increasingly found livelihood opportunities in relatively low-value added sectors, with low educational barriers to entry.⁵⁵ There are at least four reasons why some of them find work in the app-based platforms. Firstly, they are attracted in part by the higher income, and the short and relatively predictable payment cycles, in comparison with other work alternatives.

A second factor is the proliferation of mobile phones with internet connectivity. A country that had barely 1.1 telephones per 100 people in 1994,⁵⁶ had 93.27 by 2018. By 2018, 98.12 percent of all telephone connections were mobile connections, of which 34.7 percent had broadband internet access.⁵⁷ A third factor in the growth of the platform economy has been the increasing last-mile access for platform workers to reach customers. India is the world's biggest two-wheeler (motorcycle and scooter) market,⁵⁸ with a third of households owning one.⁵⁹ A fourth factor is the supply of venture capital to support the expansion of the industry.⁶⁰ Between 2013 and 2019, of a total of US\$39.7 billion in venture capital investments in India,⁶¹ of which at least an estimated US\$14.82 billion went to platforms.⁶²

4. How 'decent' is gig work in the platform economy in India?

The sectoral shift in employment has also been accompanied by changes in work conditions in the economy. Since the platforms belong to the organised economy, the conditions of work in this segment of the economy are worth examining. Although employment in the organised economy is relatively small, it accounted for 45.6 percent of total output in 2012.⁶³ In services too, the organised economy contributed 50.6 percent to the total sectoral output, whereas it accounted for only 25.6 percent of sectoral employment.

Despite the productivity of the organised economy, including in the services sector, Figure 3 shows that the proportion of regular employment with social protection has declined. This decline has been accompanied by an increasing proportion of regular employment without



Source: Generated from reference in endnote 27

Figure 3: Changing proportions of employment in the organised economy with access to social protection, 2000-2012

social protection, and non-regular employment, especially the casualisation of labour.⁶⁴ Most platform workers, euphemistically labelled "independent contractors" or "partners", belong to either the (growing) regular non-protected or the casual worker categories.

What makes it possible for platforms to get away using such euphemisms for gig workers is a murky and rapidly evolving legal and policy context in which the Indian platform economy operates. Here there are two particularly contentious issues. The first has to do with classifying workers i.e. whether they are independent contractors/ self-employed, or are they employees of the platform? Being treated as independent contractors limits the protections (minimum wages, working hours, gratuity pay), social security (Employees Provident Fund, Employees State Insurance), and collective bargaining rights available to platform workers under law, as orders or as judgements from the High Courts of Delhi⁶⁵ and Karnataka⁶⁶ have indicated⁶⁷ Some have even argued that, in the absence of regulatory clarity, worker grievances should be taken up under the Consumer Protection Act instead, by treating workers as “consumers” of the platform company.⁶⁸ For the most part, platform work has remained invisible in the context of labour laws, including labour contracting.⁶⁹

Since the state significantly shapes economic and labour market policies,⁷⁰ the Indian Federation of App-based Transport Workers, along with two gig workers, filed a Public Interest Litigation⁷¹ in September 2021 to demand recognition of gig workers as either as unorganised workers under the Unorganized Workers Act, 2008, in order to bring gig workers under the purview of the existing social security system. The litigation follows the UK ruling in the matter of Uber BV and Ors (appellants) v. Aslam and Ors (respondents),⁷² and relies on the tests of economic control, supervision, and integration to contend that there exists an employee–employer relationship between platforms and gig workers, despite the claims made by platforms.⁷³ Their appeal for recognition as unorganised workers aims to ensure the inclusion of gig workers in the social security schemes of the Central Government applicable under this Act.⁷⁴

It was only in 2020, with the passage of the Code on Social Security (CoSS), one of four labour codes, that gig work and platform work even found explicit mention in the law.⁷⁵ A gig worker is “a person who performs work or participates in a work arrangement and earns from such activities outside of traditional employer-employee relationships”, while platform work refers to “a work arrangement outside of a traditional employer-employee relationship in which organisations or individuals use an online platform to access other organisations or individuals to solve specific problems or to provide specific services or any such other activities which may be notified by the Central Government, in exchange for payment.”

This CoSS seeks to replace nine laws related to social security, including the Employees’ Provident Fund Act, 1952, the Maternity Benefit Act, 1961, and the Unorganised Workers’ Social Security Act, 2008. Social security refers to measures to ensure access to health care and provision of income security to workers. The CoSS, and the draft rules proposed to operationalise it, indicate

that platform companies will be required to make a contribution to workers' social security funds. While the CoSS extends social security benefits to gig workers on platforms, the Occupational Safety, Health and Working Conditions Code 2019,⁷⁶ the Code on Wages 2019,⁷⁷ and the Code on Industrial Relations Code 2020,⁷⁸ make no mention of either gig workers or platforms.

The second contentious issue is that platforms have long argued that they should be classified as technology companies.⁷⁹ But, over time, these companies have been brought under the purview of certain sectoral regulations pertaining to transportation, food delivery, and other services.⁸⁰ For example, the Food Safety and Standards Authority has issued specific guidelines for "ecommerce food business operators."⁸¹ Similarly, the Motor Vehicles (Amendment) Act, 2019, requires platform companies offering transport services to possess an "aggregator" classification and a state license to operate.⁸² The Motor Vehicle Aggregator Guidelines, 2020, provides a framework of rules for states to regulate the service conditions and tariffs of the aggregators.⁸³ These sector-specific regulations can also affect the earnings and working conditions of platform workers.⁸⁴ Given the range of entities potentially involved in regulating platform companies, a concern is that this may take place without a clear demarcation of accountability between these entities.⁸⁵ As a consequence, no one entity is held accountable for ensuring workers' rights or regulating their working conditions in this regulatory landscape.

5. The faith in the developmental possibilities of platforms

Even as dual uncertainties of employment status and regulatory classification allow platform companies to make their own rules with respect to wages and working conditions, platforms claim that their policies promote fairness and represent a trajectory for incorporating gig workers into the formal economy. Indeed, some claim that platforms serve as "on-ramps to formalisation in the Global South."⁸⁶

In a commentary on the issues surrounding gig work in the platform economy, a senior executive at a platform conceded that "some of the concerns are *fair*" and, as the digital platform sector grows, "and the more successful it is perceived to be, the more responsible and thoughtful it needs to be in everything it does."⁸⁷ In particular, "firms can't seek rising valuations, on the one hand, and treat their gig workers *unfairly*, on the other." With this nod to the criticism of working conditions on platforms, the remainder of the commentary justifies why "expecting a nascent industry to transform labour market conditions overnight is both unrealistic and unfair", and implores us to acknowledge that what "platforms have done is nothing short of a miracle both in terms of creating jobs as well as paying a fair wage."

The miracle is explained in terms of how the innovative application of technology to create efficient markets has enabled most platforms to extend “payouts beyond the minimum wage specified by the government.” However, there are reasons to suggest that claims about a miracle, and the frequent reference to fairness (italicised in the previous paragraph), are not fully justified. Empirically, as the research by Fairwork shows⁸⁸, the claim that most workers are paid a fair wage is not borne out by the evidence. It is worth noting here that the evidence does not compare platform work with jobs that are as “....as well paying or fulfilling as their more upmarket cousin— the IT industry.” A comparison between the labour market for college graduates (mostly from the STEM disciplines⁸⁹) in the IT industry, with that constituted by youth who have “migrated from rural farmlands” is puzzling. Rather, the evidence shows that an hourly minimum wage, after work-related expenses, is enjoyed only by a limited section of workers on certain platforms.

The inability of sections of platform labour to earn a minimum wage, points to at least two analytical limitations to the efficient markets argument. First, markets are not merely mechanisms for economic transactions; they are social institutions as well, structured by regulation⁹⁰. Indeed, firms cannot function smoothly without state regulation,⁹¹ especially when digital platforms - which are two-sided markets - generate network externalities which affect both competition and customer choice.⁹² There is little doubt that digital platforms reduce transactions costs. But to the extent that workers doing gigs, whether online or offline, cannot earn a minimum wage, it only highlights their precarity and vulnerability in increasingly flexible labour markets, and upends notions of “premature regulation”. While the nature of regulation requires debate, the need for regulation is not in question.

A second issue with the argument of efficient markets is that they fail under conditions of information asymmetry. The article indicates that platforms are “revolutionary” as customers rate every transaction, and workers with higher ratings get more assignments and an opportunity to earn more. However, there is little transparency on how the algorithms underlying platforms use client review and other metrics to determine the ratings which allow workers to build a reputation.⁹³

Another report from a platform argues that “platforms are at the heart of India’s digital economy revolution” and that “digitalisation/platformisation of work has given rise to a new classification of labour - platform labour - different from the traditional dichotomy of formal and informal labour.”⁹⁴ Despite highlighting the futility of viewing employment in terms of the formal/informal dichotomy, by highlighting platform labour as a new classification, it adds a third. Why this trichotomy offers an analytically superior means of understanding labour markets is unclear. The report

provides a justification of sorts by arguing that since the platform economy straddles the formal and the informal by, for instance, offering workers reliable payments and credit access, while allowing them to retain the flexibility of choosing their work hours, it “expedites the “formalisation” process of the economy as a whole”⁹⁵ by “plugging in a tech tool to streamline the process with an algorithmic coherence.”⁹⁶ Even assuming that platforms, wherein “innovation is the new medium of exchange”,⁹⁷ offer “a plethora of benefits to workers and the economy at large”,⁹⁸ will inevitably pave the way to formalization is far from clear.

The report’s response to how the Indian social system will transition, is to spell out five broad policy initiatives for the government. Foremost, it calls for the restructuring of “social security nets in tandem with classifications recognised in [the] Code on Social Security, 2020”.⁹⁹ While the acknowledgement of the need for social security for gig workers and platform work is welcome, the silence of the report on the other codes, especially the Code on Wages, 2019, and the Industrial Relations Code, 2020, is surprising. This is especially because wages and management control, as evidence from Fairwork suggests,¹⁰⁰ go to the heart of discontent with platform work.

The report also calls on the government to design progressively universal, portable, and inclusive social protection, upskill young workers, and ensure access to credit and finance (“platformisation of everyday must go hand-in-hand with the financialisation of everything”¹⁰¹). Thus, the report is vocal in placing the onus on government policies to unlock the potential of the platform economy. Yet, it is the silence of the recommendations on the changes to be made in management decisions, or in systems of industrial relations, that speaks more loudly.

The silence should perhaps not come as a surprise because, contrary to the sweeping claim in the report that “the platform economy leverages both demography and technology to promote financial and social inclusion across India”,¹⁰² connections in the informational economy are made “according to criteria of valuation and devaluation enforced by social interests that are dominant in these networks.”¹⁰³

6. A concluding discussion

The rise of the platform economy, due to institutional changes and emerging ICTs in the late twentieth century globally has come with possibilities and promises especially in the Global South where unemployment is widespread, as is informality. In such a situation, platforms offer work opportunities that might not otherwise exist. However, the assumption that these work opportunities offer a way to formalization is unfounded for many reasons. Firstly, from the perspective of the capitalist imperative, informal gig workers and the formal platform business entity are so inextricably linked that it is hard to imagine

the existence of the platform economy but for the availability and efforts of relatively underpaid gig workers. Indeed technological determinism and unidirectionality of this posited trajectory ignores the long history of how teleological projections of social change have come to naught - from Marxist assertions about the inevitable transition to communism,¹⁰⁴ or the assertions of modernists about the promise of mass-consumption societies.¹⁰⁵ Secondly, gig workers have(at least until now) little to no legal backing. As a result, they are treated as 'independent contractors', which allows platforms to determine their conditions of work.

