Drivers and Impacts of Structural Change in Sub-Saharan Africa

Evidence from Households and Firms

Bernardo Caldarola

Submitted for the degree of
Doctor of Philosophy in Science and Technology Policy Studies
Science Policy Research Unit (SPRU)
University of Sussex
May 2022

Statement

I hereby declare that this papers style thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree. An earlier version of Chapter 2 has been presented at the 8th Institute of Development Studies PhD Conference on 19 September 2019, and at the 7th WICK PhD Workshop on 7 January 2020. Previous and current versions of Chapter 3 have been submitted and presented at the following conferences: the 8th International Schumpeter Society Conference, 8-10 July 2021; the 1st YSI-SARChI Young Scholar Conference on Structural Change and Industrial Policy in Africa, 20-21 July 2021; the 17th Globelics International Conference, 3-5 November 2021; the 8th European Conference on Corporate R&D and Innovation – CONCORDi, 22-25 November 2021; the 5th RIE Conference and 12th MEIDE Conference, 29 November – 1 December 2021. A previous version of Chapter 4 has been presented at the 27th SPRU PhD Forum on 17 June 2021, and in research seminars at the Catholic University of Milan (23 February 2022) and at University of Bari (23 March 2022). Chapter 5 of this thesis is co-authored with Marco Grazzi, Marco Sanfilippo, and Martina Occelli, and it has been funded and prepared as part of the "Skills and transitions" research project by the Research Department of the International Labour Organization (ILO). I have taken part to the data construction, formal analysis, framework design, and writing stages of the Chapter. A previous version has been published as part of the ILO Working Paper Series, and is available online.[1] Additionally, Chapter 5 has been presented at the LABOR Workshop at the Collegio Carlo Alberto (25 February 2022), and in research seminars at University of Bari (23 March 2022), University of Sussex (27 April 2022), and UNU-Merit (28 April 2022).

Signature:

[1] https://www.ilo.org/global/publications/working-papers/WCMS_843046/lang--en/index.htm.

UNIVERSITY OF SUSSEX

Bernardo Caldarola, Doctor of Philosophy

Drivers and impacts of structural change in Sub-Saharan Africa: Evidence from households and firms

Summary

The main objective of this PhD thesis is to understand the contribution of the micro level dynamics of productive activities on structural change in Sub-Saharan Africa. The thesis combines individual, household, firm, and geospatial data from Ghana, Nigeria and Rwanda, employing quantitative methods from economics, geography and complexity science.

In the first chapter I use data from the Ghana industrial census, household surveys and exports to compare different perspectives on how its sectoral composition has changed between 2005 and 2014. In the second chapter, I investigate how such different patterns of structural change may be related. Using the industry census data from Ghana, I study the patterns of co-location of informal and formal economic activities across industries, testing different channels through which informal activities can contribute to structural change.

In the third chapter I explore how ICT can drive structural change in the informal sector. I use panel data on Nigerian non-farming household enterprises and mobile coverage data at district level (2010-2015) to measure the impact of mobile internet adoption on firms' performance and industrial composition, and the extent to which it can foster economic inclusion.

In the fourth chapter, I further explore the impact of ICT on changes in the labour market. Using data on the mobile internet roll-out in Rwandan districts combined with Labour Force and Census data (2002-2019), I estimate the impact of ICT diffusion on employment size and skill composition.

The findings of the four chapters reveal that informal economic activities are heterogeneous and interact with the economy as a whole, contributing significantly to the economic transformation of low-income countries (Chapters 1-2). Policies can also influence the pattern of structural change: for instance, ICT diffusion can benefit informal firms and labour markets, but can also lead to more inequality, further requiring policy interventions to foster processes of inclusive structural change (Chapters 3-4).

Acknowledgements

I would like to express my endless gratitude to my supervisors, Tommaso Ciarli and Maria Savona, for guiding me through this process, being my landmarks, and being a constant source of inspiration. I have always felt trusted, encouraged, and supported through turbulent times. It is no overstatement to say that this thesis would not have been possible without them, and I can't imagine this journey accompanied by anyone else. Thank you for your patience, dedication, and friendship from the deep of my heart.

I would also like to thank the University of Sussex Business School and the Science Policy Research Unit for granting me a doctoral scholarship, without which I wouldn't have been able to undertake this incredible journey.

It has been a privilege to share this journey with a fantastic community. Filippo has quickly become my friend and mentor at SPRU, guiding me through a new environment, and landing his crucial support since before applying for the PhD, with great generosity. I would like to thank my friends at Sussex, Amrita Saha, Andrea Carboni, Daniele Guariso, Emanuele De Girolamo, Graziana di Pede, Guillermo Larbalestier, Juan Manuel Del Pozo Segura, Kai Wang, Maika Schmidt, Marco Carreras, Margherita Bove, Mattia Di Ubaldo, Nicolò Tamberi and Pierfrancesco Rolla have made Brighton home and a happy place. I will always keep fond memories of the time spent in the Jubilee PhD rooms and at the IDS bar in the company of the Sexy Pint Recreational Unit, an invaluable source of peer support and learning. I would like to thank the many people who have joined the Sussex weekly football and the Pluralist Economics Society at Sussex, which have been — in very different ways — two important parts of my time in Brighton.

The SPRU research community has also been an incredibly welcoming and nurturing environment. In particular, I would like to thank Josh Siepel, Daniele Rotolo, Roberto Camerani, Alberto Marzucchi and Steve Sorrell for giving me the chance to teach and do research during my PhD. I hope I have returned half of what I have learnt with them. Amrita Saha deserves a special mention for introducing me to SPRU back in 2017 and offering me the possibility of working on many exciting projects since the beginning of my PhD, despite my little experience. I would also like to express my gratitude to Ed Steinmueller and Martin Bell, who provided comments on a very early and distant version of this thesis.

My gratitude goes to the Deep Transitions for being brave and kind enough to give me my first opportunity to sing and play music in a band. A special mention goes to the Sussex UCU and the SPRU strikers, who have supported the entire doctoral community in the struggle for fairer working conditions and a better university.

Many of the reasons I managed to complete this journey are not in Brighton. I will always be grateful to my family for their unwavering love and support, in this journey and many others, and for backing each and every decision of mine. I am grateful to my brother Francesco, for being a living reminder that there's a great deal of "magic" that escapes empirical observation and must not be overlooked. Ciccio, Davide, Gio, Gioby, Karim, Lucia and Michele have gifted me with their friendship since we first met in Pavia, more than ten years ago, and I am grateful that we keep sharing the same spirit of when we used to discuss life in the University courtyards. I would also like to thank my friends in Trani, who have always been there to make me feel at home again whenever I am back.

Most importantly, the completion of this PhD was made possible by Inma and her constant and relentless love, support, and generosity. We have shared every joy, achievement, failure, and suffering in this journey and many adventures across four countries and two continents over the last four years. I feel privileged that I could live all this during my academic journey next to an extraordinary person.

Contents

List of Tables x

List of Figures xii

1 Introduction 1

1.1 Background . . . 1
1.2 Related literature and theoretical framework . . . 4
1.3 Contribution . . . 10
1.3.1 Structural Change(s) in Ghana: a Comparison between the Trade, Formal, and Household Sectors . . . 11
1.3.2 Informal and Formal Industrial Co-location and Structural Change in Ghana . . . 12
1.3.3 Mobile Internet Adoption and Inclusive Structural Change: Evidence from Nigerian Non-Farming Enterprises . . . 13
1.3.4 Mobile Internet, Skills and Structural Transformation in Rwanda 14
1.3.5 Original contribution to knowledge . . . 15

2 Structural Change(s) in Ghana: a Comparison between the Trade, Formal, and Household Sectors 17

2.1 Introduction . . . 17
2.2 Literature review . . . 22
2.2.1 Structural change and complexity . . . 22
2.2.2 Structural change in Sub-Saharan Africa . . . 25
2.2.3 Towards an analytical framework for the analysis of structural change . . . 27
2.3 Empirical strategy . . . 30
2.3.1 Data . . . 30
2.3.2 Methodology . . . 33
2.4 Analysis and results . . . 38
2.4.1 Structural change in Ghana . . . 39
2.4.2 Export sector . . . 42
2.4.3 Formal sector . . . 48
2.4.4 Informal sector . . . 52
2.4.5 Discussion . . . 56
2.5 Conclusions . . . 60

3 Informal and Formal Industrial Co-location and Structural Change in Ghana 63

3.1 Introduction . . . 63
3.2 Background literature . . . 67
3.2.1 Structural change, industrial variety and relatedness . . . 68
3.2.2 Informality and structural change . . . 73
3.2.3 Summary and gaps . . . 76
3.3 Empirical approach . . . 78
3.3.1 Data . . . 79
3.3.2 Construction of the variables . . . 81
3.4 Results . . . 90
3.4.1 Formal and informal co-location at a glance . . . 90
3.4.2 Agglomeration patterns . . . 97
3.4.3 Drivers of co-location . . . 105
3.5 Conclusions . . . 118

4 Mobile Internet Adoption and Inclusive Structural Change: Evidence from Nigerian Non-Farming Enterprises 122
4.1 Introduction . . . 122
4.2 Analytical framework and related literature . . . 128
4.2.1 Inclusive Structural Change . . . 128
4.2.2 Internet and development in Africa . . . 130
4.2.3 Gaps and contribution . . . 132
4.3 Empirical strategy . . . 133
4.3.1 Data . . . 133
4.3.2 A descriptive picture of the data . . . 138
4.3.3 Estimation strategy . . . 143
4.4 Results . . . 152
4.4.1 Structural change . . . 152
4.4.2 Inclusion . . . 160
4.5 Conclusions . . . 164

5 Mobile Internet, Skills and Structural Transformation in Rwanda 170
5.1 Introduction . . . 170
5.2 Background literature . . . 174
5.3 Data . . . 177
5.3.1 Mobile internet . . . 177
5.3.2 Individual-level data . . . 178
5.4 Empirical specification . . . 184
5.5 Results and discussion . . . 188
5.5.1 Robustness checks . . . 191
5.5.2 Mechanisms and extensions . . . 197
5.6 Conclusion . . . 203

6 Conclusions 205
6.1 The contribution of informality to structural change . . . 207
6.1.1 Policy implications . . . 210
6.2 Technological change and inclusive structural change . . . 212
6.2.1 Policy implications . . . 215
6.3 Limitations of the analysis and avenues for further research . . . 218

Bibliography **223**

A Appendix to Chapter 2 **252**
A.1 Treemaps . . . 252
A.2 Tradable services – correspondence table . . . 257
A.3 Validating the complexity and fitness framework . . . 258
A.4 Final dataset . . . 260

B Appendix to Chapter 3 **269**
B.1 Network representation . . . 269
B.1.1 Randomisation tests . . . 273
B.2 Comparing standardisation measures . . . 275
B.3 District fitness . . . 277
B.4 Industrial complexity . . . 283

C Appendix to Chapter 4 **291**

D Appendix to Chapter 5 **294**
D.1 Figures . . . 294
D.2 Tables . . . 296

List of Tables

2.1 Changes in export shares by ISIC major groups 48
2.2 Changes in employment shares (formal sector) by ISIC major groups 52
2.3 Changes in employment shares (household sector) by ISIC major groups 56

3.1 Employment shares across macro-sectors 92
3.2 Main variables: descriptive statistics 93
3.3 Top 10 4-digit ISIC industries by complexity 103
3.4 Bottom 10 4-digit ISIC industries by complexity 104
3.5 GLM regressions: informal and formal co-location 108
3.6 GLM regression results, informal-formal co-location 114

4.1 Non-farming enterprise dataset, descriptive statistics 142
4.2 Descriptive statistics: averages by ISIC major groups 145
4.3 Electricity constraints in Nigerian households 150
4.4 Electricity constraints by lightning strike density exposure 150
4.5 OLS and 2SLS results, Log sales per worker 153
4.6 2SLS results, mechanisms 156
4.7 2SLS results, (log) sales per worker, by industry 158
4.8 2SLS results, household transition 159
4.9 2SLS results, household and non-household labour in NFEs 162
4.10 2SLS results, labour within and outside the household 163
4.11 2SLS and OLS results, NFE ownership. entry, and exit 165

5.1 Share of occupations, average across districts (2002-–19) 180
5.2 OLS and 2SLS results, employment 190
5.3 2SLS results, by type of occupation and industry of employment . . . 192
5.4 2SLS results, by education 198
5.5 2SLS results, migrant workers 200
5.6 OLS results, firm productivity 202

A.1 Correspondence between Harmonised System tradable services and ISIC industries 257
A.2 Correlation between district fitness and nighttime luminosity 259
A.3 Final dataset 260

B.1 Comparison of matrix densities 271
B.2 Ghanaian districts ordered by fitness 277
B.3 4-digit ISIC industries by complexity 283

D.1 Share of industries, average across districts 296

D.2 Descriptive statistics 297
D.3 Robustness, province-specific time trends 298
D.4 Robustness, initial conditions 298
D.5 Robustness, accounting for 4G coverage 299
D.6 OLS results, individual DHS data 300
D.7 2SLS results, migrant workers by occupation and industry 301
D.8 Production function coefficients 302

List of Figures

1.1 Co-citation network: literature on structural change in Africa. 7

2.1 De-industrialisation in Ghana and high-income countries (1960–2013) 40
2.2 Employment and value added trends in Ghana (1960–2011) 41
2.3 Trade sector changes (2003–2013) 45
2.4 Trade sector complexity (2003–2013) 46
2.5 Formal sector changes (2003–2014, only manufacturing and mining) . 49
2.6 Formal sector complexity (2003–2014, only manufacturing and mining) 50
2.7 Informal sector changes (2005–2013) 53
2.8 Informal sector complexity (2005–2013) 54

3.1 Correlation between Ghana's industrial co-location and global trade proximities. 83
3.2 Correlations between co-location measures of Ghanaian industries. . . 94
3.3 Correlations between main variables 95
3.4 Patterns of specialisation across Ghanaian districts 98
3.5 Formal and informal diversity index of Ghanaian districts. 99
3.6 Fitness of Ghanaian districts 100
3.7 Differences in ubiquity across industries 102
3.8 Empirical framework 106
3.9 Average Marginal Effects informal and formal co-location 110
3.10 Interaction between relatedness and input-out in informal co-location 111
3.11 Interaction between relatedness and input-out in formal co-location . 112
3.12 Effect of input-output linkages within and outside ISIC 1-digit industries: informal and formal co-location 113
3.13 Average Marginal Effects on informal-formal co-location 116
3.14 Interaction between input-out linkages and ISIC 1-digit industries: informal-formal co-location 117

4.1 Global access to internet 123
4.2 Inclusive Structural Change framework 130
4.3 Mobile coverage in Nigeria (2010–2015) 137
4.4 Non farming enterprises by industry (2010–2015) 140
4.5 Household transition between industries over time 141
4.6 NFE sales per worker distribution, (2010–2015) 144
4.7 Correlation between mobile internet adoption and NFE performance (2012 and 2015) 144
4.8 Correlation between lightning strikes per capita and 3G adoption . . 147

4.9 Correlation between lightning strike density and operating costs of NFEs . . . 150
4.10 Event study results, sales per worker . . . 154
4.11 Event study results, labour and sales . . . 160
4.12 Event study results, NFE ownership . . . 166

5.1 District-level mobile internet 3G coverage diffusion in Rwanda (2008–18) 178
5.2 Percentage of skilled workers in Rwandan districts (2002, 2012, 2019) 181
5.3 Share of people working in skilled occupations and mobile internet 3G coverage (2012 and 2019) . . . 182
5.4 Employment by major industrial groups (2002–2019) . . . 183
5.5 Mobile internet 2SLS coefficients, industry regressions . . . 193
5.6 Event study results . . . 196

A.1 Treemap trade (2003) . . . 252
A.2 Treemap trade (2013) . . . 253
A.3 Treemap formal (2003, only mining and manufacturing) . . . 253
A.4 Treemap formal (2014, only mining and manufacturing) . . . 254
A.5 Treemap formal (2014) . . . 254
A.6 Treemap informal (2005) . . . 255
A.7 Treemap informal (2005, agriculture excluded) . . . 255
A.8 Treemap informal (2013) . . . 256
A.9 Treemap informal (2013, agriculture excluded) . . . 256
A.10 Correlation between fitness and nighttime luminosity. . . . 259

B.1 Visual representation of the aggregate productive structure of Ghana based on domestic employment, with community structure. . . . 270
B.2 Description of the adjacency matrices . . . 272
B.3 Results of the randomisation tests . . . 274
B.4 Correlation between proximity and co-location . . . 276

C.1 Event study results, robustness checks . . . 291
C.2 Event study results, mechanisms . . . 292
C.3 Event study results, sales per worker, by industry . . . 293

D.1 Average years of education in Rwandan districts (2002 and 2019) . . . 294
D.2 Mobile phone ownership and 3G coverage (2017 and 2019) . . . 295
D.3 Home internet connections and 3G coverage (2017 and 2019) . . . 295
D.4 Migration between districts and 3G coverage (2017–2019) . . . 296

Chapter 1

Introduction

1.1 Background

Many economies in countries in Sub-Saharan Africa (SSA) have been growing at high rates since the 1990s. In some cases, sustained economic growth has been associated to reduced poverty and improved welfare among the population (Young, 2012; Arndt et al., 2016b). However, in many SSA countries, economic growth has not created the social and technological conditions, which, historically, accompany the structural transformations required to achieve a modern, sophisticated productive structure (Rodrik, 2018), able to create new jobs in highly productive and knowledge-intensive industries (Kanbur et al., 2019). Rather, in the post-colonial period, many SSA countries have relied on their natural resources endowments and increased their export specialisation in resource-intensive primary industries, which has not absorbed surplus labour from low-productivity industries, such as subsistence agriculture (McMillan et al., 2014). Several scholars have shown the limitations related to specialisation in natural resources industries, with low degrees of value added and weak technological linkages with other domestic industries, and which rely on foreign technologies. These limitations are poor development of productive capabilities and technological linkages (Hirschman, 1958; Bell and Pavitt, 1992) and limited diversification towards more knowledge-intensive industries (Prebisch, 1950; Sachs and Warner, 2001; Hidalgo et al., 2007; Lin et al., 2011; Fagerberg and Verspagen, 2021).

The dependence of many SSA countries on commodity exports, and their lack of productive and technological capabilities, are strongly associated to their structural transformation trajectory. There is a quite widespread view that, although SSA countries have been following heterogeneous directions of structural change (Bah, 2011), they have de-industrialised prematurely, as a result of weak comparative advantage in manufacturing and diffusion of labour-saving technologies (Rodrik, 2016). This is a process that is associated to industrial reallocation of jobs in the 'wrong' direction, that is, from high- to low-productivity industries, especially services (McMillan et al., 2014), resulting in the enlargement of an already sizeable and persistent informal sector (McMillan et al., 2017b; Benjamin and Mbaye, 2020). The shift of labour towards low-productivity jobs, mostly informal services, is due, in part, also to demand side factors, such as increased per capita income (Diao et al., 2019b). Such factors have prevailed on technological change factors, that were determinant in most industrialised countries, such as mechanisation and increased productivity in agriculture, manufacturing, and services (McMillan et al., 2017b).

On the one hand, the trends described above determine a process characterised by flows of informal labour from agriculture directly to services and contribute only marginally to growth-enhancing structural change. At the same time, the contribution of the informal sector to structural change could be sizeable, were the informal workers absorbed by modern emerging modern sectors, as hypothesised in the dual-sector model proposed by Lewis (1954). While this view assumes an overall passive and homogeneous informal sector, recent evidence from SSA countries shows that informal activities are highly heterogeneous in terms of sectors (Nagler and Naudé, 2017). In some cases, they have built the capabilities required to innovate and produced innovations (Fu et al., 2018; Fu, 2020; Avenyo et al., 2019, 2020) that have contributed significantly to productivity growth (Diao et al., 2018a), and industrialisation (Kraemer-Mbula and Monaco, 2020). Firms and workers in the informal sector are not a passive labour force; they carry a stock of capabilities that emerging industries can draw upon. Therefore, it is important in this context to reconsider the contribution of informality to the process of structural change in SSA, its heterogeneity, and its linkages with the aggregate economic structure.

It has been argued, also, that the current, digital technology-based technological paradigm, is providing opportunities for SSA countries to acquire new capabilities and steer structural transformation towards new, more productive, activities, especially in contexts where informality is prevalent (Kaplinsky and Kraemer-Mbula, 2022). However, the trajectory of economic transformation in the SSA countries will depend on the direction and pervasiveness of technological change. Essentially, the economic and social benefits of innovation-led structural change depend on the nature of the technologies developed and adopted. While innovation drives economic transformation by creating opportunities for the emergence of new economic sectors (Saviotti and Pyka, 2004), its disruptive nature (Schumpeter, 1934; Aghion et al., 2019) may generate exclusion (Ciarli et al., 2021b). Therefore, it is necessary to examine whether the current wave of technological change in SSA, characterised by the diffusion of Information and Communication Technologies (ICT), is driving a process of virtuous or of vicious structural change.

This thesis appraises the drivers and impacts of structural change in SSA, taking account of: i) the contribution to structural change of the formal and informal sectors, the latter amounting to 86 per cent of total employment in low- and middle-income SSA states (Chapter 2); ii) how activities in the formal and informal sectors are related (Chapter 3); iii) the nature and effect of technological change on structural change in the informal sector (Chapter 4); and iv) the inclusionary or exclusionary outcomes of innovation and structural change (Chapters 4 and 5). Section 1.2 in this introduction provides a brief review of the literature and describes the theoretical background for the empirical analysis; Section 1.3 summarises the contribution made by this thesis to the extant literature.

1.2 Related literature and theoretical framework

The contributions by McMillan et al. (2014) and Rodrik (2016), mentioned in Subsection 1.1, provide crucial evidence on the trajectory of structural change in Africa and connect different bodies of knowledge on related topics. This is depicted in the co-citation network emerging from the literature on structural change in Africa provided in Figure 1.1.[1] It can be seen that the study by McMillan et al. (2014) is the central node in the literature on structural change in Africa, and brings together several different streams of work, represented by the clusters identified by the different node colours. In particular, this paper appears to bridge between at least four groups of relatively homogeneous literature strands.

First, the green group, which includes McMillan et al. (2014), gathers together contributions from both the traditional structuralist literature (Lewis, 1954; Gerschenkron, 1962; Chenery et al., 1986) and more recent structuralist work (McMillan and Rodrik, 2011; Page, 2012; de Vries et al., 2015; Herrendorf et al., 2014; Rodrik, 2016). The early structuralist contributions examine the processes of growth and transformation in industrialised countries and identify economy-wide productivity differentials as the main driver of structural change and economic growth. As productivity increases in emerging, modern and more productive industries, the firms in these industries poach 'excess' labour from less productive, traditional sectors. Productivity gaps in the

[1]The network in Figure 1.1 is a co-citation network in which each node is a paper and the edges link papers that are co-cited. Papers are clustered based on the relative frequency of their co-citation (Small, 1973). Co-citation networks provide an accurate way to cluster a corpus of academic literature into scientific communities or topics (Boyack and Klavans, 2010). The data used to construct the network consist of all the journal articles on structural change in Africa, obtained from Web of Science, using the following search query: [`"structural change" (Topic) AND africa (Topic)`]. The search returned 208 publications; the list was refined by selecting only journal articles and book chapters in Social Sciences, which reduced the sample to 138 items. The refined results were exported to include information such as author, year, title, scientific journal, and reference list. The co-citation network was constructed by establishing a link (edge) between the papers (nodes), based on whether two papers were cited in the same journal article, for the sample of 138 articles on structural change in Africa. The strength of the connection increases with more frequent co-citation – indicated by the thickness of the edges in the network. A threshold of 5 co-citations was used to select the papers included in the network: after merging to reduce duplicate items, we obtained 66 journal articles which met this requirement. To improve visualisation, only connections with strength (number of co-citations) higher than 3 are displayed. The colour of the nodes distinguish 4 clusters of publications more likely to be cited together (Waltman et al., 2010); minimum cluster size was set to 3 items. The data and graph were elaborated using VosViewer.

economy are thus seen as temporary inefficiencies in the allocation of productive factors, such as labour, but also considered crucial to trigger structural change (Kuznets, 1966). Structuralist models, such as the dual sector model proposed by Lewis (1954) (and its later developments), constitute the theoretical basis of several empirical studies of industrial productivity differentials to determine the direction of structural change in SSA countries (McMillan and Rodrik, 2011; McMillan et al., 2014, 2017b; de Vries et al., 2015; Herrendorf et al., 2022). Most of these studies show that, since the 1990s, in many SSA countries labour shifted away from high-productivity industries – such as manufacturing (Rodrik, 2016) – towards below-average productivity industries, comprised mostly of informal jobs.

Second, the red cluster is strictly related to the green cluster, although it is more heterogeneous. In addition to some pioneering empirical works on the relationship between industrial structure and economic growth (Chenery, 1960; Kuznets, 1966), the red cluster includes several topics discussed by the structural change literature, that focus on the relative importance of specific industries. For instance, the red cluster includes examinations of the role of agriculture in structural transformation (Gollin et al., 2007; Collier and Dercon, 2014), focused, in particular, on the persistence of agricultural productivity gaps in SSA (Gollin et al., 2014). It also includes publications stressing the importance of manufacturing as the engine of economic growth (Matsuyama, 2009; Duarte and Restuccia, 2010), due to its unique characteristic of a productivity convergence that is not conditional on the so called "growth fundamentals", such as institutions, macroeconomic management, and human capital (Rodrik, 2013), and discussing the perspectives of African industrialisation (Dinh et al., 2012; Newman et al., 2016). Finally, the red cluster includes publications that look at structural change along a rural-urban axis (Christiaensen and Todo, 2014; Dorosh and Thurlow, 2014; de Brauw et al., 2014).

Third, the blue cluster appears to be peripheral to the green and red clusters, although it includes a rather homogeneous stream of work on the determinants of economic growth. On the one hand, it includes work on the role of the so-called growth fundamentals, such as macroeconomic management (Rodrik, 2008) and institutions (Acemoglu et al., 2001). On the other hand, it includes a subset of papers that are central in the cluster and engage in a debate on the role of specialisation and diversification for promoting structural change and economic growth. These contributions associate economic growth with a pattern of specialisation followed by diversification and then re-specialisation (Imbs and Wacziarg, 2003; Cadot et al., 2011) and emphasise the importance of export specialisation in determining growth trajectories (Hausmann et al., 2007). Building on the idea that a country's export specialisation reflects its productive capabilities and, therefore, determines the possible direction of structural change towards more technology- and knowledge-intensive activities, there is a strand of more recent work that uses specialisation and diversification to construct a metric – product and economic complexity – to measure the capabilities required to export the product, and the array of capabilities available in countries (Hidalgo and Hausmann, 2009). The intuition behind measurement of economic complexity is that the more complex the product, the more difficult it is to export that product because it requires more capabilities. The authors show that more 'complex' countries (that is, countries able to export more complex products) are also more diversified. In other words, over time, they have built the capabilities required to produce a diverse range of products and have managed to export more complex products. A key finding is that countries (and regions) diversify more frequently towards related products (see Hidalgo et al. 2018 for a review), and that related diversification towards more complex products (industries) is associated to higher levels of economic development (Cristelli et al., 2013; Tacchella et al., 2018).

Fourth, the yellow cluster appears to be more peripheral and heterogeneous and includes theoretical econometric contributions, such as Stock and Watson (1993); Pesaran et al. (2001); Im et al. (2003), that introduce econometric methods and estimators that have often been adopted by empirical papers studying structural change, particularly focusing on time series and panel techniques for analysing longitudinal data.

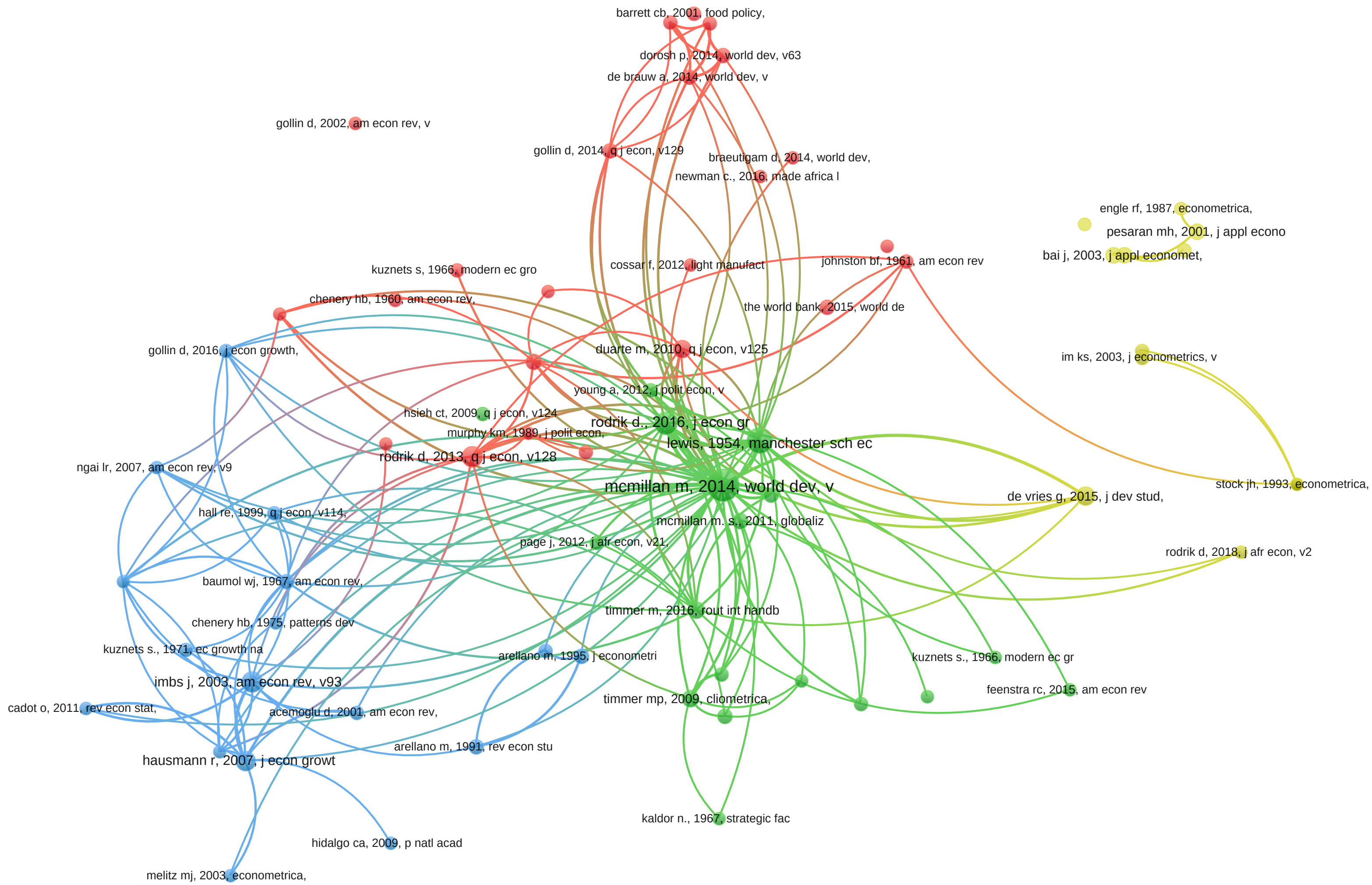

Figure 1.1: Co-citation network: literature on structural change in Africa.

Although the co-citation network is not a complete representation of the body of knowledge on structural change in SSA,[2] it covers a central body of academic research in this area. However, within this body of knowledge, there are three topics that are relatively underrepresented.

First, the literature included in the co-citation network is a good representation of the role played by formal and export activities and their industrial composition, in steering the process of structural change. However, work on the role of the informal sector – the largest employer in SSA – has framed its contribution to structural change mostly in terms of its absorption by more modern industries. Similarly, the most recent work on structural change and productivity in SSA, focus on the causes of its persistence rather than on its active contribution to structural change. However, recent evidence from Africa would seem to challenge the view that the informal sector cannot drive structural change and economic diversification and shows that informal firms are diverse (Nagler and Naudé, 2017; Diao et al., 2018b), innovative (Fu et al., 2018; Fu, 2020; Avenyo, 2018; Avenyo et al., 2019, 2020) and, in some cases, demonstrate productivity growth which accounts for a significant part of the structural change observed in the SSA countries (Diao et al., 2018a). Although they may be less productive, informal firms and workers may encompass the conditions – such as, accumulation of productive capabilities – required to trigger and sustain structural change towards more productive industries (Kraemer-Mbula and Wunsch-Vincent, 2016; Kraemer-Mbula and Monaco, 2020).

Second, the productivity and structural change literature does not provide empirical investigation of the technological channels driving productivity and the sectoral reallocation of labour[3] and, also, does not explore the process of accumulation of the capabilities underlying this process. This latter aspect is addressed, in part, by the export specialisation and complexity literature (Hidalgo and Hausmann, 2009), which

[2]Both the methodology and the source (Web of Science) used to build Figure 1.1 have some limitations. For instance, the Web of Science library does not include all the scientific journals that have published research on structural change in SSA countries.

[3]Rodrik (2016) argues that the patterns of structural change in SSA can be ascribed to labour-saving technological change, although the impact of technology is not measured directly

uses an index of capabilities that builds on trade patterns to identify a country's best developed capabilities. This study uses the underlying structure of capabilities to examine the process of diversification and economic upgrading, using the product space, an empirical tool that visualises the global production structure as a network of related products (Hidalgo et al., 2007). The "relatedness" between two products represents the measurable degree of similarity of the capabilities required to export them (Hidalgo et al., 2018). Recent advances in the product space literature have also connected the capability structure revealed by international trade patterns to technological linkages between products, further refining the measure of capabilities (Pugliese et al., 2019). However, it should be noted that, apart from some rare exceptions (Sbardella et al., 2017), a country's trade specialisation structure is often conflated with its productive structure. Evidence for SSA shows that many low- and middle-income countries are characterised by a large share of employment in non-tradable sectors, which, by definition, are excluded from the measurement of industry relatedness and complexity based on trade patterns. While export specialisation can be used to proxy for the most developed capabilities, it is a less representative proxy for the capabilities accumulated by employment and firms in non-tradable (or simply, non-traded) industries – such as the growing trade and wholesale services, but may be still relevant to national industry variety, representing the stock of capabilities on which industrial diversification strategies build.

Third, works that focus on structural change, in terms of both productivity and export specialisation, rarely measure the impact of technological change on the transformation of the productive structures in low- and middle-income countries. Moreover, despite the large literature on the relationship between economic growth and inequality (Kuznets, 1973; Bourguignon, 2004; Ravallion, 2004; Piketty, 2018), technological change and its effect on countries' industrial composition are seldom examined with respect to the inclusionary or exclusionary outcomes of this process, and whether such outcomes matter to further determine the trajectory of structural change. In fact, none of the core papers in the co-citation network on structural change in SSA deals explicitly with these issues. However, the neo-Schumpeterian literature (Freeman and Perez, 1988; Cimoli and Dosi, 1995; Aghion and Howitt, 1997), puts a strong emphasis on the role played by technology and innovation in promoting economic growth and structural change. This body of work considers that when a breakthrough technology is introduced, it pervades all aspects of the economy and society (Perez, 1986). In this view, productivity grows thanks to a better recombination of productive factors or the emergence of new products and industries (Saviotti and Pyka, 2004), and this transforms the industrial composition of the economy. However, although it enables economic diversification, innovation also replaces old economic activities, in a process that Schumpeter (1943) describes as "creative destruction". Productivity growth and the emergence of capital-intensive and more complex activities can lead to a reduction in labour input in the short term, depending on the skills required for emerging occupations (Autor et al., 2003; Buera et al., 2021). Also, technological change may lead to the concentration of economic activities in a few of the most productive firms (Autor et al., 2020; Altenburg et al., 2021). Predicting the effects of technology and innovation on the inclusion or exclusion of economic actors is difficult (Autor, 2022) and requires a deep understanding of both the nature of the technological change and the actors involved (Ciarli et al., 2021b).

1.3 Contribution

The theoretical literature reviewed above, highlights that analysis of the drivers and impacts of structural change in SSA requires a profound understanding of the effects of informality and technical change for driving structural change. It highlights, also, that their impact on inclusionary outcomes has been under-explored. This PhD thesis addresses these gaps, with a focus on SSA. First, it provides an in-depth study of the role played by informal sector activities and their positive or negative contribution to structural change. The thesis analyses the extent to which informal activities have hampered or enhanced structural transformation in SSA, directly, through the increased complexity of these activities and indirectly, through provision of inputs for formal sector activities (Chapters 2 and 3). Second, it examines the influence of technical change on structural change in SSA, where informality is prevalent (Chapters 4 and 5). Third, having accounted for the role of informal activities and technical change, the thesis analyses the distributional impacts of structural change in SSA, by investigating how workers gain and lose from technical change and structural change (Chapters 4 and 5). The thesis combines macro, meso and micro analysis and relies on the structuralist, complexity, and neo-Schumpeterian literature. It builds on their intersections, differences, and shortcomings, to contribute to analyses of the drivers and impacts of structural change in SSA. The main contributions provided by the four main chapters in this thesis, as well as the thesis' overall original contribution to knowledge, are summarised below.

1.3.1 Structural Change(s) in Ghana: a Comparison between the Trade, Formal, and Household Sectors

At the macro level, Chapter 2 uses the case of Ghana to unpack the role of informality in the process of structural change and to challenge the assumption that a country's trajectory of structural change is due entirely to its pattern of export specialisation. We combine a structuralist view of structural change as changes in the employment shares of different industries, with the insight that countries strive to diversify towards more complex industries, building on their existing capabilities and export specialisation, in pursuit of economic upgrading. Since information on exports may provide a limited view of what a country is capable of producing, Chapter 2 adopts and adapts the product space and complexity analytical frameworks to compare changes in the relative importance of industries across the trade, formal, and informal economic sectors, over a ten-year period (starting in 2003), to identify their respective structural change trajectories. To assess whether the Ghanaian labour force has moved towards more or less complex industries, the changes in relative 4-digit ISIC industry shares, are assessed against the industrial complexity index – an employment-based measure of industrial complexity that is based on the employment specialisation patterns in Ghanaian regions, to measure the complexity of finely disaggregated industries. Although there is mostly agreement in the empirical literature that exported products are usually the most competitive, the presence in SSA of a large informal sector and the relevance of non-tradable services, challenge the assumption that export specialisation can be used to capture the composition of the capabilities embedded in domestic employment. The results indicate that Ghana's export and formal sectors have moved towards more complex industries, although export specialisation has moved towards export of natural resources, characterised by high complexity, but limited employment creation. Also, while exports of manufactured goods have increased, employment in formal and informal manufacturing has contracted, although, in the former case, employment has relocated towards more complex manufacturing industries. In contrast, the heterogeneous informal sector has moved towards less complex activities, in both manufacturing and across industries.

1.3.2 Informal and Formal Industrial Co-location and Structural Change in Ghana

Chapter 3 analyses the extent to which the patterns of co-location in the informal and formal sectors and their interaction, created the conditions for structural change in Ghana. The linkages between formal and informal industries, measured by their co-location across regions, are investigated along with how different types of linkages (capabilities, technological, and complexity differentials) may explain the synergy between formal and informal industries. Chapter 3 relies on employment-based measures of informal, formal, and informal-formal co-location of 4-digit ISIC industries, explained in terms of their capabilities relatedness, input-output relationships, and complexity differentials. The results show that informal industries agglomerate more than formal industries and show a higher degree of relatedness and complementarity compared to formal sectors. The higher degree of similarity in capabilities among co-located informal industries, along with their homogeneity in terms of economic complexity, suggest that informal activities in Ghana contribute to Ghana's (related) industry variety and provide the preconditions for further economic diversification, for instance, by allowing labour flows among related industries. In contrast, formal industries are less related and present an enclave-like pattern of co-location. The patterns of interaction and co-location between informal and formal industries, suggest that informal industries are vertically integrated with unrelated formal industries, leaving few opportunities for flows of labour from informal to formal sectors.