Thirdly, access to contemporary ICTs, especially in the form of smartphones, alone does not make for inclusive development. The platform economy is based on significant power asymmetry between an atomized workforce consisting of individual contractors, and platforms that have financial muscle, thanks to venture capital backing, and the technological might to algorithmically manage workers and the work process. For these reasons, the platform economy risks of going down to the "low-road" to development as opposed to the "high road".¹⁰⁶

Notes

- ¹ Bukht, R. and Heeks, R. 2017. Defining, Conceptualising and Measuring the Digital Economy. Development Informatics Working Paper no.68, Centre for Development Informatics, University of Manchester, UK, p.1
- ² Srnicek, N. 2017. Platform Capitalism. Cambridge, UK: Polity Press. p.26
- ³ Scholtz, T. 2016. Platform Cooperativism: Challenging the Corporate Sharing Economy. New York, NY: Rosa Luxemburg Stiftung. p.4 It should be noted that as infrastructure mediating between different groups, platforms hardly need be digital. For example, shopping malls that bring together various stores, and customers seeking different products, are examples of platforms providing infrastructure that is physical. See, Rysman, M. 2009. The economics of two-sided markets. Journal of Economic Perspectives, 23(3), 125-143.
- ⁴ Schmidt, F.A. 2017. Digital Labour Markets in the Platform Economy: Mapping the Political Challenges of Crowd Work and Gig Work. Bonn, Germany: Friedrich-Ebert-Stiftung. p.5. There have emerged many terms related to the platform economy, such as the sharing economy, or the on-demand economy (to describe the overall domain), and terms such as online labor, crowd work or digital labor, to describe the type of work done. See, Heeks, R. 2017a. Digital Economy and Digital Labour Terminology: Making Sense of the "Gig Economy", "Online Labour", "Crowd Work", "Microwork", "Platform Labour", etc. Development Informatics Working Paper no.70, Centre for Development Informatics, University of Manchester, UK.
- ⁵ Harvey, D. 1990. The Condition of Postmodernity: An Enquiry into the Origins of Cultural Change. Oxford, UK: Blackwell Publishers
- ⁶ International Labour Organization (ILO). 2021. World Employment and Social Outlook: The Role of Digital Labour Platforms in Transforming the World of Work. Geneva, Switzerland: ILO.
- ⁷ International Labour Organization (ILO). 2016. Non-Standard Employment around the World: Understanding Challenges, Shaping Prospects. Geneva, Switzerland: ILO.
- ⁸ Srnicek, 2017:27-28. op.cit., pp.27-28, classifies digital platforms into four types. The first are advertising platforms, such as Facebook, which use information mined about users as basis for generating revenues by selling online advertising space. The second are cloud platforms, such as Amazon, which rent out their infrastructure to third parties. The third are industrial platforms, such as Siemens, which build the infrastructure to transform traditional manufacturing into internet-connected processes that lower production costs and transform goods into services. The fourth are platforms, such as Rolls Royce, which generate revenue by using other platforms to transform a traditional good into a service and by collecting rent or subscription fees on them.
- ⁹ De Stefano, V. 2016. The Rise of the "Just-In-Time Workforce": On-demand Work, Crowdwork and Labour Protection in the "Gig-Economy". Conditions of Work and Employment Series No.71, Geneva, Switzerland: International Labour Organization. p.1
- ¹⁰ Srnicek, 2017. op.cit.
- ¹¹ Heeks, R. 2017b. Decent Work and the Digital Gig Economy: A Developing Country Perspective on Employment Impacts and Standards in Online Outsourcing, Crowdwork, etc, Development Informatics Working Paper no.71, Centre for Development Informatics, University of Manchester, UK. p.1
- ¹² De Stefano, 2016:2. op.cit.,
- ¹³ World Bank. 2015. The Global Opportunity in Online Outsourcing. Washington DC: World Bank. p.3
- ¹⁴ ILO, 2021:19-20. op.cit.
- ¹⁵ ibid., p.19
- ¹⁶ Berg, J., Furrer, M., Harmon, E., Rani, U. and Silberman, M.S. 2018. Digital Labour Platforms and the Future of Work: Towards Decent Work in the Online World. Geneva,

-
- Switzerland: International Labour Organization. p.10
- 17 Graham, M., Lehdonvirta, V., Wood, A., Barnard, H., Hjorth, I., and Simon, D. P. 2017. *The Risks and Rewards of Online Gig Work at the Global Margins*. Oxford, UK: Oxford Internet Institute. p.2
- 18 World Bank, 2015:46. op.cit.,
- 19 ILO., 2021:19. op.cit.
- 20 NITI Aayog. 2022. *India's Booming Gig and Platform Economy: Perspectives and Recommendations on the Future of Work*. New Delhi: Government of India. pp.xvii-xviii
- 21 Gore, C. 2000. The rise and fall of the Washington Consensus as a paradigm for developing countries. *World Development*. 28(5):789-804.
- 22 Heeks, R. (ed.) (1999). *Reinventing government in the Information Age: International Practice in IT-enabled Public Sector Reform*. Abingdon, UK: Routledge.
- 23 Bardhan, P. 1989. The new institutional economics and development theory: A brief critical assessment. *World Development*. 17(9):1389-1395.
- 24 Castells, M. 2010, (2nd edition) *The Rise of the Network Society*. Volume 1 of *The Information Age: Economy, Society and Culture*, Chichester, UK: Wiley-Blackwell. p.101
- 25 Piore, M.J. and Sabel, C.E. 1984. *The Second Industrial Divide: Possibilities for Prosperity*. New York, NY: Basic Books; Jaikumar, R. 1986. Postindustrial manufacturing. *Harvard Business Review*. November- December:69-76.
- 26 Castells, 2010:290. op.cit.
- 27 Ibid.:82.
- 28 Hart, K. 1973. Informal income opportunities and urban employment in Ghana. *The Journal of Modern African Studies*. 11(01):61-89.
- 29 In this paper, we use informal sector, informal economy and informal employment interchangeably, even though we are aware there are conceptual differences among them. For a survey of the conceptualisations of informality, and the terminological differences, see Chen, M.A. 2012. *The Informal Economy: Definitions, Theories and Policies*. WIEGO Working Paper No.1. Women in Informal Employment: Globalizing and Organizing, Cambridge, MA.
- 30 Chen, 2012. op.cit.
- 31 Hart, 1973. op.cit.
- 32 National Commission of Enterprises in the Unorganised Sector (NCEUS). 2009. *The Challenge of Employment in India: An Informal Economy Perspective*. Volume I – Main Report. New Delhi: NCEUS. p.3 Accessed from http://dcmsme.gov.in/The_Challenge_of_employment_in_India.pdf. The challenge on a precisely defining the informal/unorganised sector also means that accurately measuring the size of this economy is notoriously hard. Estimates exist, but they are unlikely to capture the entire range of informal activities which, are often invisible (or 'underground'), and the engagement of individuals in such activities is unpredictable and unaccountable, See, Williams, C.C. 2019. *The Informal Economy*. Newcastle, UK: Agenda Publishing
- 33 NCEUS, 2009:3, op.cit.
- 34 Gereffi, G. 1996. Global Commodity chains: New forms of coordination and control among nations and firms in international industries. *Competition and Change*. 4:427-439.
- 35 For the rise of the informal economy in the global North, see Sassen, S. 1993. The informal economy: Between new developments and old regulations. *The Yale Law Journal*. 103:2289-2304.
- 36 Roy, A. 2005. Urban informality: Toward an epistemology of planning. *Journal of the American Planning Association*. 71(2):147-158.
- 37 Portes, A., M. Castells, M. and Benton, L.A. (eds.) 1989. *The Informal Economy: Studies in Advanced and Less Developed Countries*. Baltimore, MD: John Hopkins University Press
- 38 Murthy, S.V.R. 2019. Measurement of informal economy: Indian experience. Presented at the IMF Seventh Statistical Forum: Measuring the Informal Economy, 14-15 November, Washington D.C.,

- 39 Eisenmann, T., Parker, G., and Van Alstyne, M.W. 2006. Strategies for two-sided markets, *Harvard Business Review*. 84(10):92-101. p.95
- 40 Lee, K., Kusbit, D., Metsky, E. and Dabbish, L. 2015. Working with machines: The algorithmic and data- driven management on human workers. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, Seoul, 18-23 April, 1603-1612. p.1603
- 41 Möhlmann, M. and Zalmanson, L. 2017. Hands on the wheel: Navigating algorithmic management and Uber drivers' autonomy. In *Proceedings of the Eighth International Conference on Information Systems (ICIS 2017)*, December 10-13, Seoul, South Korea
- 42 Burrell, J. 2016. How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data and Society*. January-June, 1-12.
- 43 Ibid.:4
- 44 Ibid.:5
- 45 Lessig, L. 2006. *Code version 2.0*. New York, NY: Basic Books.
- 46 Castells, 2010:280. op.cit.
- 47 Ibid.:256.
- 48 Russell, J. 2018. Walmart completes its \$16 billion acquisition of Flipkart. Accessed from <https://techcrunch.com/2018/08/20/walmart-flipkart-deal-done/>
- 49 Data accessed from <https://databank.worldbank.org/source/world-development-indicators>.
- 50 Construction generated the most jobs in this period (35.74 percent). Bhattacharya, P. 2018. Which are the top sectors that generate employment in India? *Livemint*. 12 April. Accessed from <https://www.livemint.com/Politics/7XXmUWykSEGKoWXJqUuHM/Which-are-the-top-sectors-that-generate-employment-in-India.html>
- 51 Data accessed from <https://databank.worldbank.org/source/world-development-indicators>.
- 52 Amirapu, A. and Subramanian, A. 2015. *Manufacturing or Services? An Indian Illustration of a Development Dilemma*. Working Paper 409. Washington DC: Center for Global Development. Accessed from <https://www.cgdev.org/publication/manufacturing-or-services-indian-illustration-development-dilemma-working-paper-409>
- 53 Bhattacharya, 2018. op.cit.
- 54 Amirapu and Subramanian. 2015. op.cit.
- 55 It is suggested that, "construction is becoming a refuge employment sector as its productivity declines." See, Majid, N. 2019. *Structural Change and Employment in India*. ILO-SIDA Partnership on Employment Working Paper No.1. Geneva, CH: International Labour Organization (ILO), p.13. Accessed from https://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_735166.pdf Productivity in construction was INR 126,560 in 2106 (Bhattacharya, 2018. op.cit) and, in 2005, only 14.4 percent of its workforce had a secondary education (Amirapu and Subramanian, 2015. op.cit.).
- 56 Bell, T., Adam, J.A. and Lowe, S. J. 1996. *Communications*. *IEEE Spectrum*. 33(1):30-41.
- 57 Department of Telecommunications (DoT). 2018. *Telecom Statistics India - 2018*. Ministry of Communications, Government of India. Accessed from: <https://dot.gov.in/sites/default/files/statistical percent20Bulletin-2018.pdf?download=1>
- 58 Raman, A. 2020. *The Power of Two Wheels. Bike-Taxi: India's New Shared Mobility Frontier*. Ola Mobility Institute. Accessed from <https://olawebcdn.com/ola-institute/bike-taxi-report.pdf>
- 59 Kundu, T. and Bhattacharya, P. 2018. India runs on two wheels and animal carts, data shows. *Livemint*. 8 August. Accessed from <https://www.livemint.com/Auto/N3Ixxh0Mv3eu2vWfzeiRw3N/India-runs-on-two-wheels-and-animal-carts-data-shows.html>
- 60 Although most platforms are yet to show a profit, their ability to attract investments is explained not merely by the returns from service provision but also by the potential value

-
- of digital data that is generated “before, during, and after service provision”. See van Doorn, N. and Badger, A. 2020. Platform capitalism’s hidden abode: Producing data assets in the gig economy. *Antipode* 52(5):1475-1495.
- ⁶¹ Sheth, A., and Krishnan, T.S. 2020. Venture capital investments in India. India Venture Capital Report 2020: Perspectives on the Funding and Start-up Ecosystem. Bain & Company. Accessed from <https://www.bain.com/insights/india-venture-capital-report-2020/>
- ⁶² For details of the estimate, see Fairwork India Ratings 2020: Labour Standards in the Platform Economy. Bangalore, India; Oxford, United Kingdom. https://fair.work/wp-content/uploads/sites/131/2020/12/Fairwork_India_2020_report.pdf. Appendix IV.
- ⁶³ The share of the organised sector in total employment barely increased from 10.3 percent to 13.9 percent between 2000 and 2012. Majid, 2019, op.cit.
- ⁶⁴ The distinction here is between regular paid employment and non-regular employment. Regular paid employment has some social security or contractual protection associated with it, while regular paid (non-protected) employment has no such security or protection. Non-regular employment includes casual employment and self-employment. An organised economy has both regular and non-regular labour, while an unorganised economy is dominated by non-regular labour. See, Majid, 2019. op.cit.
- ⁶⁵ Press Trust of India. 2017. HC asks striking drivers of Ola and Uber to end stir. *The Economic Times*. 28 February. Accessed from <https://economictimes.indiatimes.com/news/politics-and-nation/hc-asks-striking-drivers-of-ola-and-uber-to-end-stir/articleshow/57393991.cms>
- ⁶⁶ Ms Ayantika Mondal v State of Karnataka WP 6757/2020 (High Court of Karnataka). Also see Raj, A. 2020. Scarce work, irregular earnings: drivers, delivery agents struggle. *The Quint*. 12 May. Accessed from <https://www.thequint.com/news/india/karnataka-gig-economy-scarce-work-irregular-earnings-drivers-delivery-agents-struggle>
- ⁶⁷ It is worth noting that existing laws, such as those on minimum wages and maternity benefits, do not apply to most segments of the Indian labour force, and not just to platform workers. Additionally, in the wake of COVID-19, the large-scale migration of workers to their home states and the scarcity of workers in urban areas, various Indian states announced that they would relax labour laws, as they pertain to either working hours, or the monitoring of working conditions for various labour segments and sectors of work, in order to grow employment. Many of these announcements have since been withdrawn following domestic and international pressure, but others have been operationalised. See for instance, *The Hindu*. 2020. Major rehaul of labour laws in Karnataka. 23 July. Accessed from <https://www.thehindu.com/news/national/karnataka/major-rehaul-of-labour-laws-in-karnataka/article32176633.ece>.
- ⁶⁸ Bapna, N. and Saxena, A. 2017. Getting Ola and Uber drivers out of the regulatory grey zone. *The Wire*. 21 March. Accessed from <https://thewire.in/business/uber-ola-drivers-classified-consumers>
- ⁶⁹ Even if platform workers are to be treated as independent contractors, there is scope for some legal coverage under the Contract Labour Act, given the definition of “contractor” under section 2(c) of the Act. 'Contract Labour (Regulations & Abolition) Act, 1970'. n.d. Chief Labour Commissioner. <https://clc.gov.in/clc/acts-rules/contract-labour-regulation-abolition-act-1970> This Act regulates the engagement of contract labour, including work done through third party contractors. But platforms do not seem to be applying this law. See Rajkumar, M. 2020. The law for gig-workers in India. *Nyaaya*. 28 April. Accessed from <https://medium.com/nyaaya/the-law-for-gig-workers-in-india-4163cfaee18b>
- ⁷⁰ Agarwala, R. 2020. Informal workers and the state in India. In M. Chen, M.A. and Carré, F. (eds.), *The Informal Economy Revisited: Examining the Past, Envisioning the Future* (1st Edition). London, UK: Routledge. pp. 239–245
- ⁷¹ *The Indian Federation of App-based Transport Workers (IFAT) & Ors v Union of India &*

-
- Ors, W.P.(C) No. 001068/2021 (Supreme Court of India)
- 72 The Uber BV ruling concluded that the respondents (platform workers) were indeed workers under the relevant laws in the UK, stating that the “true agreement” between parties may not always be representative of the terms and conditions that may be so signed or agreed by the said parties, including explicitly releasing both the parties from any obligations from an employee-employer relationship: See, Uber BV and Ors (appellants) v. Aslam and Ors (respondents), [2021] UKSC 5 (Supreme Court, UK; February 19, 2021). <https://www.supremecourt.uk/cases/docs/uksc-2019-0029-judgment.pdf>
- 73 These tests for determining an employee-employer relationship have also been propounded in several Indian precedents: See, Dhrangadhara Chemical Works Ltd. v/s State of Saurashtra & Ors, AIR 1957 SC 264 (Supreme Court of India); Hussainbhai, Calicut v/s Alath Factory Thezhilali Union, Kozhikode & Ors, (1978) 4 SCC 257 (Supreme Court of India); Ram Singh v/s Union Territory of Chandigarh, (2004) 1 SCC 126 (Supreme Court of India)
- 74 The schemes include the Indira Gandhi National Old Age Pension Scheme (<https://nsap.nic.in/login/cbtMainPage.html>); the Family Benefit Scheme (<https://transformingindia.mygov.in/scheme/national-family-benefit-scheme/#:~:text=The%20scheme%20aims%20to%20provide,bracket%20of%2018%2D64%20years.>); the Janani Suraksha Yojana (<https://www.nhp.gov.in/janani-suraksha-yojana-jsy-pg>); the Janshree Bima Yojana (<http://www.rsby.gov.in/>); the Aam Admi Bima Yojana (<https://pmmodiyojana.in/lic-aam-aadmi-bima-yojana/>); and the Rashtriya Swasthya Bima Yojana (<https://www.india.gov.in/spotlight/rashtriya-swasthya-bima-yojana>).
- 75 Accessed from: https://labour.gov.in/sites/default/files/SS_Code_Gazette.pdf and Government of India. 2020. Draft of the Code on Social Security(Central) Rules. 2020. Ministry of Labour and Employment <https://labour.gov.in/whatsnew/draft-code-social-securitycentral-rules-2020>
- 76 https://labour.gov.in/sites/default/files/osh_gazette.pdf
- 77 <https://egazette.nic.in/WriteReadData/2019/210356.pdf>;
- 78 <https://egazette.nic.in/WriteReadData/2020/222118.pdf>
- 79 Uber has argued that it is “only a provider of a technology platform to be used by drivers to offer their services to passengers and not a rental taxi provider or an agency in the sense defined by the Motor Vehicles Act.” Johnson, T.A. 2016. New Karnataka rules ‘against tech’, says Uber India. The Indian Express. 6 July. Accessed on 16 November from <https://indianexpress.com/article/technology/tech-news-technology/new-karnataka-rules-against-tech-says-uber-india-2902392/> Also see Satish N vs State Of Karnataka WP No.30917/2016 (High Court of Karnataka) where it is argued that an aggregator like Uber “merely creates a technology-based platform.” Even in official government documents, companies declare their activities in ambiguous terms. For instance, the Altered Memorandum of Association (April, 2020) of Bundl Technologies (Swiggy, a food and grocery delivery platform) includes the objective of carrying on “the business or profession of providing a platform... to facilitate transactions, commerce, electronic commerce...”. It further states that the aim is to “provide, develop, establish, run, manage, operate courier services or logistics services and manifestation for collecting and delivering whether by own arrangements, engaging third parties or through agents...” It is therefore not unreasonable to interpret Swiggy as not just a technology company but as a logistics services company managing the end-to-end courier process including the workers themselves.
- 80 The CoSS considers platform companies as “aggregators” -- a digital intermediary that connects users of a service with sellers or service providers -- and lists nine categories of aggregators, including “ride sharing services”, “food and grocery delivery services”, “logistic services”, “e-market place”, and “professional services provider.” See section 2(2)

and The Seventh Schedule of The Code on Social Security, 2020. 29 September. Ministry of Law and Justice.

- ⁸¹ For recent directives, see Food Safety and Standards Authority of India. 2018. Direction relating to FSS (Licensing and Registration) regulations. FSSAI. https://fssai.gov.in/dam/jcr:9a216a62-f67c-4ba1-93e4-9ba1ec749ca3/Direction_License_Registration_13_12_2018.pdf
- ⁸² Motor Vehicle (Amendment) Act, 2019, s. 93 r/w s.1A ('definition of aggregator'). <http://egazette.nic.in/WriteReadData/2019/210413.pdf>
- ⁸³ These guidelines have several implications for drivers, including a cap of 20% aggregator commission per trip, limiting login hours to 12 hours at a stretch followed by 10 hours of rest and 10% cancellation charge for both driver and passenger. See Sections 13(4), 7(2)d, 14(1) of the Motor Vehicle Aggregator Guidelines, 2020. http://morth.nic.in/sites/default/files/notifications_document/Motor%20Vehicle%20Aggregators2711202015004_6.pdf Even before the issue of these guidelines, there has been regulation by states. For example, the Karnataka On-demand Transportation Technology Rule 2016 capped surge pricing of app-based taxi aggregators to ensure minimum earnings for non app-based cab drivers, and to protect consumers. Yadav, U. and Variar, M. 2016. Uber ordered to limit surge pricing, comply with Karnataka rules for taxi aggregators. The Economic Times, 11 November. Accessed from <https://economictimes.indiatimes.com/small-biz/startups/uber-ordered-to-limit-surge-pricing-comply-with-karnataka-rules-for-taxi-aggregators/articleshow/55364437.cms>
- ⁸⁴ Various other laws such as competition laws, the Information Technology Act 2000, and the Digital Personal Data Protection Bill 2022 also affect platform workers. See, Ministry of Electronics and Information Technology. n.d. Information Technology Act 2000. MEITY. <https://www.meity.gov.in/content/information-technology-act-2000> and PRS. n.d. The Digital Personal Data Protection Bill, 2022. https://www.meity.gov.in/writereaddata/files/The%20Digital%20Personal%20Data%20Protection%20Bill%202022_0.pdf
- ⁸⁵ Changes to regulatory frameworks are being debated in several domains of activity in India, especially those shaped by the introduction of ICTs, and not just the platform economy. The division of labour and distribution of accountability among various regulatory authorities are a key component of this debate. See Kapur, D and Khosla, M. 2019. The reality of regulation in India. In Kapur, D and Khosla, M. (eds.), Regulation in India: Design, Capacity, Performance. New Delhi: Hart Publishing. However, platform services present an additional layer of complexity because their "users" include both service providers and end-consumers, which multiplies their regulatory requirements and the regulatory entities that are involved.
- ⁸⁶ Randolph, G., & Galperin, H. 2019. New Opportunities in the Platform Economy: On-ramps to Formalization in the Global South. Think-20 Japan 2019.(pp. 6. Accessed from <https://t20japan.org/policy-brief-new-opportunities-in-the-platform-economy/>
- ⁸⁷ Hari, T.N. 2021. India's gig economy and the question of quality jobs. Accessed from: <https://www.livemint.com/industry/retail/are-e-com-companies-creating-good-jobs-11634488029783.html>
- ⁸⁸ Ramachandran, S. and Raman, A. 2021. Unlocking Jobs in the Platform Economy: Propelling India's Post- Covid Recovery. Ola Mobility Institute (OMI). (https://olawebcdn.com/ola-institute/OMI_Platform_Economy_Report.pdf). pp13, pp.11. OMI describes itself as "a new-age policy research and social innovation think tank, focused on developing knowledge frameworks at the intersection of mobility innovation and public good...." and is funded by ANI Technologies Pvt. Ltd., the parent company of brand Ola (see: <https://ola.institute/>). Since Ola employs more workers than any other platform in India, in a domain (raid hailing) which is much larger than any other, the report is used to gain insight into how platforms conceptualise work.
- ⁸⁹ STEM refers to science, technology, engineering and management. Dey, S. 2017. The