1.3.3 Mobile Internet Adoption and Inclusive Structural Change: Evidence from Nigerian Non-Farming Enterprises

Chapter 4 deals with the impact of technological change on structural change in the informal sector, measured by firm performance, and the shift to more productive sectors. The analysis in this chapter investigates the beneficiaries of structural change driven by technological change, that is, the inclusiveness of technology-driven structural change. This micro-level analysis uses data on informal household enterprises in Nigeria. It focuses on the use of mobile internet, one of the most widely diffused digital technologies in Nigeria. Since less than 0.2 per cent of the population in Nigeria had a fixed phone line in 2015, mobile internet has provided most Nigerians with the first opportunity to access fast internet. The analysis looks at the impact of mobile internet adoption on performance, industry of activity, and demand for labour among Nigerian non-farming enterprises. To test empirically whether innovation-driven structural change leads to inclusion or exclusion in the informal sector in Nigeria, Chapter 4 provides new evidence on the effect of mobile internet adoption on the ability of entrepreneurs in the informal economy, to increase firm performance and/or to shift to more productive sectors. We also analyse the impact on inclusiveness, defined as the creation of working and entrepreneurial opportunities for Nigerian households. The findings indicate that mobile internet adoption significantly improves the performance of informal firms, almost exclusively service firms, and discourages entry to the manufacturing sector, continuing the trend towards tertiarisation. However, productivity gains are associated to both higher sales and lower demand for labour. This lower demand for labour would suggest a link between structural change and exclusion. However, our results show, also, that this reduced demand for labour is offset by higher job opportunities outside the household enterprise (possibly in the formal sector).

1.3.4 Mobile Internet, Skills and Structural Transformation in Rwanda

Finally, Chapter 5 studies the effect of the diffusion of mobile internet on Rwandan regional labour markets and, in particular, on the creation of jobs and the trajectory of structural change, over the period 2002-2019. This chapter analyses whether mobile internet has had a labour-augmenting or labour-saving impact, and whether these effects are skill-biased and favour employment creation in specific industries. The findings from the analysis indicate that the direction of structural change is associated to the diffusion of fast mobile internet and favours the service sector and low-productivity activities, such as trade, which account for most new jobs. However, more productive services, such as finance, have also expanded greatly. Rwanda's regional labour markets show that total employment has grown significantly with the diffusion of fast mobile internet. The data show, also, that skilled occupations have grown faster than unskilled jobs, suggesting labour-market inequalities in the short term. The positive effect of mobile internet diffusion on employment creation is driven, on the supply side, by a higher intake of migrant workforce in districts with higher mobile internet coverage. In the case of the rapid growth of skilled occupations, it seems that this has been driven by higher intake of highly educated labour, both local and from other districts and provinces. On the demand side, the results indicate that more productive firms are concentrated in areas with higher levels of mobile internet coverage. Taken together with the finding that employment has increased in Rwandan districts, this suggests that firms are exploiting access to mobile internet and its diffusion to improve performance and create employment.

1.3.5 Original contribution to knowledge

The four empirical chapters that compose this doctoral thesis aim to contribute to the extant literature on structural change in several different ways. First, the focus on the role of changes in value added and export shares in driving economic growth has, more often than not, excluded the contribution of the informal sector, which is at most treated as a residual of the formal economy. In this thesis, structural change is framed as the shift of employment shares (rather than value added or export shares) across industries, while also accounting for the informal sector, including both firms and workers, which are too often hidden from any considerations on economic development and structural change.

Second, the role played by the informal sector in structural transformation is analysed at a finely disaggregated level, paving the way to an exploration of the micro-level determinants of structural change. This is done by: comparing the employment composition of the informal, formal and export sectors; understanding the interactions that exist between informal and formal co-located industries; exploring the impacts of technological change on the performance, size and industrial composition of informal productive establishments. Moreover, the changes in employment shares and the employment composition of economies are analysed here through the lens of the capabilities held by informal and formal industries, using the Economic Fitness and Complexity approach, which offers a proxy for the knowledge intensity of economic activities. This is useful to identify trajectories of structural change that lead to industrial upgrading, defined beyond productivity levels.

Third, the employment-oriented view on structural change is complemented by the neo-Schumpeterian perspective on the role of technological change in changing the sectoral composition of both informal firms and regional labour markets, mostly absent from most empirical analyses on the structural transformation of SSA economies. Moreover, integrating the focus on employment and informality with the role played by technological diffusion in structural transformations, this doctoral thesis aims to shift the focus from growth-oriented structural change towards considerations on its distributional outcomes. Aided by the Inclusive Structural Change framework – defined here as a mutually reinforcing relationship between innovation and structural change that also produces inclusive outcomes, such as net employment creation – the thesis aims to assess whether the impact of the diffusion of a General Purpose Technology (fast mobile internet) fuels a mechanism that, at the same time, leads to structural change at the micro and regional level, while also producing inclusive outcomes by increasing employment levels (skilled and unskilled) and entrepreneurial opportunities.

Chapter 2

Structural Change(s) in Ghana: a Comparison between the Trade, Formal, and Household Sectors

2.1 Introduction

Since the mid-1990s, the Sub-Saharan Africa (SSA) countries have experienced substantial rates of GDP growth. However, their economic growth has often been considered jobless growth that lacks the depth that characterises structural change-driven economic growth (African Center for Economic Transformation, 2014). Several international development organisations (African Center for Economic Transformation, 2014, 2017; International Fund for Agricultural Development, 2019) and academics (Arndt et al., 2016a; Kanbur et al., 2019; Alcorta et al., 2021) have proposed different approaches and policies to promote structural transformation and long-term inclusive growth in African economies. These approaches emphasise the crucial role of increasing agricultural productivity, encouraging economic diversification and fostering employment creation. However, these objectives have to take into account that the process of structural change in African countries differs from that in other world areas (Bah, 2011) and that their pathways will depend on historical and current national and global economic conditions.

African countries somehow skipped the industrialisation stages of structural change – or prematurely de-industrialised (Rodrik, 2016) – moving directly from agriculture to non-tradable services (McMillan et al., 2017a). Also, their industry baskets are characterised by specialisation in unsophisticated products, limited economic diversification, and low productive capabilities (Bhorat et al., 2020). Nevertheless, recent evidence shows that the manufacturing sector has expanded in the last decade, contributing positively to structural change; moreover, productivity of formal manufacturing firms has grown, indicating an advancement towards the technological frontier (Kruse et al., 2021; McMillan and Zeufack, 2022).

This chapter examines structural change in one African country – Ghana – focusing on the transformation of its export, formal and informal activities to distinguish the underlying drivers of economic performance in the SSA region. This focus on Ghana is because although this country has grown very rapidly, this has not led to a pattern of structural change based, for instance, on acquisition of technological capabilities. The International Monetary Fund (International Monetary Fund, 2019) considers that Ghanaian GDP growth has outperformed that of most countries in the region. In 2017, its exports accounted for 35 per cent of GDP (World Bank, 2022), well above the average of the SSA (27.3%) and OECD countries (27.8%). While Ghana's GDP growth and exports suggest positive economic performance, this country faces several difficulties related to its productive and employment structure (Osei and Jedwab, 2017). First, the growth acceleration[1] experienced by Ghana between 2008 and 2016 was driven mainly by exports of commodities that experienced price rises,[2] especially mining sector goods (African Development Bank, 2019, p. 49),

[1]The International Monetary Fund defines a growth acceleration as an "eight years [period] with average annual growth in GDP per capita of at least 3.5 per cent and a growth rate at least 2 percentage points higher than in the previous eight years. To rule out episodes of economic recovery, real GDP must also be higher in the last year of the acceleration period than in the years preceding it" (African Development Bank, 2019). The period between 2008 and 2016 is identified as a growth acceleration on these grounds.

[2]It has been argued that the rapid industrialisation of China since the early 2000s triggered the latest "commodity price supercycle" (Erten and Ocampo, 2013; Hume and Terazono, 2021), i.e., a steady increase in commodity prices above their long-term trends, due to large shocks that have

combined with an expansion of services. Also, the primary sector accounts for only a small share of labour in Ghana and its exports are extremely concentrated, with 27 firms responsible for 62 per cent of Ghana's total exports (Sutton and Kpentey, 2012). In the absence of links with other national industries and firms (Hidalgo et al., 2007; Hausmann et al., 2008; Savona, 2021), specialisation in the export of commodities may not be beneficial for Ghana's structural change (Sachs and Warner, 2001). Second, 28.6 per cent of Ghana's GDP is generated by the informal sector, which is concentrated on low productivity activities. The informal sector accounts for 87 per cent of commercial establishments in Ghana (Ghana Statistical Service, 2015) and most informal private economic establishments are run by households. Thirdly, aggregate (formal and informal) agricultural employment is moving towards services, leaving the share of manufacturing unchanged since the 1960s (World Bank, 2022) and incapable of attracting foreign direct investments (Sutton and Kpentey, 2012). This has been the consequence not of a rural push (increased agricultural productivity), but of an urban pull (Osei and Jedwab, 2017), including the creation of jobs in urban services that has resulted in "urbanisation without industrialisation" (Gollin et al., 2016). Also, household surveys (such as the Ghana Living Standards Survey) indicate that agriculture is still the predominant source of livelihood for Ghanaian households.

If the lion's share of exported goods is made up by commodities that are poorly linked to the rest of the economy, and if employment has been moving mainly to non-tradable services, the major export sector is not leading to a structural transformation that entails an employment shift from primary to more productive activities such as manufacturing. This situation is exacerbated by the predominance of informal activities in the household sector. In this context, it is important to identify the determinants of these patterns of structural change in Ghana and what might represent a specialisation trap in the long run. The present study focuses on the main export sector, the predominant household informal sector, and the small

increased demand and to which supply has been slow to respond.

formal sector, and examines their respective trends and differences. This will provide a grounded empirical context that will give content and dimension to the nature of structural change in one of the most representative countries in SSA, with the aim of offering development policy implications.

The analysis compares the changes in the sectoral composition of Ghana's three sectors over a period of approximately 10 years, starting in 2003. The trade sector is measured by exports; the formal sector is measured by aggregating employment data from two firm censuses, the National Industrial Census (2003) and the Integrated Business Establishment Survey (2014); household informal activities are measured using household employment data from the Ghana Living Standards Survey.

In addition to studying changes in the relative shares of each industry, this study assesses the quality of these changes, using the Industry Complexity Index (ICI) – a measure of industry sophistication. Complexity metrics have been used widely to proxy for country capabilities to diversify and upgrade their productive structure towards more sophisticated and higher value added industries (Hidalgo and Hausmann, 2009; Tacchella et al., 2012). It has been argued also that actively pursuing increased complexity of industrial specialisation is an effective strategy for governments to achieve economic growth, also in Africa (Lin et al., 2020). Traditionally, the ICI builds on national export specialisation: industries or products that are less ubiquitous (i.e., that require 'rare' capabilities and, therefore, are produced by a small set of countries) and are produced by highly diversified countries (i.e., with diverse stocks of capabilities) are considered more complex. Given the importance of the employment dimension in the analysis of Ghana's structural change, industrial complexity is defined here on the basis of district employment specialisation patterns, which is a variation of the Fitness method proposed by Tacchella et al. (2012), that relies on export specialisation patterns. This employment-based measure of industrial sophistication is combined with information on sectoral changes (in terms of exported value for the trade sector, and employment in the formal and informal sectors), in order to assess whether Ghana's employment and productive structures are moving towards more sophisticated and capabilities-intensive industries.

This chapter contributes to the literature on the characteristics of structural change in African countries and the extent to which this might lead to long-term and sustainable economic development. Most of these studies have successfully examined the 'quality' of structural change in African economies, in terms of productivity (McMillan et al., 2014; de Vries et al., 2015; McMillan et al., 2017b; McMillan and Zeufack, 2022). A focus on industry sophistication (ICI), rather than productivity, makes it possible to measure the capability intensity of industries at a finer level of industry aggregation (4-digit ISIC) and to capture the estimated capabilities within each industry, which complements the existing stream of studies. A focus on the capabilities required by a particular sectoral specialisation, rather than on its productivity, helps to unpack the 'quality' of structural change, in terms not only of its economic performance (productivity) but also of the inputs (capabilities) required to achieve it. In addition, focusing on capabilities rather than productivity provides a firm foundation from which to provide recommendations for industrial policy. Targeting an upgrade in capabilities to favour a sustainable structural change is different than simply focusing on productivity increases, and involves a different range of tools.

The results show that, between 2003-2013, industry specialisation in Ghana changed in the trade and formal sectors, but not in the household sector. Ghana increased its specialisation mainly in the extraction of natural resources, such as oil, which despite their high complexity are not accompanied by significant job creation. In fact, employment creation in downstream oil processing sectors reduced over that period.

There are significant differences, also, between the changes observed in exporting and those observed in total domestic employment. While exports of manufactured products have increased, employment in manufacturing has shrunk in both the formal sector (with respect to mining) and the informal sector (with respect to all industries).

However, a part of formal manufacturing employment has relocated towards more complex manufacturing industries, setting the country's manufacturing sector on a promising trajectory of industrial upgrading. Alongside these developments, growth in exports of highly complex Information and Communication Technology (ICT) services seems to have been supported by a relevant share of employment in the formal sector, but not the informal sector, which concentrated largely in subsistence agricultural activities and has become less complex over time. We conclude that the creation of capabilities in the informal sector will be crucial to support the upgrading of Ghana's productive structure.

The remainder of the chapter is structured as follows. Section 2.2 reviews the structural change literature, focusing in particular on empirical works on structural change in Africa; it highlights the relevance of the present chapter's empirical contribution. Section 2.3 outlines the empirical strategy and the methodology used for the analysis. The results of the analysis are presented in Section 2.4. Section 2.5 concludes.

2.2 Literature review

2.2.1 Structural change and complexity

The concept of structural transformation dates back to the work of early structuralist economists, who studied the transformation of European economies after the Second World War. In the pioneering works on economic development, published in the 1950s, economic development is described as the transition of activities from traditional low-productivity sectors to modern productive sectors, with the traditional sectors supplying "unlimited labour" (Lewis, 1954) to the modern sectors, and favouring higher productivity gains. Chenery's (1979) pioneering empirical research confirmed that structural transformations were characterised by a decrease in the share of labour employed in subsistence agriculture and a shift of labour to novel and more productive sectors.

Observation of the industrialisation process in contemporary high-income countries reveals the pivotal role of productivity gains for consolidating urban labour markets and spurring diversification (Kuznets, 1966). In particular, the role of diversification emphasises the importance of the emergence of new industries through the creation of backward and forward linkages (Hirschman, 1958): new economic activities (especially manufacturing) represent a new source of intermediate and final demand, which allows the creation of new productive activities in upstream and downstream industries. From a the structuralist perspective, industrialisation and manufacturing, via capital accumulation, are considered crucial for transforming economies, absorbing unproductive labour (Lewis, 1954; Rostow, 1959) and acting as an "engine of growth" due to the presence in the modern, capital intensive sector characterised by increasing returns. Kaldor's first two empirical laws (1966) formally test the engine of growth hypothesis, showing that output growth in manufacturing has a positive and more than proportional effect on economic growth and productivity in manufacturing industries. Such positive feedback was earlier defined by Myrdal (1957) as "cumulative causation".

As the shares of employment and value added in manufacturing in high-income countries have shrunk, scholars have turned to analysing the role of manufacturing in creating opportunities for high value added services. Manufacturing activities lead to the creation of highly productive services, speeding up the process of technological accumulation (Su and Yao, 2017). The evidence in Rodrik (2013, 2018) provides further support for the importance of manufacturing activities for sustaining economic growth and shows that manufacturing productivity converges unconditionally across countries, meaning that productivity in manufacturing depends less on the so-called "growth fundamentals" (such as institutional quality and human capital) and more on the intrinsic dynamics of manufacturing productivity which lead to increasing returns and make it a powerful engine of growth – especially in low- and middle-income countries.

More recently and also from a different disciplinary perspective than the structuralist tradition, Hidalgo et al. (2007) provided a more granular view of the part played by different products in the process of economic development. The authors developed an analytical tool – the "product space" – which allowed them to identify which baskets of exported products are most conducive to economic development. They show that the most difficult to produce products are those that require high levels of knowledge and, therefore, are a rare component of the product basket. Also, the countries with Revealed Comparative Advantage (RCA) in these products are those that are more developed. Therefore, this framework, which is a-theoretical and completely data driven, assumes a correlation between a high level of knowledge intensity in the country, a high level of production capabilities and a high level of product sophistication. The more sophisticated the product, the more it is an indication of the complexity of the capabilities required to produce it and of the high knowledge intensity and level of development of the country that produces it. This framework focuses on product characteristics rather than the inputs needed to produce it, including employment. The authors use this information to measure the stock of capabilities embedded in the particular country's export basket, that is, the Product Complexity Index (PCI) (Hidalgo and Hausmann, 2009).

Their analysis uses trade data and starts from the assumption that national capability to export a specific product depends on national capacity to export the related product, since these activities require similar institutions, infrastructure, physical factors and technology. Countries tend to accumulate the capabilities to produce and export products that are 'similar' to those that they already export. The degree of similarity between product A and product B is based on evidence showing that countries that export product A are also more likely to export product B. By moving from less sophisticated products, which require lower capabilities, to more sophisticated products, which require higher capabilities, a country can transform its economic structure in such a way that it sustains growth. In this framework, the capabilities embedded in the country's productive structure are seen as a precondition for economic development and diversification, which involve the mobilisation of knowledge and productive capabilities across sectors (Hausmann et al., 2014). This conceptualisation of economic complexity paved the way to the development of several measures of economic complexity of products, industries and geographical areas.[3]

However, countries do not always and not easily move towards more complex products. There are several factors that can push countries to specialise in low productivity sectors; these include their natural resource endowments, an overvalued currency, a rigid labour market (McMillan et al., 2014) and the industrial specialisation and competitiveness of other countries (Baldwin, 2013). Achieving specialisation in complex products will be harder for countries specialised in the production of low-complexity goods that are not 'related' to higher complexity products (Hidalgo et al., 2007). Basically, the concepts of complexity and relatedness are closely coupled to the concept of high development linkages in the Hirschman framework referred to above. This means that (products) sectors that are unrelated (poorly linked) to other more productive sectors (complex products) might be detrimental to the upgrading of industry specialisation.

2.2.2 Structural change in Sub-Saharan Africa

Despite the high GDP growth experienced since the 2000s, many African countries have failed to see a substantial shift in employment from agriculture to manufacturing. Instead, their positive economic performance has been driven, most often, by a boom in commodity prices, although this has not created new jobs (Valensisi and Davis, 2011) and has driven investment away from non-resource intensive industries and hampered diversification (Harding and Venables, 2016). This is a phenomenon that has been described as the natural resources curse (Sachs and Warner, 2001).

Structural change in African countries tends to be characterised by peculiar trajectories. Most African countries have somehow skipped the industrialisation phase. Although a large share of employment and value added is concentrated in agriculture,

[3] For a broad review of complexity measures and algorithms, see Freire (2021).

most of the labour leaving agriculture has transitioned towards services (Bah, 2011; Szirmai, 2012; McMillan et al., 2017b). Moreover, the relative shares of employment and value added in manufacturing have reached a peak at low levels of GDP per-capita, compared to current high-income countries (Rodrik, 2016).

According to Rodrik (2016), one of the factors that has led to this "premature de-industrialisation" is the change in the relative prices of manufacturing compared to non-manufacturing goods, due to the shift of global manufacturing to low wage countries in South and East Asia (Haraguchi et al., 2017). This scenario is consistent with the model developed by Kaldor (1970), which shows that countries producing goods subject to decreasing returns (such as agricultural goods) that trade with countries with comparative advantage in the production of goods subject to increasing returns (such as manufactured goods), are likely to increase specialisation in the former types of good. In addition, labour costs in Africa have remained high (Newman et al., 2016) compared to those in other emerging economies (Benjamin and Mbaye, 2020), which has had a negative impact on competitiveness. As a result, high shares of labour in SSA countries have moved towards less-than-average-productivity services (McMillan et al., 2014) with negative or below-average productivity growth rates (de Vries et al., 2015; Diao et al., 2019b).

It has been argued that manufacturing is a necessary step in structural change, including in low- and middle-income countries (Haraguchi et al., 2017), based on the still valid hypothesis that manufacturing is an engine of growth (Rodrik, 2013). Nevertheless, the role of services in low-income countries' economic growth has increased (Szirmai, 2012; Owusu et al., 2021; Baccini et al., 2021) and, in some cases, has contributed to increasing the aggregate productivity of the economy, especially if related to manufacturing (Di Meglio et al., 2018). However, the experience of service-led development in African countries such as Rwanda, for instance, reveals that despite the high productivity of the service sector (10 times more than agriculture in Rwanda) and the movement of labour from agriculture to high productivity industries (Ggombe and Newfarmer, 2017), the lack of integration of modern services with the existing labour markets, and the failure to generate the necessary skills and technological linkages in modern services has led to disappointing results in terms of development (Behuria and Goodfellow, 2019).

2.2.3 Towards an analytical framework for the analysis of structural change

The previous two sections summarise two contrasting, but complementary views of how African economies have evolved. On the one hand, the structuralist literature examines changes in employment to describe the process of structural transformation in SSA, focusing on productivity as the main engine of economic transformation. In this view, the shift of labour towards the most productive activities is paramount to a country's developmental outcomes. On the other hand, the complexity literature highlights the role of capabilities, proxied by export specialisation, for determining the possible ways to transform the national economic structure. In a complexity framework, the knowledge embedded in a country's productive structure determines the path of further diversification towards related, more complex industries and, ultimately, its process of structural change.

The present study builds on and attempts to bridge between these two frameworks, putting the employment dimension of structural change at the centre of the analysis. Following the structuralist tradition, this chapter looks at changes in the relative importance of finely disaggregated industries in order to identify and explain the trajectory of structural change in Ghana. This approach is complemented by insights from the complexity literature, which points to the role of capabilities in creating the conditions for virtuous structural change – exemplified by an upgrading of the productive structure.

Based on the frameworks outlined above, this chapter offers multiple conceptual and methodological advancements. First, the structuralist tradition is complemented with the more recent findings from complexity literature, building upon the similarities (sectoral linkages and relatedness, technological capabilities and complexity) and differences that can be combined to achieve a more comprehensive assessment of the pathways of structural change in Ghana. The differences relate to the consideration of the production structure by the structuralists, and the export specialisation by the complexity economics literature. These two are necessarily interrelated yet taken into account separately in these two approaches, while here they are reconciled in a unified production/export view. Second, the pathways of structural change are identified and explained by emphasising the role of employment. Again, this is done acknowledging that capabilities are importantly embedded in employment and are in turn characterized by sectoral specificities. Third, the empirical analysis specifically includes the role of the informal sector, which, as highlighted in the introduction, predominates in Ghana and other SSA countries. The structuralist and the complexity literature most often disregard the role of informal activities in structural change. There are some exceptions: Lewis (1979) highlighted the importance of what he describes as the "in between sector", represented by the most productive and innovative informal firms, for leading the shift in specialisation from traditional to modern sectors. Similarly, Diao and McMillan (2018) highlight the importance of the expansion of the non-tradable sector for bringing the productivity gains required to trigger a virtuous process of structural change. It has been pointed out, also, that informal firms can be the agents of bottom-up industrialisation (Kraemer-Mbula and Monaco, 2020), given their considerable innovation activities (Fu, 2020) and contribution to economy-wide productivity growth (Diao et al., 2019a). Fourth, the above contributions are enabled by our employment of a novel empirical strategy and proposal of novel indicators of structural change.

The existing measures of complexity are constructed using data on traded products. It is argued here that exports alone may not be a good representation of a country's industrial composition, especially if there is a large share of labour employed in non-tradable activities, which, by definition, cannot be exported. The empirical analysis in the present chapter uses an alternative measure of complexity that accounts for national capabilities, building on the country's employment – rather than export – specialisation. While a few studies use the employment structure to measure the complexity (Sbardella et al., 2017) or the industrial specialisation of regions (Cicerone et al., 2020), there are no empirical works that account for the informal sector in their measurement of complexity in low- and middle-income countries. Providing information on whether informal labour is shifting towards sectors requiring higher/lower capabilities could shed light on the direction in which the less productive workforce is moving and whether structural change is virtuous/vicious.

The present chapter addresses these gaps in an analysis of ten years of structural change in Ghana – a rather successful country in SSA. The aim is to identify the type of structural change that occurred in Ghana in the time period analysed, that is, towards higher capabilities or not, and the economic sectors affected. To provide a clear picture of the changes that have occurred in Ghana's economic structure, the analysis considers three economic sectors, each of which matters, but in different ways. These sectors are:

- the trade sector, measured by exports, whose importance rests on the fact that export specialisation (although with some limitations which are discussed in the course of the analysis) can be considered a signal of the most advanced capabilities available in the country (Lin, 2012);
- the formal sector, measured by domestic formal firms, given their importance in the process of economic diversification via forward and backward linkages (Hirschman, 1958, 1977; Andreoni, 2019);
- the informal sector, whose importance was discussed earlier in this section.

Also, to reflect the importance of the employment dimension of structural change, the relative dimension of industries is defined in terms of employment shares for the formal and the informal sector. Due to the unavailability of employment restricted to exporting industries, industry shares in the trade sector are measured using their relative exported value. Combining information on sector changes in relative employment shares, with information on industry complexity, provides an understanding of whether the patterns of Ghana's structural transformation have benefited its workforce in terms of the capabilities available in the country.

This approach has some limitations. First, the first year of data (2003) for the formal sector includes only information on manufacturing and mining firms, which limits examination of the changes that have occurred in the formal sector to these two industries. Second, not all industries are observed in all of the three sectors. Since the industrial complexity measure is built using the Integrated Business Establishment Census (IBES) (see Section 2.3.1.2), only the industries included in this dataset are used to measure the Industry Complexity Index. This limitation is discussed further in Sections 2.3.2 and 2.4.2. Third, as mentioned earlier, while the formal and informal sectors are measured in terms of employment shares, the trade sector is measured in terms of relative shares of export value. Although this does not allow to describe the trends in employment generated by exported activities, measuring the change in the composition of the trade sector is a useful reference to identify whether the changes in Ghana's export specialisation are reflected by changes in the formal and informal employment structures.

2.3 Empirical strategy

2.3.1 Data

The analysis exploits four main sources of data, which result in unique information on the trade, formal and informal sectors of the Ghanaian economy. Measurement of the relative size and composition of exports relies on data from the 2003 and 2013 UN Comtrade Database; measurement of the formal sector relies on the 2003 Ghana National Industry Census (NIC) and the 2014 IBES (2014); and measurement of household informal activities relies on the 2005 and 2013 World Bank Ghana Living Standard Survey (GLSS).

The dataset derived from aggregation and harmonisation of these data sources considers the industry as the unit of observation. Industries are defined by the 4-digit International Standard of Industry Classification, rev. 3.1 (ISIC). The relative size of each industry is computed for each of the three sectors between 2003 and 2013/2014.[4] For the export sector, relative size is given by the share of each industry's exports over total exports; for the other two sectors (formal firms and informal household activities) relative size is measured as the share of employment in each industry over total employment.

2.3.1.1 Export sector: UN Comtrade

The UN Comtrade dataset includes fairly detailed information on exported products, classified according to the Harmonised System (HS). The original Comtrade data were processed following the Bustos-Yildrim method[5] (Hausmann et al., 2014), which compares exported and imported values between partners, in order to account for the additional costs of freight and insurance, and controls for the reliability of the importing and exporting countries. Information is available for the period 1995 to 2020; the data were filtered for years 2003 and 2013 to allow comparability with data on formal firms' and household informal activities. At the 4-digit HS classification level, the dataset includes 1,248 products in 2003 and 1,246 products in 2013. The product-level data were aggregated by summing the export values of HS products belonging to the same industry, according to the 4-digit level ISIC (revision 3.1), to allow comparability with the other two sectors, where the information is also at the 4-digit ISIC level.[6] Products belonging to more than one industry were counted fractionally for each industry involved. This resulted in an industry level dataset, where each observation is a 4-digit ISIC 3.1 industry, and which provides information

---

[4] The year of observation differs depending on the data source: starting in 2005 instead of 2003, for household informal activities; ending in 2014 rather than 2013 for formal firms.

[5] For a detailed description of the Bustos-Yildrim cleaning procedure applied to import-export data, see the Atlas of Complexity: http://atlas.cid.harvard.edu/data.

[6] Correspondence between HS products and ISIC industries was enabled by the `concordance` package in R (Liao et al., 2020), which allows conversion across HS, Standard International Trade Classification (SITC) and ISIC classification systems.

on relative size over total exports in 2003 and 2013.

2.3.1.2 Formal firms: NIC 2003 and IBES 2014

The Ghana NIC 2003 and 2014 IBES data cover the formal domestic productive establishments. Both sets of data were collected in the respective years by the Ghana Statistical Office, and they represent censuses of economic establishments in Ghana. The NIC 2003 covers all non-household establishments – all units of production whose physical location is fixed and can be described and traced – engaged, primarily, in mining and manufacturing, for a total of 3,519 firms. IBES 2014 covers all non-household establishments from all industries, for a total of 638,743 firms. The large difference in the number of observations between the 2003 and 2014 censuses is due to the fact that services (and all sectors other than manufacturing and mining) were included only in 2014. This is a shortcoming since all considerations about the formal productive structure must necessarily be limited to the manufacturing and mining sectors. Formal firms are identified based on the Ghana Statistical Service definition (Ghana Statistical Service, 2015), that is, that they are registered at the Ghana Registrar-General Department (RGD) and maintain formal accounts. This applies to 81,262 firms (12.73%), employing 54.45 per cent of the total labour force. While some of these firms may be involved in exporting and, therefore, overlap with the export sector, the census deliberately excludes all economic establishments based within households and, so, excludes household informal activities.

2.3.1.3 Household informal activities: the Ghana Living Standard Survey (GLSS)

The third source of data is a nationally representative sample of Ghanaian households, which provides information on both farming and non-farming activities carried out by households. The data are collected at the household level and re-sampled in 2005, 2013 and 2017 (with only the first two waves used for the present study). The non-farming enterprise module of the survey provides detailed information on many aspects of household firms and, particularly, their industry (according to ISIC rev. 3.1 for 2005 and ISIC rev. 4 for 2013) and employment size. An industry level dataset was constructed using GLSS 2005 and 2013, following the 4-digit ISIC 3.1 classification and aggregating information on household enterprises (which include non-farming enterprises and agricultural plots) for each industry, along with information on agricultural plots and employed labour. Total employment in each ISIC industry is measured as the total number of people employed in household activities belonging to a given ISIC industry.

2.3.1.4 Additional data

The analysis also uses some additional data to describe long term structural change in Ghana. The World Bank World Development Indicators (WDI) provide information on Ghana's GDP between 1960 and 2010. These data are complemented by the Groningen Growth and Development Centre African Sector Database, from which information on aggregate employment and value added relative shares (agriculture, mining, manufacturing and services) over the same time period is extracted. Finally, to validate the complexity and fitness framework, the analysis in Appendix A.3 uses VIIRS time series data (Elvidge et al., 2017) on Nighttime Light imagery, to measure economic development at the subnational level.

2.3.2 Methodology

The four data sources described above were combined to provide a unique dataset of 295 observations, corresponding to the 4-digit ISIC Rev. 3.1 industries observed for the three sectors of exporting, formal firms and household informal activities. The complete, harmonised data are presented in Appendix A.4. This new dataset serves as a comprehensive source of quantitative information allowing comparison of the patterns of structural change in Ghana, in the three sectors considered, over a period of around 10 years (2003-2013/14), based on industry shares. For each industry, the dataset includes information on: i) export shares (2003 and 2013); ii) formal sector shares (2003 and 2014); iii) household sector shares (2005 and 2013); iv) industrial complexity.

The direction of the industrial transformations in the three sectors considered relies on a measure of the 'sophistication' of each industry in each of the three sectors: that is, the Industrial Complexity Index (ICI). The ICI is a synthetic measure of the level of productive capabilities likely to be required for a country/region to specialise in a given industry (Hidalgo and Hausmann, 2009). That is, to become competitive in a more complex industry, the country/region would need first to become competitive in a less complex industry. Therefore, an increase in the complexity of the industry mix in which the country/region specialises, suggests a pattern of virtuous structural change.

Since our analysis focuses on a single country, it uses districts (Admin-2 geographical units in Ghana) and industries, to compute the ICI. In the formulation proposed by Tacchella et al. (2012), two elements contribute to industrial complexity. First, non-ubiquity of industries, that is, only a few countries/regions are specialised in the particular industry. Thus, non-ubiquitousness is used to proxy for the "rareness" of the capabilities required to specialise in a given industry. Second, high level of regional (or national) diversification in a large number of industries, where diversification signals presence of a wide range of capabilities. Therefore, the "fitness" of an entire country/region can be measured by aggregating information on its diversification and the complexity of the industries in which it is specialised. Information on industry ubiquitousness and country/region diversification can be combined iteratively to obtain measures of industrial complexity and regional fitness.

There are several existing measures of industry and country complexity and fitness such as the Product Complexity Index[7] based on Hidalgo and Hausmann (2009) or the fitness,[8] based on Tacchella et al. (2012). However, in the context of the present analysis, these measures present a number of shortcomings, namely that they are

[7] Available at: https://oec.world/en/blog/post/2020-trends-in-economic-complexity.
[8] Available upon request: http://www.economic-fitness.com/.

only based on product-based export specialisation. For this reason, the measure of industrial complexity used in the analysis – the Industry Complexity Index – builds on employment-based regional specialisation. The analysis employs the fitness algorithm proposed by Tacchella et al. (2012), but using an employment- rather than export-based specialisation matrix, following an approach similar to Sbardella et al. (2017). We computed a static ICI, using the rich data provided by the IBES for all 4-digit ISIC industries (249) in 2014 – that is, based on one year of data, since the 2003 census includes only the manufacturing and mining industries. The computation uses information related to all 638,743 (formal and informal) Ghanaian productive establishments; the relative specialisation of a district is better computed if all activities are included in order to capture a complete snapshot of the capabilities available in the Ghanaian workforce. The analysis uses the ICI built for 2014, over the overall industry space, to measure changes in the complexity of exports, formal firms and informal household activities, between 2003 and 2014.

The ICI is computed by aggregating employment in productive establishments by district (216), to obtain a cross-tabulation of employment by districts and industries. The first step involved in constructing the ICI consists of building a district-industry matrix, to identify district specialisation across industries. Thus, district d is specialised in industry i if its employment share in that industry is higher than the national average employment share in the same industry. The district-industry specialisation matrix M_{di} is defined as follows; the matrix cells take the value 1 if the district is specialised in an industry, and 0 otherwise:

$$M_{di} = \begin{cases} 1 & \text{if} \quad \frac{x_{d,i}}{\sum_i x_{d,i}} \Big/ \frac{\sum_d x_{d,i}}{\sum_{d,i} x_{d,i}} \geq 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

where $x_{d,i}$ is employment in district d and industry i. The first term (numerator) of the first equation corresponds to the share of employment of industry i in district d, while the second term (denominator) quantifies the average share of employment in industry

i across districts. If the former is larger than the latter, district d is specialised in industry i. The M_{di} matrix provides information on district diversification (i.e., the number of industries in which a district is specialised, given by the row sums of the specialisation matrix) and on the industry ubiquity (i.e. the number of districts that specialise in a given industry, given by the column sums).

This information is used to compute, iteratively, a measure of district fitness (F_d), which depends on the district industry complexity (Q_i), i.e. the ICI[9] which in turn depends on the district fitness, F_d. F_d is proportional to district d's diversification weighted by the complexity of the industries in which the district is specialised. Q_i is inversely proportional to the ubiquity of the industries in which districts are specialised, but is positively correlated with their fitness.

$$\begin{cases} \tilde{F}_d^{(n)} = \sum_i M_{di} Q_d^{(n-1)} \\ \\ \tilde{Q}_i^{(n)} = \dfrac{1}{\sum_d M_{di} \dfrac{1}{F_d^{(n-1)}}} \end{cases} \quad \text{with} \quad F_d^{(0)} = 1 \ \forall d \ \text{ and } \ Q_i^{(0)} = 1 \ \forall i \tag{2.2}$$

The tilde indicates that the two measures are combined iteratively until stability. After the first iteration ($n = 1$), the fitness F of district d corresponds simply to its diversification, while the complexity Q of industry i is equal to the inverse of its ubiquity. At each iteration, the measures are weighted progressively, using the complexity of the industries in which the district d is specialised to refine district fitness and using the fitness of the districts specialised in industry i to weight complexity. After 20 iterations of the algorithm, both measures reach stability in our data.[10] It was not possible to compute the ICI for all the industries in the trade sector. According to the trade data, Ghana exports products in industries that do

[9]The complexity of industry i is referred to as Q_i, mainly in the mathematical notations, and as ICI_i – Industry Complexity Index – ICI – in the rest of the analysis. Fitness (F_d refers to the complexity of geographical areas (in this case, the districts).

[10]The algorithm was implemented using the R package `economiccomplexity`, available via R's CRAN repository (Vargas et al., 2020).

not include any domestic firms. This issue will be discussed later in the results Section (2.4). In the case of tradable services (such as ICT, finance, transport, and travels and tourism) the Harmonised System classification does not have a direct correspondence in terms of ISIC 4-digit industries. Therefore, it has been operated a correspondence between HS services and ISIC 4-digit industries, which is shown in Table A.1 in Appendix A.2.[11] The complexity of those services was then computed as the average complexity of industries related to them.

To assess the validity of the employment-based measure of industry complexity (ICI) empirically, Appendix A.3 reports the results of an analysis of the correlation between district fitness and economic development, measured using average nighttime luminosity. The results show a strong correlation between district average ICI and district economic development. This is in line with a growing stream of work showing that economic complexity is a good predictor of economic development (Cristelli et al., 2017; Tacchella et al., 2018).

To summarise, the final industry-level dataset, which includes the start and end years for the three sectors, contains: i) the relative share (exported value or employment) of each industry; and ii) the ICI of each industry. The former is used to study and compare structural change in the three sectors: that is, the percentage change in the share of each industry between 2003-2013. The latter is used to study and compare the direction of structural change in the three sectors: that is, whether the variation in industry shares has led to higher complexity of the industrial composition. We propose a weighted measure of change in relative size within the major ISIC groups j:[12]

[11] The sectoral re-aggregation follows the special grouping guidance provided by the International Standard Industrial Classification of the United Nations (United Nations Statistical Division, 2002).

[12] Major industry groups are defined as the lowest industrial level of disaggregation provided by ISIC 3.1 (1-digit level). They are: A- Agriculture, hunting and forestry; B- Fishing; C- Mining and quarrying; D- Manufacturing; E- Electricity, gas and water supply; F- Construction; G- Wholesale and retail trade; repair of motor vehicles, motorcycles and personal and household goods; H- Hotels and restaurants; I- Transport, storage and communications; J- Financial intermediation; K- Real estate, renting and business activities; L- Public administration and defence; compulsory social security; M- Education; N- Health and social work; O- Other community, social and personal service activities; P- Activities of private households as employers and undifferentiated production activities of private households; Q- Extraterritorial organizations and bodies.

$$\Delta X_{jk}^{ICI} = \sum_{i \in J} \left(\left(X_{ijk_{t+1}} - X_{ijk_t} \right) \times ICI_{ij}^{std} \right) \tag{2.3}$$

where ΔX_{jk}^{ICI} is the complexity-weighted size change in the major group j in the sector k; $(X_{ijk_{t+1}} - X_{ijk_t)}$ is the variation in the share of industry i in the major ISIC group j and sector k over the time period considered; and ICI_{ij}^{std} is the ICI of industry i, standardised between 0 and 1. The complexity-weighted change in the industry share indicates whether the relative changes within each sector k and each industry group j contribute positively ($\Delta X_{jk}^{ICI} > 0$) or negatively ($\Delta X_{jk}^{ICI} < 0$) to the aggregate complexity of each sector. A numerical example could be of aid to illustrate how this measure is constructed. Let's imagine that two industries h (with $ICI_i^{std} = 0.5$) and i (with $ICI_i^{std} = 1$) belonging to the same ISIC major group j (say, manufacturing), in sector k (say, exports). The change in exports shares between 2003 and 2013 for both industries is + 0.1. The complexity-weighted change for ISIC major group j (manufacturing) in sector k (exports) is equal to $(0.1 \times 0.5) + (0.1 \times 1) = 0.15$. The same is done for all ISIC major groups in each of the three sectors (exports, formal and informal).