-
- Evolution of the IT Engineer in India: Looking Back to See Ahead. Thesis submitted for the Master of Science by Research degree, International Institute of Information Technology, Bangalore
- ⁹⁰ Swedberg, R. 2005. Markets in society. In Smelser, N.J. and Swedberg, R. (eds.). *The Handbook of Economic Sociology* (2nd ed.). Princeton, NJ: Princeton University Press. pp.233-253.
- ⁹¹ Evans, P. 1995. *Embedded Autonomy: States and Industrial Transformation*, Princeton, NJ: Princeton University Press.
- ⁹² Reysman, M. 2009. The economics of two-sided markets. *Journal of Economic Perspectives*.23(3):125-142.
- ⁹³ ILO, 2021.op.cit.
- ⁹⁴ Ramachandran, S. and Raman, A. 2021. *Unlocking Jobs in the Platform Economy: Propelling India's Post- Covid Recovery*. Ola Mobility Institute (OMI). (https://olawebcdn.com/ola-institute/OMI_Platform_Economy_Report.pdf). pp13, pp.11. OMI describes itself as “a new-age policy research and social innovation think tank, focused on developing knowledge frameworks at the intersection of mobility innovation and public good....” and is funded by ANI Technologies Pvt. Ltd., the parent company of brand Ola (see: <https://ola.institute/>). Since Ola employs more workers than any other platform in India, in a domain (raid hailing) which is much larger than any other, the report is used to gain insight into how platforms conceptualise work.
- ⁹⁵ *ibid.*:11
- ⁹⁶ *ibid.*:30
- ⁹⁷ *ibid.*:38
- ⁹⁸ *ibid.*:13
- ⁹⁹ *ibid.*:16
- ¹⁰⁰ <https://fair.work/en/ratings/india/>
- ¹⁰¹ Ramachandran and Raman, 2021. op.cit :p.70
- ¹⁰² *ibid.*:72
- ¹⁰³ Castells, M. 2002. Preface. In Evans, P. (ed.). *Livable Cities? Urban Struggles for Livelihood and Sustainability*. Berkeley and Los Angeles, CA: University of California Press.
- ¹⁰⁴ Marx, K. and Engels, F. 1848. *The Communist Manifesto*. <https://www.marxists.org/archive/marx/works/download/pdf/Manifesto.pdf>
- ¹⁰⁵ Rostow. W.W.1991 [1960] 3rd ed. *The Stages of Economic Growth: A Non-Communist Manifesto*.Cambridge, UK: Cambridge University Press.
- ¹⁰⁶ Segenberger, W. and Pyke, F. 1991. Industrial districts and local economic regeneration: Research and policy issues. *Labour and Society*, 16(1): 1-24

Engaging with the fourth industrial revolution: Implications for south economies

Sujit Bhattacharya
CSIR - NIScPR, India

Abstract

The paper situates within the broader context of structural and dynamic changes that are happening in economy, society, the way we work and live as The Fourth Industrial Revolution (4IR) technologies are getting embedded in new processes and products. These new technologies draw from convergence between the physical, digital and biological systems, and challenges our understanding of technologies and their impact. Along with radical changes in product core components surrounded by patents and standards, these technologies calls for new types of negotiation between human and machines. These new negotiations are already visible in factory floors, products and concepts. A new manufacturing ecosystem called 'IIoT: Industrial Internet of Things' is taking shape with the adoption of 4IR technologies.

Creating 4IR enabled manufacturing systems calls for adoption of new technologies that are complex, capital intensive with high entry barriers, requiring new types of factories and demanding sophisticated digital system and new job skills sets among others. Being part of the value chain of the new technologies gives huge opportunities for countries including emerging and low income countries. They can transit from current economic levels to middle/higher income levels and address developmental challenges. Even progressive deployment of new 4.0 technologies have the potentiality of enabling significant transformation in their industry, a strong positive change in host of sectors from transport to healthcare, and can enable more rapid financial inclusion.

OECD economies are developing new strategies and policies for transition to this new manufacturing ecosystem. The challenges are much higher for South economies that are at different stages of economic development, struggling with various developmental challenges compounded by Covid pandemic and climate change. The wherewithal for adopting 4IR technologies are largely missing in these economies. The ever growing pervasive and ubiquitous nature of the new technologies will give countries no choice but to incorporate 4IR with consequences dependent on how well they negotiate with it. The concern is that 4IR will further exacerbate existing social and economic inequalities, as some countries and individuals benefit more than others from the new technologies.

The paper draws from a close examination of 4IR to infer some of its salient aspects and their implications. It draws in particular attention to the changing manufacturing and production system, the human-machine interface aspects, changing dynamics of patenting and standard development. The implications of 4IR especially for developing economies are examined. The paper draws from this to argue for a new conceptual framework that is not driven by competitiveness framework alone. It calls for equity, inclusion, risk, regulation, governance, society need to be embedded within the conceptual framework. The paper concludes by discussing the policy challenges and plausible pathways.

Keywords: Fourth Industrial Revolution, Industrial Internet of Things, Human-Machine Interface, Anticipatory Standards, Equity and Inclusion

Introduction

Emerging technology breakthroughs in artificial intelligence, big data, robotics, 3D printing etc. are changing the manufacturing ecosystem in advanced countries. 3D printing leading to additive manufacturing, global digital networks opening up access to manufacturing processes globally, Internet of Things (IoT) providing dynamic information exchanges between networked devices that are allowing to remotely connect, track and manage products, systems and grids; new possibility of changing hardware systems through software modifications, development of biological computational interfaces, etc. are defining new types of convergence between digital, physical and biological technologies (see for example Aegerter and Benziger, 2015; Baweja et al., 2016; Liu, 2017; Deloitte, 2019, Unido-industry-4.0, 2023)

The new technologies challenges our understanding of technologies and their impact. Along with radical changes that are competence destroying, the new technologies calls for new types of negotiation between human-machine interfaces. These new negotiations are already visible in many novel products and concepts such as autonomous cars, implanted sensor enabled medical devices, AI controlled agriculture, smart grids, etc. The radical shifts that these new technologies are bringing in the industry, economy and society is now being called as the 'Fourth Industrial Revolution (4IR)' (Schwab, 2015). The convergence of various disparate technologies makes them highly disruptive. These transformative technologies are already exhibiting far reaching consequences affecting industry, technology development and absorption, creating new capital intensive sophisticated infrastructures, disrupting existing industry supply and value chains and require much higher levels of personal/society engagements with them.

The promising possibilities that 4IR brings are huge and is already been seen in the new class of products that are coming in the market. The new 4IR driven manufacturing ecosystem is leading to high efficiency in the whole production process. It is helping in creating products that are highly efficient and with varied functionalities that can be controlled through IoT. The the whole product landscape is changing with the IoT enabling and can be seen across range of customer products from watches to cars. 4IR technologies are being adopted for addressing sustainable development goals and achieving circular economy through optimization of the products and materials cycles, etc.

Massive Modularity in products could be seen in the ICT industry that provided opportunities for many countries to be part of the global production and supply chain (see for example World Bank, 2023). An example of Massive modularity is the smartphone handset which can be broken into modules sharing common traits: (i) they are interconnected with each other according to standard interfaces; (ii) innovation can take place independently in each module, as long as

the interface standard is adhered to and continually updated; and (iii) they can be broken down into smaller, more specialized modules. This modularity of the different products helped in developing global production manufacturing and value chains. South

Korea, Singapore, Hong Kong, Taiwan, the so called Asian Tiger economies and China were able to take advantage of this product modularity and could become major partner in the global production and value chain. It also helped them in learning in various ways, along with technology, the institutions, the policies and manpower that needed to be created to become technology developers. These countries invested heavily in research, created a new culture of entrepreneurship in their universities, focused on creating an STI ecosystem that facilitated public-private partnerships, developed strong international partnerships among others. These multiple factors were instrumental in these countries becoming technology creators, challenging the dominance of advanced OECD economies in various product classes in different sectors.

4IR provides new ways of innovation driven production process. It threatens to disrupt these production chains and provide further barriers to emerging and developing countries like India, Brazil which are trying to replicate the success of these economies. The challenges will be more difficult for lower income developing economies. The paper situates within the broader context of structural and dynamic changes that can disrupt the innovation ecosystem with radical technologies that underscores the Fourth Industrial Revolution.

The paper is an exploratory study drawing from a large body of scholarly work across different strands of research. The main intention of the paper is to draw an analytical assessment of the structure and dynamics of technologies that underscore 4IR. It attempts to fill the research gap by drawing a more coherent view of 4IR and its implications. It proposes a revisit of the dominant conceptual framework that suggests plausible pathways for engaging with the 4IR.

Methodology

The analytical framework of the study is drawn from a wide cross disciplinary strands of research; primarily innovation system, STS framework and competitiveness. This analytical framework has helped in selecting influential research that have examined the technology shifts and their implications from different perspective. The research method used is systematic literature review and discussion with diverse stakeholders. The literature review was not restricted only to research papers. Close reading was done from various reports from different international bodies, white papers, country policy documents, company reports etc. The study is supplemented with wide discussions across different platforms to bring out the critical aspects. The study has also used the Derwent Innovation Index, and websites of IEEE standards and 3GPP and 5G Infrastructure public-private partnership that are involved in developing

standard for 5G for constructing data and assessment of the development that is happening.

The research papers were identified using the SALSA (Search, Appraisal, Synthesis and Analysis) approach (Booth et al. 2016). The broad keywords 'Fourth Industrial Revolution' or 4IR was applied which was systematically refined to develop the research paper corpus. The citation, download and the sources in which these research papers have been published helped to perform the quality assessment. Integration was done by assigning the selected papers in broad categories. Analysis was undertaken keeping the context in perspective. These keywords were also applied for searching relevant documents from other sources.

The paper is structured as follows. Section 2 underscores the salient characteristics of the Fourth Industrial Revolution. It draws the technology space of the 4IR and its implications. Section 3 discusses on patents and standards. Section 4 underscores the implications of the study. Section 5 covers discussion and conclusion including possible pathways. It also draws attention to conceptual framework.