2.4 Analysis and results

This section begins with an overview of aggregate historical trends in structural change in Ghana and their relevance for each of the three sectors. It proceeds by describing the changes in productive structure between 2003-2013 for the trade sector, between 2003-2014 for the formal sector and between 2005-2013 for the household sector. These changes are qualified in terms of shifts in the complexity of the industry composition in each sector. As argued in Section 2.3.2, the employment-based ICI captures the intensity and diversity of the capabilities required to specialise in a given industry. Change towards a relatively larger share of more complex industries (in each of the three sectors) are interpreted as an upgrading of the productive structure

– that is, a virtuous structural change.

2.4.1 Structural change in Ghana

Since the 1980s, Ghana has recorded a steady annual growth rate of around 5 per cent of the GDP, peaking at 14 per cent in 2011 and then slowing to reach 0 growth at the beginning of 2020 as a result of the Covid-19 pandemic (World Bank, 2022). Exports also increased more or less steadily up to the early 2000s and accounted for 35 per cent of the total GDP in 2017 (African Development Bank, 2019). The composition of exported products reflects Ghana's dependence on commodities: in 2017, gold represented 31 per cent of total exports, crude petroleum 12 per cent; cocoa beans, butter, and paste 27 per cent (The Growth Lab at Harvard University, 2022).

Figures for shares of employment and value-added in manufacturing are consistent with the premature de-industrialisation narrative (Rodrik, 2016). Compared with a sample of seven high-income countries (Japan, South Korea, Italy, Spain, France, UK and US), the share of manufacturing value added and employment in Ghana peaked at lower levels of GDP per capita and never reached the levels achieved by these other countries (Figure 2.1). Figure 2.2 shows that the decline in the shares of employment and value added in agriculture was accompanied by an increase in services and, more recently, an increase in mining since 2008, following the discovery of oil in Ghana. In terms of employment, manufacturing shares have remained at low levels since 1960 and in terms of value added, have even decreased.

Detailing the causes of premature de-industrialisation in Ghana is beyond the scope of the present chapter (although in the succeeding chapters, we discuss some of these causes). Here, the objective is to demonstrate that understanding the process of structural change in emerging countries such as Ghana, requires a more granular view to identify those sectors that present the most favourable opportunities for diversification. The aggregate figures presented above provide a general description

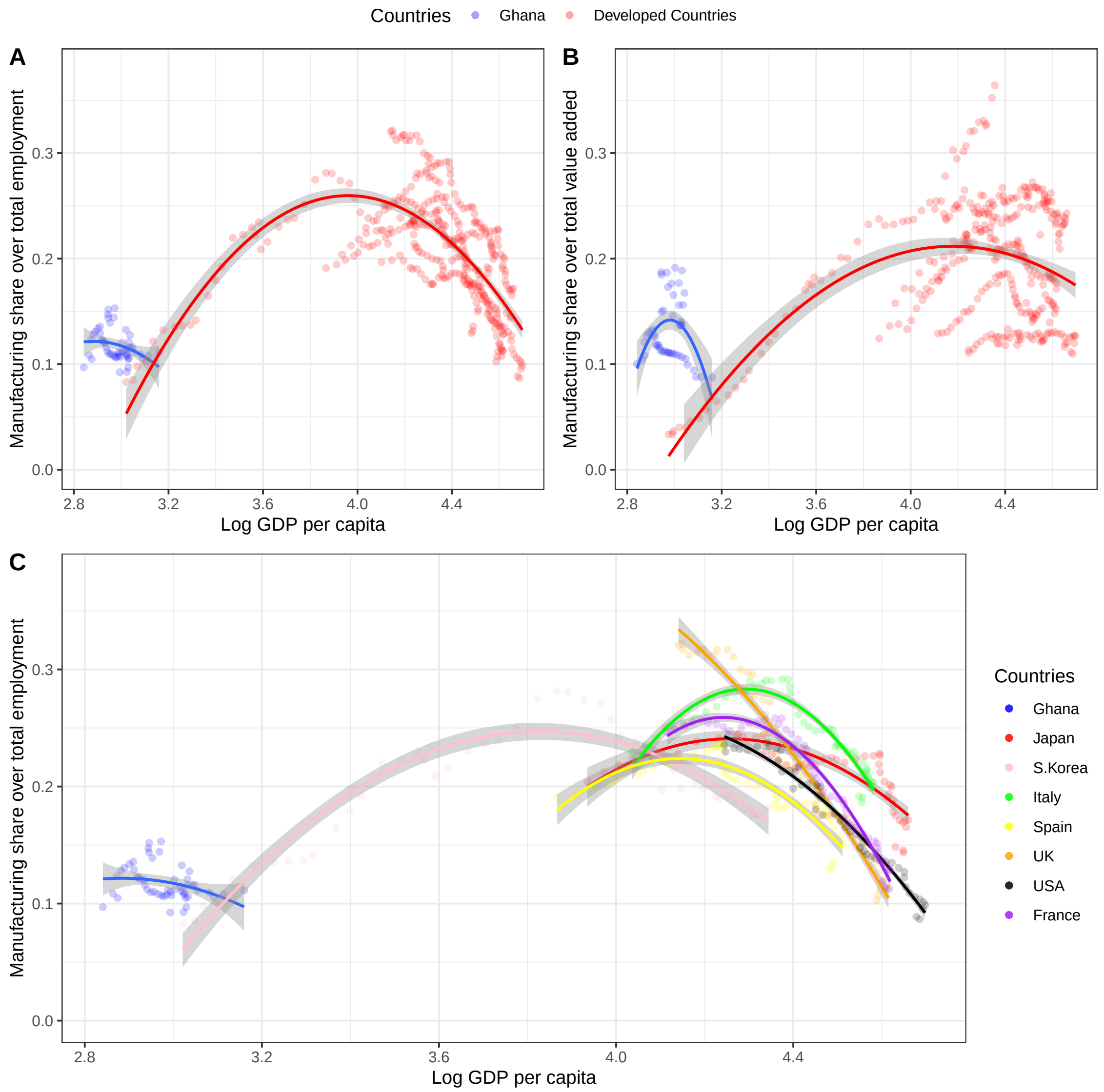

Figure 2.1: This graph plots GDP against the share of manufacturing in total employment (panel A) and total value added (panel B), comparing Ghana with a sample of 7 high-income countries (Japan, Korea, Italy, Spain, France, UK and US). The figure is based on data on GDP per capita, and manufacturing employment and value added between 1960 and 2013. Panels A and B compare Ghana with the average high-income country in the sample; panel C compares Ghana with the 7 countries individually. Fitted lines represent the quadratic fit between the two variables on the X and Y axes. Sources: African Sector Database, Groningen Growth and Development Centre (manufacturing employment and value added); WDI (GDP per capita).

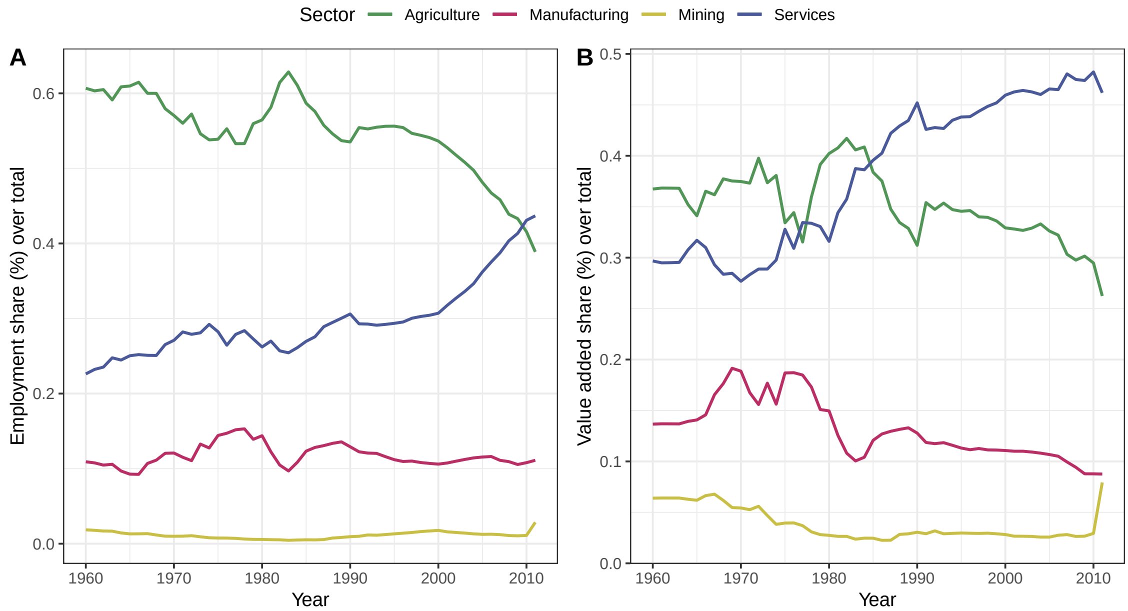

Figure 2.2: Employment and value added trends in agriculture, manufacturing, mining and services in Ghana, 1960-2011. Source: WDI.

of the composition of employment and value-added in formal activities but may not apply to all sectors of Ghana's economy.

The reason why formal production cannot be used, on its own, to measure Ghana's economic structure, is that the weight of the informal sector is high. The most recent data from the Ghana Statistical Service suggest that the informal sector (defined as all businesses not registered at the national Registrar-General's Department) accounts for 28.6 per cent of GDP (Ghana Statistical Service, 2015), which is comparable to the weight of exports 35% according to the World Bank 2022). In terms of employment, the informal sector accounts for 46.45 per cent of the active workforce employed in non-household establishments (Ghana Statistical Service, 2015). However, note that these figures should be considered a lower-bound estimation since they exclude household farming and non-farming enterprises, which are not included in the IBES survey (see Section 2.3.1.2).

The existence of a large informal sector highlights the limitations of Figures 2.1 and 2.2 to depict structural change. These figures show only the evolution in the composition of formal activities. Also, the contribution of non-farming household enterprises has increased over time in low- and middle-income countries (Haggblade et al., 2010; Nagler and Naudé, 2017), although they have been moving to services rather than manufacturing (Jayne et al., 2018) and their creation is often motivated by geographical factors, such as proximity to a city (Christiaensen et al., 2013; Djido and Shiferaw, 2018).

To account more fully for industry composition, in the following subsections we compare three distinct, but interrelated economic sectors: i) the trade sector, measured by export activities, which correspond to more internationally competitive activities and, therefore, industries holding a comparative advantage; ii) formal domestic activities, defined here as all activities carried out in formal establishments, which account for activities aimed at the domestic market; and iii) household enterprises, that is, activities carried out by households – mainly informal, both on- and off-farm. The next section describes the results for each of the three sectors.

2.4.2 Export sector

To facilitate the comparison of export shares between 2003 and 2013, Figure 2.3A compares the share of export value in 2003 (X axis) with the share of export value for the same industry in 2013 (Y axis) (see the figure notes for a more detailed explanation). Figure 2.3 shows that, while most activities account for less than 5 per cent of the exports in both years, Ghana's export composition has changed substantially towards the extractive industries. We observe a large increase of 20 percentage points (pp) in the share of exports of unrefined petroleum and natural gas, which jumped from 0 per cent in 2003 to 20 per cent of total exports in 2013, and a noticeable growth of 7 percentage points in the share of manufactured basic precious metals (from 18% to more than 25%). Among the larger export sectors, ICT shares increased from 3 per cent to 5 per cent, while the share of fruit, nuts, beverage and spice crops in total exports, decreased from 27 per cent to 13 per cent, and travel and tourism decreased from 14 per cent to 5 per cent. This would seem to confirm a picture of an economy whose reliance on agriculture is decreasing and is shifting to another primary industry that relies on export of extractive natural resources.

Figures 2.3A and 2.3B show that some of the growing manufacturing activities seem to be related to the processing of natural commodities, hinting at an opportunity for functional upgrading in their respective value chains. One such example is the gold value chain, given the increase in the export of manufactured basic precious metals (Figure 2.3A). Another is the cocoa value chain, which experienced a decrease in the export of the raw commodity (Growing of fruits, nuts, beverage and spice crops sector, Figure 2.3A), coupled with an increase in manufactured products (Manufacture of cocoa, chocolate and sugar confectionery, Figure 2.3B). However, the manufacture of refined petroleum products shrank by almost a half over the same period, indicating that export-oriented firms in Ghana have specialised in the extraction rather than the processing of oil.

Overall, the relative share of traditional manufacturing activities in exports has been declining. While in Africa as a whole, current trends show that the decrease in the relative share of agriculture has been coupled with the increased relevance of services, this does not apply to the Ghanaian export sector. Although the relative weight of ICT increased between 2003 and 2013, this has been more than offset by the reduced shares of other tradable services (travel and tourism). Overall, the share of the service sector in exports has decreased in Ghana.

To investigate whether these changes in the structure of exports implies a virtuous or a vicious structural change, the shift in the export industries experiencing increases or decreases over the 10 years analysed, is combined with the ICI. Recall, that our measure of industrial complexity is computed using information from the 2014 census of all firms in Ghana: this means that if an industry does not appear in the census, the ICI cannot be computed for that industry. In Figure 2.4, the empty bins with a red contour indicate the percentage points change in the relative share of export value, for industries not present in the census; the bins with a black contour indicate the same for industries for which the ICI is available. The colour filling of the bins indicates the ICI of each industry, from green (high) to red (low), passing through white (mean value). The highest increase in this group is observed for the manufacture of basic precious and non ferrous metals, for which we do not have an ICI. The fact that Ghana exports products from this industry might indicate that: those products are imported and then re-exported by companies specialised in other industries; that production activities in this industry are carried out by multinational corporations; or that the products are made by firms whose secondary sector of activity is the manufacture of basic precious metals. In all these cases, the benefits to local firms and workers may be non-existent or reduced. First, re-export of products does not provide any information on the country's productive capabilities. Second, in the case of production conducted by multinational corporations, the benefits, in terms of capabilities and technology transfer, will depend on each firm's practices, and its use of domestic skilled and unskilled workforce. Third, if the manufacture of basic precious metals, such as gold, is not the main sector of activity of the domestic formal firm, it will represent only a minor contribution to the domestic productive structure. Despite this, it must be noted that evidence from Ghana suggests the existence of backward linkages feeding into the gold extraction industry from consulting services, which provide inputs to the mining industry (Bloch and Owusu, 2012) and represent a valuable path for diversification towards knowledge-intensive activities for mining-intensive countries. However, the level of disaggregation of the data used in the

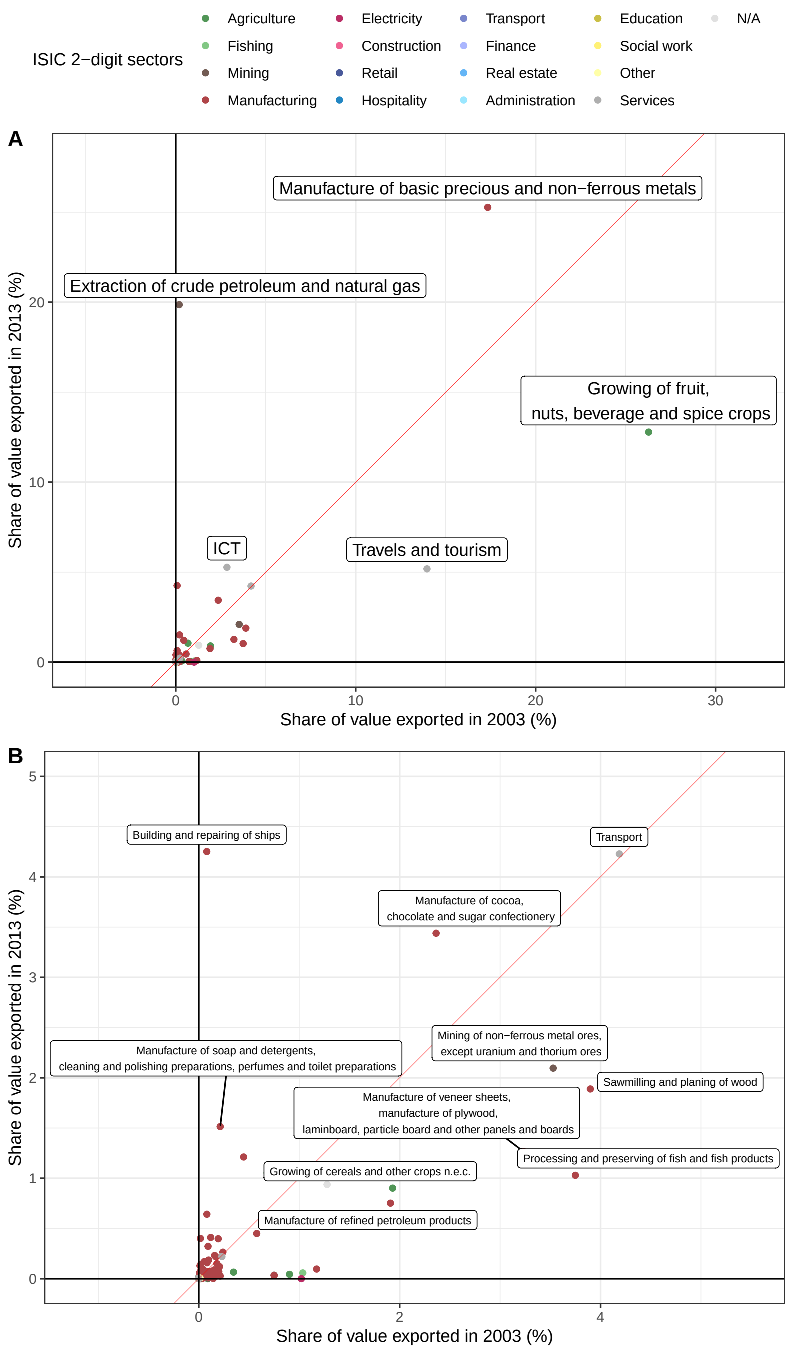

Figure 2.3: The dots in the scatterplot represent a different 4-digit ISIC industry in the trade sector; the different colours indicate the major ISIC group. Industries with a constant relative weight over time lie on the 45° line; industries closer to the X axis and far from the Y axis have experienced a decrease in their relative weight between 2003 and 2013. Industries that are closer to the Y axis (2013) and far from the X (2003) axis are those that have experienced a change in relative weight over time. Panel A includes all 4-digit industries; panel B zooms in on those industries with a share lower than 5% in both years.

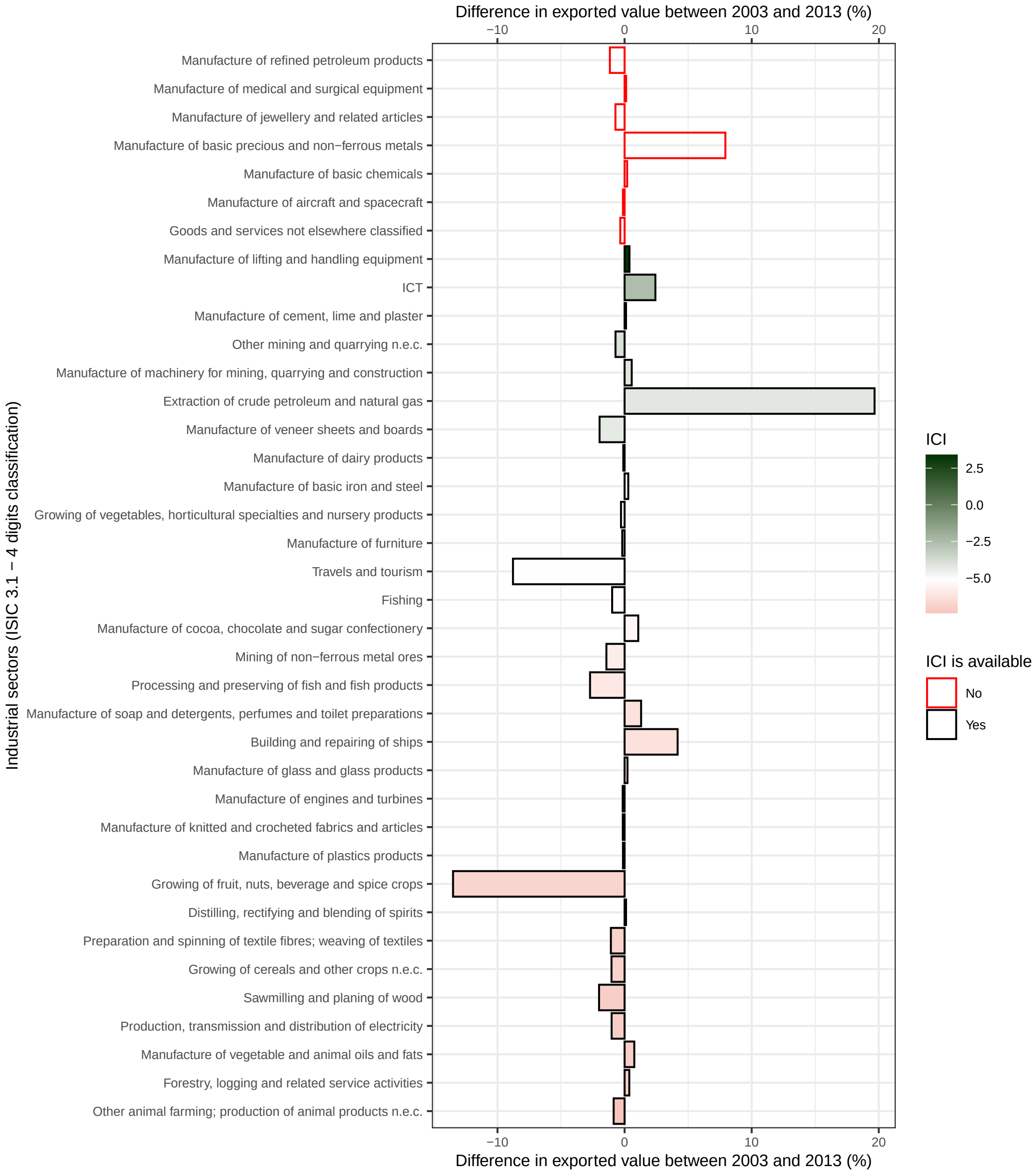

Figure 2.4: The length of the bins represents the difference (in pp) in the relative shares of each export industry between 2003 and 2013. Positive values indicate an increase in exports in 2013 and negative values indicate a decrease in exports in 2013. The colours indicate the complexity of each industry, from low (red) to high (green); values around the mean are in white. Empty red bars indicate industries where the ICI is not available. Industries varying by less than 0.1 percentage points are excluded from the diagram.

analysis doesn't allow to pinpoint mining-specific services, preventing the possibility to test this hypothesis.

Figure 2.4 provides evidence of a large increase in the share of industries whose complexity is higher than the sample mean (in green), such as ICT[13] and oil and gas extraction. At the same time, the industry showing the biggest decline in exports share in exports is growing of fruits and nuts, beverage and spice crops, a low-complexity industry, accompanied by a small increase in the share of manufactured goods linked to these crops (manufacturing of cocoa, chocolate, and sugar confectionery), which is also a low complexity industry.

Table 2.1 presents the direction of structural change in Ghana's export sector based on the percentage points change in relative industry shares, aggregated by ISIC major groups (column 2) and the (logged and standardised) ICI, and the sum of their product for all the 4-digit ISIC industries, by industrial group (column 3). This measure captures the change in relative share, weighted by the average complexity of the macro sector, as per Equation (2.3). Table 2.1 shows that there was an increase in exported manufactured products that was much lower than in mining, but higher than in services. However, this might be because, as we discuss below, exports do not include non-tradable services. Moreover, the three ISIC major groups which show the highest level of growth (manufacturing, mining, and ICT) are also among those sectors with the highest average complexity; the only high-complexity sector that has been decreasing between 2003 and 2013 is transport. Overall, the trade sector appears to be moving towards more complex activities (especially in mining and manufacturing), and away from the export of agricultural commodities. The overall complexity-weighted export change, obtained by summing all the available complexity weighted changes per 4-digit industry, is positive (+2.49). We next investigate whether the pattern displayed by exports applies to the other two sectors.

[13] See Appendix A.2 and A.4 for more detail on ICT-related industries.

Table 2.1: Changes in export shares by ISIC major groups

ISIC group	Export difference	Avg. complexity	Weighted export change
Administration	0.000	0.144	0.000
Agriculture	-15.385	0.162	-2.251
Construction	0.000	0.220	0.000
Education	0.000	0.086	0.000
Electricity	-1.021	0.177	-0.132
Finance	-0.010	0.252	-0.003
Fishing	-0.978	0.262	-0.256
Hospitality	0.000	0.116	0.000
ICT	2.424	0.461	1.116
Manufacturing	6.604	0.325	0.385
Mining	17.453	0.445	6.062
Other	-0.007	0.170	-0.001
Real estate	0.000	0.304	0.000
Retail	0.000	0.187	0.000
Social work	0.000	0.127	0.000
Transport	-8.741	0.358	-2.425

Note: Total relative export difference measured as percentage points change between 2003 and 2013, average complexity of the 4-digit industries in each group, and 4-digit level change in exports, weighted by the ICI (also 4-digit ISIC) and summed by ISIC group are indicated for each 1-digit ISIC group.

2.4.3 Formal sector

The analysis of the formal sector is restricted to the manufacturing and mining activities since the 2003 National Industry Census focused only on these two industries. This means that industry share refers only to the employment percentages in these two groups and not all industries. However, Figure A.5 in Appendix A.1 depicts the full composition of domestic formal employment in 2014, for all industries, and shows the wide range of economic activities beyond the manufacturing and mining ISIC groups. Visual inspection of Figures 2.5A and 2.5B reveals an increase in mining activities, such as mining of non ferrous metal ores, mining and agglomeration of hard coal and extraction of crude petroleum and natural gas. In the case of oil extraction, while its relative share in formal employment grew between 2003 and 2013, its relative employment size is smaller compared to the share of export value. This suggests that, despite the large revenues generated by oil exports, extraction of this natural resource has only a comparatively modest capacity to absorb labour

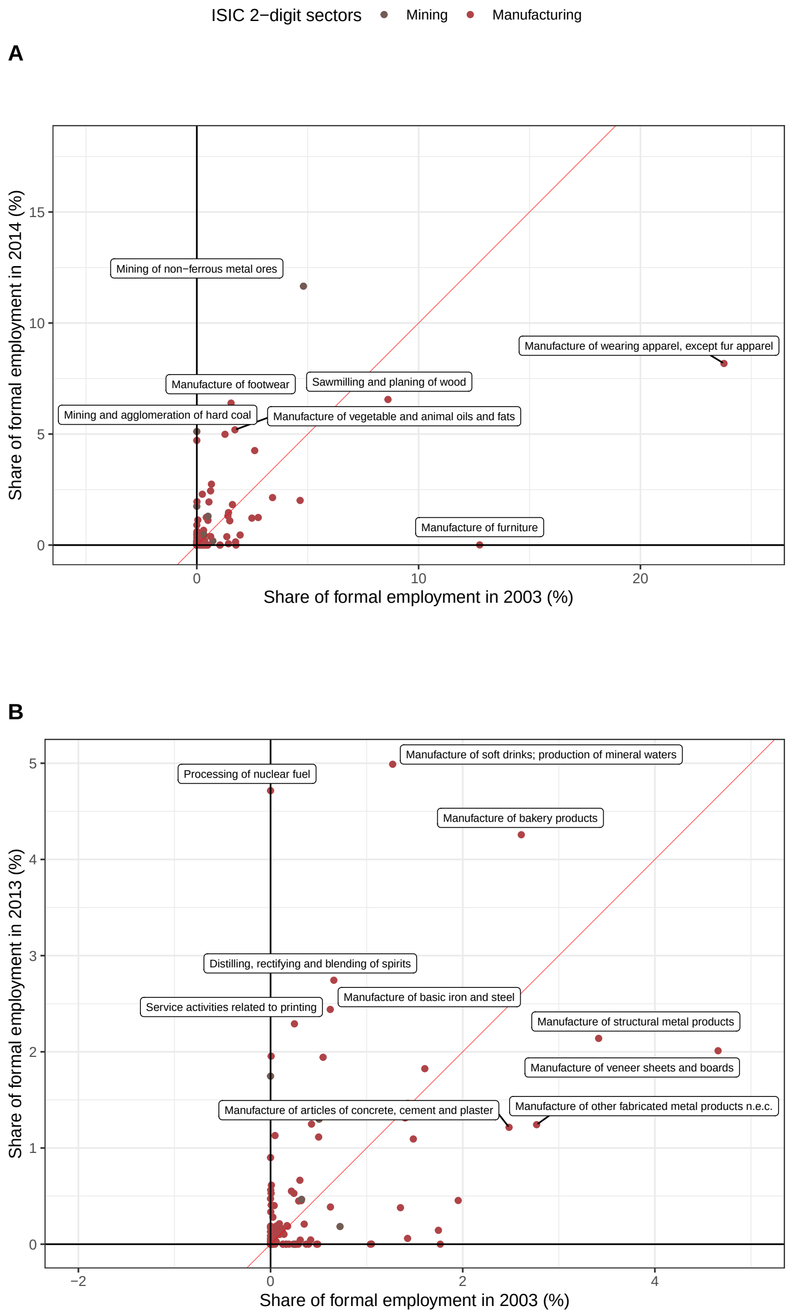

Figure 2.5: The dots in the scatterplot represent a 4-digit ISIC industry in the formal sector; the colours indicate the major ISIC group. Industries with a constant relative weight over time lie on the 45° line; industries closer to the X axis and far from the Y axis have experienced a decrease in relative weight between 2003 and 2014. Industries that are closer to the Y axis (2014) and far from the X (2003) axis are those that experienced a change in relative weight over time. Panel A includes all 4-digit industries; panel B zooms in on industries with a share lower than 5% in both years.

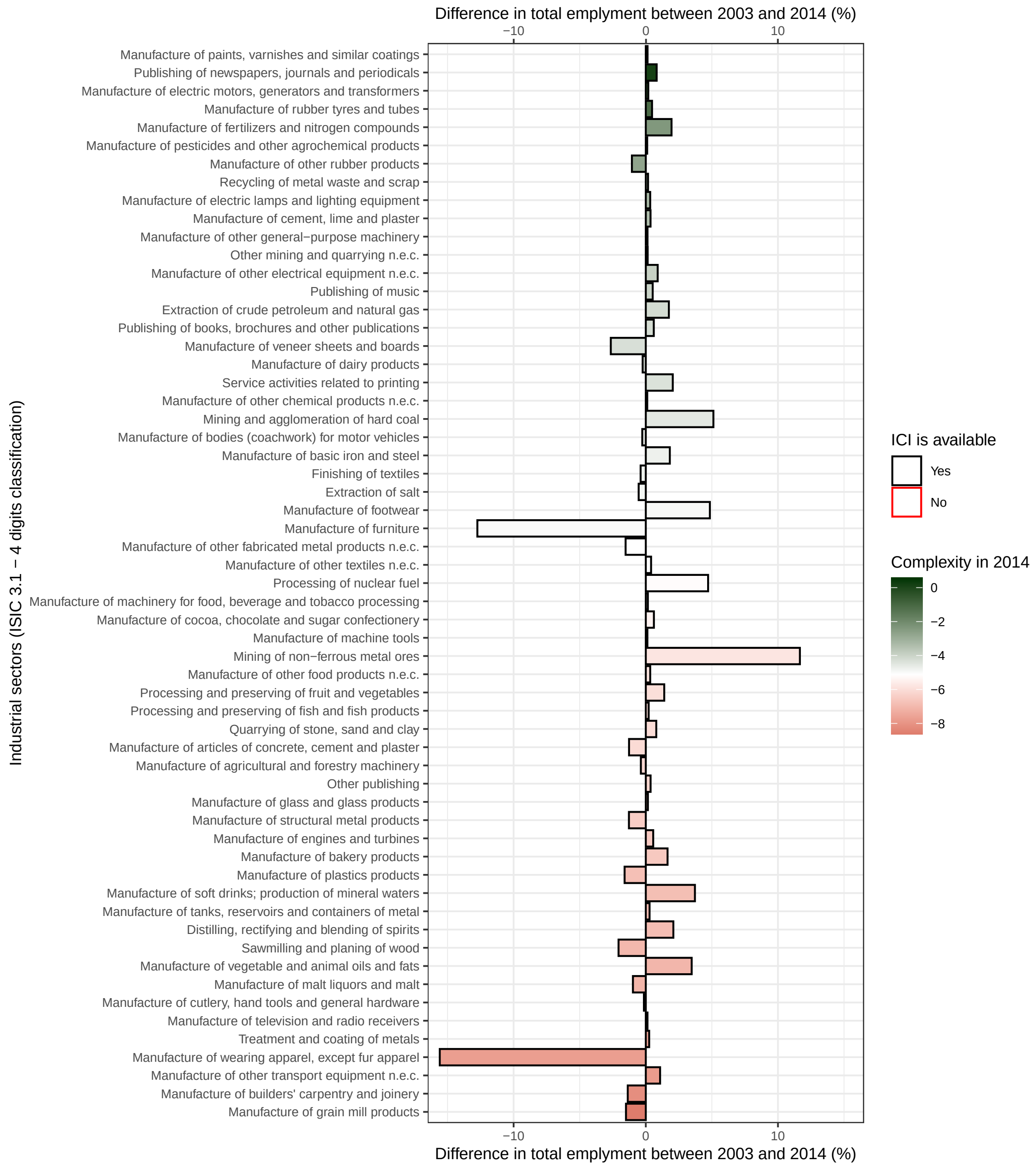

Figure 2.6: The length of the bins represents the difference (in pp) between the relative shares of manufacturing and mining formal industries between 2003 and 2014; positive values indicate an increase in formal employment in 2014 and negative values indicate a decrease. The colour filling of the bin indicates the complexity of each industry, from low (red) to high (green); values around the mean are in white. Empty red bars indicate industries where ICI is not available. Industries that vary by less than 0.1 are excluded.

and is not likely to drive structural change involving higher levels of employment.

Similar to the trade sector, the formal sector experienced a decrease in the relative share of some traditional manufacturing activities, such as manufacture of wearing apparel, and in activities linked to wood processing (sawmilling and planing of wood, and manufacture of furniture). However, other traditional manufacturing has grown in relative terms, including the manufacture of footwear, and the manufacture of vegetable and animal oils and fats. It is clear that the nuclear fuel processing sector has also grown – and quite considerably – between 2003 and 2013, probably as the result of the development of several nuclear projects in Ghana. However, there are some concerns about their long-run sustainability (Ramana and Agyapong, 2016).

To assess these changes, we use the ICI. Figure 2.6 shows that more complex industries (in green) tend to show an increase in their relative shares, although the largest growing industry is the low complexity mining of non ferrous metal ores industry (including gold). The industry that has shrunk the most is manufacture of apparel, which is among the least complex industries. Another industry that has declined is manufacture of furniture, which is more complex than the mean industry in the sample. Table 2.2 summarises the changes in (complexity-weighted) employment by major ISIC groups (manufacturing and mining only). On the one hand, the manufacturing employment figures (column 1) seem to confirm a continuing trend towards de-industrialisation, with mining activities gaining weight relative to manufacturing. On the other hand, complexity-weighted employment change (column 3) indicates that, although employment has generally moved away from manufacturing, more complex industries have grown, making the measure positive. This suggests that although aggregate labour has decreased, it is showing a shift from employment in low-complexity manufacturing activities towards more complex activities. In the formal sector, overall complexity-weighted employment change is +6.45. However, while these figures suggest a trajectory of growth-enhancing structural change, it should be remembered that the mining and manufacturing sector accounted for only 17 per cent of formal employment in Ghana in 2014 and that the aggregate employment figures indicate that, over the time period considered, the share of labour in low-complexity service-related industries increased.

Table 2.2: Changes in employment shares (formal sector) by ISIC major groups

ISIC group	Emp. difference	Avg. Complexity	Weighted emp. change
Mining	18.926	0.439	4.684
Manufacturing	-7.407	0.318	1.762

Note: For each ISIC group (1-digit ISIC, only manufacturing and mining), this table indicates the total difference in employment shares measured as the percentage points change between 2003 and 2014, the average complexity of the 4-digit industries in each group and the 4-digit level change in employment shares, weighted by the ICI (also 4-digit ISIC) and summed by ISIC group.

2.4.4 Informal sector

The composition of household activities, revealed by the GLSS data collected in 2005 and 2014, shows considerable variety in activities across 4-digit ISIC industries and groups. However, this sector shows very little change over the eight year period considered and is dominated by agricultural activities (see Figure 2.7A). Note that the aggregate employment figures for the household sector may include some double counting of employees that work on more than one activity since the aggregate figures sum employees (who cannot be uniquely identified) working on agricultural plots and in non-farming enterprises, divided by sector of activity.

To improve the visualisation of employment shares in non-farming activities, Figure 2.7B and Figures A.7 and A.9 in Appendix A.1 exclude agriculture. We observe that in both 2005 and 2013, the majority of household manufacturing activities were related to processing of agricultural products and production of food, although there was some contribution from other traditional manufacturing activities (such as manufacture of apparel). However, these activities have decreased over time, while employment has increased in non-tradable and low-productivity activities such as wholesale and retail, repairs, hairdressing and restaurants.

Figure 2.8 presents the changes in complexity in the household sector. While the

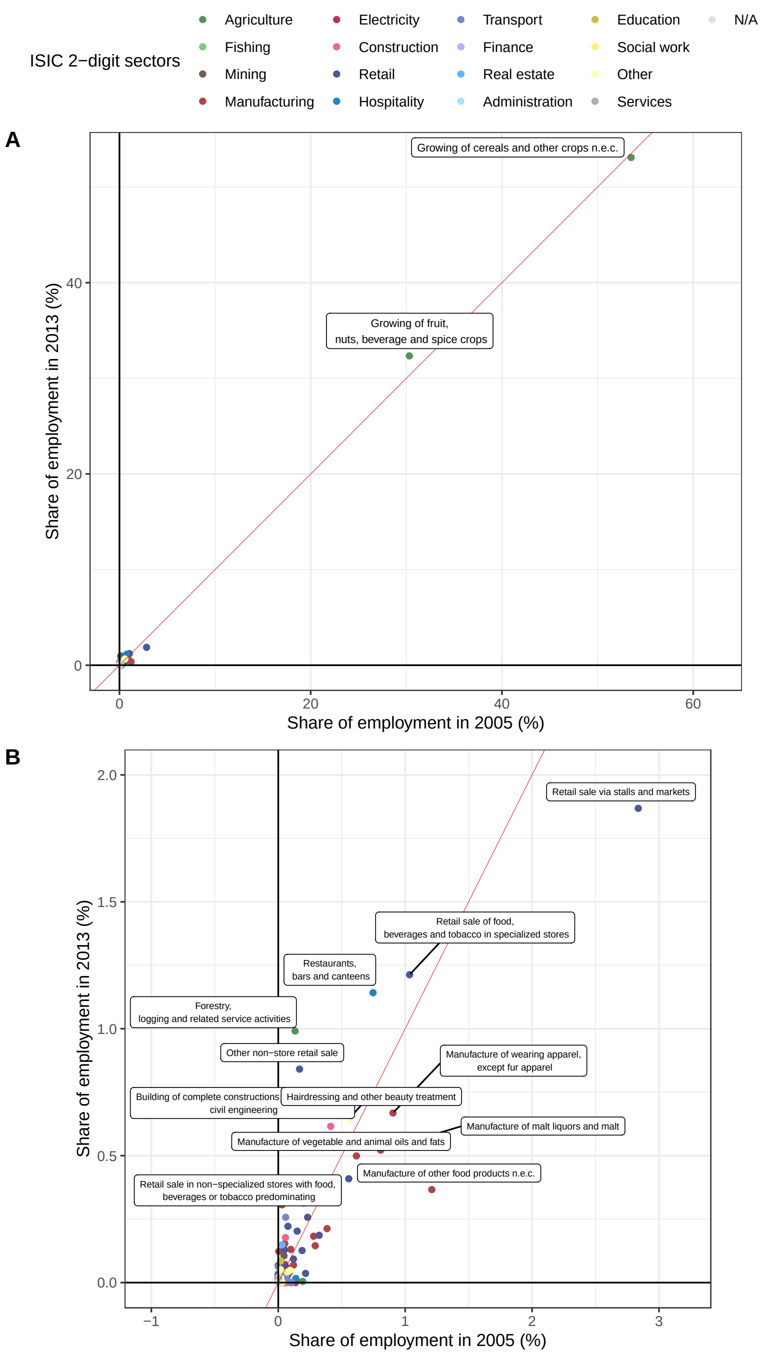

Figure 2.7: The dots in the scatterplot represent a different 4-digit ISIC industry in the informal sector; the colours indicate the major ISIC group classification. Industries with constant relative weight over time lie on the 45° line; industries closer to the X axis and far from the Y axis have experienced a decrease in relative weight between 2005 and 2013. Industries that are closer to the Y axis (2013) and far from the X axis (2005) are those experiencing a change in relative weight over time; panel B zoom into industries with a share lower than 2% in both years.