The fourth industrial revolution: Some salient aspects

Understanding 4IR

There are important learning lessons from industrial history that have implications for understanding the 4IR. A rich body of scholarly works underscore each of the industrial revolutions in terms of "critical technologies", the role of leading sectors in economic growth that were instrumental in driving and distinguishing each industrial revolution (see for example, Rostow 1975). A rich understanding of the industrial revolutions and innovation through time can be discerned using the lens of Innovation System Concept, storytelling based on historical sketches, patent evidence (see for ex Macleod, C, 1988; Bruland and Mowery 2005; Lundvall and Borrás 2005; Dodgson and Gann 2010). Along with leading technologies they draw attention to the role of large number of smaller technologies, the coexistence of a range of innovation modes, institutional processes, and organizational forms.

Some early evidence of role of science in technology innovation (Glass making, for ex), organizational innovation, and germination of concepts such as strategic linkages, open innovation can be discerned from the first industrial revolution itself (1780s onwards). Josiah Wedgwood (1730-95) innovations in ceramic industry demonstrates the role of monopoly, and political linkages (Dodgson and Gann, 2010). The importance of public policies for example creating conditions for attracting skilled manpower i.e. technology diffusion through immigration, social innovation that created conditions for technology diffusion can be traced even before the advent of the first industrial revolution, 1IR which

became more established in the middle of the nineteenth century (see for example, Lundvall and Borras 2005). Key role of individual inventors was a distinguishing feature in the first and in the early half of the second industrial revolution, 2IR (1870s onwards). Innovations covered different sectors of the economy and also brought new aspects that have key relevance in the present context. *This is true when we look at later industrial revolutions!* On the other hand, patents as a tool to control monopoly and kill competition has been inherent in key technologies that shaped industrial revolutions. James Watt patent on “method of diminishing the consumption of fuel” in steam engine showed the clever as well as manipulative way of gaining monopoly and stopping technology diffusion. It was a broad claim that prevented other follow up innovations. Edison cleverly used his patents to build an empire and stop technologies that was superior to his. The ICT revolution that underscored the third industrial revolution 3IR further created monopoly structure and cartelization with a few firms dominating the global economy. The 4IR is now situating itself in the structures and dynamics that have been created during the 3IR.

Another useful perspective of the 4IR comes from the study of UBS white paper for the World Economic Forum (Baweja et. al., 2016). This study underscores connectivity and automation as the two key factors that defined each industrial revolution; each of the later industrial revolution showing increasing load on these two factors. They characterize 4IR as one of ‘Extreme Connectivity’ and ‘Extreme Automation’. Key inflexion points that were defining moments of each of the industrial revolution were highlighted which corresponded to the two factors. The key inflexion point for *automation* in the 1IR was identified as the ‘Puddling process’ that turned pig iron into wrought iron (Henry Cort Invention, 1784). Steam power and locomotive as key inflexion point for *connectivity*. They also underscore some important innovations particularly in the agriculture innovation such as crop rotation and selective breeding. Higher levels of *automation* via the development of mass production was identified as the key inflexion point for the 2IR. More efficient *connectivity* in production via distribution of labor, the advent of “Supply Chains”, Boosting Agriculture production by industrial fertilizer, refrigeration contributing to food storage Standardization, more Legal and trade protections. The 3IR (1960s onwards) saw the advent of science based technologies which were more complex and pervasive across sectors. Internet, integrated circuits, microprocessors, global manufacturing and supply chains were some of the key outcomes. They have significant impact on automation and connectivity. Fourth industrial revolution is moving towards maturation of key technologies (artificial intelligence, human-human, human-machine, Cyber-Physical Systems, Big Data, Cloud Computing, IoT) and creating a new class of products. Advanced Robotics which are different from conventional Robots whose actions are completely predetermined, meaning that they are deterministic and not autonomous. 4IR technologies enable Robots to become flexible, intelligent and autonomous

adjusting itself to the environment that is dynamic and subject to change. Autonomous vehicles and drones which allows them to operate and navigate with little or no human controls, Cloud technology and big data, Blockchain a distributed electronic ledger that uses software algorithms to record and confirm transactions with reliability and anonymity, Virtual & augmented reality Computer-generated simulation of a three-dimensional image overlaid to the physical world (AR) or computer environment (VR) define some of the new 4IR technologies.

The two perspectives complement each other i.e. the role of technological and organisational innovation, monopoly and strategy with the factors of automation and connectivity complement each other in understanding the industrial revolutions. *The 4IR however is radically different from the earlier industrial revolutions.* It brings the possibility of far reaching changes effecting how we work and live. The new interfaces between the digital and biological systems are showing unimaginable human-machine interfaces (Agerter and Benzigarm 2015). It is making 'Human' also as a 'Thing' connected to the digital ecosystem. Technologies are being developed that 'read' the brain to record brain activity and decode its meaning, *and those that 'write' to the brain to manipulate activity in specific regions and affect their function.* KAIST and University of Washington for example have built a novel device capable of *remotely controlling the brain circuitry of mice, via a smartphone.* It is the first wireless neural device that can continuously deliver drugs and colored light beams to control brain circuits. It uses a *plug and play interface* with LEGO like replaceable drug cartridges, a probe with the thickness of a human hair, and powerful low-energy bluetooth to deliver the drugs and light, which turn neurons on or off without hurting the rodents. Scientists are exploring this to the study of diseases such as Parkinson's, Alzheimer's addiction, depression and pain (Qazi et al., 2022). New technology CRISPR/Cas9 allows Genome editing. It is a way of making specific changes to the DNA of a cell or organism. An enzyme cuts the DNA at a specific sequence, and when this is repaired by the cell a change or 'edit' is made to the sequence. Jennifer A Doudna and Emmanuelle Charpentier got Noble Prize in Chemistry in 2020 for developing CRISPER technology. As an example Gene editing is increasingly applied for COVID-19 diagnostics and therapeutics. CRISPER has improved our ability to edit genes and modulate gene expression. Already there are over 2000 patent applications in gene editing as search in Derwent Innovation Index reveals! It shows how this technology is getting increasingly restrictive. Another new platform messenger RNA, mRNA has emerged that allows to control message to be sent to DNA for making proteins. This platform was used for making mRNA based vaccines during Covid pandemic. The designed mRNA was used as a blueprint to build foreign protein replicating that of the covid, stimulating an adaptive immune response. It taught the body to identify and destroy the corresponding pathogen i.e. the covid. This completely changes the vaccine production process making it more safe and is being used to create new

possibilities for cancer and other treatments.

The 4IR is leading to the development of a new manufacturing ecosystem called the 'Industrial Internet of Things' (Serpanos and Wolf 2018). The IIoT goes beyond the normal consumer devices and internetworking of physical devices usually associated with the IoT. It enables integration of the information technology (IT) with operational process in the production value chain including industrial control systems. It is providing system integration through automation and optimization in monitoring and controlling physical infrastructure in industrial operations, in supply chains and logistics. It is leading towards a new way of 'decision making' with real-time data from sensors and other information sources helping industrial devices and infrastructures to take specific actions without human interface. IIOT is integral to the '4IR' with machines enabled to take on and automate tasks that previous industrial revolutions could not handle.

Data-driven manufacturing concept called "Digital Twins" is an example of this new manufacturing paradigm. It enables manufacturers and enterprises to simulate products, building faster, cost-effective, and quality benchmarked products. It can for example create virtual replica of an entire car, software, mechanics, electrics, and physical behavior of the vehicle. All real-time performance, sensor, and inspection data, as well as service history, configuration changes, part replacement, and warranty data, are stored in the digital twin. A product's digital twin aids in determining its performance and quality by virtually experimenting with different compounds and raw materials to enhance the design and optimize the product's performance. Companies can create a digital twin of a factory's infrastructure and train workers remotely without physically installing the equipment. To detect fault recurrences and causes, digital twins use real-time data extracted from IoT devices and sensors during the manufacturing process. Sustainability and circularity in the product and material life cycle and supply chain is being designed through digital replica.

4IR is being increasingly seen as one of the key drivers for addressing sustainable development goals (SDG). One of the major study undertaken to examine the relevance of 4IR technologies for addressing SDG goals was by the World Economic Forum in collaboration with PoW research (World Economic Forum, 2020). It mapped the 4IR technologies impact on the 17 SDG Global Goals and their 169 targets. They found that 70% of the targets could be enabled by 4IR technology applications already in deployment. The applications were found to be playing an important role in 10 of the Global Goals in particular, with Health (Goal 3), Clean Energy (Goal 7) and Industry, Innovation and Infrastructure (Goal 9). This includes Precision agriculture to optimize inputs and returns and early detection of diseases and issues, Drones for remote delivery of medicines, medical equipment and samples, Smart open educational

resources to increase affordability and accessibility, AI-enabled digital footprint for mobile money access targeted at women consumers and entrepreneurs.

Other good example of the influence of 4IR technologies contributing to SDG goals can be discerned. Additive manufacturing (3D printing) process does not use water for cooling or lubrication. Less water is thus used in production, and consequently, less wastewater is produced which addresses the SDG 6 goal of better water management and availability of clean water and sanitation (Stock et al. 2018). Internet of Things and Smart Grids can significantly contribute to energy savings (SDG 7) in the production process. Real-time monitoring of the production and energy use enables matching energy consumption with production and information-based decision making about the cost-efficient use of renewable energy sources. In combination with Big Data analytics, the previous two technologies enable implementation of life cycle assessment, which in turn can significantly increase the optimisation in the production process as well as reducing carbon footprint (Bonilla et al. 2018).

Monopoly through intensive patenting and standards

One of our key characteristic of 4IR technologies is the high level of patenting and standards surrounding them. The new industrial revolution now brings a new war among countries and firms that holds the patents and standards in 4IR. Standards with patents embedded inside them bring a new dimension to when we think of standards. A new form of monopoly is being created through this. Anticipatory standards is the new normal in standard development. Various studies coming from industry groups/associations are showing the industry assessment based on technology trends of 4IR technologies see for example Tech Trends 2023 (Deloitte, 2023), Insights (2023). Deeper insight is available from recent exhaustive study undertaken by the European Patent Office on patenting activity in 4IR¹. The study is based on international patent families (IPFs; these are inventions filed in two or more territories) across 350 distinct technology fields that cover the technologies of 4IR. The study identified 264565 IPFs in 4IR with filing dates from 2000 to 2018; average annual growth rate in patenting close to 20% from 2010 to 2018, compared with 12.8% during the period 2000 to 2009. The annual increase in IPFs for 4IR technologies was found to be nearly five times greater than the growth of IPFs in all fields since 2010. This showed how intensively patents are being filed in 4IR. USA is dominating with Japan, and European Union, fast losing ground to South Korea and China. The top 10 applicants in the period 2010-2018 together account for 23.8% of all IPFs for 4IR technologies; four firms from USA, two each from Korea and European Union, and one each from Japan and China. A few countries and firms are thus dominating the 4IR technology space, shaping the 4IR technologies and making entry barriers high for new entrants. The large number of technology classes covering 4IR as can be observed from this study shows wide sectoral dispersion covering technology classes from disparate technology groups. It thus

underscores the pervasive impact of 4IR technologies across the sectors.

Patenting activity in key 4IR technology driven products provides deeper insights. Patenting in drones showed how the monopoly is being created. 73,350 patents have been filed covering three key domains that were identified namely flight control and safety, sensor signal, and battery storage. China has 20403 patents, USA 15141 patents, South Korea 7760 patents and Japan 6421 patents filed. The packing order does not change much with Germany in some key technologies for example aftreatment technologies among the top three. Primarily it is a very restrictive space with a few advanced western economies, South Korea and China among the dominant players.

Thus to participate in the technology development surrounding industry 4.0 one has to negotiate along with monopoly challenge of intensive patenting, the entry barriers created by standards that underpin these technologies. One can see emergence of cartels created by big firms who are in the thick of things in developing 'anticipatory' standards covering many technologies under the umbrella of technologies of 4.0. Strong country involvement in pushing their firms to emerge as leaders is another key aspect visible. *The anticipatory approach is unlike the traditional way in which historically standards have been developed in different sectors.* Traditionally standards main purpose was to achieve interoperability and create order in the markets, make production/products more structured and ensure safety. The 'anticipatory' approach implies development of the standards before the development of the products. Standards are now increasingly covered by patents known as a standard essential patent, or a 'SEP'. It is a patent that protects technology which is essential to implementing a standard. Thus a new type of monopoly is created as complying with a standard needs access to SEPs within them. The regulatory authority apply fair, reasonable and non-discriminatory (FRAND) licensing principle so that innovators have more reasonable terms for obtaining license to essential patents from SEP holders. However, contentious battles are visible in the pre and later stages between SEPs and innovators. Patenting on tools and techniques are especially problematic as it leads to anti-commons (Heeler and Eisenberg, 1998) and also blocks future research and innovation. *SEPs brings in a new form of anti- commons.*

The 'anticipatory approach' to standard development can be traced to the development of IEEE standard for wired and wireless communications which resulted in creating a suite of standards under the board umbrella of the 802 standards (Rashba and Gamota, 2003). The anticipatory approach is a proactive strategy to in which a product and market requirements are anticipated and competing parties are influenced to come together to develop standards. Thus in this approach standards are developed before the products. The anticipatory approach was applied in the development of Wi-Fi or IEEE 80211b standard by IEEE. It created a working group of interested parties and with them developed

Wi-Fi standards in 1999, much ahead of the technology cycle. The disruptive effect of this standard is well known as it revolutionized the telecommunication industry and helped develop a large class of products. Thus unlike standards that are created typically late in the development cycle of a technology, the IEEE took a proactive approach, anticipated the product and the market requirements and convinced competing parties to come together to develop standards under the suite IEEE802 incorporating provisions for vendors to have flexibility within the standard development suit their competency strength.

The above process of anticipatory standard development was largely a positive and enabling approach. However, this is now being stretched when one observes in the development of new standards surrounding technologies of the 4IR. The case in point is the development of 5G standards. 5G networks will be the key enabler for the digital ecosystem, the internet of things, autonomous robots, the smart utility grids, and autonomous vehicles and drones. Ugly standard wars among countries and between firms are being played. Cartels are getting developed with tacit support of countries that want to capture the technology space. Two major spilt can be seen between US and OECD countries on one side and China on the other. Drawing from Derwent innovation index, we observe a large number of patents and patents in standards (essential patents) are held by US major ICT firms (QUALCOM, Intel dominating) and Chinese firms (ZTE and Huawei). The approach of standard development is anticipatory. But here the anticipatory development shows various shades of technology cartalisation unlike the positive approach one had observed in the developed of IEEE 802 standards.

Different big industry players in cellular communications are involved; QUALCOMM actively involved in developing products for 5G primarily computer chips for wireless devices and base wireless technologies, like waveforms. Its key competitor is another US firm Intel. QUALCOMM which solely developed the CDMA system for 3G created strategic partnership with South Korean government to commercialize this standard. ETRI joined as a partner to develop the central monitoring system of this network. Korea did not own core technologies and had taken a risk to develop its cellular network on CDMA whereas GSM was the more visible standard being adopted globally (Lee and Lim, 2001). As case study of Choung et al. (2011) highlights the role the adoption and partial ownership of this global standard helped South Korea getting recognition as a key player in cellular mobile network for 3G. The learning by doing, the tacit knowledge transfer, and developing value added products surrounding this standard made Korean firms especially SAMSUNG as a key player in communication system and networks. Both Samsung and Qualcomm has strong presence in 4G and this success also helped them to emerge as major players in 5G. Along with Intel as highlighted above, the other key players are Finland's Nokia, Sweden's Ericsson, South Korea's Samsung and two ICT firms from China ZTE and Huawei. The two Chinese firm's dominance

in antenna and base station architecture, particularly with Huawei leadership in successful demonstration and implementation of 5G, its emergence as leader in radio access network, large portfolio of patents and strong Chinese government push has led US and European firms to consolidate, create partnerships and raise security issues and unfair practices of Chinese firms to stop them from emerging as leader in 5G. The country support is strongly visible in this ugly war with Chinese firms.

As can be seen, network effect becomes very important in the development of a standard. A firm that controls a technology that becomes established as a standard can have an extremely profitable market position, what Ferguson and Morris (1993) call an "architectural franchise. This can be seen in the development of the different generations of cellular networks. Because the prize is so tempting, sponsors may compete fiercely to have their technologies become the standard, and this competition will generally dissipate part—perhaps a large part—of the potential gains. Competition to become the standard may also delay market growth by encouraging buyers to wait to see what the standard will be, that is, what other buyers will do. This happened during the war between the two standards GSM and CDMA. Standard development is a very strategic and depends on various types of configurations as can be seen in the issues highlighted. There is potential standard war when the firms are symmetric in their market and technology positions, when the standards battle does not greatly delay the adoption of the technology by consumers, and when intra-technology competition is especially likely to dissipate potential industry profits. Each of the competing firm try to build an early lead, establishing a large installed base quickly, attracting the suppliers and creating a strong base. The firm which is able to do so can successfully imbed its standard in the market. This is happening in 5G to some extent. However, the other new dimensions are emerging which has not been explicitly articulated in research on standards. On the other hand, the market acceptance of new products or vendors selling new products is dependent on standard then firms will increasingly go for cooperative agreement in the development of standard. It may include adopt a hybrid standard that combines the technologies of various competing firms, involving independent third party. This happened in the development of IEEE 802 wireless standard. The standard development provides some insight of how a few firms will dominate the digital ecosystem in 4IR.

4IR and its implications for SOUTH economies

The 4IR offers opportunities to developing countries to leapfrog traditional development paths and accelerate economic growth by adopting and integrating new 4IR technologies, improve productivity, enhance competitiveness, and create new industries and jobs. It can help to address some of their major developmental challenges such as poverty, inequality, and

environmental degradation. The analytical assessment draws attention to the complexity and challenges in engaging with the 4IR technologies. It calls for building a completely different manufacturing ecosystem. It calls for developing various new types of regulations, new ways of doing things and has wide ranging disruptive impacts.

One of the key prerequisite for implanting 4IR technologies is providing digital access. A huge digital divide is there between the North and the South economies and also within a South economy. The 4IR technologies need a more sophisticated broadband internet system which leads to a wider digital divide. The contentious issues are cyber security, cross-border data flows, and various treaty and frameworks that are developed by OECD economies and are being pushed to be adopted by emerging and developed economies. India's digitalisation efforts which to a large extent epitomises that of the emerging economies can provide some useful insights. India has given a significant push for digitalisation in the last 15 years or so. A major mission mode programme 'The Digital India programme' was enacted in 2015. This was followed up by National Digital Communication Policy 2018. They are umbrella programmes and involve different ministries and to the local level of the legislative mechanism in India, the Gram Panchayat i.e. total of 2,50,000 dispersed through 6600 blocks and 641 districts. The success of digital India would be to provide digital connectivity to this level. One of the successful initiative has been to implement Aadhar; a unique identification number, a programme cited as the world's largest digital ID programme. It is estimated that over 1.34 billion people i.e. 99 percent of India's population have been registered under Aadhar. Various reports have highlighted India's successful digitisation efforts, see for example McKinsey Global Institute report (McKinsey, 2019). This report highlighted that India is digitalising faster than all other countries but Indonesia. The recent trends show India has surpassed Indonesia. One of the successful initiative is the development of a standard for 5G through joint collaboration between university-industry-government involving IIT-Madras, IIT-Hyderabad, TSDSI (Telecommunications Standards Development Society) and the Centre of Excellence in Wireless Technology (CEWiT). This standard helps to enhance signal transmission that can lead to better coverage of 5G in rural areas. This can help to make broadband services to go to the last mile connectivity i.e. Gram Panchayat. It has been a very difficult process to get this finally accepted by 3GPP, the global governance body of 5G. This would allow product development to match with global standards.

In spite of these positive developments, India lacks an overall framework covering the entire digital ecosystem i.e. from Social Media platforms, OTT platforms, online apps, to blockchain based apps, to the metaverse and cybercrimes. It does not have a sound legislation for data protection that includes personal digital data. Framework like OSAKA Track calls for unhindered cross-border data flows, Budapest treaty on cybercrime calls for

difficult mechanisms to implement them (G-20 Osaka Summit, 2019; Seger, 2016). These international provisions open the digital ecosystem that can be difficult for emerging and developing economies to manage and prevent unauthorised access by global firms of important protected data that includes personal data. There is intense pressure to accept these international provisions from OECD countries. This is reflected in the Personal Data protection Bill enacted by the Indian government in 2019 that had to be withdrawn on August 2022 (Digital personal Data Protection Bill, 2022). The Indian government is now in the process of introducing a revised bill which if passed by the legislative mechanism will be enacted as a law which all Indian citizens and entities working in India have to adhere. On the other hand the ambitious target for a stable digital access and translating to higher broadband needs huge capital outlay, significant technology upgradation among others. Challenges like this confronts the developing countries which are much severe for low income and marginalised economies.

Industry 4.0 means transformation from labor-intensive toward autonomous production. One of the major concern is the impact on jobs with the extreme automation and connectivity (see for example Baweja et al., 2016). One of the consequence of this would be companies will have fewer workforce which are highly skilled. Near term polarization in labor force thus would lead to greater income inequality. Automation has already made a significant impact globally on low skilled jobs. Extreme automation can lead to more severe impact. It is difficult to get an estimate of the job that would be lost due to these new transformations. Broadly the research shows that not only manual jobs will be lost but also many cognitive tasks like routine accounting, ledger matching will be lost. On the other hand special tailor made manual tasks will remain. The demand for highly skilled cognitive tasks will increase (Brynjolfsson and McAfee, 2014). Some of the low-skilled parts of the labor force (for instance assembly line work) that have already been heavily affected by basic automation could be further impacted. Thus broadly it is envisaged that low-skill employment will likely to contract, and an increasing range of middle-skill jobs will become venerable as extreme automation is rolled out. *With the advent of new generative AI algorithms, many non-routine cognitive tasks will increasingly get affected. Thus, the earlier assumption of highly skilled cognitive tasks will not be affected may be increasingly challenged as AI and ML technologies further mature and new products develop from them.*

The shift towards higher skilled labour with high levels of cognitive skills is already visible as big firms in developed economies are increasingly adopting 4IR. The middle sized firms in these economies particularly the German Mittelstand have been one of the fastest adopters of the 4IR. The industry 4.0 framework was first brought forward from Germany plausibly draws from the early engagements their industry was having with 4IR technologies. Strong shift towards new forms of organization that create intangible assets in form of

intellectual property and these assets then become the source of attracting venture capital or lucrative exit strategy is built (FMEE, Germany, 2016). Startups are emerging as the key player in creating these types of assets as their competitiveness comes from bringing a novel idea to the market. 4IR technologies are giving them opportunity to create novel products and processes, business models. This is primarily seen in developed economies and to some extent in emerging economies. The difference however in emerging economies is that only a few startups are in 4IR based product development and are mainly using 4IR for strengthening supply chain, e-commerce, and in sector level interventions. Sector level interventions includes drones for crop management, surveillance etc, IoT based sensors for electronic products, AI and ML for selecting target population for product-customer mapping, delivery etc. China, South Korea and other Asian tiger economies are outliers as they had early engagement with 3IR technologies and their firms have been able to create global presence in product developed through 3IR technologies. This has enabled their industry and startups to be involved similar to OECD economies in 4IR based product development.