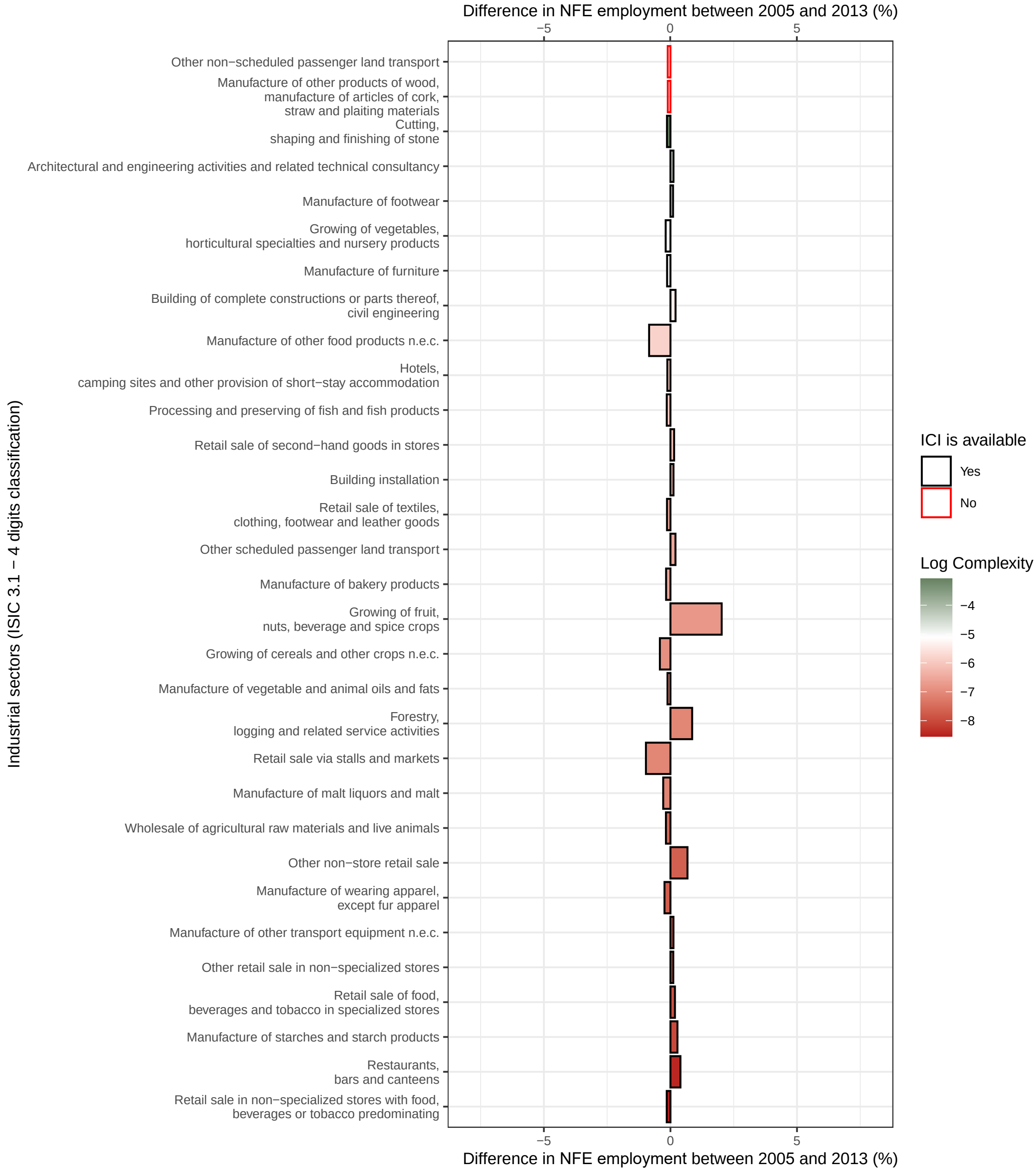

Figure 2.8: The bin's length measures the difference (in percentage points) of the relative share of each informal industry between 2005 and 2013. Positive values indicate an increase in informal employment in 2013, while negative values indicate a decrease. The colour filling of the bin indicates the complexity for each industry, from low (red) to high (green); values around the mean are shown in white. Empty, red bars indicate industries for which the ICI is not available. Industries with variation lower than 0.1 are excluded for visualisation purposes.

employment composition of household enterprises remained constant in the period of observation (2005-2013), we observe that the more complex industries are those that show the smallest changes in the relative size of employment. We can see more variation across the less complex industries, with the largest change in cultivation of fruit, nuts, beverage, and spice crops – a low complexity industry. Similarly to the trade and formal sectors, we observe in Table 2.3 the extent to which the ISIC major industrial groups have changed in terms of employment and how those changes contributed to the sector's aggregate complexity. If we compare the agricultural and manufacturing sector, we see that the former has grown in size more than the latter; however, complexity-weighted employment change indicates that the increase in aggregate complexity of household employment in agriculture has been more than offset by the reduction of employment in manufacturing. In a few industries (hospitality and other services), employment has grown, but still contributes negatively to the aggregate complexity of employment in the economy. Despite the small changes in the informal sector between 2005 and 2013, its complexity-weighted employment change is negative (-0.04). Although smaller in magnitude compared to the trade and formal sectors, this indicates that the small changes in relative labour shares were in low-complexity sectors.

Table 2.3: Changes in employment shares (household sector) by ISIC major groups

ISIC group	Emp. difference	Avg. complexity	Weighted emp. change
Agriculture	2.249	0.162	0.291
Fishing	-0.036	0.262	-0.009
Mining	0.086	0.445	0.029
Manufacturing	-2.067	0.325	-0.391
Electricity	-0.054	0.177	-0.010
Construction	0.248	0.220	0.070
Retail	-0.233	0.187	-0.028
Hospitality	0.273	0.116	-0.016
Transport	-0.006	0.364	-0.012
Finance	0.004	0.246	0.003
Real estate	0.112	0.304	0.038
Administration	0.011	0.144	0.002
Education	0.064	0.086	0.000
Social work	-0.086	0.127	-0.012
Other	0.039	0.170	-0.002

Note: For each ISIC group (ISIC 1-digit), total difference in employment shares measured in percentage points change between 2005 and 2013, average complexity of 4-digit industries in each group and 4-digit level change in employment shares, weighted by the ICI (also 4-digit ISIC), and summed by ISIC group are indicated.

2.4.5 Discussion

The mapping exercise in the previous subsections provides detailed insights into Ghana's industrial composition, across sectors. Based on the above evidence, we next discuss the results, comparing among the three sectors. We propose a set of stylised facts related to Ghana's structural transformation, regarding: i) sectoral specialisation; ii) de-industrialisation; iii) the service economy; and iv) the overall transformation trajectory.

Specialisation. The analysis of the trade and formal sectors provides evidence on Ghana's specialisation in the extraction and mining of natural resources. These activities have a high ICI, due, likely, to the fact that most are highly geographically localised. While the high complexity of these activities might suggest that they are facilitating related industrial diversification, employment changes in the downstream activities linked to these activities suggest that it might not be feasible for Ghana to upgrade along the respective value chains. For instance, growth in the processing of metal ores appears to be led by re-exporting, production by non-Ghanaian firms or secondary production by domestic firms (see Section 2.4.2). In addition, despite the high growth in exports of raw petroleum, employment in the manufacture of processed oil products shrank by almost a half, while oil extraction represents only 0.1 per cent of total formal domestic employment and 0.09 per cent of household employment in Ghana (compared to 20% of total exports). However, the increase in exports of precious metals products seems to have been accompanied by increasing employment in metal ore extraction. In terms of agricultural production, the trade sector shows a decrease in exports of agricultural products while agricultural employment in the household sector shows continuous growth. This suggests that household agriculture is oriented more to the domestic market and that Ghana has not abandoned its agricultural activities, which continue to represent an important source of employment.

(De)industrialisation. We began by acknowledging the trend in Ghana towards premature de-industrialisation (Rodrik, 2016) (Figure 2.1) between 1960 and 2010. The analysis provided further evidence on the composition of the manufacturing sector in Ghana, which has been observed across three different sectors. Trade sector data seem to indicate that the relative importance of manufacturing exports is growing, although a break down of the data across sectors presents a different picture. First, the continued importance of some manufactured goods exports is not likely to benefit local employment: we have shown that production of the fastest growing exports of manufactured goods (basic precious and non-ferrous metals) may not be carried out in Ghana, as there are no firms in the country that operate in that industry. Second, in the formal sector, the relative share of manufacturing has been shrinking in terms of aggregate employment (although, recall, that consideration of formal manufacturing is restricted to a comparison among the mining industries since data on the other industrial groups is unavailable). However, Section 2.4.3 showed that labour has been moving away from low complexity to higher complexity manufacturing industries.

This pattern is consistent with recent evidence on African manufacturing, which indicates that the productivity of formal manufacturing firms is growing, suggesting an advancement of African manufacturing towards the technological frontier (McMillan and Zeufack, 2022), in some cases, reversing the de-industrialisation trend (Kruse et al., 2021). However, the importance of manufacturing as an employer of household labour has decreased considerably, compared to other industries in the sector. This evidence is apparently in contrast with Kruse et al. (2021) and McMillan and Zeufack (2022), which show that informal manufacturing has been growing and contributing positively to structural change. It must be noted however that these authors look at informality as the residual from total employment after excluding formal employment, while here informality is represented by employment in household enterprises – a specific subset of the residual agglomerate identified by Kruse et al. (2021) and McMillan and Zeufack (2022).

Services. This analysis of Ghana's sectors provides some useful insights for debate on the growing importance of services in African economies (Owusu et al., 2021; Baccini et al., 2021). The composition of Ghanaian exports indicates that the relative share of tradable services has decreased overall, although the ICT sector, which is likely to be one of the sectors crucial for sustaining catch-up by low- and middle-income economies (Kaplinsky and Kraemer-Mbula, 2022), has grown. The growth in ICT exports seems to have been accompanied by some employment in ICT-related industries (1.2% in the formal sector in 2014, but zero in the household sector) and a growth in their high level of complexity, suggesting that ICT has the potential, simultaneously, to become a source of economic diversification and employment creation. Nevertheless, household sector employment is moving towards low-productivity services, which is weighting negatively on the aggregate complexity of the household productive structure.

In sum, the results of this analysis suggest a trajectory of structural change in Ghana between 2003 and 2013, characterised by growing specialisation in the extraction of natural resources, with mixed effects on employment creation. In addition, Ghana has experienced a drastic reduction in export of agricultural commodities, despite the continued importance of agricultural activities for household livelihoods; a reduction in the size of the manufacturing sector, which has experienced some reallocation of labour towards more complex industries; and a reduction in the relative weight of tradable services (with some exceptions, such as ICT), accompanied by employment growth in low-productivity services. However, findings for the formal manufacturing sector are limited to its relative weight with respect to mining since data on the other sectors were not available for 2003.

There are two final implications of the structural change trajectory identified by this study. First, the analysis shows that using the industrial composition of traded activities to proxy for national productive capabilities could be misleading. We showed that some growing export industries are not reflected in domestic employment and, therefore, cannot be considered as signalling the strength of domestic economic activity or as indicators of the stock of locally available capabilities. These considerations have a methodological counterpart; in Section 2.3.2 we argued that using trade-based measures of complexity might be misleading since they do not account for employment composition at the country level and, also, assume that exports represent those activities in which the country is most competitive. While this latter assumption might not be misplaced, the analysis in this chapter shows that, in some cases, export activities do not match domestic capabilities (in terms of industry employment) and, therefore, a measure of industrial complexity based on subnational employment data is more appropriate.

Second, we argued that the informal sector is a crucial element in the process of structural transformation. However, Section 2.4.4 indicates that while the trade and formal sectors experienced some transformation of their industrial composition between 2003 and 2013, the industrial composition of the household sector remained substantially unchanged. There is some evidence suggesting that the informal employment structure is moving towards less complex industries, driven mainly by the continued (and growing) importance of agriculture and a contraction in manufacturing activities. On the one hand, this is not necessarily related to a lack of dynamism. In Chapter 3, we investigate the participation of the informal sector in the process of structural change, through interaction with the formal sector. On the other hand, the availability of capabilities in household and informal activities that match the aggregate productive structure, remains crucial to foster inclusive and growth enhance structural change.

2.5 Conclusions

This chapter contributes to the literature on structural change and industrial transformation for development, by mapping and discussing the trajectory of Ghana, one of the most successful countries in the SSA region. We add to the synergy between the structuralist and the complexity literatures, while reconciling potential differences and proposing a conceptual and empirical framework. The analysis conducted in this chapter considered production alongside Ghana's export structure, using an employment-based measure of industrial complexity – the ICI – and, most importantly, accounting for the (predominant) informal sector as well as the export and domestic formal sectors. The analysis provides a rich and exhaustive map of structural change in Ghana, which suggests novel angles and perspectives that are more nuanced than suggested in the current scholarly literature. Most important, it offers concrete evidence which can serve as the basis for industrial policy aimed at sustainable development and structural transformation.

This study provides several results related to the trajectory of structural change in Ghana. These are discussed in detail in Section 2.4.5 and can be summarised as follows.

First, by comparing the trade, formal and informal sectors, the analysis shows that Ghana is specialising in the extraction of natural resources, such as oil and gold, although neither activity has created employment in the downstream sectors.

Second, the analysis of structural change using employment-based complexity, complements (and in some ways contrasts with) the structuralist view that African labour has moved to less productive sectors. In fact, this chapter shows that, in some cases (particularly in the formal sector), labour has moved to more complex industries, while contracting elsewhere (in the trade and informal sectors).

Third, it has been observed that, in aggregate, Ghanaian labour has moved away from agriculture towards services, skipping the industrialisation phase of structural change. However, the findings from our analysis indicate that, in terms of trade composition, the relative share of services has shrunk, although the importance of more complex services has increased. On the other hand, employment in low-complexity services has grown in the informal sector.

Fourth, the relative importance of agricultural exports has decreased, although their relative importance in informal employment has grown. This suggests that while total employment in agriculture may have reduced, as suggested by Figure 2.2, the commitment towards subsistence agriculture by Ghanaian households (in terms of household members who are dedicated to household agriculture) has not reduced, and has even increased. This challenges the view that African labour is leaving the agricultural sector and highlights yet another mismatch between levels of production in Ghana.

The above evidence offers opportunities for industry policy related to the informal sector. Most policy tools aimed at high-income countries and adapted to low- or middle-income contexts, ignore the informal sector. The results indicate that a large room for action is due, not necessarily by favouring a shift from informal to formal, but to create synergies between the upgrading formal sector and the large, low-complexity and yet heterogeneous informal sector.

More specifically, bridging between the structuralist and complexity frameworks allowed an emphasis on the importance of aligning the productive capabilities of the informal sector with the Ghana's productive structure. This is crucial to allow the participation of households in the process of structural transformation, and to ensure the availability of skills for the emerging activities, in particular those that will require the inflow of new workers while also facing the challenge posed by emerging breakthrough technologies such as automation-related ones, which may reduce the labour intensity of these activities. Finally, the analytical and empirical approaches proposed have some limitations. Some of these are discussed in this chapter. One limitation is that the analysis conducted here does not allow examination of the direct linkages among the three sectors and we need more research in this direction. The following chapter addresses this issue in part by investigating the links that exist between the formal and informal industries in Ghana.

Chapter 3

Informal and Formal Industrial Co-location and Structural Change in Ghana

3.1 Introduction

Economic development is inherently linked to the concept of structural change. As countries develop, not only do their incomes grow, but deep changes permeate through the social and economic spheres. The concept of structural change is one that well describes such transformation, and is defined as long-term change in the composition of economic aggregates, including the relative importance of industrial sectors in the economy, the physical location of economic activity, and other concomitant aspects of industrialization that accompany economic growth (Syrquin, 2010). With the emergence of new, modern industries (Schumpeter, 1943; Saviotti and Pyka, 2004), and increasing productivity in traditional ones (Lewis, 1954), firms can shed unproductive labour that will be absorbed by modern expanding industries, allowing the economy to upgrade towards more complex activities (Hidalgo et al., 2007) by transforming their comparative advantage (Lin et al., 2011). However, not all countries have experienced this mechanism linking industrial changes to economic growth: despite persistently high GDP growth rates, African countries have failed to transform their economies away from subsistence and low-productivity industries (Bah, 2011; African Center for Economic Transformation, 2014). In fact, until the 2000s African labour has shifted towards less productive industries following a pattern of growth-reducing structural change (McMillan et al., 2014) accompanied by a premature decrease in relative shares of formal manufacturing, and by the proliferation of informal activities (Rodrik, 2016). The link between de-industrialisation and the growth of informal industries has contributed to the framing of informal activities as an obstacle to development and structural change, and has shifted the debate on industrial strategies towards a reduction in size of the degree of informality (Lundvall and Lema, 2014), without considering the potentially active role that could be played by informal activities in the process of structural change (Diao et al., 2018b). Against this backdrop, this chapter proposes to investigate whether informality could instead represent an untapped opportunity to support structural change and long term growth, as its contribution has yet to be unpacked.

Including informality in the study of structural change is important for several reasons. To start with, the exclusion of informal activities from considerations on industrial development leads unavoidably to an underestimation of a country's industrial potential, as many informal activities operate in modern industries (Lewis, 1979; Fu et al., 2013; Diao et al., 2018a), thereby contributing to the industrial variety of countries – although they fall below the radar of official statistics. Secondly, the informal sector constitutes a potentially vast reserve of capabilities at the level of firms, industries and productive ecosystems; provided that informal firms do innovate (Avenyo, 2018; Fu et al., 2018; Fu, 2020), with the right conditions, informal production and innovation capabilities could be used to sustain growth-enhancing economic transformation (Kraemer-Mbula and Monaco, 2020). Thirdly, informal activities can benefit from interaction with formal industries in various ways (Arimah, 2001; Chen, 2012; Meagher, 2013; Avenyo et al., 2020); whether through supply chains or competition, the co-location and interaction between informal and formal firms represents a potential source of spillover and externalities for both informal firms and the whole economy. Nevertheless this conjecture – that informal activities can contribute to structural change, participating in a country's industrial variety and co-locating with formal firms – is under-explored in the extant literature. This chapter provides a first contribution in this direction.

This chapter examines the case of Ghana: despite its steady and positive economic growth performance (International Monetary Fund, 2019), the sustainability of the economic growth experienced by Ghana and many other countries in the African continent has been questioned as it has failed to fuel growth-enhancing structural change (McMillan et al., 2014), and to enable the upgrading of informal activities (African Development Bank, 2019), which employ around a half of the active workforce (Ghana Statistical Service, 2015). In order to assess the potential contribution of informal industries to structural change, this study analyses Ghana's industrial variety by looking at the patterns of formal and informal co-location and at its degree of relatedness. As it is easier to for new industries to become established when building upon related industrial variety (Boschma, 2021), structural change in a context dominated by informality crucially depends on the geographic structure of the industrial composition, and on the opportunities for agglomeration externalities between informal and formal industries.

Drawing upon the extant literature on relatedness and structural change, this study uses three different measures of industrial relatedness to describe the channels that drive the co-location of formal and informal industry pairs. First, it measures the degree to which industries share the same capabilities, proxied by their shared occupations (Neffke and Henning, 2013). Secondly, it measures technological relatedness using the input-output relationship between industries (Bahar et al., 2019). Lastly, and based on the evidence that knowledge-intensive activities tend to co-locate in the same areas (Balland et al., 2020), a major innovation proposed in the analysis is the adoption of complexity differentials between informal and formal industries as an explanatory factor of their co-location, with the aim of testing whether a concentration of highly complex industries can also be found in the interaction between informal and formal industries. This endeavour is made possible by the recent publication of a census of formal and informal firms – the Integrated Business Establishment Survey 2014 (IBES) – which has increased the coverage of informal activities in Ghana to an unprecedented degree. Using the IBES in combination with the Ghana Labour Force Survey (GLFS) (Ghana Statistical Service, 2016) and the Eora input-output tables for the Ghanaian domestic economy (Lenzen et al., 2012), this chapter proposes an empirical framework to identify the potential contribution of informality to structural change via informal-to-formal industrial co-location.

Firstly, the characteristics of informal and formal industrial agglomeration across Ghanaian districts are analysed. Secondly, the study aims to unpack the factors that explain the co-location of informal, formal and informal-formal industry pairs in order to assess the extent to which these co-locate following their similarity in terms of capabilities, technology and knowledge intensity; such relatedness represents a main pre-condition for diversification and structural change. To achieve this, a network of co-locating industries is created using information on the economic specialisation of Ghanaian districts. It was therefore necessary to investigate the factors that drive the co-location of industries (informal, formal and informal-with-formal), using measures of formal and informal geographical co-location, as well as indicators of relatedness in capabilities, input-output linkages and complexity differentials.

The main results of the analysis indicate that the co-location of informal and formal industries follows a different set of forces, and that informal industries tend to agglomerate more than formal ones, against the common wisdom that they would be more dispersed. Consistently with the extant literature on the drivers of industrial co-location (Ellison et al., 2010; Diodato et al., 2018), we find that input-output linkages have a positive (although non-linear) effect on all measures of co-location. Furthermore, informal industries show higher relatedness and complementarity as compared to formal ones, and appear to be horizontally integrated, unlike formal ones which show vertical integration in clusters characterised by unrelated variety.

However, when informal industries co-locate with formal ones, they show low levels of relatedness, forming clusters mixed in complexity. This finding suggests, on the one hand, a degree of complementarity between co-located formal and informal industries, and on the other hand, points towards the absence of labour flows between geographically co-located informal and formal industries. On the basis of these findings, we argue that informal industries show potential for agglomeration externalities and economic diversification, if adequately supported by intentional policy. The remainder of the chapter is organised as follows; Section 3.2 reviews the relevant background literature, and identifies the gaps which the analysis conducted here aims to address; Section 3.3 illustrates the empirical strategy adopted in the analysis, and describes the data and methods; Section 3.4 presents the results. Section 3.5 concludes.

3.2 Background literature

The aim of this section is twofold. The first subsection provides an overview of the relationship between structural change, industrial variety and relatedness, arguing that the latter – in its various forms and channels – represents a crucial precondition for industrial diversification and upgrading. The second subsection reviews both theory and empirical evidence on the role of informality in structural change, backing the argument that informality should be considered as part of the industrial variety to identify the stock of capabilities that makes diversification and structural change possible. The last subsection identifies the gaps in the literature on structural change, relatedness and informality which will be addressed in the empirical section.

3.2.1 Structural change, industrial variety and relatedness

3.2.1.1 Defining structural change

Economic transformation lies at the very core of development. The concept of structural change has been used to indicate the process of reallocation of labour and/or value added associated with economic development along different lines: from agriculture to manufacturing, and then to services (Fischer, 1938; Clark, 1940; Fourastié, 1949); from traditional to modern industrial sectors (Lewis, 1954); from rural to urban areas (Harris and Todaro, 1970). While there is a common denominator across those different perspectives – the sectoral unbalancedness of growth associated to structural change – explanations of the process of labour reallocation that underlies structural change differ remarkably, given the absence of a general theory. On the supply side, structural change can be driven by productivity differentials between industries (Kuznets, 1966; Baumol, 1967), innovation and technical change (Schumpeter, 1943; Nelson and Winter, 1982; Dosi, 1988), technological linkages between industries (Hirschman, 1958, 1977), comparative advantages (Hausmann and Klinger, 2006; Lin, 2012), productive factor proportions (Acemoglu and Guerrieri, 2008) and the location choices made by firms (Krugman, 1991). On the demand side, the direction of structural change is also influenced by income trends, as individual preferences move away from agricultural goods as income grows. This regularity, known as Engel's law, constitutes the intuition that is fundamental to the interaction between differing income elasticities of demand across industries (Kuznets, 1973) and technological dynamics (Kaldor, 1966). In open economies, external demand also plays a role in the patterns of economic growth (Thirlwall, 1979). The interaction of supply- and demand-side factors eventually determines the direction of structural change (Krueger, 2008) and patterns of economic growth (Lorentz et al., 2019).

3.2.1.2 Structural change, industrial variety and co-location

One major aspect of growth-enhancing structural change is its relationship with economic diversification (Kuznets, 1973) and industrial upgrading (Lin et al., 2011);

technological advances play a crucial role allowing the emergence of new industries in the economy, which in turn sustain economic growth (Saviotti and Pyka, 2004). In order to upgrade their productive structure, countries continuously strive to specialise and diversify their industrial portfolio towards more and more sophisticated activities, the same happening at the subnational level (Iammarino, 2005; Boschma and Iammarino, 2009). In this process, the technological and industrial varieties determine both constraints and opportunities for diversification and structural change (Boschma and Frenken, 2009). Competing theories have tried to disentangle the relationship between industrial variety and economic growth. On the one hand, some theories have focused on the knowledge spillovers that take place across firms in the same industry in geographically specialised districts (like the so-called Marshall-Arrow-Romer (MAR) theories – Marshall 1920; Arrow 1962; Romer 1986), due to lower transportation costs, labour pooling, and the fact that knowledge spillover is easier in localised industrial clusters. On the other hand, some theoretical propositions stress the importance of externalities arising in highly diversified geographical units, due to the higher possibility of cross-fertilisation between industries (Jacobs, 1969): with a focus on cities, the seminal work by Glaeser et al. (1992) provides evidence of the prevalence of the Jacobs type externalities over the MAR type, owing to knowledge spillovers between, rather than within, industries. However, more recent applied work has also provided evidence on within-cluster externalities (MAR type, as in Ellison et al. 2010), where learning is fostered by the embeddedness of firms within business networks which involves both economic and non-economic relationships (Keeble and Wilkinson, 2000), and where shared informal networks can channel the diffusion of knowledge across firms and encourage a process of socialisation that boosts innovation (Saxenian, 1996).

Due to the presence of economic externalities, the co-location and geographical proximity of industries remain crucial to the achievement of economic diversification at the regional and national level despite the reduction in transportation and transaction costs brought by the Information and Communication Technology (ICT) revolution

(Boschma, 2005). In fact, knowledge and technological spillovers across co-located industries are the basis of the mechanism linking industrial variety and structural change, as geographical factors interact with knowledge, capabilities, institutions and technology, often in a non-linear way (Boschma, 2005). At the firm level, for instance, the success of knowledge local diffusion also depends on the knowledge base of firms in clusters, where the knowledge is unevenly distributed (Morrison and Rabellotti, 2009) and on their position in the knowledge network (Giuliani, 2007).

3.2.1.3 Structural change and relatedness

The geographical aspects of structural change go well beyond the physical proximity of industries (Boschma, 2021). Similarity of capabilities, technologies, and knowledge in the industrial variety are required for spillovers to happen (Frenken et al., 2007; Boschma and Frenken, 2009, 2011) and for the absorption of external knowledge (Boschma and Iammarino, 2009), and even more so in the early stages of economic development (Petralia et al., 2017). In order to conceptualise, isolate and measure (Boschma, 1999) the degree of proximity between technologies and industries, the Economic Geography literature has put forward the concept of "relatedness", which is empirically testable and aims to measure the factors that, along with geographical proximity, drive the geographical concentration of knowledge (for a review see Hidalgo et al. 2018). Evidence on the study of industrial relatedness shows that a related industrial variety is an important pre-condition for entering new (and related) industries (Boschma, 2021), while unrelated diversification is much more unlikely (Neffke et al., 2018), but more frequently observed in larger and higher-income regions (Galetti et al., 2021). However the effect of relatedness on knowledge diffusion may vary according to several factors such as industry maturity, modes of competition, innovation and learning intensity (Neffke et al., 2011).

Empirical approaches based on the concept of relatedness often rely on network techniques to represent productive structures as the interaction between, rather than the aggregation of different industries (some examples are Hidalgo et al. 2007; Neffke and Henning 2008; Caldarelli et al. 2012). This empirical strand, often referred to as the product space literature, takes an agnostic approach to the relatedness between products or industries, aiming to provide an outcomes-based measure without any assumption *a priori* of the level of similarity between them. In the work by Hidalgo et al. (2007), this measure is called "proximity" and is built using information on the co-location of industries with a Revealed Comparative Advantage in the exports of products, relying on the assumption that a country's ability to export a specific product depends on its ability to export another related product, as they require similar institutions, infrastructure, physical factors, technology and capabilities. Consequently, countries should strive to jump on the 'closest' and most complex product in the space in order to transform their economic structure in such a way as to sustain growth. Using this framework, changes in the export patterns (for instance, a move towards more complex and sophisticated products) are the core of economic transformation, which is seen as a country's journey through a 'space' of related activities.

3.2.1.4 Structural change and complexity

The relatedness literature is tightly linked to the growing literature on economic complexity: even if still at its early stages, the latter has contributed "to the study of structural change by addressing questions related to how the diversification of countries and the patterns of trade specialization affect and are affected by economic growth, and how diversification and trade specialization change over time, and the determinants of such change" (Freire, 2021, p 214), providing useful tools to study disaggregated aspects of structural change.

Complex products and industries are defined as those resulting from the combination of non-ubiquitous knowledge and skills. However, measuring complexity of products, technologies or industries is not straightforward, and has recently drawn considerable research interest (Hausmann and Hidalgo, 2009; Tacchella et al., 2012; Balland and Rigby, 2017; Gomez-Lievano, 2018; Mealy et al., 2019; Freire, 2021).While the debate revolves mainly around methodological issues, the core intuition behind most complexity measures is focused on trade patterns: product complexity depends on the ubiquity of export specialisation, and the degree of diversification of exporters. More ubiquitous products are assumed to be less complex as the knowledge required to specialize in such products is widespread; at the same time, if the countries that export that same product are more diversified, the complexity of the product increases because it is exported by countries with a more diversified productive structure. The same reasoning can be applied on levels lower than the country, such as the city (Gomez-Lievano and Patterson-Lomba, 2019; Balland et al., 2020), region (Sbardella et al., 2017; Cicerone et al., 2020), technology (Sbardella et al., 2018a) or industry level (Mealy et al., 2019).

To achieve growth-enhancing structural change, countries (regions) strive to diversify towards more complex and related products/industries. Evidence suggests that as countries grow and develop, they begin to export more complex products and to abandon less complex ones (Felipe et al., 2012), and that as countries move towards more complex, knowledge intensive industries, such activities tend to concentrate consistently across firms, technologies, occupations and scientific fields as the knowledge required by more complex activities is often tacit and subject to local spillover (Balland et al., 2020).

3.2.1.5 Drivers of industrial co-location

To summarise, network and complexity-based approaches provide a granular account of a country's productive structure, which is described as the result of the interaction (for instance, via co-location) between different products/industries that ultimately represent the structural stock of capabilities available in a country (Sbardella et al., 2018b). In this way, the emerging productive structure is more than a mere aggregation of employment in, and output from, different economic activities. In fact, network-based analyses allow us to analyse complex structures and to infer the level of relatedness across industries based on their geographical patterns of agglomeration.

It must be acknowledged that the product space approach provides a useful empirical framework for mapping links between industries and identifying viable diversification strategies based on a country's productive frontier (Hausmann et al., 2014; Hausmann and Chauvin, 2015). However, the mixed nature of technological relatedness – which depends on a broad set of economic and technological factors – represents a major obstacle to disentangling the mechanisms that trigger agglomeration externalities such as knowledge spillovers (Rosenberg, 1979; Rosenberg and Frischtak, 1983).

It has been suggested that more effort should be directed into testing the channels that drive the co-location of industries or products (Hidalgo et al., 2018; Bahar et al., 2019), placing the study of relatedness high on the research agenda. This has yielded a growing body of evidence gathered from investigation and comparison of the determinants of geographical co-location of industries. By comparing different measures of relatedness at the industry pair level, Ellison et al. (2010) find that all of the MAR type channels (skills, knowledge and inputs) matter in driving industrial agglomeration, with a major role played by input-output relationships. An extension of this approach by Diodato et al. (2018) shows that, in the US, skill similarity is the major force driving the agglomeration of services, while the co-location of manufacturing activities still depends on the input-output relationship between firms. Similarly, Neffke and Henning (2013) use information on cross-industry labour flows to find that similarity in skills is a good predictor of the type of diversification occurring in firms.

3.2.2 Informality and structural change

In African countries, most of the domestic workforce and capabilities reside in informal firms and industries. The following subsection focuses on two of the main theoretical bodies that have framed the role of informality in structural change, and reviews the evidence that identifies its potential contribution to economic transformation and growth.

3.2.2.1 Theoretical approaches on informality

The concept of informality has been defined using many different criteria across the social sciences (for a review of the different perspectives and definitions, see Chen 2012). Likewise, the it has also evolved along with developmental paradigms since modernisation theories à-la-Lewis (see McCormick et al. 2020 for an overview), wherein informality was seen as a 'problem to solve' with industrialisation; a problem that resulted from the fragmentation of labour markets and structural rigidities (Chenery, 1975). This perspective also assumes implicitly that low-productivity and informal activities would be replaced by labour in modern industries as industrialisation unfolds, with informal firms disappearing. In this view, the informal sector is not seen as a source of entrepreneurial or productive capabilities, as its only contribution to structural change is that it enables the passive absorption of informal employment by formal, modern industries.

The paradigmatic shift to neo-liberal economic policy in the 1970s also caused a revolution in the way mainstream economics framed informality, which was now seen not as the result of economic dualism, but of rigid regulations that exclude workers from participation in the formal sector. Informal entrepreneurs are deemed to operate cost-benefit analyses to decide the extent to which they want to be engaged with the institutions that regulate labour and entrepreneurship on the basis of the expected benefit brought by entering (or leaving) the formal sector, paying taxes and obtaining protection from formal institutions (Maloney, 2004; Perry et al., 2007; Bosch and Maloney, 2010). Seldom is the trade-off expected to lean towards formalisation (La Porta and Shleifer, 2014).

Despite the wide differences in the two approaches, the underlying hypothesis in both is that informal firms are a homogeneous group of low-productivity profit-maximising entities, that can only contribute to economic growth through formalisation. Neither of the two theoretical paradigms considers the potential stock of capabilities and technology of informal firms and workers as a potentially active source of economic opportunities. Rather, they take a reductionist approach by considering informality only as a survival strategy, that will be passively absorbed by modern formal industries. In so doing, these theories overlook any potentially active role of informal firms in structural change and development, as they are assumed not to contribute to the industrial variety of a country.

3.2.2.2 Evidence challenging theory

These theories have been challenged by a growing body of evidence on informality in Africa. Recent studies on African structural change have relied on dual models à-la-Lewis, emphasising the role of productivity gaps in triggering the process of structural change (McMillan et al., 2014; de Vries et al., 2015). Those studies find that, until the 2000s, most African countries have followed growth-reducing patterns of structural change, with labour moving towards less productive and informal activities, inducing the process that Rodrik (2016) has defined "premature de-industrialisation". However, not all informal activities are 'traditional' even if they result from a contraction of modern formal industries. Many of these activities "[are] useful in [their] own right, meeting genuine market needs, and providing a lot of employment in the process" (Lewis, 1979, p 222). This is in line with the evidence that informality co-exists with formality and that informal activities are highly heterogeneous (Nagler and Naudé, 2017), unveiling a dualism within modern industries (Diao et al., 2018a).

Informality may represent an untapped opportunity for a number of mechanisms. To start with, the growth of labour intensive informal activities can contribute to labour intensive growth (Diao et al., 2018a): evidence from Tanzania (Diao et al., 2018b) shows that informal firms account for the bulk of new jobs in non-farming activities, representing an important – albeit only potential – contribution to structural change. A second mechanism relies on their productivity and ability to innovate, as these are two key determinants of structural change. Since the pioneering studies on informal firms (Hart, 1973), they have been described as highly dynamic, showing the capacity to generate income in unfavourable environments. Although informality in Sub-Saharan Africa shows considerable heterogeneity both in terms of industrial sectors of activity and productivity (Nagler and Naudé, 2017), recent studies on Ghanaian informal firms (Fu et al., 2018; Avenyo, 2018; Avenyo et al., 2020) reveal that they do innovate, although with a lower impact on productivity when compared to innovative formal firms, as they rely more on adaptation and imitation than traditional mechanisms such as research and development (Kraemer-Mbula and Monaco, 2020).

Moreover, it has been argued that informal activities should be considered as part and parcel of a broader system rather than as an isolated sector (Kraemer-Mbula and Monaco, 2020), as they contribute to the aggregate economic variety, also participating in economic activity carried out by formal firms (Arimah, 2001; Meagher, 2013). In fact, traditional firm-to-firm links can also occur between formal and informal organisations: such links can be defined on a purely market basis, as in the case of formal-informal competition (Avenyo et al., 2020), or within industrial networks as well as along value chains in which both formal and informal firms participate (Chen, 2012). For instance, formal-to-informal sub-contracting has been found to increase the size of the more modern segment of Indian informal activities (Moreno et al., 2012). The location of informal firms is one important determinant of their performance. In the specific case of African countries, recent evidence suggests that close geographical proximity to urban centres is associated with higher informal productivity (Nagler and Naudé, 2017), and clustering of highly productive firms is conducive to productivity gains for other individual firms (Naudé, 2015), facilitating interactive learning and diffusion of technology (Nakano et al., 2018).

3.2.3 Summary and gaps

The first subsection of the literature review has described the links between structural change, industrial co-location and relatedness, showing how structural change and economic diversification are tightly linked not only to industrial concentration, but also to the degree of relatedness of agglomerating industries and to their complexity. While the body of literature exploring the drivers of industrial co-location and the nature of relatedness is continuously growing and is producing increasingly detailed insights, it has mostly focused on formal domestic production (see for example Neffke and Henning 2008; Diodato et al. 2018), implicitly assuming the passive or negligible role of informality, under the assumption that industries driven by local demand do not agglomerate (Delgado et al., 2016).

By the same token, the product space literature relies on the global production structure emerging from trade activities (Hidalgo et al., 2007; Tacchella et al., 2012). Although exported activities are assumed to pin down a country's most developed capabilities, many low- and middle-income countries often export an extremely modest share of their GDP, as most of their workforce is informal. This leads to the conclusion that formal domestic activities – let alone exports – do not portray satisfactorily their employment structure, neglecting a large stock of (less developed) capabilities, which could still feed into the process of economic transformation. Therefore, the first gap that this chapter aims to address is the need to consider the potential contribution of informal activities to structural change, looking at the co-location patterns of informal and formal activities (separately and together) and their drivers, such as relatedness, input-output linkages and complexity, which have been reviewed in this section, highlighting their role in triggering structural change and sustaining economic growth.

Moreover, the evidence on the co-location and interaction between informal and formal firms presented in the second subsection challenges the idea that informality cannot contribute actively to employment growth and upgrading within the Ghanaian productive structure, showing that informal activities are heterogeneous both in terms of industries and productivity, and interact with the productive structure of regions and countries. This gap calls for a more in-depth understanding of the way in which informal activities interact with the productive structure: analysing co-location patterns of informal and formal industries could shed light on the conditions under which informal firms are best placed to contribute actively to growth-enhancing structural change. While the mechanisms that drive the co-location of formal industries have been explored in the literature reviewed in the previous subsection, there is a dearth of evidence on whether the co-location of formal and informal firms responds to the same driving forces that apply to formal industries (with respect to the latter, see Diodato et al. 2018). The closest contribution in this sense was made by Mukim (2015). Focusing on within-industry co-location of formal and informal activities in Indian manufacturing, the study finds that buyer-seller and technological linkages are the most important drivers of informal-to-formal co-location within industries, in turn attracting new informal firms in more agglomerated areas. However, Mukim's study focuses exclusively on within-industry co-location of manufacturing activities, while in countries like Ghana the majority of informal enterprises are active in the service sector. Evidence on the drivers of informal-formal co-location across different industries in African countries is still substantially missing. The next section illustrates the empirical strategy adopted to study informal-formal co-location and its drivers.

3.3 Empirical approach

This chapter argues that industrial co-location and relatedness are central factors in structural change. However it has also been highlighted that informality, despite its dominance and variety in Africa, has been denied an active role in this process – despite growing evidence of dynamism and connectedness with the whole economy. This section will illustrate the empirical strategy implemented to identify the patterns of informal, formal and informal-formal co-location, along with their drivers, in order to measure and isolate the channels through which informality is most likely to contribute to structural change. It begins with a description of the agglomeration patterns of informal and formal industries, in order to highlight differences and similarities. Secondly, it tests the extent to which the drivers of industrial co-location – such as relatedness of capabilities, input-output relationships and complexity, along with industry-specific characteristics – correlate with the formation of formal, informal and informal/formal clusters.

3.3.1 Data

The analysis integrates information from three different data sources: the Integrated Business Establishment Survey (IBES) (Ghana Statistical Service, 2015), the Ghana Labour Force Survey (GLFS) (Ghana Statistical Service, 2016) and Ghana Input-Output tables for 2015 (Lenzen et al., 2012). Each data source is described below.

3.3.1.1 Integrated Business Establishment Survey (IBES) 2014

The IBES 2014 is a dataset extracted from a census of Ghanaian firms (Ghana Statistical Service, 2015), and it provides information on basic characteristics of all non-household Ghanaian firms including their location, employment size by gender, and nature of business. The Ghana Statistical Service (GSS) identifies as formal firms those that are registered with the Registrar-General Department (RGD) and operate a formal account; formality is defined accordingly throughout the analysis carried out in the present study. Out of a total of 638,232 business establishments in the dataset, 60,312 (9.45%) correspond to this definition. However, the GSO's definition excludes many state-owned enterprises, which should be considered formal, regardless of whether they are registered with the RGD or whether or not they keep formal accounts, based upon the fact that they exist because of a public initiative, or with the support of the state and can hardly be considered spontaneous entrepreneurial firms. Such firms are therefore regarded as formal. If we add state-owned enterprises to the set of formal firms, the number grows to 81,266 firms, 12.73 per cent of all Ghanaian non-household business establishments; the remaining 577,920 are informal firms, where informal refers to privately owned firms which are not registered with the RGD and that keep either informal or no accounts at all. This definition of informality is consistent with the idea that informal firms are those unprotected by formal institutions (Perry et al., 2007), or that avoid paying taxes even if they register some of their workers or sales, or by relying on informal accounts. The IBES data is used to construct the measures of geographical co-location of industries, as discussed in Section 3.3.2.1, and industrial complexity, as discussed in Section 3.3.2.4.

3.3.1.2 Ghana Labour Force Survey (GLFS) 2015

The Ghana Labour Force Survey 2015 (GLFS) (Ghana Statistical Service, 2016) includes indicators on the nature and type of employment along with other workers' individual characteristics, and it was collected by GSS to assess the labour force situation in Ghana. The survey covers information on both public and private industries and other relevant institutions. The sample is nationally representative, following a two-stage sampling strategy. In the first stage, 402 enumerator areas are selected, stratified by regions, as well as by urban and rural location; in the second stage, 15 households per enumerator area are sampled, ensuring that each household in the country had the same probability of being selected. The final sample consisted of 6,030 households and 9,604 workers. In the section of interest of the survey, workers reported what had been their main job in the last 7 days. The characteristics of the workers' main jobs were collected using information on the occupation of the worker (ISCO08 classification, 4-digit level of disaggregation) and the industry in which they were employed (ISIC Rev.4 classification, 4-digit level of disaggregation). The GLFS was used to compute a measure of relatedness between industries on the basis of shared occupations, as explained in greater detail in Section 3.3.2.2.

3.3.1.3 Eora Input-Output tables

In order to capture technological relatedness between each industry pair, the study relies on Ghana's domestic Input-Output Tables (IOTs). IOTs describe the supply-use relationships between producers and consumers within an economy by industrial sector, providing a measure of the degree of connection between two industrial sectors based on the volume of goods traded between them. Industries that trade extensively are assumed to have either similar or compatible technologies, as the outputs of one industry are used as input by another. IOTs for Ghana's domestic production have been made available by Eora (Lenzen et al., 2012);[1] these have been compiled using a serial iterative procedure based on reconciling raw data across years in a stepwise manner. However, the Eora IOTs are only available for 2-digit ISIC industries (59 in total).

3.3.2 Construction of the variables

3.3.2.1 Co-location matrices

The information provided by the IBES 2014 census is used to create four adjacency matrices based on co-location of ISIC 3.1 industries: i) co-location of all firms at the district level; ii) co-location of formal firms at the district level; iii) co-location of informal firms at the district level; iv) co-location of informal with formal firms at the district level. As we are interested in the contribution of informal industries to structural change, where the employment dimension is of paramount importance, all co-location matrices are constructed using a measure based on employment in co-locating industries, rather than value produced or exported by them – similarly to what Sbardella et al. (2017) do for US counties. One advantage of using an employment-based indicator is that it provides a good description of the size of an industry without overestimating small firms or underestimating large firms (Lazzeretti et al., 2012; Cicerone et al., 2020).

These matrices represent the starting point from which to identify sectoral patterns of formal and informal industrial variety. The IBES data can be used to cross tabulate districts and employment in ISIC 3.1 industries, which yields an incidence matrix of 216 rows (districts) and 249 columns (industries), and where every cell corresponds to the number of workers for each district-industry pair. Three incidence matrices of this kind are constructed for formal, informal and all activities separately, differing in

[1] Documentation and material on the Eora national input-output tables are available at https://worldmrio.com/countrywise/.

terms of whether they report the district-industry employment for formal, informal or all activities.

In order to select only the more meaningful tendency of industries to co-locate a specialisation index based on the Balassa method (Balassa, 1965) was used. Two industries are defined as co-located in a district if the same district holds an employment-based specialisation in both industries; the total co-location measure for an industry pair is given by the sum of districts which specialise in both industries. In order to obtain this measure of co-location between industry pairs, three binary matrices are constructed using district-industry cross tabulations, where the elements of the matrix take value 1 if a district d is specialised in a given industry i; and 0 otherwise. if the relative share of their employment in that industry is higher than the average share of employment in the same industry across districts, and therefore if the following relation holds:

$$\frac{x_{d,i}}{\sum_i x_{d,i}} \Big/ \frac{\sum_d x_{d,i}}{\sum_{d,i} x_{d,i}} \geq 1 \tag{3.1}$$

where $x_{d,i}$ is the share of employment of district d in industry i. The value on the right-hand side of the inequality above is an arbitrary cutoff value indicating the strength of specialisation; if $RCA_{d,i}$ is equal to 1, it follows that the employment share of district d in industry i is equal to the average employment share across the country in the same industry. Figure 3.4 shows the differences in specialisation patterns of districts for three different specialisation cutoffs (1, 2 and 3) and for all, formal and informal activities separately.

As a term of comparison for the co-location measure based on Ghana's domestic industries, the formal co-location measure presented above is compared to the pioneering measure of proximity elaborated by Hidalgo et al. (2007).[2] Proximity is based on the same grounds as the co-location measures used here, as it identifies pairs

[2]The most up-to-date data can be retrieved from the Atlas of complexity website: https://intl-atlas-downloads.s3.amazonaws.com/index.html.

of exported products in which countries are likely to hold a comparative advantage at the same time. In order to make the two measures comparable, the product[3] proximity measures for the year 2015 have been averaged by industrial sectors following the ISIC Rev. 3.1 classification.[4] Figure 3.1 illustrates the relationship between the two variables, which are only weakly and positively correlated ($\rho = 0.08$, p-value < 0.001); many traded industries are not observed in Ghana -– these are pointed out by the dots that lie on the Y axis of the figure. Despite the positive correlation, the high variability described by the cloud of points indicates that industries that may co-locate on average (based on global exports data), may not co-locate in a low-income country where production capabilities may be lacking or underdeveloped. One of this chapter's methodological arguments is that using a measure based on domestic production provides a better picture for analysing context-specific industrial variety.

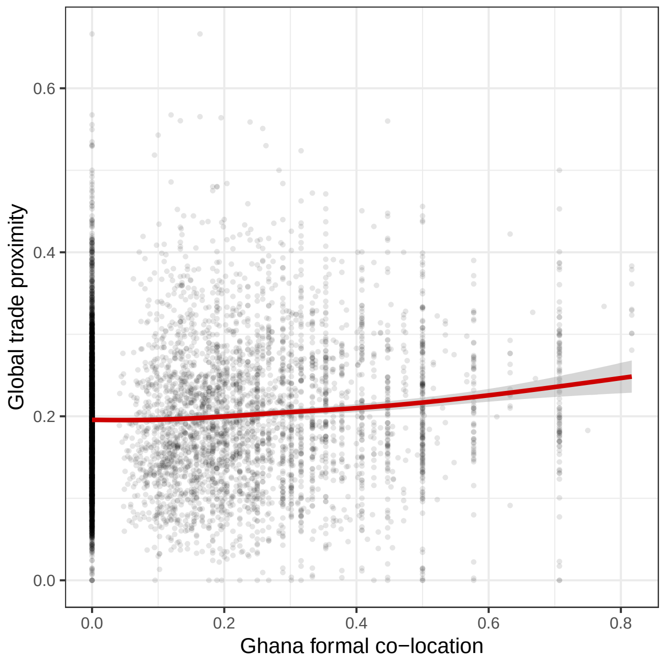

Figure 3.1: Correlation between Ghana's industrial co-location, measured using IBES data, and global trade proximities, obtained averaging the proximities (Hidalgo et al., 2007) of HS products by 4-digit ISIC industries. The red line is the Generalised Additive Model fit line between the two variables.

[3]Products are classified using the 6-digit Harmonised System 92.

[4]Concordance product-industry has been made possible by the `concordance` package (Liao et al., 2020) available on the CRAN repository for R: https://cran.r-project.org/web/packages/concordance/index.html.

The industry co-location matrices were constructed based on the district-industry specialisation matrices: formal co-location (symmetric), informal co-location (symmetric), informal-formal co-location (asymmetric, with non-trivial diagonal values), and formal-informal co-location (asymmetric, with non-trivial diagonal values). The first two indicate the frequency with which two industries i and j co-locate if they are both formal or informal, and where co-location is measured by the number of districts that are specialised in both industries. The remaining two matrices, on the other hand, indicate the frequency of co-location of industries i and j, when the first is informal (formal) and the other is informal (formal); these four matrices identify the phenomena – the co-location patterns of industry pairs – that we aim to explain and compare. A generalised approach to obtaining the co-location adjacency matrices can be summarised by the equation below:

$$\underset{p\times p}{\mathbf{II^{CL}}} = \underset{p\times n}{\mathbf{DI^{T}}} \times \underset{n\times p}{\mathbf{DI}} \tag{3.2}$$

where $\mathbf{DI^{T}_{p,n}} = \mathbf{ID_{n,p}}$, the informal (formal) industry-district specialisation matrix with p rows (industries) and n columns (districts); $\mathbf{DI_{n,p}}$ the transpose of the former; and $\mathbf{II_{p,p}}$ is the cross-product of the former two (or the bipartite projection of $\mathbf{DI^{T}}$). Therefore, in order to obtain the informal-formal co-location matrix, informal industry-district specialisation matrix is cross-multiplied by the formal district-industry matrix. As the two matrices are conformable, their products return a square matrix, an industry co-location adjacency matrix of the form:

$$\mathbf{II^{CL}} = \begin{bmatrix} CL_{1,1} & CL_{1,1} & ... & CL_{1,p} \\ CL_{2,1} & CL_{2,2} & ... & CL_{2,p} \\ \vdots & \vdots & \ddots & CL_{1,1} \\ CL_{p,1} & CL_{p,2} & ... & CL_{p,p} \end{bmatrix} \tag{3.3}$$

where the cross elements indicate the frequency of co-occurrence in the same district of industries i (informal) and j (formal). The other three co-location matrices are obtained following the same method but changing the product terms (formal-informal, only formal, only informal). Visually, an adjacency matrix resulting from the bipartite projection of an incidence matrix can be represented as a network of connected nodes, the strength of the connection being the value of the matrix corresponding to the nodes i and j. Figure B.1 in Appendix B represents the adjacency matrix of all domestic activities in Ghana as a network of related industries. Further description of the network structure and a comparison for different thresholds of the specialisation index are provided in Appendix B.1, The randomisation tests presented in Appendix B.1.1 provide evidence on the non-randomness of the patterns described by the co-location matrix.

All the matrices above are then transformed to standardise the weights in such a way as to have comparable measures across the co-location and the occupation similarity matrices. One approach to constructing a similarity matrix between industrial sectors is to use the cosine of the angle between all pairs of columns in the industry-district incidence matrix $\mathbf{II^{CL}}$. In the new matrix, the individual elements i_{ij} have been replaced by ρ_{ij} such that:

$$\rho_{ij} = \frac{x_i \vec{x}_j}{\sqrt{\vec{x_i^2}}\sqrt{\vec{x_j^2}}} \tag{3.4}$$

where x_i and x_j are distinct column vectors of the $\mathbf{II^{CL}}$ matrix, with $i \neq j$. A comparison between the cosine similarity and an alternative measure (minimum conditional probability, as in Hidalgo et al. 2007) is proposed in Appendix B.2.

3.3.2.2 Relatedness in capabilities

This measure of relatedness is created using the Ghana Labour Survey 2015. After performing a concordance exercise to transform the industry classifications from ISIC Rev. 4 to ISIC Rev. 3.1 (consistently with the IBES 2014 data), a cross-tabulation of jobs and industries was created in order to assess the similarity between industries based on the number of occupations they share, which provides the grounds for a relatedness measure based on shared capabilities (proxied by occupations) across industries (as in Diodato et al. 2018). This matrix provides a useful measure of the extent to which industries share similar skills, which make possible the movement of labour across them and can help explain the patterns of industrial co-location. The process is identical to the one followed for the co-location matrices: starting from a cross-tabulation of industries and occupation retrieved from the Ghana Labour Force Survey 2015, an industry-industry matrix $\mathbf{II^{OC}_{p\times p}}$ is created, where connections between industries are determined by the extent to which such industries share the same occupational characteristics:

$$\underset{p\times p}{\mathbf{II^{OC}}} = \underset{p\times k}{\mathbf{OI^{T}}} \times \underset{k\times p}{\mathbf{OI}} \tag{3.5}$$

where $\mathbf{OI^{T}_{k\times p}}$ is the industry-occupation incidence matrix, which is then standardised using the cosine similarity approach introduced in the previous subsection.

3.3.2.3 Input-Output

As mentioned in Section 3.3.1, the technological linkages for the industry pairs are captured by their input-output relationship, drawing from the Ghanaian national IOTs compiled by Lenzen et al. (2012). As suggested by Timmer et al. (2018), the Leontief inverse (1953) was used to quantify the input-output linkages between two industries i and j:

$$\underset{s\times s}{\mathbf{II^{IO}}} = (\mathbf{I} - \underset{s\times s}{\mathbf{A^{D}}})^{-1} \tag{3.6}$$

The equation above is the Leontief inverse, where $\mathbf{I}$ is a square identity matrix with as many rows (columns) as the number of 2-digit ISIC industrial sectors s; and

$\mathbf{A^D}$ is the $s \times s$ matrix of domestic coefficients, where each element a_{ij} measures the amount of domestic value added from industry i which is used to produce one unit of output by industry j. The Leontief transformation allows us to factor in all domestic output, both direct and indirect, required to produce one additional unit of value added from a given industrial sector (Timmer et al., 2018). Unlike the co-location and relatedness measures, which are available at the 4-digit level of disaggregation, the IO-derived measure of technological linkages is only available at the 2-digit level of the same industrial classification (ISIC 3.1), and applies to 56 industries. Assuming a certain degree of homogeneity of technological linkages between 4- and 2-digit industries, every 4-digit industry pair is associated with the I-O measure relevant to the respective 2-digit industry.

3.3.2.4 Measuring complexity

The intuition behind the measurement of economic complexity is to combine iteratively product ubiquity and country diversification to construct a metric of product/industry complexity that depends (negatively) on the ubiquity of a product and (positively) on the diversification of the geographical area specialising in that product (in this case, the Ghanaian districts). Existing methods and metrics, such as the Economic Complexity Index proposed by Hausmann and Hidalgo Hausmann and Hidalgo (2009) or the Economic Fitness Index (Tacchella et al., 2012), present a number of shortfalls. Firstly, these metrics cover only a small number of service activities, as they rely on export specialisation. However, Ghanaian employment in services is predominantly in non-tradable services, and given the focus on informality in this study, the exclusion of such activities would represent a major loss of nuance in the co-location patterns. Related to the first point, the study considers specialisation in employment rather than exports, and it pursues a measure of industrial complexity that reflects that intention, consistently with the co-location measures presented earlier in this section.

The present study adopts the Economic Fitness method to compute an Industrial

Complexity Index (ICI) and the fitness of Ghanaian districts (Tacchella et al., 2012), using the information contained in the **DI** employment specialisation matrix of all sectors (formal and informal). Unlike its most known competing method (the method of reflections proposed by Hausmann and Hidalgo 2009), the former constructs metrics of fitness (for countries/regions) and complexity (for products/industries) which take into account the non-linear interactions that exist between the diversification of a region and the ubiquity of the industries in which they specialise. Applied to the context of districts and industries, the fitness method allows us to estimate a measure of industrial complexity which is more than just the average fitness of the regions specialising in that industry, but reflects the diversification of such districts and their fitness (for a detailed discussion of the differences between the two methods, see Pietronero et al. 2017; Freire 2021). The fitness and complexity algorithm can be illustrated by the equations below:

$$\begin{cases} \tilde{F}_d^{(n)} = \sum_i M_{di} Q_d^{n-1} \\ \\ \tilde{Q}_i^{(n)} = \dfrac{1}{\sum_i M_{di} \dfrac{1}{F_d^{n-1}}} \end{cases} \quad \text{with} \quad F_d^{(0)} = 1 \ \forall d \ \text{and} \ Q_i^{(0)} = 1 \ \forall i \tag{3.7}$$

where F_d is fitness of district d, which depends on industrial diversification of district d (extracted from the **DI** employment specialisation matrix) weighted by the industrial complexity of all industries in which the district specialises and Q_i is the industrial complexity (henceforth referred to as ICI), which depends inversely on the diversification of the districts specialising in industry i, weighted by the inverse of their fitness. The algorithm is run for $n = 20$ iterations, after which the fitness and complexity metrics remain stable.[5]

Using the ICI obtained for all 249 4-digit ISIC Rev. 3.1 industries available in the

[5]The algorithm has been implemented using the R package `economiccomplexity` (Vargas et al., 2020), available on the R CRAN repository: https://cran.r-project.org/web/packages/economiccomplexity/index.html.

IBES, the complexity differentials for each industry pair are derived as follows:

$$\Delta \mathrm{ICI_{ij}} = \mathrm{ICI_i} - \mathrm{ICI_j} \quad \forall\, i \neq j \tag{3.8}$$

where $\Delta \mathrm{ICI_{ij}}$ is the complexity differential between industries i and j. The ICI differential aims to capture the difference in industrial complexity between each industry pair, in order to unveil the role of industrial sophistication in the co-location of industries.

3.3.2.5 Employment ratios

Along with the main factors discussed above, industrial co-location can also be driven by the ubiquitousness of both activities across the country for unobserved reasons (such as physical endowments, infrastructure, favourable regulations). For this reason, in order to moderate the possible omitted variable bias, the estimated model features an additional control variable which attempts to capture the average predominance of the informal industry i over the formal j:

$$\frac{inf_i}{for_j} = \frac{\sum_d (\sigma_d^{inf} \times emp_{id}^{inf})}{\sum_d (\sigma_d^{for} \times emp_{jd}^{for})} \tag{3.9}$$

where σ_{id} is the share of employment of district d over the total, and emp_{id} is the absolute size of employment of district d in industry i, informal or formal. Ratios are then standardised between 0 and 1.

3.3.2.6 Final dataset

The final dataset consists of 35,812 industry pairs based on the 4-digit ISIC rev. 3.1 using the three data sources illustrated above. For each industry pair the following were computed: measures of formal, informal and informal-formal co-location (IBES 2014); a measure of relatedness based on shared capabilities between industries

(GLFS 2015); a measure of the input-output relationship between industries as a proxy for technological linkages (Eora); the complexity differential of the industry pair based on the aggregate productive structure (IBES 2014); a dichotomous variable identifying industries belonging to the same industrial macro-sector (1-digit ISIC); and a district-weighted ratio of informal-to-formal employment in the pair (IBES 2014).

3.4 Results

The results will be presented in three parts. Initially, a description of the variables identifying agglomeration patterns and their covariates is provided. We then move on to a deeper analysis of the aggregate formal and informal co-location patterns, as well as of the complexity of industries and districts, in order to identify those characteristics of industrial variety that can influence the broader process of structural change. Finally, the analysis will focus on explaining formal, informal and informal-formal co-location in terms of relatedness of capabilities, input-output relationships, complexity differentials and industry/employment factors.

3.4.1 Formal and informal co-location at a glance

To begin with, Table 3.1 provides an overview of the Ghanaian employment composition based on the macro-sector of the employing firm, as reported by the Integrated Business Establishment Survey (Ghana Statistical Service, 2015). It is important to recall that the IBES 2014 only covers non-household establishments; in fact, agricultural employment here appears to be relatively low, although it has been shown that agricultural activities represent the vast majority of household establishments, as shown in the previous chapter of this thesis (Chapter 2). Looking at the aggregate employment figures, service activities are the largest employer; this is especially true among informal activities, where wholesale and retail trade along with other services reach almost 60 per cent of total informal employment. As already suggested by existing evidence (Moreno et al., 2012; Kraemer-Mbula and Wunsch-Vincent, 2016; Diao et al., 2018a), the share of manufacturing employment is higher in the informal sector than in the formal, indicating that informal firms are also active in sophisticated industries. Overall, all activities in mining, utility supplies, ICT, finance and provision of basic services like education and health are predominantly formal. The formal sector accounts for 54.45 per cent of all employment, and 9.45 per cent of all firms.

Table 3.2 provides a description of the variables introduced in Section 3.3.2; the sample includes 35,812 industry pairs. Regarding co-location measures, it appears that these have a long right tail, with the majority of values being concentrated around zero. This is particularly a concern for the dependent variable, as its left-skewed distribution makes regressions to the mean less reliable. The high number of zeros hints towards the fact that not all industries co-locate, share capabilities or participate in the same supply chain; in particular the most skewed variable appears to be relatedness in capabilities, indicating that the majority of industries do not share any occupations.

Table 3.1: Employment shares across macro-sectors

ISIC Group	Tot. (%)	For. (%)	Inf. (%)	Tot. (#)	For. (#)	Inf. (#)
Agriculture, forestry and fishing	1.96	2.98	0.75	57,594	47,677	10,039
Mining and quarrying	1.38	2.17	0.43	40,551	34,718	5,756
Manufacturing	10.74	8.11	13.88	315,592	129,751	185,796
Electricity and gas supply	0.34	0.58	0.05	9,991	9,279	669
Water supply, sewerage, waste management	1.19	1.85	0.4	34,968	29,598	5,354
Construction	2.86	4.82	0.52	84,040	77,115	6,961
Wholesale and retail trade	23.22	11.64	37.06	682,314	186,227	496,081
Transportation and storage	2.46	3.67	1.02	72,286	58,716	13,654
Accommodation and food service	5.69	2.72	9.23	167,199	43,517	123,552
Information and communication	1.26	1.84	0.56	37,025	29,438	7,496
Financial and insurance activities	3.81	6.01	1.2	111,956	96,153	16,063
Real estate activities	0.32	0.47	0.14	9,403	7,519	1,874
Professional, scientific and technical	2.54	3.91	0.91	74,637	62,556	12,181
Administrative and support service	3.35	5.29	1.04	98,439	84,636	13,923
Public administration and defence	5.46	9.72	0.37	160,441	155,508	4,953
Education	15.54	22.15	7.63	456,639	354,375	102,134
Human health and social work	4.32	6.82	1.33	126,942	109,111	17,801
Arts, entertainment and recreation	0.54	0.49	0.6	15,868	7,837	8,031
Other service	12.91	4.59	22.85	379,357	73,430	305,867
Activities of households as employers	0.01	0.01	-	294	160	-
Extraterritorial organizations	0.1	0.15	0.03	2,938	2,400	402

Note: ISIC groups correspond to the ISIC 1-digit classification of industries. The first three columns indicate the percentage of overall, formal and informal employment in each industrial group. The last three columns break down employment across industries in absolute terms. Source: Integrated Business Establishment Survey 2014.

Table 3.2: Main variables: descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Pctl(25)	Median	Pctl(75)	Max
Informal-formal co-location[a]	35,812	0.108	0.121	0	0	0.085	0.191	1
Formal co-location[a]	35,812	0.128	0.140	0	0	0.105	0.214	0.894
Informal co-location[a]	35,812	0.130	0.124	0	0	0.118	0.219	0.816
Rel. in capabilities[a]	35,812	0.021	0.105	0	0	0	0	1
Informal-formal ratio[c]	35,812	0.0004	0.012	0	0	0	0	1
Input-Output Eora[b]	35,812	0.1	0.261	0	0.001	0.005	0.03	1
Complexity differential[c]	35,812	−0.008	0.071	−1	−0.0001	−0.00001	0.00001	1
Same ISIC 1d	35,812	0.187	0.390	0	0	0	0	1
Ubiquity informal	35,812	37.764	34.395	1	13	27	49	184
Ubiquity formal	35,812	25.089	27.389	1	7	15	34	184

[a] Variable standardised using the cosine transformation
[b] Coefficients from the inverse Leontief matrix
[c] Variable standardised between 0 and 1

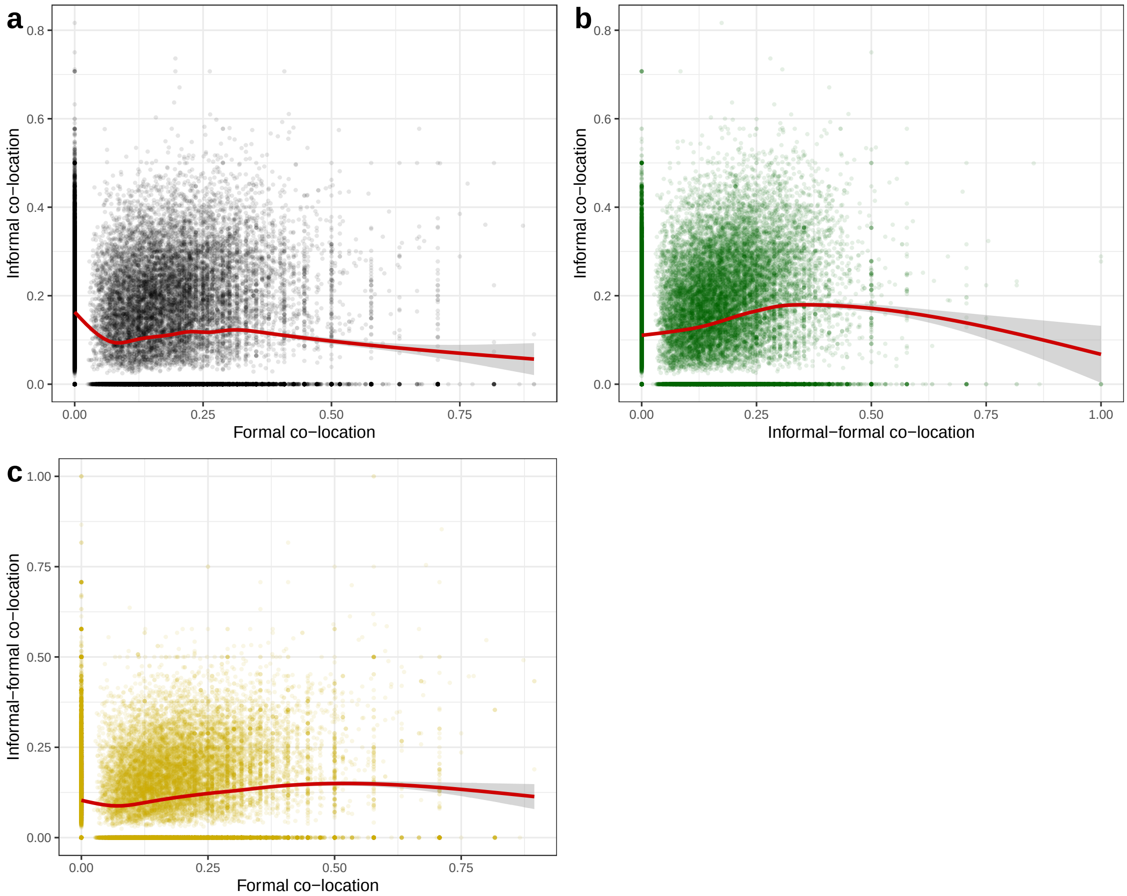

Figure 3.2: Correlations between co-location measures of Ghanaian industries. Panels a, b, and c in the figure plot the correlation between all the possible combinations of co-location measures (formal, informal, and informal-formal). The red line is the Generalised Additive Model fit line between the two variables in each plot.

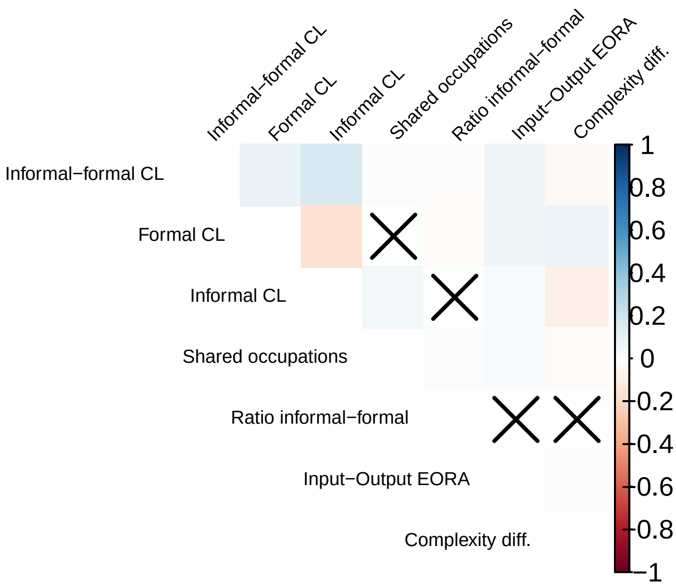

Figure 3.3: Correlations between main variables. Each cell in the correlation matrix indicates the magnitude (red if negative, blue if positive) of the correlation between the variables in each row and column of the matrix. Crossed cells indicate statistically insignificant correlations.

Despite their similar distribution, the co-location variables present relevant differences. The scatterplots in Figure 3.2 describe the pairwise relationship between these three co-location variables, in order to show the extent to which industry pairs co-locate similarly (or not) when the pair is made of informal, formal or informal-formal industries. Each plot shows the correlation patterns between each possible pair of co-location measures, enriched by a non-parametric fit that summarises each correlation. The data is fitted using a Generalised Additive Model, which shows similar correlatiomal-formal industiesn patterns – although with some differences – across panels a), b) and c). Extreme cases are identified by the high number of values lying on both axes of the figure; in panel a), the dots lying on the horizontal axis represent industry pairs that co-locate when firms are formal, but do not in the opposite case; the same applies to the variables plotted in panels b) and c). However, it must be noted that the majority of the values are different from zero, as can be easily inferred from Table 3.2.

Overall, the high variation between the three co-location measures represents a first exploratory finding on the heterogeneous nature of industrial variety: the factors that drive the co-location of informal industries may not be the same as the drivers of formal industrial clusters, and may also differ from the forces that lead informal industries to co-locate with formal ones. This is partly corroborated by Figure 3.3, which shows the correlations with p-value < 0.05 between all continuous variables from Table 3.2. The top-left region of the matrix confirms the varying degrees of correlation between co-location variables, with only the formal/informal pair being negatively correlated. This last coefficient further confirms that industries that co-locate when formal do not necessarily do the same in the informal sector, and vice versa – highlighting that co-location is driven by additional factors to those characteristics inherent in industry.

Figure 3.3 also unveils some other interesting relationships: the relatedness measure based on shared occupations appears significantly but weakly correlated to informal, but not to formal, co-location; informal co-located industries tend to share capabilities more than formal ones, which do not appear to share capabilities at all. This could be due to the fact that formal industries tend to emerge as enclaves rather than parts of the industrial ecosystem, and provides initial grounds to identify the relatedness of informal and formal industrial variety. By the same token, the low correlation between shared occupations and informal-formal co-location suggests a low degree of relatedness of the mixed (informal and formal) industrial variety, meaning that informal-formal industrial clusters mix heterogeneous levels of complexity. In addition, input-output measures correlate with all measures of co-location, although only weakly, suggesting some degree of integration of co-located industries. Another interesting relationship is the one between complexity differentials (taken in absolute values) and co-location variables: a larger complexity differential appears to be in a negative relationship with informal co-location, suggesting that formal firms may create clusters which are homogeneous in terms of complexity; while formal industrial clusters move in the opposite direction, as the larger the complexity differential of an industry pair, the higher its co-location. Finally, the informal-formal ratio has only feeble relationships with all covariates, indicating a limited role of the ratio of informal to formal activities within districts in driving the co-location measures.

3.4.2 Agglomeration patterns

We move now to a geographical analysis of economic diversification, in order to describe the industrial variety of districts, which are measured using their DIVERSITY, i.e. the row sums of the district-industry specialisation matrix (**DI**) introduced in Section 3.3.2.1:

$$DIVERSITY_d = \sum_i M_{di} \tag{3.10}$$

where M_{di} are the row elements of the district-industry specialisation matrix. The measure of DIVERSITY corresponds to the total number of industries in which each district specialises. Measuring economic specialisation and diversification of formal and informal activities separately reveals useful to obtain a more nuanced although general insight on the patterns of economic agglomeration across districts.

The visual representation of the specialisation matrices introduced in Section 3.3.2.1 (Figure 3.4) shows the patterns of specialisation in formal and informal activities for different cutoff levels of the specialisation index, computed following Equation (3.1), where the rows and columns have been reordered according to the sum of values in the rows/columns from left to right (top to bottom). The triangular shape of the matrix indicates a pattern of specialisation of Ghanaian districts, which is however more evident using the lowest cutoff ($Spec.Index = 1$), adopted in the upcoming analysis. In general, industries appear to be "nested" (Bustos et al., 2012), giving the pattern of specialisation a triangular shape; even more diversified districts (columns on the right) maintain specialisation in industries in which less diversified districts also specialise, with the exception of the most diversified districts which lose specialisation in the most ubiquitous industries (as shown by the black spaces in the top-right-hand corner of the matrix). Across types of activities, informal ones show a more uniform pattern of specialisation, while the formal specialisation matrix reveals that a small number of districts specialises in a large number of industries, while the majority specialises only in a few.

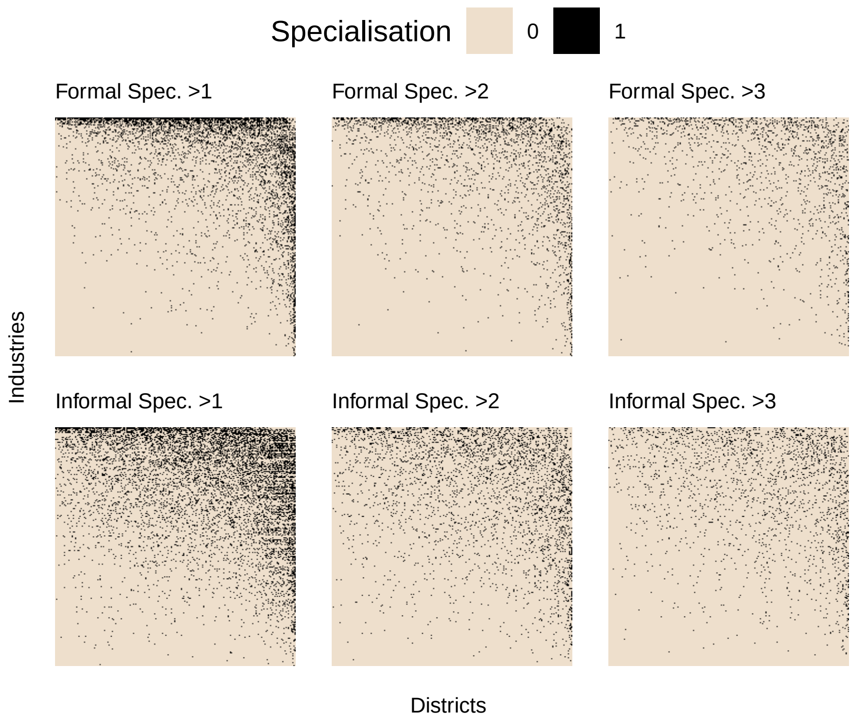

Figure 3.4: Patterns of specialisation across Ghanaian districts. The graph plots the formal (top row) and informal (bottom row) specialisation patterns of Ghanaian districts, for increasing specialisation cut-off values (from left to right). Rows indicate industries, columns indicate districts; cells coloured in black indicate district specialisation in a given industry. To unveil the triangular shape of the matrix, columns have been ordered from the least diversified (left) to the most diversified (right) district.

Moreover, informal activities tend to cluster more than formal ones, as can be inferred from summing up the diversity of each district in the two scenarios ($DIVERSITY_{inf}$ = 6,970; $DIVERSITY_{for}$ = 4,200). Figure 3.5 provides a more fine-grained description of formal and informal concentration of industries: while formal activities are more concentrated in a small number of districts (left panel, particularly in the Greater Accra and Tema metropolis areas), informal activities appear to be concentrated in a higher number of districts (right panel). This result points towards the fact that districts show, on average, higher informal diversification than formal. This is an important finding, as it has been argued that industries linked to localised demand – such as informal activities – tend to agglomerate less than formal industries serving external demand (Hidalgo et al., 2007; Delgado et al., 2016). This high tendency of informal firms to concentrate suggests that the analogy between informality and insignificance could be misleading, as informal industries may seek agglomeration to benefit from the same externalities that have been observed in formal clusters, contributing to the related variety as shown by the positive (although weak) role of shared occupations between informal co-located industries. Building upon this finding, the drivers of informal co-location will be analysed in the next subsection, and compared with those of informal-formal co-location.