Discussion

The 4IR technologies have far reaching impact which challenges even the existing manufacturing processes in advanced OECD countries. The new production process allows a wide variety of enterprises to be connected, integration of manufacturing and smart technologies that would allow manufactured products to interact with any necessary object in the network. The actors in the network are smart products that manage to control their own production in independent and decentralized way using sensors and communication systems. IoT, artificial intelligence embedded in machines are redrawing a new form of communication, moving from human-machine communication to machine-machine communication! Moving into New Technologies, 4IR technologies demands New Skills, New Ways of Thinking and Innovative Approaches.

Many advanced North OECD economies are creating these new manufacturing ecosystems. South Korea, China and Asian Tiger economies are also significantly changing its manufacturing ecosystem by embedding IIOT in the production process. The other emerging economies such as India, Brazil, and Thailand are placing IoT enabling in some aspects of their production process. The development of the new manufacturing ecosystem in advanced OECD countries, in South Korea, China has been possible due to their strengthening their research and innovation ecosystem through a long sustained process. This includes investing in R&D with GERD ranging from 2 to 4 percent, creating various institutional mechanisms for developing synergetic partnerships with university and industry, regional clusters, international partnerships. Policies have been created with long future outlook with a competitive framework

model that connects to industry 4.0. NBIC- convergence Programme in the USA shows the systematic convergence of nano, bio, information and cognitive technologies. It has also enacted an Advanced Manufacturing Partnership programme, European Union expression “convergent technologies” for the European Knowledge Society, “Public-Private Partnership for Factories of the Future” are now finding space in EU policy documents and implementation models. The concept of Industry 4.0 germinated in Germany. Germany exhibits first mover advantage with a strong policy implementation framework centered on their SMEs (FMEE, Germany, 2016). South Korea “Manufacturing Innovation 3.0”, Singapore “Super Smart Society” initiative, China “Made in China 2025” initiative have embedded 4IR as key strategy as visible from these policy documents (see for example Roco and Bainbridge, 2003; Liao et al. 2018).

Emerging economies are constrained by the developmental challenges and weak institutions. Competitiveness of emerging economies are based on low cost labour. India a typical example of challenges faced by emerging economies in the new transition. It is in an early stages of developing an enabling ecosystem for innovation and entrepreneurship. Deindustrialization has led to huge impact on jobs. Following leapfrogging strategies of China and Asian Tiger economy may provide only limited success in the 4IR. New scenario of automation, AI enabled systems (already visible) showing early signs of further job losses. Ambitious programmes for skill development, reinvigorating the MSME sector, university entrepreneurship shows the difficulty in translation of policy to action Sustainability challenges impede India and developing countries to be a more equal partner in the new transition. Emerging and developing economies are developing digital ecosystem but transition to key enabling technology of 5G that would make IoT a reality is confronted with various challenges. The big digital divide, the system to react to cyber-attacks, regulatory challenges are huge as can be seen in these new technologies

Concerns on new technologies range from digital privacy, regulation to ethics. Some proactive actions are being taken in advanced OECD economies by funding research in these areas, bringing specific regulations, white papers etc. A new area has emerged within ethics specifically bioethics called Neuroethics – to ensure that technologies that directly affect the brain are developed in an ethical manner. These new areas will evolve within other technologies that define 4IR as the contentious issues evolve in human-machine interfaces and security aspects among others. Ethics need to be integrated into the initial design and development stages of such technologies, to maximize their benefit and to identify and minimize their potential harm – whether to individuals or to wider society. AI tools also introduces ethical issues of which regulators have little experience. Machine-learning software learns to analyse data by generating algorithms that cannot be predicted and that are difficult, or impossible, to comprehend. New research motivated and supported by big technology giants are on areas such as bidirectional coupling in which

computers respond to people's brain activity and insert information into their neural circuitry. *These new technologies are being transferred to emerging, developing and low income economies without the safeguards that are applied in countries where technology is being created. The South economies are becoming testing labs for these new technologies.*

Disciplinary shift can clearly be observed. New Biology for example is now scientific integration of chemical and biological sciences, computer science, biology, engineering, mathematics. It is now a systems approach: From DNA Sequence to Omics (the study of complex interactions between genes, proteins, and metabolites within the resulting phenotype) New areas to study the various layers: *Genomics joined by Epigenomics, Transcriptomics, Proteomics, and Metabolomics.* These disciplinary shifts are visible in different areas which has led to new technologies as highlighted. The universities in emerging and developing economies do not have the wherewithal to bring these types of disciplines in their universities. The progressive lack of research capacity can make it difficult to absorb the new technologies and do follow up innovation. This has been one of the key challenges in earlier technology transfers which will be exacerbated as 4IR technologies need a deeper engagements with new research.

Conclusion: The plausible pathways

As the transformative shift towards 4IR is happening, there are huge challenges that are before us. Even now when we look at the debate on climate change, pollution, genetic engineering, exposure and toxicity among others we increasingly find that new models are required to address these contemporary challenges. Accelerating Technology Change and increasing competitive pressure is plausibly the reason behind introduction of new products in the market without the extensive testing required. New 4IR Technologies have far reaching Socio-Economic Impact as highlighted in this paper. Patenting in genetic engineering and nanotechnology had led to contentious ELSI debates. The ELSI challenges will be extremely difficult to handle in 4IR technologies as human-machine interfaces are redefined. The industry 4.0 rests on technologies emerging from convergence between emerging technologies i.e. NESTs (New and Emerging Science and Technologies) reaching maturity stage and increasing convergence between them. Thus the puzzles and challenges for research and exploiting emerging technologies becomes integral to these new technologies. Innovation failures of many promising products emerging from exploiting emerging technologies are happening due to reasons that cannot be simply linked to earlier innovation failures.

The dominant conceptual narrative of 4IR is built on competitive framework. One of the framework that is proposed is by Liu (2017). This framework is constructed from theories of competitiveness and development; linking

competitiveness as a systemic interaction between macroeconomic and microeconomic factors. In developing this conceptual framework, different theoretical strands have been critically examined that provides clarity to the concepts and some useful overlap between the different strands. This provides robustness to the framework. One useful understanding brought forth is the linkage between two dominant models namely that between resource-based view of firm's competitiveness with international competitiveness as defined by Porter's diamond model. Resource based view underscores firms competitiveness as a combination of resources and a firm's capabilities to use these resources. Liu posits that country-specific determinants of Porter's model of competitiveness namely factor endowment, demand conditions, and clusters can be seen as resources, and firm's strategy, structure, rivalry, and interaction in its cluster as capabilities to use the resources.

Inspite of the usefulness of these conceptual frameworks, they are primarily centered on how firms can compete in the new 4IR. There is a need to build new models to bring the society as the dominant actor. The models like this also disconnects itself from the larger challenges of 4IR, the regulation and ELSI challenges. They are also strongly centered on NORTH economies. *Conceptual limitations of the dominant narrative can be evidenced from the failures of this type of framework during Covid 19 that highlighted the vulnerability even of the rich OECD countries.* Their ecosystem was not resilient enough to address the health and socio-economic challenges. Collective global response, local networks, open sharing, exploring new daring pathways that in some ways the pandemic compelled to take, redirecting efforts for meeting societal challenges, novel public-private partnerships, etc. provided solutions and helped to mitigate the severe impact of this pandemic. Various innovative solutions emerged also from low and middle income countries inspite of the structural weaknesses of their innovation system, institutional voids and resource constraints (Bhattacharya, 2023). These interventions were driven by 'social welfare', challenging the capitalistic economic model(s) as they were not driven by profit motive, the central tenet of this/these model(s). The successful interventions that has made a paradigm shift in research- innovation ecosystem, the production and consumption system among others showed they cannot sustain unless policy framing is reset to bring development, inclusiveness and equity in the economic model(s).

The role of interactive learning, learning by doing, and other forms of learnings such as through user-producer interactions, context specific interactions/learning between the diverse values and knowledge systems has become more important than ever before (Lundvall 2017, von Hippel 1988). The learning however has been restricted with North acting as the main actor and South as a recipient. This can be seen in the context of North-South Technology Transfer. These have been contentious and have not adhered to the TRIPs agreement of easier terms and support for technology transfers. The successful

technology transfers on deeper investigation shows the adverse socio-economic impact that have been created. A case in point is that of Chlorofluorocarbons (CFCs) which was a ubiquitous cooling agent and became a global concern as increasing scientific evidence showed them to be the major contributor to ozone-layer depletion. The contentious replacement of this cooling agent with huge technology transfer cost for developing countries (May and Sell, 2008). Many of the refrigerant firms in developing countries could not cope with this costly technology replacement and closed down. South economies that comprises largely developing, low income and marginalised economies lack absorptive capacity to assimilate and modify many of the technological interventions and undertake follow up innovations. This is primarily 'capability failure' that happens due to inadequate education and lack of training to develop skill sets. Countries like South Korea, Taiwan, and China were successful as they had invested in education and research which made them take advantage of the technology transfer. Not investing enough in R&D may be one of the plausible reasons that emerging economies like India have that impedes the follow up innovations and translate to new technology products. The comparison between India and China shows the stagnation in GDP on R&D. As Nagaraj (2017) draw from World Bank indicators to highlight this, both the countries had 0.6% share of their GDP on R&D in 1996. For China it tripled to 1.8% of GDP, whereas for India it was 0.8% in 2011. The latest figure shows China is spending 2.55% of GDP whereas India is spending around 0.6%. The share of R&D in GDP is much lower for majority of the South countries. Thus there is need to give priority to research investment for R&D and South economies can create a more cooperative mechanism by which the R&D costs can be shared. This is further discussed.

We argue for developing more engagement between South economies in the new transformative change that is happening due to 4IR. It can address many of the barriers that can be seen coming up between North-South technology engagements and also many challenges that are similar between them can lead to more collective innovative approach. Cooperation between South economies can help to develop 4IR technology solutions more directed to their major developmental challenges such as poverty, inequality, and environmental degradation. Digital technologies can improve access to education and healthcare, financial inclusion including benefit sharing, and enable more efficient and sustainable use of natural resources. Technologies developed by a South country can be transferred in easier terms with supporting the other country for successful adoption and build follow up innovations. It can be useful as a particular South country technology development can be more useful to the other South country. A case in point is India's unique identification number 'Aadhar' which is helping in financial inclusion, providing services which are reaching the target beneficiary through this id validation. Similarly the standard developed for 5G that helps to enhance signal transmission which is so useful for major of South economies that have poor digital infrastructure.