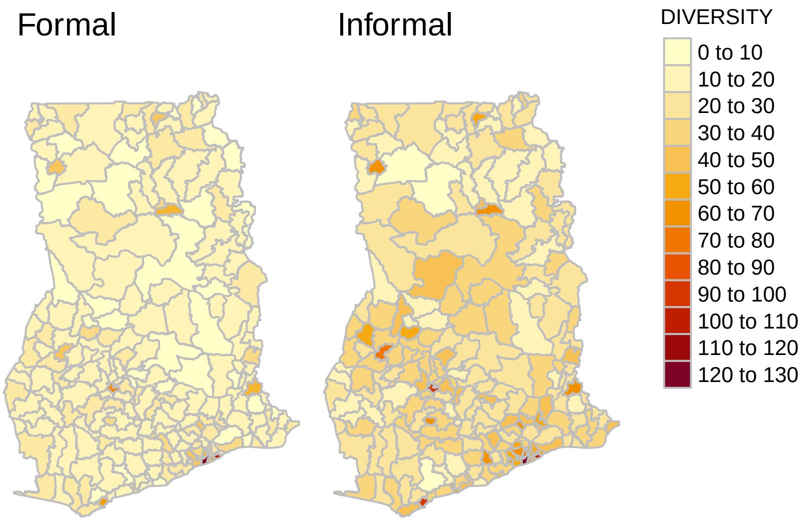

Figure 3.5: Formal and informal diversity index of Ghanaian districts. The colour of each district indicates the number of formal (left) and informal (right) industries in which each district is specialised. Interactive versions of these maps are available at https://bernacalda.github.io/formalghana/ (formal) and https://bernacalda.github.io/informalghana/ (informal).

A general description of the fitness of each district is offered by Figure 3.6. Considering specialisation in both formal and informal activity, district fitness is measured using the first element of the fitness algorithm, F_d, presented in Equation (3.7). It is worth recalling that the fitness of a district is a measure of the complexity of its economic structure, which is higher if the district specialises in a high number of industries with high complexity (and lower in the opposite case). From the map it emerges clearly that metropolitan areas are the ones with highest fitness due to higher agglomeration of economic activity, both formal and informal, resulting from the co-location of many non-ubiquitous industries. A full list of districts and their fitness is provided in Appendix B.3,[6] and it shows that some districts owe their fitness to their informal activities. For instance, Kma (Kumasi Metropolitan Assembly), a large metropolitan area situated in the centre-south of the country and distant from the coast and the capital city of Accra, scores high in aggregate and informal fitness, but lower fitness – in terms of ranking – if the measurement is restricted to only formal activities, corroborating the importance of informality in providing a full picture of the industrial variety.

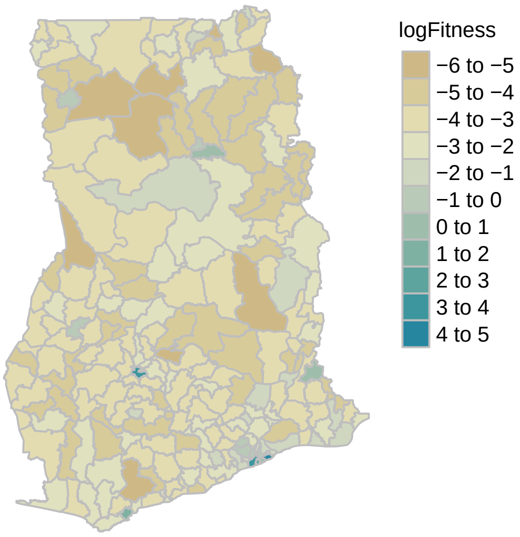

Figure 3.6: Fitness of Ghanaian districts. Each district is coloured based on the magnitude of its (logged) fitness. A dynamic version of this map is available at https://bernacalda.github.io/complexityghana/.

To further highlight the difference between formal and informal industries, Tables 3.3 and 3.4 compare the (log of) ICI and the UBIQUITY (number of districts specialised in a given industry i) for each formal and informal industry.[7] For the full ranking, see Appendix B.4. We start by noting that the most complex industry in Ghana (Mining of uranium and thorium ores) is absent in the informal sector, likely due to their high level of technological sophistication and trade intensity. Moreover,

[6]It should be noted that the fitness and ICI magnitude cannot be compared between formal, informal and all industries; comparisons can only be made with respect to the rankings of each industry across definitions.

[7]See previous note.

some industries appear to be relatively more complex when formal than informal (Advertising, Manufacture of builders' carpentry and joinery) and the other way around (Sea and coastal water transport). It could be hypothesised that such difference may be due to the very nature of the tasks and specific type of production carried out in the industry when activities are formally registered. However there is no way to test it empirically using the data available.

In addition, Figure 3.7 shows the relationship between the ICI for formal (X axis) and informal (Y axis) industries, highlighting the overall complexity (colour scale) and difference in formal and informal ubiquity (size of the dots). The graph shows that overall, and despite the differences in ubiquity across formal and informal sectors, the measures on the axes tend to correlate. It must be noted that industries with higher overall complexity tend to have consistent measures of formal and informal ICI, with lower differences in ubiquity (also because those are the least ubiquitous); higher differences in ubiquity and formal-informal ICI tend to be observed for industries with lower overall complexity.

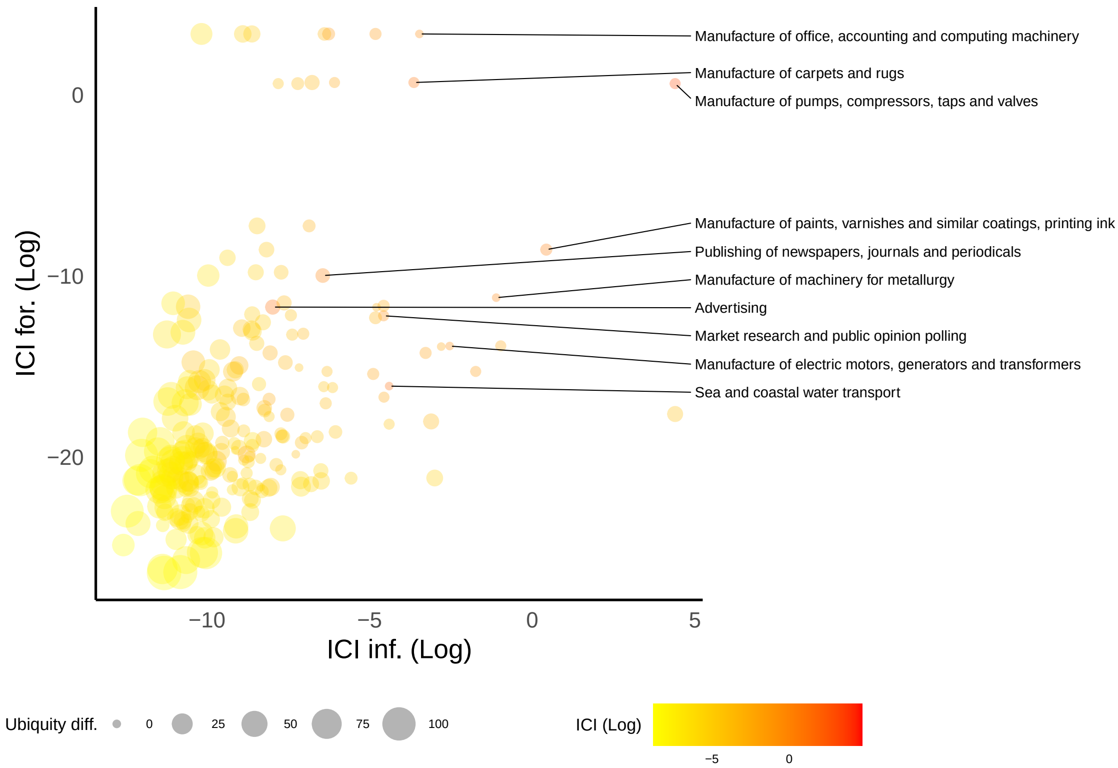

Figure 3.7: Differences in ubiquity across industries. The graph plots the informal (X axis) log ICI against the formal (Y axis) log ICI. The colour of the dots (industries) indicates the overall log ICI of each industry; the size of the dots represents the difference in formal and informal ubiquity of each industry.

Table 3.3: Top 10 4-digit ISIC industries by complexity

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
1	1	-	1200	Mining of uranium and thorium ores	4.70	1	3.34	1	-	-
2	-	1	3694	Manufacture of games and toys	4.02	1	-	-	4.39	1
3	3	-	2915	Manufacture of lifting and handling equipment	3.41	2	0.6	2	-	-
4	3	1	2912	Manufacture of pumps, compressors, taps and valves	2.97	3	0.6	2	4.39	1
5	7	-	1310	Mining of iron ores	2.43	2	-9.17	2	-	-
6	49	13	6110	Sea and coastal water transport	1.47	3	-16.09	4	-4.4	4
7	15	60	7430	Advertising	1.37	4	-11.73	5	-7.98	12
8	2	12	1722	Manufacture of carpets and rugs	1.10	2	0.66	1	-3.64	2
9	5	2	2422	Manufacture of paints, varnishes and similar coatings, printing ink and mastics	0.58	3	-8.56	3	0.43	1
10	9	33	2212	Publishing of newspapers, journals and periodicals	0.11	2	-9.99	2	-6.44	8

Table 3.4: Bottom 10 4-digit ISIC industries by complexity

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
204	206	126	7511	General (overall) public service activities	-8.04	84	-25.30	113	-10.03	21
205	208	190	8511	Hospital activities	-8.12	118	-26.18	144	-11.37	66
206	81	198	2022	Manufacture of builders' carpentry and joinery	-8.16	91	-18.64	10	-11.98	81
207	144	202	5040	Sale, maintenance and repair of motorcycles and related parts and accessories	-8.38	109	-21.31	25	-12.15	96
208	110	199	5520	Restaurants, bars and canteens	-8.45	139	-19.91	19	-12.01	117
209	209	166	8021	General secondary education	-8.50	144	-26.35	168	-10.82	62
210	143	200	5211	Retail sale in non-specialized stores with food, beverages or tobacco predominating	-8.55	149	-21.28	29	-12.06	121
211	180	203	1531	Manufacture of grain mill products	-8.64	138	-22.99	26	-12.45	123
212	210	185	8010	Primary education	-8.73	184	-26.40	195	-11.32	85
213	204	204	9191	Activities of religious organizations	-8.78	160	-24.86	116	-12.57	148

It is also useful to compare the measure of complexity used here with more general measures, such as the Product Complexity Index (PCI) (The Growth Lab at Harvard University, 2019), which is calculated using global trade data. As complexity measures at the industry level are not publicly available, the complexity of products ranked by the PCI in 2014 have been grouped and averaged by industrial sectors.[8] The industry-averaged PCI shows a positive correlation with the Industry Complexity Index ($\rho = 0.056$, p-value = 0.53). The PCI is not available for non-tradable services, so the ICI values for service activities have been excluded in this comparison. Moreover, a perfect correlation between the two measures was not expected; as the PCI is computed without information about services, the lack of the latter influences the computation of the measure for the non-service industries.

3.4.3 Drivers of co-location

As the main element of interest here is to unpack the process of informal, formal and informal-formal co-location between industry pairs, their driving forces have still to be understood; it has been shown that informal industries agglomerate more than formal ones, and that industries that co-locate when informal may not do so when they are formal. This section analyses the drivers of formal, informal and informal-formal industrial co-location, in order to identify the degree of relatedness of the informal, formal and aggregate industrial variety of Ghana. Coherent and related co-location of industries can be indicative of favourable preconditions for agglomeration externalities which, in turn, may favour the growth of industrial clusters, labour flows across industries and economic diversification - three fundamental drivers of structural change. We investigate this for formal, informal, and informal-formal industry pairs.

Below, the econometric setting is introduced, along with the estimation results. Using the unique dataset at the industry pair level described in Sections 3.3.1 and

[8] Once more, the product-industry correspondence exercise has been aided by the `concordance` package (Liao et al., 2020) available on the CRAN repository for R: https://cran.r-project.org/web/packages/concordance/index.html.

3.3.2, where observations correspond to all the possible pairings drawn from the 249 industries in the dataset, we start by estimating the effect of geographical concentration, relatedness, and complexity differentials (plus controls) on informal and formal co-location separately, according to the empirical framework illustrated by Figure 3.8. The econometric specification is described by the equation below:

$$CL_{ij}^{i,f} = \alpha + \beta CL_{ij}^{i,f} + \gamma REL_{ij} + \delta IO_{ij} + \zeta |\Delta ICI_{ij}| + \eta X_{ij} + \epsilon_{ij} \tag{3.11}$$

where $CL_{ij}^{i,f}$ is the co-location matrix for informal (i superscript) or formal (f superscript) industry pairs; REL_{ij}^{r} is the relatedness measure for the industry pairs based on shared occupations/capabilities; IO_{ij} is the measure of technological linkages based on the input-output relationship between the two industries; $|\Delta ICI_{ij}|$ is the absolute value of the difference in logged complexity between industries i and j; and X_{ij} is a vector of industry dummies (one for each industry in the pair, and one identifying industries belonging to the same ISIC 1-digit category).

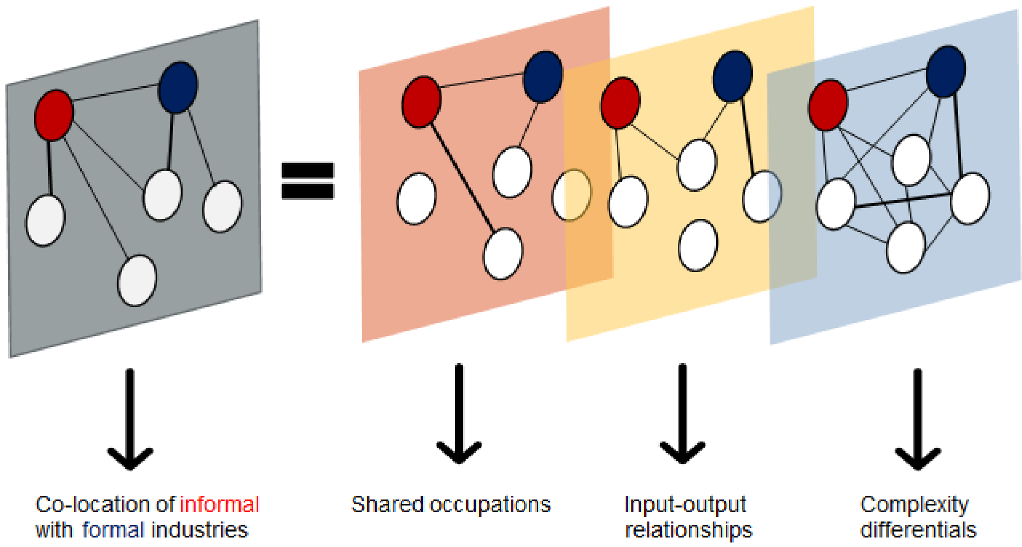

Figure 3.8: Empirical framework of the analysis. Each measure (co-location, relatedness of capabilities, input-output relationships and complexity differentials can be seen as an individual network, where dots represent industries, and edges measure the intensity of each measure. The co-location network (grey figure on the left) can be seen as the result of the overlap of the other three networks.

Under a Generalised Linear Model (GLM) approach (Nelder and Wedderburn, 1972), the model coefficients are estimated adopting the strategy proposed by Papke and Woolridge (1996) for fractional data, which accounts for the bounded nature of the dependent variable. The authors propose a functional form that assumes a binomial distribution for the bounded dependent variables and a logistic linkage between dependent and independent variables, using a quasi-likelihood estimation method. Gourieroux et al. (1984) show that for quasi-likelihood estimations, all parameters estimated maximising a likelihood belonging to the linear exponential family are consistent among themselves. Square and interaction terms for the measures that capture specific mechanisms of relatedness (capabilities, input-output and complexity) are also added to the specification.

The regression coefficients estimated by Equations (3.11) are presented in Table 3.5, which confirms expectations about the relationship between informal and formal co-location as they move in opposite directions: industries that co-locate when informal do not do the same when they are formal. However, given the non-linearity of the GLM fit, it is necessary to resort to marginal effects to obtain a meaningful interpretation of the results. Average Marginal Effects (AME) for the variables of interest, computed for models 2 and 6 in Table 3.5, are presented in Figure 3.9.

Relatedness of capabilities shows a reversed U-shape for informal co-location, but is not significant in its formal counterpart. With respect to the former, shared capabilities begin to have a negative effect on co-location for values higher than around 0.4, indicating that when informal industries share 'too many' occupations, they co-locate less. It is interesting to note that Boschma (2005) finds the same non-linear relationship for cognitive proximity and innovation, suggesting that 'too much' proximity may hamper the ability of firms to benefit from agglomeration externalities. This is also consistent with evidence from informal contexts which shows that, for interactive learning to happen, similarity in capabilities is fundamental (Fu et al., 2013). This finding also allows us to hypothesise that there may be room for labour flows across informal industries: this could be one of the mechanisms at the root of their greater tendency to agglomerate (as per Figure 3.5). The same effect is not

Table 3.5: GLM regressions: informal and formal co-location

	Dependent variable:							
	Informal co-location				Formal co-location			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Formal CL	-1.253^{***} (0.052)	-1.155^{***} (0.050)	-1.157^{***} (0.050)	-1.165^{***} (0.050)				
Informal CL					-1.545^{***} (0.063)	-1.449^{***} (0.062)	-1.464^{***} (0.063)	-1.461^{***} (0.062)
Cap. rel.	1.246^{***} (0.155)	0.959^{***} (0.151)	0.256^{***} (0.058)	0.921^{***} (0.150)	0.291^{*} (0.173)	0.177 (0.170)	-0.075 (0.062)	0.135 (0.170)
Cap. rel. sq.	-1.125^{***} (0.204)	-1.096^{***} (0.199)		-1.036^{***} (0.197)	-0.276 (0.223)	-0.363^{*} (0.219)		-0.400^{*} (0.218)
I-O.	0.142^{***} (0.022)	-0.089 (0.110)	0.162^{***} (0.026)	0.234^{***} (0.036)	0.284^{***} (0.024)	1.069^{***} (0.127)	0.111^{***} (0.033)	0.545^{***} (0.037)
I-O. sq.		0.217^{**} (0.095)				-0.912^{***} (0.114)		
|ICI diff.|	-0.192^{***} (0.015)	-0.008 (0.038)	-0.138^{***} (0.016)	-0.127^{***} (0.015)	0.185^{***} (0.011)	0.362^{***} (0.035)	0.182^{***} (0.013)	0.184^{***} (0.012)
ICI diff. sq.		-0.040^{***} (0.012)				-0.054^{***} (0.011)		
Same ISIC 1d	0.103^{***} (0.015)	0.137^{***} (0.019)	0.134^{***} (0.019)	0.164^{***} (0.021)	0.219^{***} (0.019)	0.392^{***} (0.023)	0.371^{***} (0.023)	0.503^{***} (0.025)
cap*IO			-0.382^{***} (0.095)				-0.444^{***} (0.134)	
IO*ICI			0.100^{*} (0.057)				0.005 (0.052)	
cap*ICI			-1.672 (1.172)				0.480 (0.557)	
IO*same1d				-0.140^{***} (0.045)				-0.755^{***} (0.051)
Constant	-1.769^{***} (0.008)	-2.220^{***} (0.042)	-2.199^{***} (0.041)	-2.211^{***} (0.042)	-1.849^{***} (0.010)	-2.809^{***} (0.047)	-2.828^{***} (0.047)	-2.809^{***} (0.047)
Pseudo R^2	0.031	0.12	0.12	0.12	0.041	0.102	0.1	0.105
2d ISIC i	NO	YES	YES	YES	NO	YES	YES	YES
2d ISIC j	NO	YES	YES	YES	NO	YES	YES	YES
Observations	35,812	35,812	35,812	35,812	35,812	35,812	35,812	35,812

NOte: The table reports the results of the estimation of Equation (3.11). The models are estimated via a quasi-likelihood estimation method, using a Generalised Linear Model approach. Models 1-4 use informal co-location as dependent variable; models 5-8 use formal co-location. The Pseudo R^2 is obtained by calculating the correlation coefficient between the dependent variable and its predicted values in the relevant model. Heteroskedasticity-robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01.

found in formal industries (column 6), where the effect of relatedness of capabilities has no role in driving co-location, seemingly supporting the hypothesis that many formal industries arise as enclaves, in isolation from the productive ecosystem.

The effect of input-output relationships is positive and significant across specifications: this finding confirms the evidence in the extant literature on co-location (Diodato et al., 2018), which consistently attributes a major role to input-output linkages in driving the co-location of industries. While the existing evidence is based on US data, this study further consolidates it by testing the same mechanism on formal and informal industries in Ghana – a middle-income country. Moving to the analysis of complexity differentials for informal industries, they appear to have a negative effect on co-location, indicating that informal clusters are homogeneous in terms of complexity: the smaller the difference in complexity, the higher the likelihood of co-location. On the other hand, in formal industries the opposite result is found – the greater the difference in complexity, the higher their degree of co-location. This challenges the evidence found by Balland et al. (2020), which finds that complex industries always agglomerate; this study finds the same for informal industries, but not for formal ones that present themselves in mixed clusters in terms of complexity.

The analysis further explores the marginal effects of relatedness on informal and formal co-location, looking at models 3 and 4 (informal), 7 and 8 (formal) in Table 3.5 where the three relatedness measures are interacted with each other. Across specifications, the only significant interactions are the ones between the measures of shared capabilities and input-output, along with a dummy identifying pair that belong to the same ISIC group (1-digit). For informal activities, Figure 3.10 shows that predicted co-location is highest for lower values of relatedness of capabilities and higher values of input-output relationship. In combination with results from models 2 and 7 (Figure 3.9), this finding points towards a story of complementarity among informal industries, which co-locate only with industries which perform different tasks (different occupations) but with whom they are horizontally integrated, given the homogeneous complexity of informal clusters. On the other hand, for formal

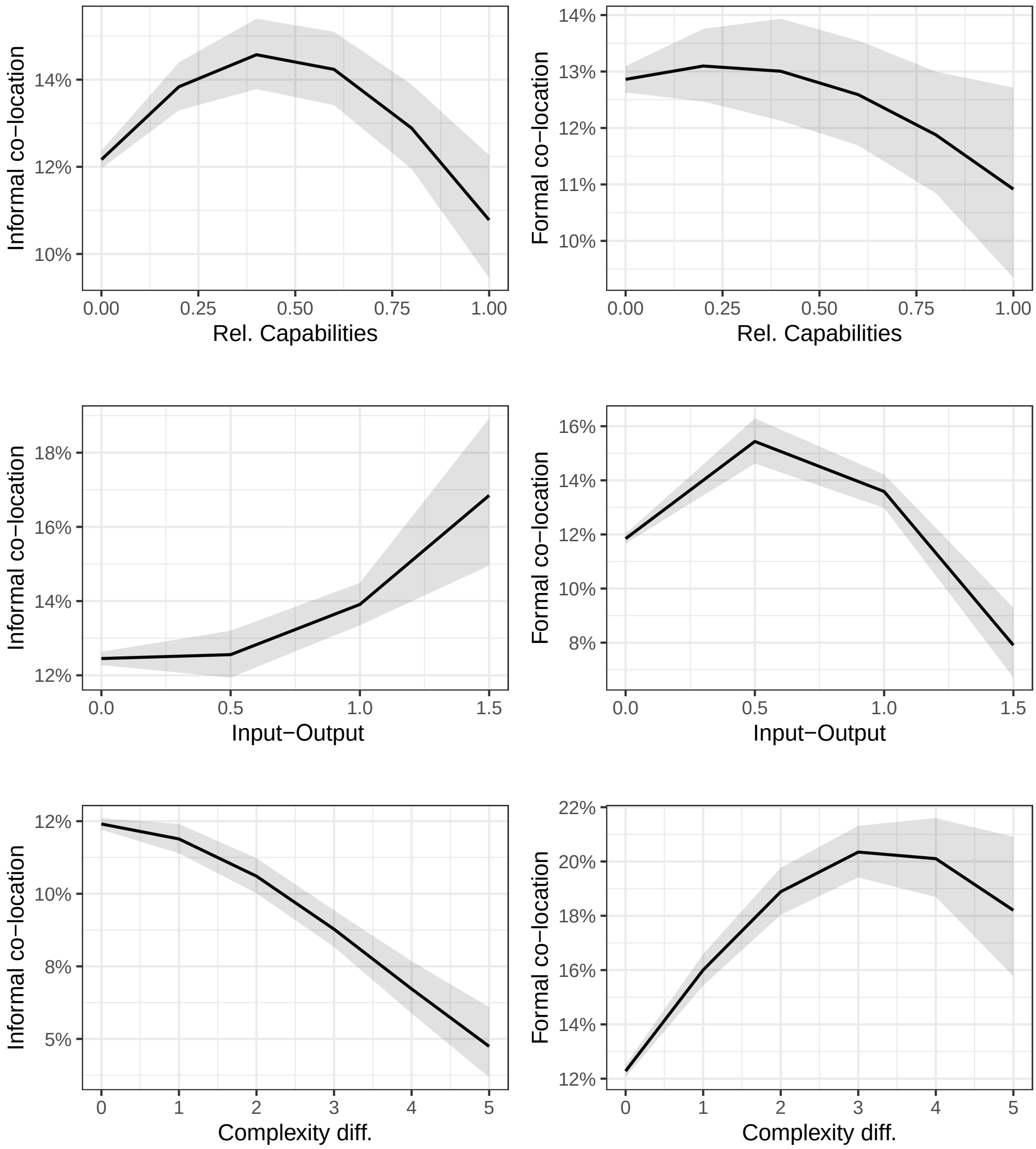

Figure 3.9: Average Marginal Effects of squared relatedness, input-output and complexity differentials on informal and formal co-location. Each graph in the figure plots the AME between the two variables indicated on the axis labels, estimated using model 2 for informal co-location and model 6 for formal co-location, summarised in Table 3.5.

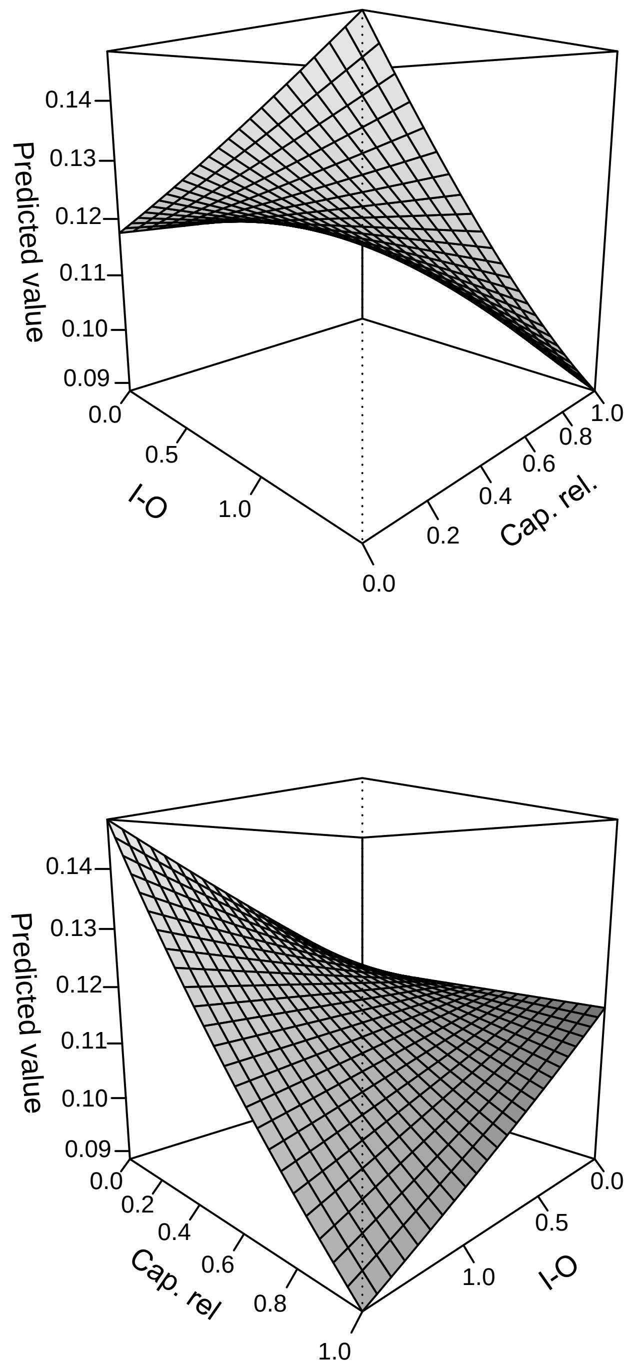

Figure 3.10: Interaction between relatedness of capabilities and input-out relationships in informal co-location, plotted in 3 dimensions, from two different angles. The Y axis indicates the informal co-location predicted values, obtained from the interaction between relatedness and input-output relationships. The graph is constructed using the model results from column 3 in Table 3.5.

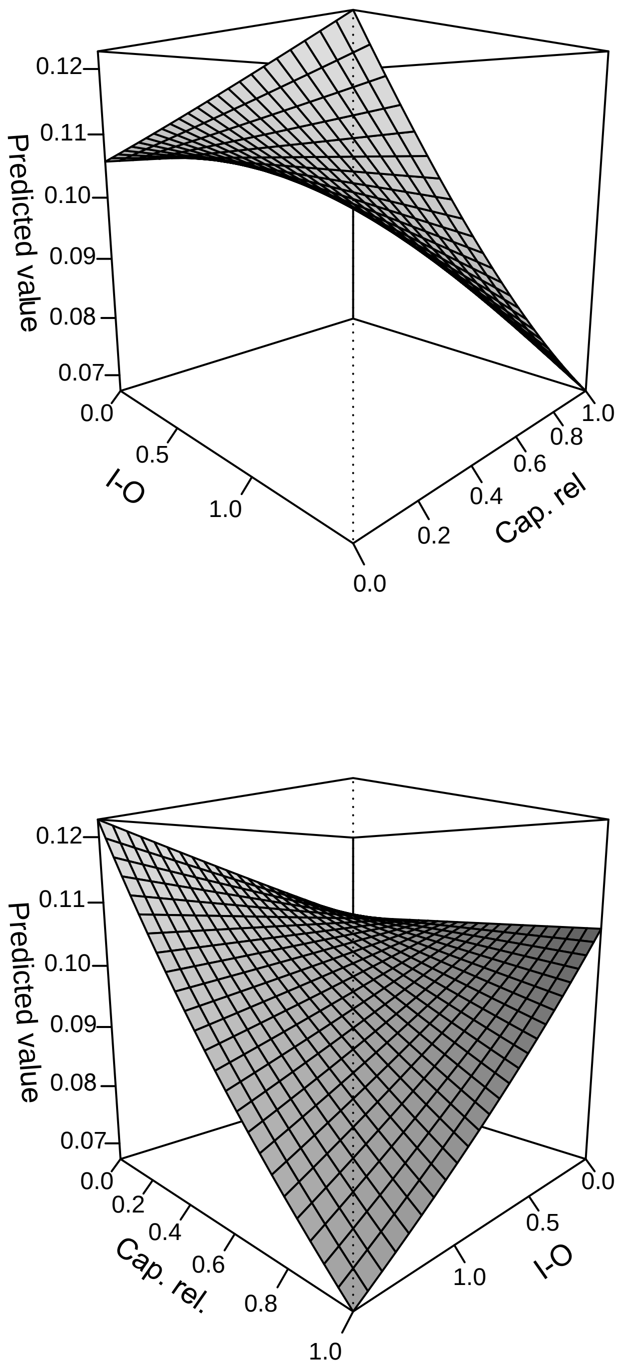

Figure 3.11: Interaction between relatedness of capabilities and input-out relationships in formal co-location, plotted in 3 dimensions, from two different angles. The Y axis indicates the formal co-location predicted values, obtained from the interaction between relatedness and input-output relationships. The graph is constructed using the model results from column 7 in Table 3.5.

activities (Figure 3.11), the highest co-location is achieved for the lowest values both of capabilities and input-output, supporting the formal enclave hypothesis. Looking at the interaction between input-output pairs within the same macro-sector (Figure 3.12), informal co-location increases with input-output linkages both within and outside macro-sectors, supporting the evidence on horizontal integration. On the other hand, co-location decreases for formal industry pairs with high input-output linkages within the same 1-digit sectors, and increases outside, making more plausible the hypothesis of vertical integration instead.

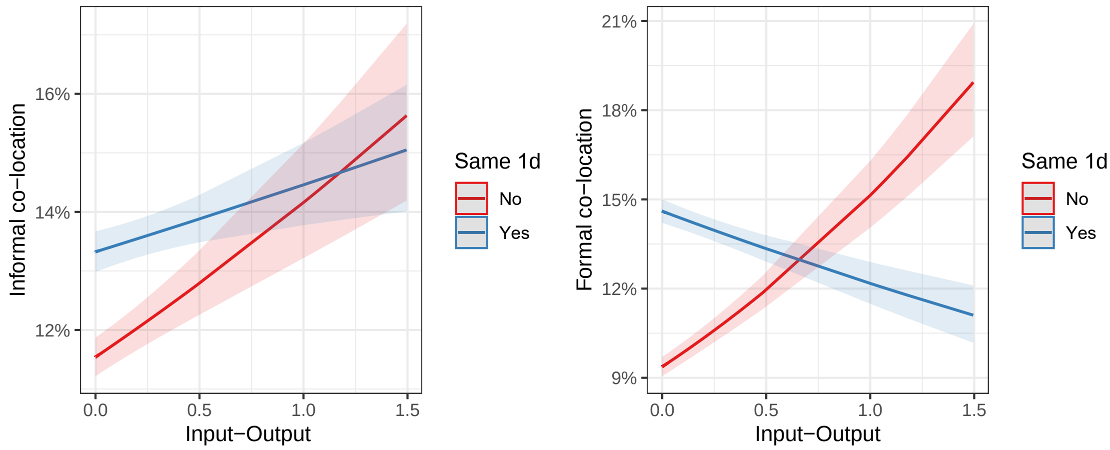

Figure 3.12: Effect of input-output linkages inside and outside the same macro-sector in informal (left panel) and formal (right panel) co-location. The graph plots the interaction between input-output relationships and a variable identifying industry pairs within (blue line) and outside (red line) the same ISIC 1-digit industry. The results for informal industries are shown in the left panel, using the model results from column 4, Table 3.5. The results for formal industries are shown on the right panel, using the model results from column 8, Table 3.5.

Moving on now to analyse the drivers of informal-formal co-location, the following equation is estimated:

$$CL_{ij}^{if} = \alpha + \beta CL_{ij}^{i,f} + \gamma REL_{ij} + \delta IO_{ij} + \zeta \Delta ICI_{ij} + \eta X_{ij} + \epsilon_{ij} \tag{3.12}$$

where CL_{ij}^{if} is the measure of informal-formal co-location; the vector of controls X_{ij} now also features the informal-formal employment ratio explained in subsection 3.3.2.5. Table 3.6 summarises the regression results for Equation (3.12).

From models 1 and 2 in Table 3.6 it emerges clearly that the high co-location of

Table 3.6: GLM regression results, informal-formal co-location

	Dependent variable:				
	Informal-formal co-location				
	(1)	(2)	(3)	(4)	(5)
Formal CL		1.070***	1.237***	1.238***	1.213***
		(0.050)	(0.048)	(0.048)	(0.048)
Informal CL		1.820***	1.137***	1.143***	1.135***
		(0.050)	(0.050)	(0.050)	(0.050)
Cap. rel.	0.520***	0.185	0.250*	−0.027	0.197
	(0.169)	(0.159)	(0.150)	(0.061)	(0.150)
Cap. rel. sq.	−0.532**	−0.222	−0.395**		−0.318
	(0.223)	(0.206)	(0.196)		(0.195)
I-O.	0.315***	0.256***	−0.094	0.173***	0.341***
	(0.023)	(0.022)	(0.114)	(0.026)	(0.032)
I-O. sq.			0.249***		
			(0.094)		
|ICI diff.|	−0.068***	−0.068***	0.097**	0.033*	0.034**
	(0.015)	(0.016)	(0.040)	(0.018)	(0.017)
ICI diff. sq.			−0.021		
			(0.014)		
Same ISIC 1d	−0.020	−0.065***	0.143***	0.147***	0.202***
	(0.019)	(0.018)	(0.022)	(0.022)	(0.025)
Ratio inf-for	−3.656	−3.102	−0.511	−0.440	−0.517
	(3.237)	(2.848)	(1.213)	(1.190)	(1.219)
cap*IO				0.108	
				(0.076)	
IO*ICI				0.008	
				(0.060)	
cap*ICI				−1.761	
				(1.635)	
IO*same1d					−0.281***
					(0.045)
Constant	−2.126***	−2.507***	−2.906***	−2.895***	−2.891***
	(0.008)	(0.013)	(0.057)	(0.057)	(0.057)
2d ISIC inf.	NO	NO	YES	YES	YES
2d ISIC for.	NO	NO	YES	YES	YES
Pseudo R^2	0.006	0.045	0.157	0.157	0.158
Observations	35,812	35,812	35,812	35,812	35,812

Note: The table reports the results of the estimation of Equation (3.12). The models are estimated via a quasi-likelihood estimation method, using a Generalised Linear Model approach. Columns 1-6 include the results of regressions using informal-formal co-location as outcome variable. The Pseudo R^2 is obtained by calculating the correlation coefficient between the dependent variable and its predicted values in the relevant model. Heteroskedasticity-robust standard errors in parentheses. *p<0.1; **p<0.05; ***p<0.01.

informal activities in general does a better job in explaining the informal co-location around formal industries across specifications. In other words, if two informal industries i and j are strongly co-located, the co-location of industry i (informal) with industry j (formal) is more likely than when i and j are both formal. This result also holds after controlling for 2-digit ISIC industry dummies of both informal and formal industries (columns 3-5). Looking at column 3, relatedness of capabilities becomes insignificant after controlling for informal and formal macro-sectors and for the co-location measures. Even if two industries are related because they share similar occupations, this does not increase their probability of co-locating, suggesting that there may be few or no grounds for exchange in employment across informal and formal industries.

At the same time, this result hints towards a certain degree of complementarity between informal and formal industries; if industries co-locate when they do not employ similar workers, they may perform tasks that require different skills in the industrial ecosystem of a given geographical area. The story of complementarity is partially strengthened by the positive effect of input-output coefficients, which correlates with the co-location of industries, although only weakly (Figure 3.13, top-right panel). The complementarity of informal activities to formal ones seems to indicate that the former may be integrated, to some extent, in the Ghanaian productive structure and is consistent with the findings of Diodato et al. (2018) who show that input-output relationships are the most important driver of co-location of firms. The results for differences in complexity of co-located industries indicate that they do not play a role in informal-formal co-location: informal-formal clusters appear to be mixed in complexity, again providing a new insight with respect to the extant evidence that more complex clusters tend to attract similarly complex ones (Gomez-Lievano and Patterson-Lomba, 2019; Balland et al., 2020). The evidence presented here suggests that this does not occur in the co-location of informal and formal industries.

To understand more about the nature of informal-formal co-location, the input-output

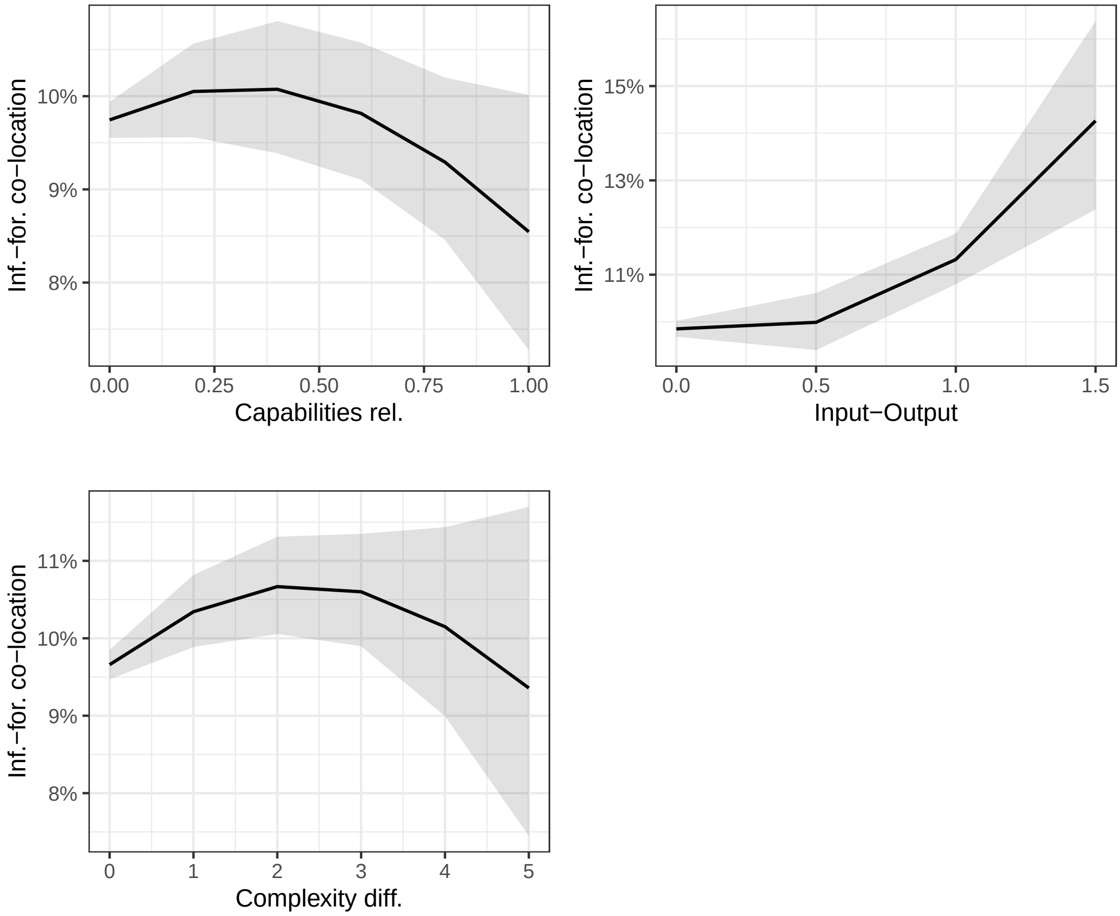

Figure 3.13: Average Marginal Effects of squared relatedness, input-output and complexity differentials on informal-formal co-location. Each graph in the figure plots the AME between the two variables indicated on the axis labels, estimated using model 3, summarised in Table 3.6.

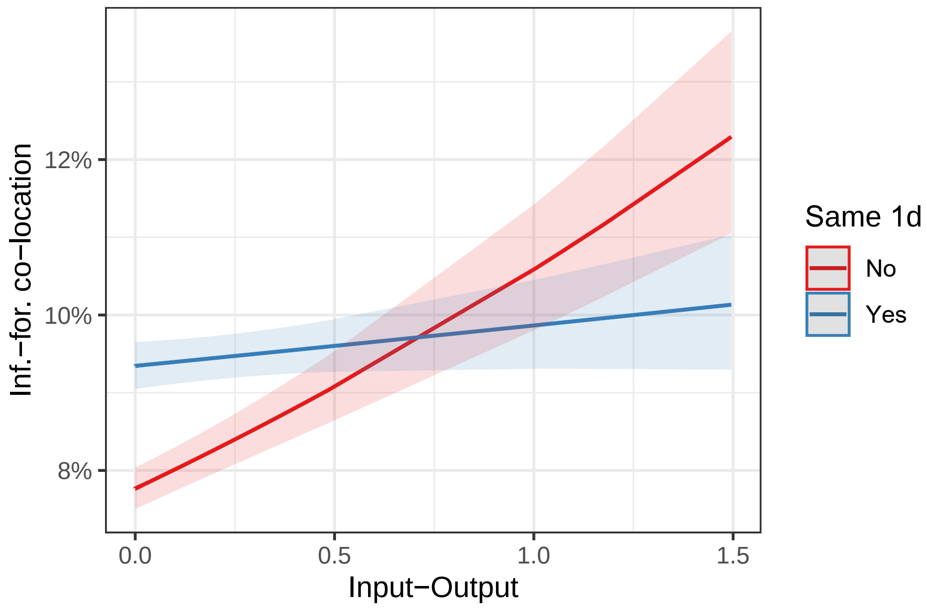

Figure 3.14: Interaction between input-out linkages inside and outside the same macro-sector in informal-formal co-location. The graph plots the interaction between input-output relationships and a variable identifying industry pairs within (blue line) and outside (red line) the same ISIC 1-digit industry. The results for informal industries are shown in the left panel, using the model results from column 5, Table 3.6.

variable is interacted again with the same macro-sector dummy: Figure 3.14 reveals that the marginal effect of input-output linkages on informal-formal co-location grows faster outside macro-sectors, indicating that this dimension of relatedness may be more relevant for vertical integration of informal and formal industries than within groups of similar industries. However, effects are rather weak.

Finally, the low coefficients and significance of the employment ratios seem to confirm that the co-location of informal and formal firms is not driven by the higher number of informal, compared to formal, firms — especially after controlling for both informal and formal industry dummies. It is worth noting that the increase in goodness of fit after controlling for informal and formal industry dummies suggests that the dynamics of co-location may follow directions which are dependent on the characteristics of specific industries, which differ widely in type and level of technology, capabilities, and knowledge. Overall, the low goodness of fit across models indicates that the process of industrial agglomeration in Ghana is difficult to explain in full.

To summarise, the results of the GLM estimation have shown that relatedness in capabilities is positively related to the co-location of informal firms, following an inverted U-shaped relationship. In the case of formal and informal-formal co-location, shared capabilities do not correlate to higher industrial co-location. Looking at input-linkages, the results indicate that higher buyer-seller linkages between 2-digit industries are positively associated with industrial co-location of informal and informal-formal 4-digit pairs, although they correlate negatively for formal industries. An analysis of the interaction between input-output linkages and 1-digit ISIC dummy demonstrates that informal industries co-locate similarly within and outside ISIC major groups, while formal and informal-formal industry pairs co-locate more outside, and less inside ISIC major groups. Complexity differentials show that informal industries co-locate with similarly complex industries, unlike formal ones, which co-located with industries with higher (lower) levels of industrial complexity. In the case of informal-formal industry pairs, complexity doesn't seem to explain their co-location patterns.

3.5 Conclusions

The persistence of informality has often been considered a major hurdle in the economic transformation of many low- and middle-income countries. However, given its large size and industrial heterogeneity, as well as the knowledge and capabilities 'in stock' within informal industries, it may represent an untapped opportunity. The dominance of informal activities in the economic landscape of low- and middle-income countries, as well as their embeddedness in the economic system, highlight the importance of research efforts oriented towards a deeper understanding of the mechanisms through which informal activities can feed into growth-enhancing structural change.

The analysis conducted here tests the hypothesis that the participation of informality to structural change may stem from its contribution to industrial variety, which hosts the conditions that allow inter-sectoral spillovers and agglomeration externalities. Another channel of participation is represented by the linkages with formal firms, from which informal industries – and the whole economy – can benefit in terms of firm and industrial upgrading. Using the case of Ghana, this chapter represents a pioneering attempt to explore the patterns of co-location between informal and formal industries in terms of relatedness. It does so by providing an extensive account of the characteristics of the Ghanaian productive structure, which is represented as a network of related (co-located) industries, and by estimating the role played by formal and informal geographical co-location, input-output relations, similarity in capabilities, and differences in complexity on the co-location of informal and formal industries. Despite the descriptive character of the results, which is imposed by the cross-sectional nature of the data, the empirical analysis carried out here provides some insightful evidence to expand our understanding of the contribution of informal activities to structural change through the lens of their co-location patterns.

To start with, the exploratory descriptive analysis of co-location patterns shows that industries follow different patterns of co-location when they are both informal (formal) and when industry i is informal and industry j is informal. Such variation indicates that industries that appear to be related when considering the formal sector may not be so in their informal counterpart, justifying an empirical approach that takes into account the specificity of informality. Secondly, the analysis unveils the high diversification of informal activities, and their higher tendency to agglomerate with respect to formal industries, pointing towards an under-exploited potential for inter-industry spillover. This finding also challenges the evidence that industries dominated by local demand do not agglomerate (Delgado et al., 2016); not only do they agglomerate more than formal ones, they also show a higher degree of complementarity – sharing capabilities and input-output linkages in clusters that are homogeneous in terms of complexity – when compared to formal industries, whose co-location is driven mostly by input-output linkages, forming clusters of mixed complexity. In fact, the results have shown that: i) relatedness in capabilities explains only informal co-location; ii) high input-output linkages are associated to higher informal and informal-formal co-location, but lower formal co-location; and iii) informal industries co-locate with similarly complex industries, while formal industries do the opposite, and informal-formal co-location is not correlated with differences in ICI. Overall, the fact that informal industrial variety appears to be more related than its formal counterpart represents a step forward in terms of our understanding of informality, where potential agglomeration externalities and contribution to economic diversification may have been hitherto underrated.

In exploring the drivers of informal-formal co-location across pairs of industries, we find that the only factor that significantly drives the co-location of informal with formal industries is their input-output relationship, and that this effect is stronger between, rather than within, macro-sectors. This result, in combination with the insignificant role of capabilities, outlines a form of complementarity where informal firms may be integrated in supply chains with formal firms, whereby they perform different tasks requiring different capabilities. Nevertheless, the fact that co-located industries do not share similar occupations could represent an obstacle to the upgrading of the informal labour force towards more sophisticated and formal activities. This finding is corroborated by the fact that informal and formal firms tend to form industrial clusters that are heterogeneous in terms of complexity, preventing both labour pooling and knowledge spillovers. On the basis of these findings, it can be argued that the informal industrial variety of Ghana represents much more than a store of homogeneous activities that should be absorbed by modern and more sophisticated formal sectors; on the contrary, their degree of relatedness and the linkages found in the analysis performed in this research show that informal industries hold many of the conditions that could trigger economic externalities leading to diversification, upgrading and labour flows across industries, which in turn can be seen as micro-level determinants of structural change. Furthermore, the findings summarised above raise an important point on formalisation. Becoming formal has often been considered (by academics and policy makers) the final goal for informal activities. While it can be desirable *per se*, formalisation of informal industries may not offer any contribution to structural change if formal industries remain unrelated to each other, and to informal ones. If co-located industries are not related, agglomeration externalities will not take place: according to the evidence presented here, these are more likely to happen among informal industries. This issue can only be tackled by policies that support, on the one hand, the formation of capabilities for small and informal producers, and that take advantage of the relatedness of informal activities to foster industrial development; and on the other hand, the development of both formal and informal industrial clusters that are related to the local industrial variety, in a way that allows formal industries to develop linkages from which both formal and informal can benefit.

These findings could also help to inform policy decisions oriented towards the transformation of Ghana's economic structure. While industrial policy may have the benefit of creating new employment opportunities (such as the 'One District, One Factory'[9] policy implemented by the Ghanaian government in 2016 to support job creation through industrialisation) it will have to take into account the effect of formal diversification and growth on informal activities. If the informal sector and small firms are not directly addressed by such policies, they will remain at the margins of the economy even when interacting with formal industries in the same supply chain. In line with suggestions by Kraemer-Mbula and Monaco (2020), the lack of support for small producers is likely to lead to a failure of development related, for example, to employment creation and industrial development. One way to pursue such goals and to draw upon the stock of resources brought by informality — especially in a context lacking external investment and technology transfer -– would be to ensure that a certain degree of relatedness exists between formal and informal industries to support the process of upgrading of the productive structure, and to create a formal sector that is able to absorb excess labour from informality.

[9]Source: https://1d1f.gov.gh/.

Chapter 4

Mobile Internet Adoption and Inclusive Structural Change: Evidence from Nigerian Non-Farming Enterprises

4.1 Introduction

While the internet has diffused rapidly in high- and middle-income countries since the '90s, some areas in low- and middle-income countries have lagged behind. In particular, Sub-Saharan Africa (SSA) and South Asia have experienced diffusion of the internet only since the mid-2000s (Figure 4.1). Increased access to the internet in Africa has been made possible by the rapid diffusion of mobile phones,[1] which, in addition to having radically and rapidly provided opportunities to connect distant individuals and markets in African economies (Aker and Mbiti, 2010), have now become the main means of accessing fast mobile internet services.

[1]In 2015, 16% of the population in SSA was using internet, but there were only 1.1 fixed telephone subscriptions per 100 people, against 70 mobile phone subscriptions ("World Development Indicators", accessed on 8 May 2022).

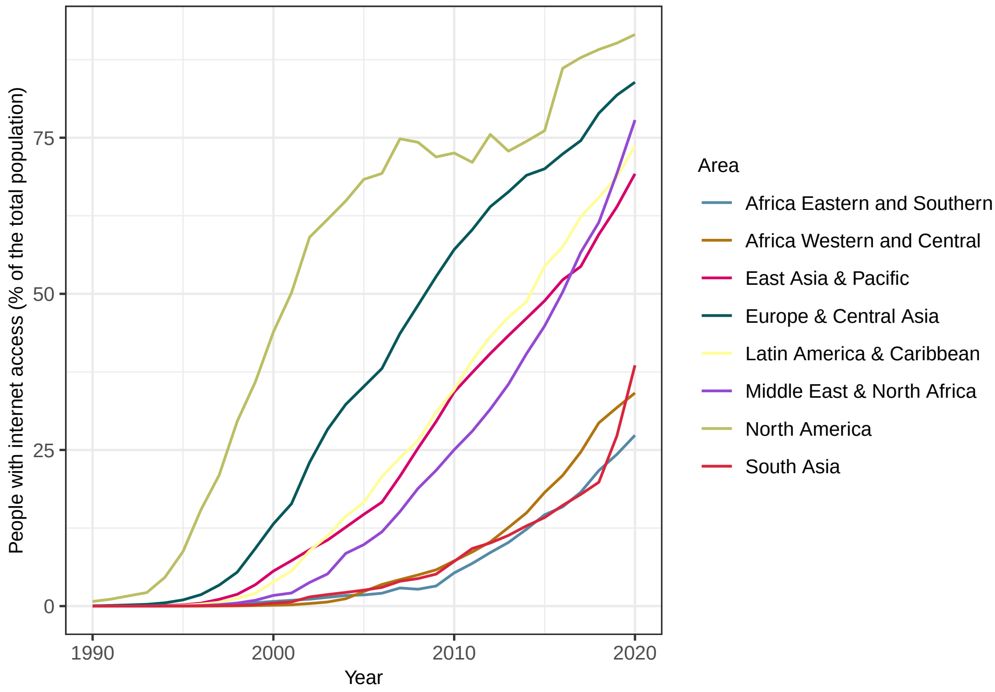

Figure 4.1: Percentage of population with access to internet services across the globe (1990-2020). Source: World Development Indicators.

It has been argued that the increased adoption of digital technologies, which allows access to internet services through mobile phones, creates a window of opportunity for low- and middle-income countries to undergo a process of structural change (Fagerberg et al., 2021; Kaplinsky and Kraemer-Mbula, 2022). However, the structural change experienced by SSA countries since the rapid diffusion of the internet has not always been virtuous (McMillan et al., 2014), at times associated with labour moving away from more productive industries (de Vries et al., 2015), such as manufacturing (Rodrik, 2016), and adding to the pervasive presence of informal services and activities.

On the one hand, there is evidence that access to fast internet in Sub-Saharan Africa has had a positive effect on employment (Hjort and Poulsen, 2019; Bahia et al., 2020, 2021; Caldarola et al., 2022), labour productivity growth (Hjort and Poulsen, 2019; Hjort and Tian, 2021), and firm entry (especially in services), due to lower operation costs (Houngbonon et al., 2022). On the other hand, it has been acknowledged that digitalisation may lead to the concentration of benefits in larger firms, excluding smaller players (Altenburg et al., 2021), and that digital technologies can be labour-saving or skill-biased technologies, favouring skilled workers over unskilled ones (Autor et al., 2003; Buera et al., 2021). Moreover, technical change and innovation can create new working opportunities in modern, emerging industries, making 'old' jobs obsolete, through a process of "creative destruction" (Schumpeter, 1934; Aghion et al., 2019). Ultimately, as the trajectories of digital technologies co-evolve with the skills required to operate them, and with firms' innovation routines (Ciarli et al., 2021a), predicting the effect of their adoption on employment and productivity remains empirically challenging. It is for policy makers to single out the conditions under which the diffusion of digital technologies could yield a process of structural change that is inclusive for all (Ciarli et al., 2021b).

This chapter contributes new evidence regarding the effect of internet access on firm performance and employment in low- and middle-income countries. The study is particularly interested in the extent to which such effects on performance and employment may lead to structural change and inclusion. Internet access can influence structural change through at least two channels. First, via an increase in labour productivity: complementing workers' skills, fostering human capital development, or improving firm-worker matching (Hjort and Tian, 2021). When firms become more productive, while also growing in size, they become able to absorb less productive labour; the shift of labour from less to more productive activities is one of the essential components of structural change. Second, the uptake of mobile internet can influence industrial composition, pushing labour and firms towards technologically intensive industries. However, the effect of internet-driven structural change will depend on: whether it leads to increased job opportunities, and therefore on whether its adoption will complement or replace labour; and on whether it allows the entry of new firms, for instance by reducing entry costs, or instead, by raising market barriers and favouring the incumbents. In order to answer these questions, this study estimates the effect of internet adoption at the industry level on structural change (measured in terms of firms' performance and industrial dynamics) and inclusion (measured by the creation/destruction of jobs and entrepreneurial opportunities). In so doing, the chapter empirically investigates whether internet adoption results in inclusive structural change (Saha and Ciarli, 2018; Ciarli et al., 2021b), that is, to a process where the impact of technology, leading to elements of structural transformation at the micro level, also lead to inclusive outcomes, such as the creation of employment and entrepreneurial opportunities. This concept is better explained in Section 4.2.1.

Because most non-agricultural employment in SSA is concentrated in the informal sector (Haggblade et al., 2010; Nagler and Naudé, 2017), the focus here is on household Non-Farming Enterprises (NFEs). These firms conduct a considerable amount of innovation activity (Fu et al., 2018; Avenyo et al., 2020), playing a potentially central role in the process of structural change, given their potential to upgrade (Kraemer-Mbula and Monaco, 2020). However, most non-farming informal activities have low productivity levels (Diao and McMillan, 2018); in order to contribute actively to the process of structural change, informal activities would need to increase their productivity (Diao et al., 2018a).

Within the focus of this thesis on structural change in SSA, the case of Nigeria is particularly relevant to the analysis of the impact of mobile internet adoption on informal economic activities for a number of reasons. First, the country is characterised by a negative or "growth-reducing" structural transformation (McMillan et al., 2014; Benjamin and Mbaye, 2020), due to its increased dependency on the oil industry, and to the shift of labour from high to low productivity industries. This process has resulted in the enlargement of the informal sector: according to the International Labour Organisation, 89 per cent of its non-agricultural employment was informal in 2018 (International Labour Organization, 2018).

Second, in 2015 Nigeria ranked sixth[2] amongst Sub-Saharan African countries for mobile services penetration, with 34 per cent of unique subscribers to either 2G or 3G services over the total population. Thirteen per cent of the population had a 3G subscription in 2015, whereas only 0.2 per cent had access to a fixed phone line in the same year (GSMA, 2015). Anecdotal evidence also indicates that the arrival of

[2]After, in descending order, the Seychelles, South Africa, Kenya, Namibia.

mobile internet has wiped out cyber-cafes – businesses that offered a shared space in which to connect to the internet and represented the only opportunity to do so before the diffusion of mobile internet. Most cyber-cafes either converted or closed down after the fast paced diffusion of mobile internet services.[3]

The analysis in this chapter relies on data from a nationally representative panel survey of Nigerian households, collected in three rounds between 2010 and 2015, namely the Living Standard Measurement Survey data. The survey includes relevant information on household non-farming enterprises, most of which are informal, along with information on socio-demographic characteristics, household employment and asset ownership, including possession of mobile phones. The dataset is complemented by information on fast mobile internet coverage (3G)[4] across Nigerian Local Government Areas (LGA). Combining information on geographical mobile internet coverage and mobile phone ownership at the household level, a measure of mobile internet adoption is constructed.

The roll-out of mobile internet is likely to be correlated with a number of factors that may influence both mobile operators' choice of where to build their mobile cell towers, and informal business activity. In order to address the endogeneity between the availability of mobile internet and entrepreneurial decisions that remains after controlling for household and industry unobserved heterogeneity, year fixed effects, and a number of social, demographic, and geographical control variables, the analysis adopts two distinct identification strategies. First, a Two-Stage Least Squares (2SLS) estimation exploiting the differential and time-invariant exposure of Nigerian Local Government Areas (LGAs) to lightning strikes (following Andersen et al. 2012; Manacorda and Tesei 2020; Guriev et al. 2021). Second, because the instrumental variable (IV) approach illustrated above will correctly identify the relationship between mobile adoption, productivity, and labour only in the absence of parallel trends, it is complemented by an event study that further corroborates

[3]From Quartz Africa: https://qz.com/africa/412745/how-mobile-internet-killed-off-cyber-cafes-in-nigeria.

[4]From now on, the terms fast mobile internet and 3G internet will be used interchangeably.

and expands on the findings provided by the 2SLS estimation.

The results indicate that the adoption of mobile internet at the household level has a positive and significant effect on the sales-to-workers ratio of non-farming enterprises, showing that the extant empirical evidence on the effect of the internet on labour productivity holds also for the informal sector.

However, interestingly, in terms of the effect on inclusive structural change, this chapter finds that the positive impact on sales per worker is likely to occur only in service industries, providing an incentive to invest particularly in retail and wholesale trade activities. As a result, we find that internet adoption reduces the probability that households will switch towards manufacturing industries.

In relation to inclusion, the results first indicate that the growth in sales per worker driven by internet access is due to both higher sales and lower employment. More specifically, households are likely to reduce the number of household employees working in the NFE and to retain – or even increase – the number of workers from outside the household. This could be because NFEs need skills which may not be available within the household, or because household workers move to better jobs in the formal sector. However, the reduction of household labour in NFEs is likely to be offset by increased waged opportunities outside the household, neutralising the labour-saving effect on household labour within NFEs.

Secondly, mobile adoption does not encourage the entry of new households into the market, indicating that only incumbent firms benefit from access to mobile internet. This last finding represents one possible exclusionary outcome of structural change: given the importance of entrepreneurship for innovation, employment creation, and structural transformation, the entry of firms will have to be encouraged and facilitated in order to foster employment creation and to sustain growth in productivity.

The remainder of the chapter is organised as follows. Section 4.2 illustrates the Inclusive Structural Change framework adopted in the study to motivate the analysis (Section 4.2.1) and reviews the extant literature on the economic effects of internet access on low- and middle-income economies (Section 4.2.2); Section 4.3 describes the data and the empirical strategy adopted in this study, along with the two identification strategies; Section 4.4 summarises the results; Section 4.5 concludes.

4.2 Analytical framework and related literature

4.2.1 Inclusive Structural Change

Technical change is a powerful engine of structural transformation (Kuznets, 1973; Dosi, 1988). The diffusion of new technologies enables the emergence of new, modern industries (Saviotti and Pyka, 2004), leading to productivity growth and the consequent reallocation of labour towards more productive industries, sustaining economic growth and creating new opportunities to innovate. The mutually reinforcing relationship between innovation and structural change constitutes a well established empirical and theoretical concept. However, innovation is also disruptive (Schumpeter, 1934) due to its effect on the organisation of the productive structure. Technical change directly affects income distribution (Paunov, 2013) by changing the labour demand towards skilled occupations. This can lead to a process of skill-biased structural change (Buera et al., 2021), from which only the most productive firms benefit, and to higher market concentration (Autor et al., 2020). Therefore, in an unequal society, innovation might cause structural change that benefits only a few, perpetuating exclusion (OECD, 2015; Aghion et al., 2019). However, these outcomes are not inevitable, as innovation can also be inclusive (Chataway et al., 2014). For instance, structural change towards labour-intensive technologies may increase the labour share of the economy, and create working opportunities for those who previously did not participate in the workforce. The direction of the effect of innovation, especially when it is diffused via technology transfer in low- and middle-income countries, will also depend on its degree of appropriateness to the specific context (Fu et al., 2011; Hanlin and Kaplinsky, 2016).

Against this backdrop, the empirical evidence shows that labour-saving (and skill biased) technical change has been one of the major culprits in premature de-industrialisation in African countries (Rodrik, 2016), pushing labour away from modern industries into the informal sector. Nevertheless, it has been argued that the transformative potential of digital technologies represents a major window of opportunity for low- and middle-income countries and informal firms/workers to leapfrog from their historic productive structure (Kaplinsky and Kraemer-Mbula, 2022). While digital technologies – such as the internet – have often been considered General Purpose Technologies (GPTs) and linked to productivity growth (Cardona et al., 2013), it is not possible to know *a priori* whether they will result in inclusive outcomes based on their diffusion and adoption (Saha and Ciarli, 2018), and how the skills required to master new technologies will co-evolve with the technology itself (Ciarli et al., 2021a).

The relationship between innovation, structural change and inclusion should be analysed in a way that acknowledges the trade-offs and synergies that exist between them. The direction of structural change determined by the adoption of a new technology, and its effect on inclusion, will depend upon the underlying economic and social conditions, such as the structure of inequality, market concentration, and power relations, alongside the specific nature and trajectory of the technology itself. Ciarli et al. (2021b) have proposed a framework that offers analytical support to empirically testable hypotheses, with the aim of unpacking the conditions under which innovation leads to inclusionary or exclusionary structural change in low- and middle-income countries. The Inclusive Structural Change (ISC) framework posits that the impact of innovation on structural change and inclusion will depend on the actors involved in the process of diffusion or adoption of the innovation; on their interactions; and on the initial conditions within which the process unfolds. Focusing on the role of innovation in inclusion and structural change (Figure 4.2), the framework makes it possible to test the effect of innovation on structural change and inclusion, without making any assumptions about the direction of that effect. The adoption of a new technology may generate a trade-off between structural change and inclusion. For instance, the new technology may lead to labour productivity growth across industries, while replacing labour with capital, which in turn will exclude part of the (especially unskilled) workforce. Conversely, the adoption of the same technology may create working opportunities that allow previously unemployed individuals to take a labour-intensive, low-productivity job, reducing the productivity levels in the economy.

Using the lenses of the ISC framework, the next section reviews the available empirical evidence on the impact of the adoption of fast internet on inclusion and structural change in low- and middle-income countries.

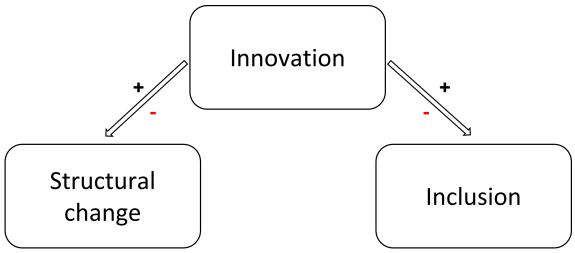

Figure 4.2: Inclusive Structural Change framework.

4.2.2 Internet and development in Africa

Internet adoption is associated with structural change via its positive effect on labour productivity across industries (Akerman et al., 2015; DeStefano et al., 2018; Hjort and Poulsen, 2019; Barrero et al., 2021), including agriculture (Gupta et al., 2020). The mechanisms underlying the relationship between internet adoption by firms and/or workers and labour productivity can be: direct, that is, via increased sales and/or labour replacement; via human capital development, allowing workers to acquire new skills; via better matching between firms and workers, allowing firms to identify the best and more skilled workers for their technology (Hjort and Tian, 2021). While this relationship appears to be rather uncontroversial in the empirical literature, the effects on inclusion of increased labour productivity induced by access to the internet are more mixed, as they seem to depend on the context, especially in low- and middle-income countries (Hjort and Tian, 2021) where evidence on the economic effects of internet diffusion is still scant.

Employment is one possible channel through which productivity gains induced by the diffusion of the internet can be associated with inclusion: if firms amend their labour requirements after adopting the internet, this will lead to exclusionary outcomes for the population in terms of working opportunities. Even with employment growth, firms may start to search for more skilled workers, leaving behind unskilled workers. Hjort and Poulsen (2019) exploit the staggered connection of fast internet submarine cables to the backbone networks of 12 African countries, and the differential timing in the connection of individual locations to those networks to estimate the effect of both internet access and speed on the individual probability of being employed. They find that individuals are more likely to be employed if they live in connected locations, although the effect on unskilled employment is only significant for individuals who had completed their primary education.

Focusing on mobile internet (3G), Caldarola et al. (2022) found a positive effect on aggregate employment across Rwandan districts. While both skilled and unskilled occupations grow, the former grow at a faster pace, while the latter grow more in absolute terms. Overall, the diffusion of fast mobile internet appears to steer structural change towards the creation of employment in services, and to discourage employment creation in manufacturing. A positive effect of mobile internet diffusion on labour market participation has also been observed across Nigerian households (Bahia et al., 2020), who further benefit in terms of welfare and poverty reduction. A similar approach has been adopted to study Tanzanian households (Bahia et al., 2021), showing that labour force participation is increased for young and highly educated male individuals.

Other possible effects of internet access on inclusionary/exclusionary outcomes are the creation of business opportunities in new industries, or the lowering of entry costs in existing ones. There is evidence that living in areas connected to broadband internet fixed lines has favoured the entry of firms in African countries, both formal (Hjort and Poulsen, 2019) and informal ones, although the latter mainly in services, where operating costs are lower (Houngbonon et al., 2022). Moreover, the diffusion of fast mobile internet in Tanzania has allowed highly skilled women to leave farming to take up self-employment activities (Bahia et al., 2021). Overall, the literature suggests that, unlike other forms of structural change, the diffusion of the internet in SSA has had positive effects on inclusion, favouring both employment creation and entrepreneurship, and that these effects are more likely to occur in services.

4.2.3 Gaps and contribution

A number of questions remain unanswered in the existing literature. First, there is little evidence of the effect of broadband internet on the performance of informal non-farming enterprises, which is where most of the population of SSA work. Houngbonon et al. (2022) found that, across African countries, the diffusion of mobile internet increases the entry of non-farming household enterprises. However, to the best of our knowledge, there is no evidence on the impact of the internet on two major aspects of structural change, those of productivity growth and industrial composition, in contexts where informality is prevalent over other forms of economic activity. In fact, the productivity of NFEs and their industrial specialisations are a key driver of structural change in areas where informality is prevalent (Nagler and Naudé, 2017; Diao et al., 2018a), because in these firms reside the capabilities that may or may not allow them to move to the productive formal sector and feed into the emergence of new economic industries.

Secondly, most of the firm-level evidence on the economic effect of internet access in Africa uses empirical strategies that look at the diffusion of land line internet. However, figures from Sub-Saharan Africa show that the number of fixed phone subscriptions decreased from 1.6 per 100 inhabitants in 2009 to 0.7 in 2019. On the other hand, the penetration of mobile internet has grown exponentially: in 2015, 34 per cent of the Nigerian population subscribed to 2G mobile internet services, and 12 per cent had 3G subscriptions (GSMA, 2015). These figures indicate that mobile internet has had a much broader outreach in African countries than the land line.

Thirdly, the available evidence focuses on the diffusion, rather than the adoption of the internet. While the presence of the internet (whether land line or mobile) may encourage economic activity, more evidence is needed on the direct effect of internet adoption on the performance of firms, the industries in which they operate, and how these may impact structural change and inclusion.

In order to address these gaps, this chapter first studies the effect of the adoption of fast mobile internet (3G services) on: i) the performance of Nigerian non-farming enterprises, measured as sales per worker; ii) the industry composition of such activities and their switches, or transitions, to other industries. Second, it studies whether these effects on structural change (sales per worker and industry composition) positively or negatively influence inclusion measured in terms of: i) NFE labour, ii) household labour, and iii) firm entry. The next section illustrates the data and methods employed to achieve these goals.

4.3 Empirical strategy

4.3.1 Data

The analysis performed in this chapter relies on a household panel dataset constructed by merging data from three sources – 1) the Nigeria Living Standard Measurement Survey (LSMS), for household and firm level data; 2) the Global System for Mobile Communications Association (GSMA) mobile coverage maps, for mobile internet diffusion; and 3) WWLLN Global Lightning Climatology and timeseries (WGLC). This is used to construct an instrumental variable based on lightning strikes (for more details, see Sections 4.3.1.3 and 4.3.3.1). Each data source is briefly discussed below.

4.3.1.1 Household data

The main source of information used for the construction of the dataset is the Nigeria Living Standard Measurement Survey – Integrated Survey for Agriculture

(LSMS-ISA). The LSMS-ISA dataset is a nationally representative, geo-referenced panel data source at the household level, gathered as part of a broad data collection effort coordinated by the World Bank in several low- and middle-income countries in order to strengthen their household survey systems and to better inform development policies. Nigerian households were interviewed repeatedly across three waves (in 2010/11, 2012/13 and 2015/16), and each wave was collected in two rounds – a post-planting visit and a post-harvesting one.

The main reason for adopting this data source is that it includes a non-farming enterprise module. NFEs are defined as all economic establishments owned and run by household members, including self-employed individuals and excluding agricultural activities such as pre- and post-harvest activities on agricultural plots. Household NFEs can employ both household and external labour, and are active in many industrial sectors. In the survey, households were asked to self-report information on their household businesses, such as their sales, costs, capital, labour, and industry of activity. With respect to the latter, the dataset adopts the International Standard Industry Classification (ISIC) rev. 3.1. In the analysis, firms will be classified using the ISIC major groups: agriculture, business and real estate services, construction, education, electricity and water, finance, fishing, health, hospitality, manufacturing, mining, other services, public services, trade services,[5] and transport. The analysis will often refer to 'services', a category that includes business and real estate, finance, hospitality, public, retail/wholesale trade, and other services. Following Behuria and Goodfellow (2019), the study also groups together a subset of services that have exhibited particularly high productivity growth in the same African context, such as finance, insurance and real estate services (henceforth, FIRE services).

The household modules of the survey include individual and household socio-demographic information, such as the age, gender, education, and employment status of individuals; and information on household-level asset ownership, consumption, expenditure, and non-farming businesses. The agricultural module includes

[5]Henceforth, 'trade' refers to retail and wholesale trade services.

information on all household agricultural plots. The geographical module holds information such as the coordinates of the Enumerator Area where the household is located (essentially the size of a village, or an urban neighbourhood), distance from the closest markets, roads and urban areas. Each NFE in the dataset can be associated with an owner and/or a manager, making it possible to merge firm-level data with relevant individual-level data about the NFE's owner, and related household-level information.

4.3.1.2 Mobile internet data

The data on mobile internet coverage is provided by the Global System for Mobile Communications Association (GSMA) in partnership with Collins-Bartholomew, a firm that specialises in map production.[6] The data used in this analysis gives the percentage of the population with access to 2G, 3G, and 4G mobile services in each of the 775 Local Government Areas (LGA).[7] While 2G services do not allow to use fast mobile internet, the first technology that allowed to access fast mobile internet broadband services is the 3G technology, followed by the 4G, which improved the quality and speed of mobile internet. However, in the time period under consideration, there was no 4G coverage in Nigeria, and for this reason, the analysis will focus on the adoption of 3G mobile internet technology.

The original data comes from submissions by the four mobile operators offering 3G services in Nigeria: MTN, Airtel, Etisalat, and Glo (GSMA, 2015). In order to construct the measure of mobile services coverage in Nigeria, the analysis follows the approach elaborated by Mensah (2021) to obtain measures at the LGA level, as follows. The original disaggregated data takes the form of a raster map with separate layers for each technology (2G, 3G, and 4G). Each cell in the raster represents a 1 x 1 km area, with dichotomous information on whether the cell is covered by a mobile signal. The GSMA raster is overlaid with the Global Population of the World (GPW) data (Centre for International Earth Science Information Network CIESIN

[6]The author is extremely grateful to Justin Tei Mensah for facilitating access to the data.
[7]LGAs correspond to the Admin-2 subnational boundaries of Nigeria.

Columbia University, 2018),[8] in order to capture the number of people with access to each mobile service. Then cells are aggregated at the LGA level to compute the total population covered by mobile internet in each LGA; this number is divided by the total population of the LGA to obtain the share of population covered by each mobile technology in each LGA and year. The results of this exercise are shown in Figure 4.3, which describes the staggered roll-out of the 3G mobile internet across Nigerian LGAs and over time.

The main focus of this chapter, however, is not to measure the effect of mobile 3G internet diffusion, but to capture the effect of technological adoption on firms and households. In order to identify the adoption of mobile internet at the household level, the district-level measure of 3G coverage has been weighted to reflect the number of mobile phones per capita in each household, obtained from the LSMS household data described in the previous subsection. In this way, a household in a district with high mobile internet coverage but where no one owns a mobile phone, is not counted as an adopter. This made it possible to refine the previously rather coarse measure of mobile internet diffusion at the LGA level, and to single out the effect of adoption of the new technology, as it is likely that some households will be covered by the infrastructure, but will not be able to access it due to income constraints.

4.3.1.3 Additional data

Two additional sources have been added to the household data and the mobile internet coverage data. The first is the WWLLN Global Lightning Climatology and Timeseries (WGLC) data (Kaplan and Lau, 2021),[9] which is employed to construct a variable to instrument the main measure of interest – mobile internet adoption – to pursue the identification strategy described in Section 4.3.3.1. The WGLC data

[8] Population rasters are only available for 2010 and 2015. Therefore, the raster for 2012 is obtained from linear interpolation between the 2010 and 2015 measures. The population density data is publicly available at https://sedac.ciesin.columbia.edu/data/set/gpw-v4-population-density-rev11/data-download#close.

[9] Open access data, available from https://zenodo.org/record/4882792#.YnIo89rP1hG.

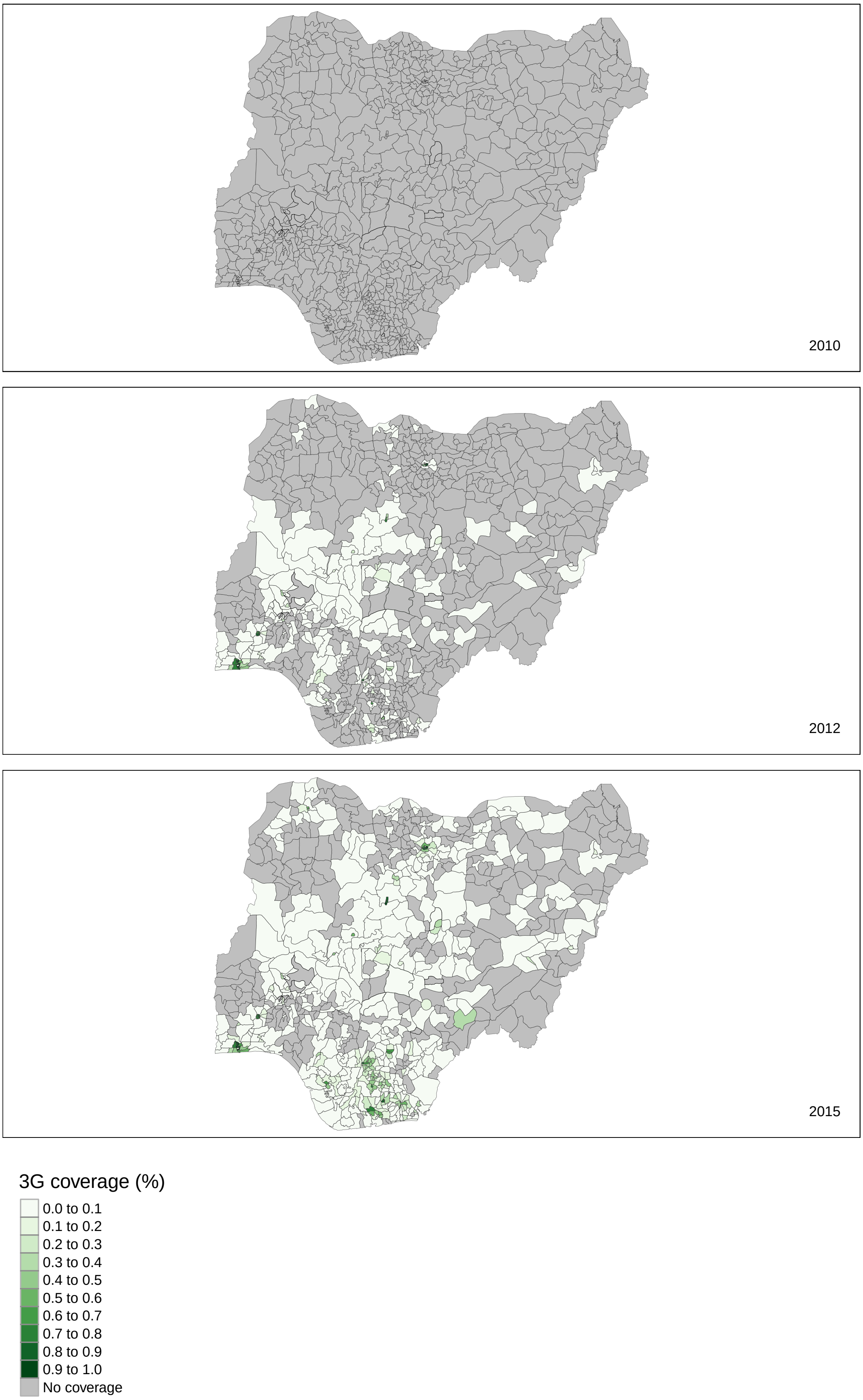

Figure 4.3: Mobile coverage in Nigeria in 2010-2015, based on GSMA Admin-2 (LGA) data. The colour indicates the percentage of 3G mobile internet coverage in each Nigerian LGA, in 2010, 2012, and 2015.

for Nigeria constitutes of rasters of 5 arc-minute resolution (around 8 km x 8 km at Nigeria's latitude) measuring the mean daily and monthly strike density (number of strikes per km^2) in each cell, for the period 2010-2020. In order to capture the exposure of each geographical unit to lightning, the series is averaged over time, resulting in a time-invariant measure of strike density. Similarly to the treatment of the mobile internet data, the strike density measure is weighted in terms of the GPW data (Centre for International Earth Science Information Network CIESIN Columbia University, 2018) for 2010, to obtain a time-invariant measure of strikes per capita, which is averaged for each LGA to obtain a subnational level measure of strikes per inhabitant. The basic intuition behind this instrument is that lightning strikes directly affect the presence of 3G cell towers by releasing electrostatic discharges which make it more costly for mobile phone operators to install the mobile signal broadcasting devices. More details are provided in Section 4.3.3.1, along with data showing that the incidence of lightning strikes does not affect economic activities directly.