Precision agriculture has shown how effective it is in making agriculture sustainable, productive and provide long term benefit for soil health, effective water usage, reducing wastage etc. Many South countries are dependent on agriculture and facing the above challenges. Blockchain has helped in connecting farming community produce in the supply chain with proper monitoring and tracking of the farm product and also protecting from quality infringement. Drones have been used to deliver medicines in remote places. Many South countries are developing applications surrounding 4IR technologies. The study examined the 'Technology 4SME' data base of UN ESCAP-APCTT (Asia Pacific Technology Transfer) that has 1168 technologies primarily developed by South economies listed therein (<https://apctt.org/technology-offer>). The technology transfer facilitation is easier and also supported by APCTT. We estimated around 400 technologies that have incorporated 4IR technologies and address developmental challenges.

Another of our argument is for South countries to develop institutional mechanisms to develop strong engagements between them. India for example has strong bilateral STI partnerships with majority of the advanced OECD countries. An example of this is Indo-French bilateral STI institution CEFIPRA through which various interventions for addressing developmental challenges have been made (Bhattacharya, et al. 2014). On the otherhand there is no significant bilateral STI partnership of India with South countries. Plausibly this is similar in other South economies. It is also useful to engage more extensively with multilateral institutions that focus on South countries partnerships in STI, entrepreneurship and technology transfer. The UN-ESCAP (The Economic and Social Commission for Asia and the Pacific) is a case in point. It operates in the Asia-Pacific region and has 53 member states and 9 associate members. ESCAP is headquartered in Thailand. Along with APCTT, the other regional institutes of ESCAP are APCICT- Asian and Pacific Training Centre for Information and Communication Technology for Development in *Incheon City, Korea*; APDIM - Asian and Pacific Centre for the Development of Disaster Information Management in *Tehran, Islamic Republic of Iran*; CSAM- Centre for Sustainable Agricultural Mechanization in *Beijing, China*, and SIAP- Statistical Institute for Asia and the Pacific *Makuhari, Chiba Prefecture, Japan*. The ESCAP along with its regional institutions provides support for inclusive, resilient, and sustainable development in the region. They all have close cooperation with other United Nations entities and intergovernmental organizations in the region. The author has done a broad study of ESCAP and a detailed study of APCTT. The study has highlighted the various institutional arrangements that are there to strengthen the cooperation between the Asia-Pacific region through the regional institutes. Many initiatives have been taken to develop capacity of the member countries in 4IR. South Korea, China, India have been actively involved in this. As many member countries face climate challenges that compound their developmental challenges, the disaster management and sustainability through 4IR is getting focus. The Study however finds that the

extent of engagement can be increased multi-folds. The technology transfer for example as seen through APCTT has been few and also the technologies do not actively updated by member countries in the databases.

The covid-19 has exposed the fault lines. However, drawing from the stories of resilience, unprecedented response and some extraordinary moves from countries to fight this pandemic have also suggested new ways forward. Development scholars have been highlighting the need for long term transformative change in our industrial society. They argue that there is need to expand the criteria of “success” in economic development beyond growth and profit. The thesis for post-transformation makes the argument that growth cannot be the sole objective, there are other human values such as equity or sustainability that are critical to be addressed. Trust and sense of social contract are important in this new framework. *This is highly relevant in our engagement with 4IR.*

The promise of 4IR can only come when we revisit and address the challenges which we are facing. The imperfections in our systems, the global governance without addressing them the Promise of 4IR will only be for few! It will lead to a more polarized world and within even rich countries a highly skewed inequality gap. Challenge is not only in creating an enabling ecosystem; *It further calls for confronting the uncertainties of converging technologies, reconfiguring Outmoded restrictive regulations (Maynard, 2015) and moving into post-normal science (Funtowicz & Ravetz, 1993) and to prepare human (society) for radical human-machine interface (Kraneneburg, 2008; Aegerter and Benziger, 2015, Bhattacharya, 2020).*

References

- Agerter, A. and benziger, A.S. (2015). Study of internet on the Human Abilities in Future. *International Journal of Advanced and Applied Sciences*, 2(5), 1-5.
- Baweja, B., Donovan, P., Haefele, M., Siddiqui, L., and Smiles, S. (2016). Extreme automation and connectivity: The global, regional, and investment implications of the Fourth Industrial Revolution, UBS White Paper for the World Economic Forum Annual Meeting 2016. <https://www.ip-watch.org/weblog/wp-content/uploads/2017/09/ubs-vierte-industrielle-revolution-2016-01-21.pdf>
- Bhattacharya, S. (2020) Science and Innovation in the 21st Century: New Paradigms and Challenges for Policy Design, *Current Science*, 18 (3), 348-349.
- Bhattacharya, S., Kaul, A., Shilpa, Fayaz, S. (2014). CEFIPRA 25 Years: Strengthen Scientific Competence & Enable Industrial Competitiveness of Indo-French S&T Ecosystem. Published by CEFIPRA (Indo- French Centre for Promotion of Advanced Research): New Delhi. 2014.
- Bhattacharya, S (2023). Covid-19: Opportunities, New Expectations and challenges for Innovation Policy In lakhwinder Singh & K.J. Joseph (Ed.) *Reimagining innovation systems in the Covid-19 and Post Covid-19 world* (pp 121-1310, , New Delhi & Abingdon, Oxfordshire: Routledge
- Booth, A., Sutton, A., Papiouannou, D. (2016). *Systematic approaches to a successful literature review*. London:Sage.
- Bonilla SH, Silva HRO, Terra da Silva M, Franco Gonçalves R, Sacomano JB. (2018). Industry 4.0 and Sustainability Implications: A Scenario-Based Analysis of the Impacts and Challenges. *Sustainability*. 10(10):3740.
- Bruland, K. and Mowery D.C. (2005). Science, Technology and innovation Policy In: *The Oxford Handbook of Innovation* Ed. Jan Fagerberg, David C. Mowery and Richard R. Nelson. Oxford University Press: UK
- Brynjolfsson, E. and A. McAfee (2014), *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W.W. Norton & Company Inc.
- Choung J.Y., hameed, T. and Ji I. (2011). Role of formal standards in transition to the technology frontier; Korean ICT Systems. *Telecommunication Policy* 35:269-287
- Christopher May and Susan K. Sell (2008). *Intellectual Property Rights: A Critical History*. Lynne Rienner Publishers: USA
- Deloitte Insights. Tech Trend, 2019
- Digital Personal Data Protection Bill, 2022. <https://www.meity.gov.in/content/digital- personal-data-protection-bill-2022> Access on 4th January 2023
- Dodgson, M. and Gann, D. (2010) *Innovation: A Very Short History*. Oxford university Press: UK
- Drew. L. (2019). The ethics of brain-computer interfaces. *Nature* 571(7766):S19-S21

- E von Hippel, (1988). Sources of Innovation. New York: Oxford University Press.
- Federal Ministry for Economic Affairs and Energy, Germany (2016) Future of the German Mittelstand Action programme.
- Ferguson, C. H., & Morris, C. R. 1993. Computer wars. New York: Random House.
- Fujimoto T. (2007). Architectural based competitive advantage- A design information view of manufacturing. *Evolutionary and Institutional Economics Review*, 4(1):55-112
- Funtowicz, S. and Ravetz, J. R., (1993). Science for the post-normal age. *Futures*, 25: 735-755.
- "G-20 Osaka Summit." *IndianExpress*, June 2019, <https://indianexpress.com/article/india/g-20-osaka-summit-narendra-mod-india-declaration-on-free-flow-of-data-across-borders-shinzo-abe-5805846/> Access on 3 December 2022
- Heller, M. A. and Eisenberg, R. (May 1998). Can Patents Deter Innovation? The Anticommons in Biomedical Research". *Science*, 280 (5364): 698-701.
- Henderson R.M. and Calrk K.B. (1990). Architectural Innovation: The reconfiguration of existing product technologies and the failure of established firms. *Administrative Science Quarterly*, 35(1):9-30
- Kranenburg R. Van (2008). The Internet of Things: A Critique of ambient technology and all-seeing network of RFID. Amsterdam
- Lee, K. and Lim, C. (2001). Technological regimes, catching-up and leapfrogging: findings from the Korean industries. *Research Policy* 30, 459-483
- Liao, at al. (2018). The impact of the fourth industrial revolution: A cross-country/region comparison. 10.1590/0103-6513.20180061
- Liu. C (2017). International competitiveness and the Fourth industrial Revolution. *Entrepreneurial Business and Economics Review*, 111-133
- Lundwall, B.A. and Borrás. S. (2005). Science, Technology and innovation Policy In: *The Oxford Handbook of Innovation* Ed. Jan Fagerberg, David C. Mowery and Richard R. Nelson. Oxford University Press: UK
- Macleod, C (1988). *Inventing the Industrial Revolution: The English Patent System, 1660-1800*. Cambridge University Press: UK.
- Marcus, C. (2015). Architectural Innovation in China: The Concept and its Implications for Institutional analysis. *Econstar Working paper on east Asian studies*, No. 104/2015
- McKinsey Global Institute (2019). <https://www.mckinsey.com>. Access 4th December 2022.
- Maynard, A. (2015). Why we need risk innovation. *Nature Nanotech* 10, 730-731.
- Nagaraj, R. (2017). Economic Reforms and Manufacturing Sector Growth. *Economic and Political Weekly*, LII (2). Press Information Bureau (2014) Digital India – A programme to transform India into digital empowered society and knowledge economy. <https://pib.gov.in/newsite/printrelease.aspx?relid=108926>. Access 11th December 2022.

- Qazi, A. M., Mahmood, S. H., Haleem, A., Bahl, S., Javaid, M., & Gopal, K. (2022). The impact of smart materials, digital twins (DTs) and Internet of things (IoT) in an Industry 4.0 integrated automation industry. *Materials Today: Proceedings*, 62, 18-25.
- Rashba, E. and Gamota, D. (2003). Anticipatory standards and commercialization of nanotechnology. *Journal of nanoparticle Research* 5, 401-407
- Ravetz, J. R., Post-normal Science and the Complexity of Transitions Towards Sustainability. *Ecological Complexity*, 2006, 3 (4), pp. 275-284.
- Roco M and Bainbridge W (2003). Converging technologies for improving human performance: Nanotechnology, Biotechnology, Information Technology, and Cognitive Science. Ed. by and. Dordrecht; Kluwer Ac. Publ
- Rostow W.W. (1975). Kondratieff, Schumpeter and Kuznets: Trends Periods Revisited. *The Journal of Economic History*, 35(40, 719-753.
- Segar, Alexandar "India and Budapest Convention." *Digital Frontier* Oct 2016, <https://www.orfonline.org/expert-speak/india-and-the-budapest-convention-why-not/> Access on 3 December 2022
- Serpanos, D., Wolf, M., Serpanos, D., & Wolf, M. (2018). Industrial internet of things. *Internet-of-Things (IoT) Systems: Architectures, Algorithms, Methodologies*, 37-54.
- Schwab, K. (2017). The fourth industrial revolution. Currency.
- Unido.org/UNIDO_industry_4.0
- Weibe E. Bijker, Thomas P. Hughes and Trevor J. Pinch (eds.) *The Social Construction of Technological Systems: New Directions in the Sociology and History of Technology*, London: The MIT Press, 1987 pp. 17-50.

Acknowledgement

The author expresses thanks to INDIALICS wherein discussion on 4IR and opportunity to speak in the lecture series has helped to frame this paper.

Note

¹ <https://www.epo.org/news-events/in-focus/ict/fourth-industrial-revolution.html>