The second additional source is the Armed Conflict Location and Event Data Project (ACLED) (Raleigh et al., 2010),[10] an extensive dataset that holds geo-coded information on conflict-related events around the world. This study relies on the Nigerian data on conflict events (battles, explosions, remote violence, and violence against civilians) to construct an LGA-level indicator of number of conflicts per capita. Based on the evidence that conflict events do affect economic activity (Collier and Duponchel, 2013; Ciarli et al., 2015), the information on conflict events is used as an additional control variable.

4.3.2 A descriptive picture of the data

The final dataset holds data for 3,459 NFEs in 2010, 4,446 in 2012, and 4,005 in 2015, owned by 3,443 unique households. All firms were operating at the time of the interview. A general description of the dataset is provided in Table 4.1, which shows

[10] The data is publicly available from https://acleddata.com/#/dashboard.

that around 59 per cent of the NFEs are operated by male members of the household. Owners are approximately 42 years old on average. Only 5 per cent of the NFEs are formal (officially registered with the government), indicating that informality is the rule rather than the exception among Nigerian NFEs. Figure 4.4 shows a degree of industrial heterogeneity across NFEs: the most common industry is retail and wholesale trade ("trade"), followed by manufacturing, and other services ("other"). Trade and manufacturing firms grew in both absolute and relative terms between 2010 and 2015.

While Figure 4.4 does not show any dramatic changes in the aggregate industrial composition, Figure 4.5 indicates that both entry into a new industry and exit from the market are not infrequent among Nigerian households. The alluvial plot shows the aggregate number of households entering a new industry (all flows starting from any industry) or leaving the market (flows pointing to "none") over the time period available in the data.[11] The "None" block in 2010 groups together households that were not in the market in 2010, but entered at some point in either 2012 or 2015. To facilitate the visualisation, the flows starting from trade in 2010 are highlighted in blue, while those starting from manufacturing in the same year are highlighted in red; this makes it possible to trace the path of the households that started in the two largest industries in 2010. It can be observed that households active in manufacturing could enter in the trade sector in the following time period (for instance, between 2010 and 2012), more than offsetting entry from trade into manufacturing. However, by 2015 almost half of the households that were active in trade services in 2010 had left the industry (mostly leaving the market, but also changing type of activity). Nevertheless, the new entries more than offset the exits in 2012, and almost did so in 2015.

Figure 4.6 describes the distribution of NFE sales per worker, over time. The log-

[11] Households that owned more than one NFE in different industries at the initial time period are shown twice; for instance, if a household owned a manufacturing firm and a retail firm in 2010, and it started a transport firm in 2012 without shutting down the firms active in 2010, that household would be counted in both flows going from manufacturing and trade in 2010 to transport in 2012. The figure excludes households that did not own an NFE across any of the survey waves.

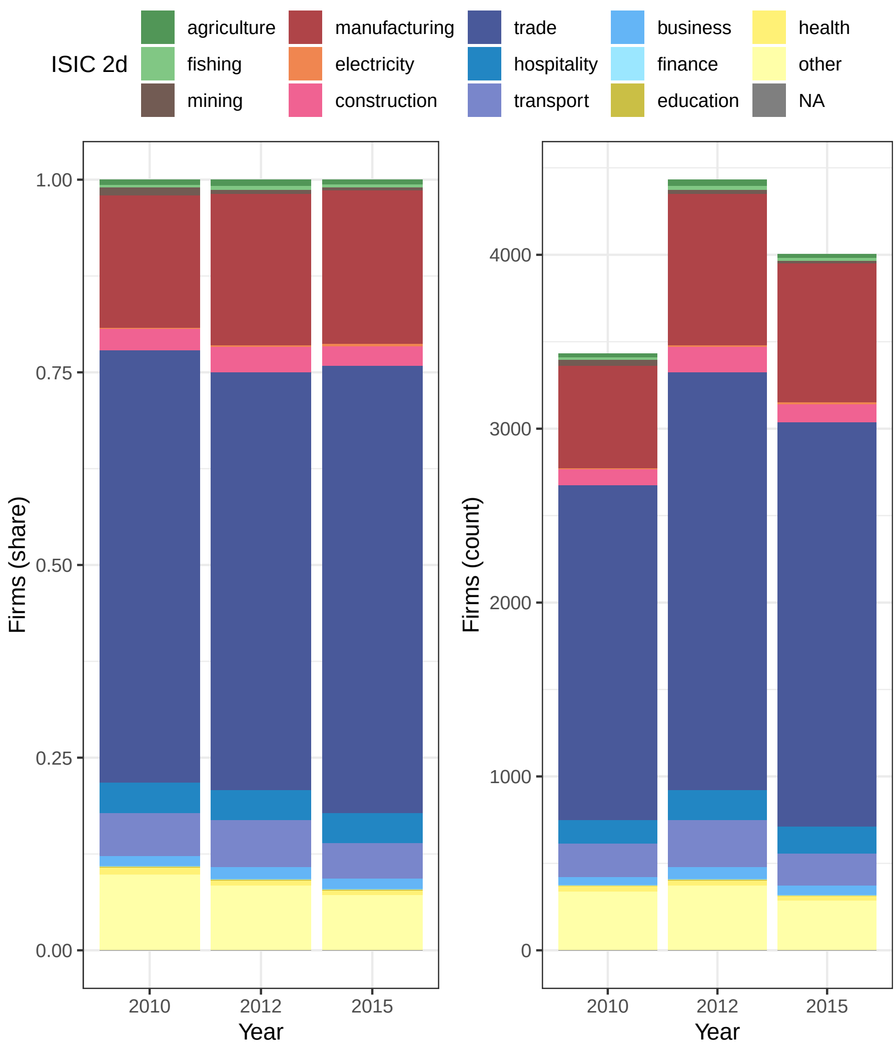

Figure 4.4: Non farming enterprises by industry over time: relative and absolute shares. Source: LSMS Nigeria.

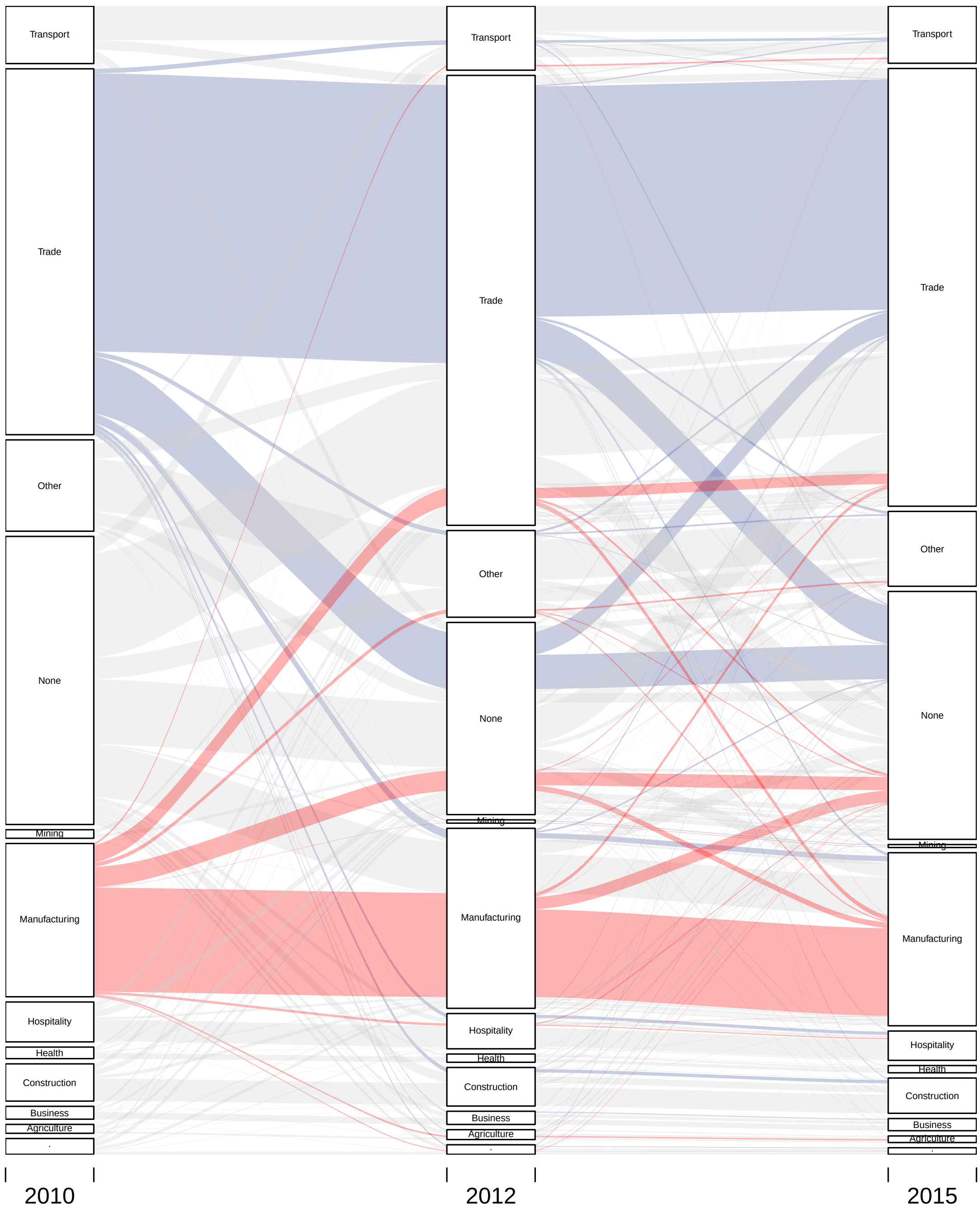

Figure 4.5: Household transition between industries over time. The alluvial plot shows the flow of households across sectors; red (blue) flows identify households who had an NFE in the manufacturing (trade) sector in 2010, tracking their path over time. Source: LSMS Nigeria.

Table 4.1: Non-farming enterprise dataset, descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
3G Adoption	11,900	0.037	0.126	0.000	1.852
Male owner	11,868	0.588	0.492	0	1
Age (owner)	11,807	40.199	14.054	15	65
# Members	11,910	7.441	3.521	1	35
# Members in work age	11,910	3.487	1.953	0	22
Household head is female	11,910	0.117	0.321	0	1
Household has agri. plot	11,910	0.568	0.495	0	1
Owner can read	11,716	0.633	0.482	0	1
Owner attended school	11,706	0.688	0.463	0	1
Formal	11,037	0.052	0.221	0	1
Dist. from road (miles)	11,910	7.863	11.620	0.000	102.100
Dist. from pop. centre (miles)	11,910	20.365	18.698	0.060	101.500
Dist. from market (miles)	11,910	66.636	44.253	0.370	214.340
Dist. from capital (miles)	11,910	61.484	52.209	0.180	442.700
3G coverage (%)	11,910	0.061	0.173	0.000	1.000
Mobile phones per cap.	11,900	0.430	0.402	0.000	12.000
# Conflict events (LGA)	11,910	0.694	4.380	0	77
Strike density	11,873	0.002	0.002	0.0004	0.014

Note: Based LSMS Nigeria, GSMA, ACLED, and WGLC data.

normal shape persists across the three years; an interesting phenomenon is the fact that the mean logged measure of sales per worker moves slowly but steadily to the right between 2010 and 2015, and the right tail of the distribution – where the most productive firms are found – becomes progressively thicker, indicating that they have increased in number compared to previous years. Additionally, Table 4.2 reveals some interesting features of the data, which emerge after grouping firms according to the major categories of the International Standard Industrial Classification (ISIC rev. 3.1). Focusing on the two most numerous groups, trade and manufacturing, it can be observed that the former exhibits higher levels of productivity, while the latter is one of the least productive industries in the sample. This description clashes somewhat with the idea that manufacturing firms are more productive than those active in services, which here have significantly higher values of sales per worker; however, it must be noted that trade activities sustain higher costs (for instance, inventory costs). Moreover, it should be noted that the measure of NFE performance used here is sales per worker, a very coarse approximation of productivity that doesn't take into account operating costs; this explains the lower performance of manufacturing NFEs in the data.

Looking at the industrial differences in average mobile internet adoption (first column in Table 4.2), the data shows some heterogeneity across industries, with education and the FIRE services exhibiting higher average mobile internet adoption when compared to other industries. Overall, this heterogeneity highlights some possible correlations between the type of economic activities and the degree of mobile internet adoption. For instance, the presence of mobile internet infrastructure may attract industries that can benefit from its adoption. This issue will be dealt with in the identification strategy, described in Section 4.3.3.1, by controlling for unobserved heterogeneity across major ISIC groups. Finally, Figure 4.7 provides grounds for an analysis of the effect of mobile internet adoption on NFE performance, showing a positive correlation between 3G adoption and the (logged) sales per worker of Nigerian NFEs. Bearing in mind that a causal interpretation of this relationship is not possible at this stage, the fitted lines indicate that the degree of positive correlation increases with time, presumably as a result of the progressive diffusion of 3G technology which allows higher adoption and an increasingly marked effect on economic activity. The next section illustrates the empirical strategy which will be employed to single out the effect of mobile internet adoption on NFE performance, industrial dynamics, and on the inclusiveness of the transformation that accompanies the roll-out of fast mobile internet.

4.3.3 Estimation strategy

The main goal of this study is to estimate the effect of mobile internet adoption by Nigerian households on structural change and inclusion, looking at their non farming enterprises. The baseline models can be outlined by Equation (4.1):

$$Y_{iht} = \alpha + \beta mobile_{ht} + \gamma X_{it} + \zeta Z_{ht} + \sigma_h + \tau_t + \psi_j + \epsilon_{iht} \tag{4.1}$$

where Y_{it} is a placeholder for the outcomes of interest at the NFE (i) level, belonging

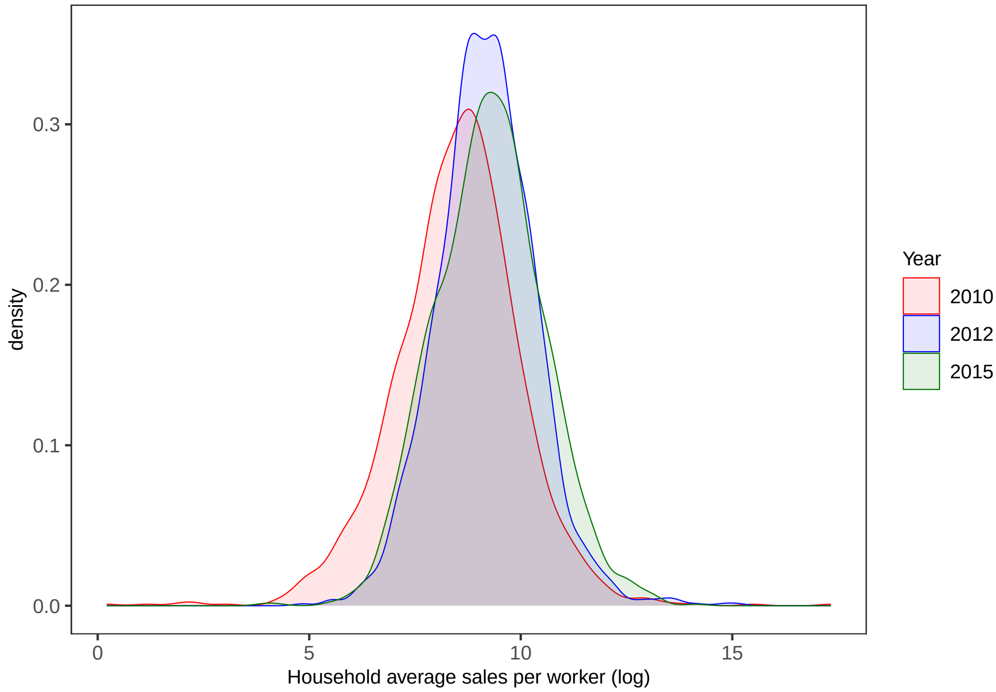

Figure 4.6: Kernel distribution of NFE sales per worker (logged), 2010-2015. Source: LSMS Nigeria.

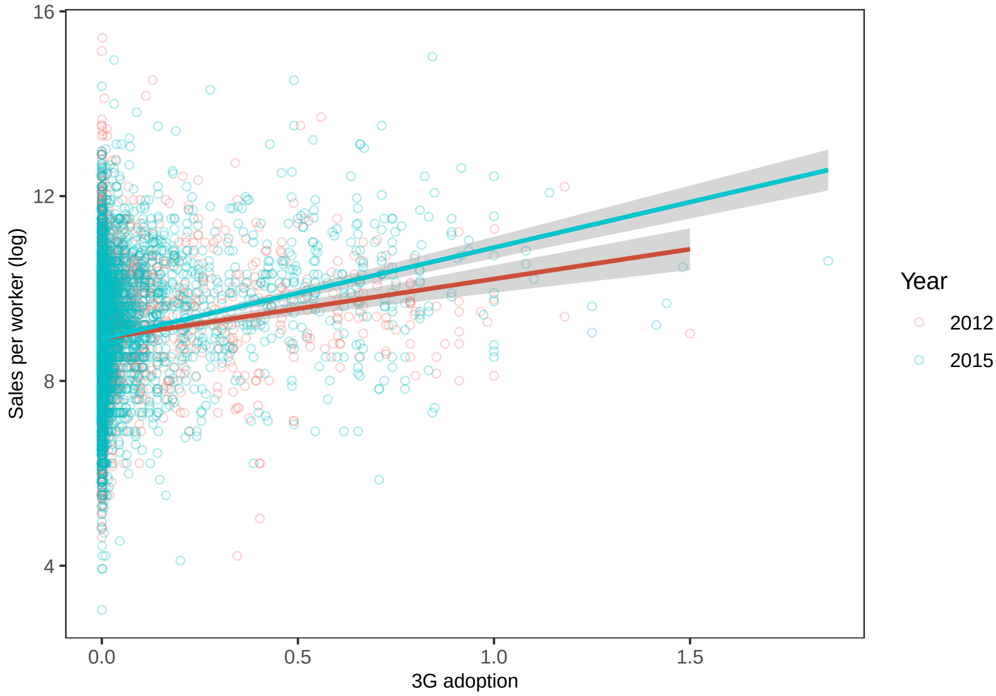

Figure 4.7: Correlation between mobile internet adoption (X axis) and logged sales per worker (Y axis). The fitted lines separately indicate the degree of correlation between the two variables, in 2012 (red) and in 2015 (blue). Source: LSMS Nigeria and GSMA.

Table 4.2: Descriptive statistics: averages by ISIC major groups

Industry	3G adop.	Sales PW	Costs	Capital	Sales	Labour	Formal	N
Agriculture	0.03	61.41	100.50	214.47	155.31	2.57	0.08	84
FIRE	0.09	42.18	99.06	900.54	124.93	2.98	0.22	175
Transport	0.03	35.33	117.53	872.40	157.57	2.55	0.23	642
Construction	0.03	25.89	34.67	191.49	96.28	3.00	0.07	348
Trade	0.04	23.55	59.50	182.36	56.37	2.58	0.03	6,646
Electricity	0.07	20.03	80.75	286.12	42.75	2.38	0.04	26
Mining	0.00	19.67	35.89	132.99	56.77	2.48	0.04	69
Health	0.06	19.55	35.20	525.87	55.78	2.64	0.27	85
Other	0.04	13.73	22.51	142.85	66.13	2.73	0.07	992
Education	0.15	12.95	89.48	2,566.82	164.32	9.41	0.65	17
Manufacturing	0.03	11.04	11.27	97.99	31.88	2.73	0.03	2,259
Hospitality	0.03	9.22	12.19	307.39	23.49	2.65	0.01	471
Fishing	0.01	7.96	15.91	76.44	19.98	2.79	0.17	52
Public	0.00	5.83	1.60	9.33	12.00	2.67	0.00	3

Note: All values are industry-wise means. 3G adoption is expressed as mean 3G coverage weighted by mobile phones per capita; sales per worker is expressed as average sales (in thousand Nairas) per worker; costs, capital, and sales are expressed in thousand Nairas; labour in number of workers; and formal as the share of formal firms in the industry. FIRE indicates financial, insurance, and real estate services.

to household h (sales per worker, and industrial transition for structural change; firm labour and firm entry for inclusion); $mobile_{ht}$ is the percentage of 3G coverage in the household's LGA, weighted by mobile phones per capita at the household level (β is therefore the main coefficient of interest). X_{it} is a vector of time-varying controls at the firm level, such as owner's age, sex and education; Z_{ht} is a second vector of controls, this time at the household h level, which includes number of household members, agricultural activities, distance from markets, towns, and main roads. As mentioned in Section 4.3.1.3, the vector Z also includes the number of conflicts in the household district. In order to account for the higher conflict number in more populated LGAs, the variable is added as a per capita measure. Finally, household (σ_h), time (τ_t) and industry[12] (ψ_j) fixed effects are added to capture household and industry unobserved heterogeneity, along with time-related fixed effects. In all estimations, the standard errors are clustered at the enumerator area level (equivalent to the size of a village in rural areas).

Nevertheless, the estimation of Equation (4.1) is likely to be biased by endogeneity, as the roll-out of 3G mobile internet is not likely to be 'as good as random'. Endogeneity

[12] Industries are grouped according to the ISIC major groups, as per Table 4.2.

issues can have several causes, including measurement errors, omitted variable bias (for example, the presence of unobserved, time-varying geographical or societal features that may influence both 3G adoption and structural change/inclusion) and reverse causality (households with NFEs have higher income and may attract more mobile operators). Moreover, Equation (4.1) will be correctly identified only in the absence of pre-trends in both structural change and inclusion. In order to causally estimate the effect of mobile internet adoption on structural change and inclusiveness, the analysis resorts to two different identification strategies; the first is based on an instrumental variable approach, the second on an event study design. Both approaches are described below.

4.3.3.1 Instrumental Variable: lightning strikes

The first part of the identification strategy consists of an Instrumental Variable (IV) approach, implemented using a Two-Stage Least Squares (2SLS) estimation. The instrument exploits the geographical variation in lightning strikes across subnational areas, based on the evidence that the mobile phone infrastructure (mainly the cell towers responsible for 3G signal emission, installed by mobile operators) is affected by frequent storms, as cloud-to-ground lightning strikes cause power surges making cell tower maintenance more costly (Manacorda and Tesei, 2020). While this instrumental variable approach has been previously used in empirical settings studying the effect of fast mobile internet on social phenomena, such as citizens' political mobilisation (Manacorda and Tesei, 2020) or trust in government (Guriev et al., 2021), a growing body of literature has employed the variation in lightning strikes to break the endogeneity between mobile internet diffusion and economic outcomes (Andersen et al., 2012; Mensah, 2021; Caldarola et al., 2022).

The core argument is well described by Andersen et al. (2012), who use US data to demonstrate that lightning strikes can influence growth and productivity via their impact on the diffusion of IT technologies, due to the fact that the power surges caused by cloud-to-ground strikes disrupt the functioning of sensitive electronic equipment used to diffuse mobile signals. The implementation of the 2SLS approach used in this chapter follows the one suggested by Guriev et al. (2021): an instrument based on a strikes per capita measure at the LGA level is constructed (details of the variable construction are described in Section 4.3.1.3). The rationale behind using a per capita measure of lightning strikes is that this will reflect the mobile operators' decision criterion when installing a mobile cell tower: the higher costs due to frequent power surges, which may discourage the operator from installing a mobile cell tower in an area highly exposed to storms, are moderated by a larger market (proxied by higher population density) which may offset the additional maintenance costs of the cell towers in areas with higher strike density. Moreover, Guriev et al. (2021) use a time-invariant measure of strikes per capita, constructed as the average strike density per capita throughout the available series. This choice, also adopted here, is based on evidence that lightning strikes follow a stationary process, so the instrument will capture the intrinsic exposure of geographical areas to strikes, on the grounds of which the mobile operators make the decision whether or not to build a cell tower. In order to reflect the increasing mobile internet roll-out, the time-invariant measure is interacted with a linear time trend.

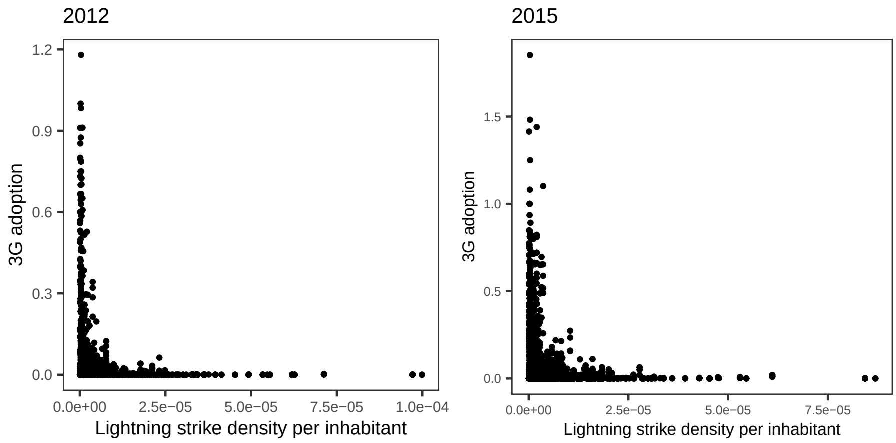

Figure 4.8: Relevance restriction: correlation between lightning strikes per capita in Nigerian LGAs (X axis) and the adoption of 3G mobile internet services by Nigerian households. Year 2010 is excluded due to the absence of mobile internet coverage. Source: GSMA and WGLC data.

The relevance of the instrument – lightning strikes per capita in Nigerian Local Government Areas – with respect to mobile internet adoption is shown clearly by

Figure 4.8, which reveals a neat, negative, and hyperbolic correlation between the two variables. It is interesting to note that the pattern is consistent over time, as the areas with the highest number of lightning strikes per capita show the smallest (in some cases null) rate of 3G adoption both in 2012 (left panel in Figure 4.8) and in 2015 (right panel). The particular shape of this relationship suggests implicitly that an over identified first stage could be used in the 2SLS estimation to increase its precision, using the square of the *strikes* variable, as shown by Equation (4.2):

$$mobile_{ht} = \gamma_0 + \gamma_1 strikes_l \times t + \gamma_2 (strikes_l \times t)^2 + \gamma_3 X_{it} + \gamma_4 Z_{ht} + \sigma_i + \tau_t + \psi_j + u_{ht} \quad (4.2)$$

where $strikes_l$ is the measure of lightning strike density in the LGA, interacted with a linear time trend t. The second stage estimation is defined by Equation (4.3):

$$Y_{iht} = \theta_1 \widehat{mobile}_{ht} + \theta_2 X_{it} + \theta_3 Z_{ht} + \sigma_h + \tau_t + \psi_j + \epsilon_{iht} \quad (4.3)$$

where the coefficient θ_1 can be interpreted as a continuous, Locally Averaged Treatment Effect (LATE, Angrist 2004) of mobile internet adoption on the outcomes of interest related to structural change (NFE performance, and household industrial transition) and inclusion in terms of entrepreneurial and working opportunities (NFE ownership and labour).

A possible concern is that lightning strikes may directly affect the performance of non-farming enterprises by disrupting access, quality and cost of electricity. The study tests this, using self-reported household information on the constraints to starting and operating a business, included in the LSMS data. Each household was asked to report the three major constraints to starting and operating a business, and this analysis focuses on whether access to, the quality of, or the cost of electricity represent an obstacle to starting or operating a business in areas with higher exposure to lightning strikes. First, Table 4.3 shows that living in areas with frequent lightning does not affect firm entry on the basis of electricity-related hurdles: although areas where NFE owners are located experience lower average lightning density, the difference is very small with respect to the full variable range. Moreover, this does not affect electricity access, quality, or cost; in fact, electricity seems to be a more pressing constraint for NFE owners as compared to non-owners. Secondly, Table 4.4 shows the percentage of households indicating electricity access, cost, and quality as obstacles to running household businesses, by different levels of lightning strike density. Again, there is no clear association between the frequency of strikes, and electricity-related obstacles to running a business. Finally, Figure 4.9 plots the relationship between lightning density and total NFE operating costs.[13] It shows that there is no positive correlation: higher lightning density does not seem to be related to higher costs for NFE, on average. If anything, the correlation between lightning density and average firm costs by household is negative. This descriptive evidence is consistent with the argument and evidence proposed by Andersen et al. (2012), that lightning strikes affect the economy only via their effect on the diffusion of ICT technologies, and not directly.

4.3.3.2 Event study

The 2SLS estimation with fixed effects presented above works only under the assumption that there are no pre-trends before the treatment (that is, that the outcomes of interest did not show the same trend before and after adopting mobile internet) and no anticipation (meaning that individuals could not have anticipated the arrival of the new mobile internet technology, for instance taking decisions on how to operate their enterprises from the perspective of imminent access to the new technology, thus affecting the outcome variables of interest before treatment. Although violation of

[13] Costs include (in the last month before each interview): salaries and wages, purchase of goods, transport, insurance, rent, interest on loans, raw materials, and other costs.

Table 4.3: Electricity constraints in Nigerian households

Variable	No NFE	Has NFE	Difference	T stat.
Lightning density (mean)	0.0025	0.0023	2e-04	4.08***
Lightning density (sd)	0.0017	0.0018	-1e-04	-
Electricity cost (mean)	0.0227	0.065	-0.0423	-7.501***
Electricity cost (sd)	0.0227	0.065	-0.0423	-
Electricity quality (mean)	0.0312	0.0839	-0.0527	-8.01***
Electricity quality (sd)	0.1739	0.3983	-0.2244	-
Electricity access (mean)	0.0244	0.068	-0.0437	-7.443***
Electricity access (sd)	0.0244	0.068	-0.0437	-
Households (N)	3443	5233	-	-

Note: The table compares the extent to which electricity (access, cost, and quality) represents a constraint on starting or operating a business for Nigerian households, comparing the share of households that identify these as one of the three obstacles across NFE-owning and non-owning households. The t-stat column is the result of an t-student inference test on the statistical significance of the difference between the two groups. Source: LSMS Nigeria and WGLC.

Table 4.4: Electricity constraints by lightning strike density exposure

Lightning quintile	N	Lightning density (mean)	Electricity cost (mean)	Electricity quality (mean)	Electricity access (mean)
1^{st}	1783	7e-04	0.0333	0.0972	0.0365
2^{nd}	1804	0.0012	0.0197	0.0685	0.042
3^{rd}	1762	0.002	0.0908	0.07	0.0876
4^{th}	1783	0.003	0.0561	0.0417	0.0451
5^{th}	1776	0.0051	0.0542	0.057	0.0611

Note: The table compares the share of households indicating electricity (cost, quality, and access) as a major constraint to operating or starting a business in Nigeria, comparing households by the lightning strike density of the LGAs in which they leave (divided by quintiles of the lightning strike distribution). Source: LSMS Nigeria and WGLC.

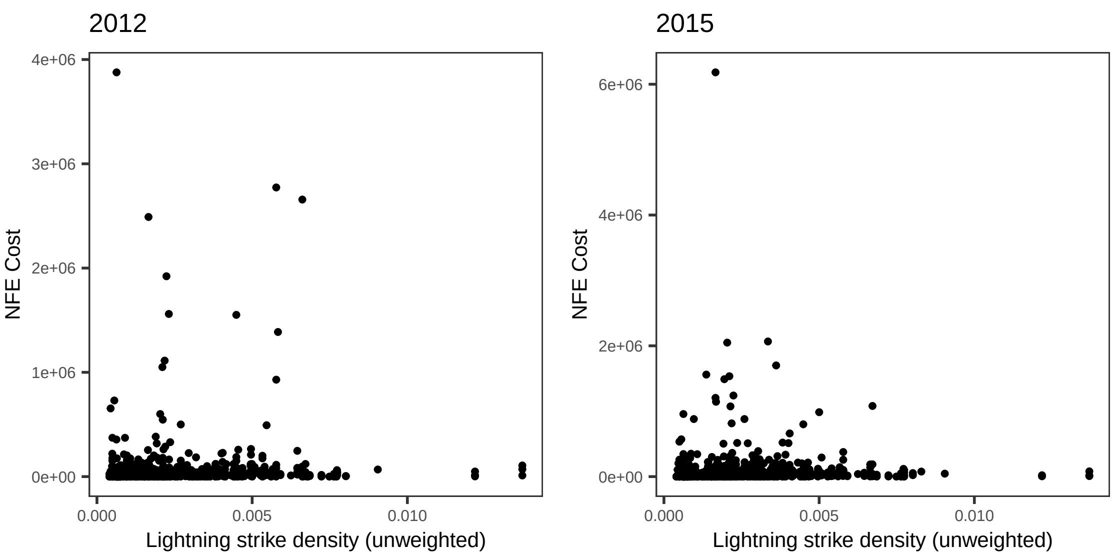

Figure 4.9: Exclusion restriction: correlation between lightning strike density (X axis) and operating costs of non-farming enterprises (Y axis). Source: LSMS Nigeria and WGLC.

these assumptions can be controlled for (Rambachan and Roth, 2021), the analysis also implements an event study design, which particularly recommended in cases where an experimental setting is absent (Clarke and Tapia-Schythe, 2021) in order to test whether such assumptions hold in the data. This approach requires the definition of a dichotomous treatment (or event) through a set of variables measuring whether a household has or has not been 'treated' in all time periods. This is implemented by adding lags and leads to Equation (4.1) to measure the time distance from mobile internet adoption:

$$Y_{iht} = \sum_{\tau=-2}^{2} \phi_{\tau} treated_{ih} + \phi_2 X_{it} + \phi_3 Z_{ht} + \sigma_i + \tau_t + \psi_j + \epsilon_{iht} \tag{4.4}$$

where the variable *treated* is a dichotomous variable that is activated only for the time periods τ in which a household h adopts mobile internet, and in subsequent time periods. Given that the dataset covers only three time periods, the event design disposes of a maximum of two lags (periods before treatment) and two leads (periods after treatment). To define the dichotomous variable, the following arbitrary threshold is adopted: a household is 'treated' by adoption of mobile internet when $mobile > 0.2$. In practice, this means that a household is treated if, for instance, it is located in an area fully covered by 3G internet access and 20 per cent of households in the area have a mobile phone; or, when a household has one phone per member of working age, and is located in an area where at least 20 per cent of the population in the LGA have access to 3G internet.[14] As a robustness check, event study regressions are also run for different treatment thresholds (0.15, 0.2, 0.3, and 0.4). The main advantage of this approach is that, although it requires the use of a categorical treatment variable, it makes it possible to check for parallel trends (in the absence of which the causal interpretation of the parameters cannot be claimed) and to observe the persistence of the treatment effect over time. The baseline period is $\tau = -1$

[14] It can be expected that households living in areas with good mobile internet coverage will have an incentive to purchase mobile phones to access the mobile internet service.

– the year before the treatment. For the 'no parallel trends' and 'no anticipation' assumptions to hold, it is required that the coefficient of the event variable before the baseline period is not significant. The event study results will be presented along with the 2SLS results in the next section.

4.4 Results

4.4.1 Structural change

Table 4.5 reports results of estimating Equation 4.1 using both OLS and 2SLS. The 2SLS estimation results (columns 3 and 4) show that mobile internet adoption leads to an increase in sales per worker for the average NFE. Comparison with the OLS results (columns 1 and 2) suggests that these estimations suffer from a downward bias (consistent with similar studies, such as Manacorda and Tesei 2020; Mensah 2021; Caldarola et al. 2022). Column 5 in Table 4.5 displays the results of the first stage of the 2SLS estimation in column 4, indicating that the instrument enters significantly with a negative sign in its base form, and with a positive sign when squared. The exclusion of both instruments has a Kleibergen-Paap rk Wald F statistic of 27.56, safely above the $F > 10$ criterion suggested by Stock and Yogo (2005) to test against weak instruments. The Sargan-Hansen test yields a p-value above 0.05, indicating that the null hypothesis of valid over identifying restrictions cannot be rejected.

The quantification exercise at the bottom of Table 4.5 gives a clearer idea of the size of the effect of mobile internet adoption on NFE performance, showing that a change from the 25^{th} to the 75^{th} percentile along the mobile internet adoption distribution increases a firm's sales per worker by 3.6 per cent. The positive effect of 3G adoption on NFE performance is corroborated by the results of the event study estimation (Figure 4.10), which show a statistically significant positive effect of mobile adoption on sales per worker after two time periods from adoption, and not before. The statistically insignificant coefficient before the treatment confirms that there are no parallel trends. The robustness checks in Figure C.1 also support

Table 4.5: OLS and 2SLS results, Log sales per worker

	Dependent variable:				
	Log sales per worker				3G adop.
	OLS FE	OLS FE	2SLS	2SLS	First stage
	(1)	(2)	(3)	(4)	(5)
3G adoption	0.260	0.189	1.947**	2.071**	
	(0.224)	(0.213)	(0.804)	(0.860)	
Owner is female		-0.827***		-0.822***	-0.00304
		(0.0443)		(0.0449)	(0.00221)
Age		0.00408***		0.00411***	-4.48e-05
		(0.00147)		(0.00147)	(8.93e-05)
N. members		0.00779		0.0257	-0.00953***
		(0.0192)		(0.0221)	(0.00267)
Female head		0.0799		0.0660	0.0130
		(0.110)		(0.110)	(0.0134)
Agriculture		-0.0534		-0.0466	-0.00617
		(0.0811)		(0.0808)	(0.00556)
Owner can read		0.0582		0.0651	-0.00516
		(0.0546)		(0.0546)	(0.00317)
Owner attended school		0.0538		0.0455	0.00329
		(0.0675)		(0.0681)	(0.00380)
Dist. from main road		0.000717		-0.00244	0.00118***
		(0.00340)		(0.00369)	(0.000262)
Dist. from town		-0.00426		-0.00339	-0.000467*
		(0.00288)		(0.00293)	(0.000255)
Dist. from market		-0.00290		-0.00129	-0.000149
		(0.00429)		(0.00465)	(0.000315)
Dist. from capital		0.00132		0.00216	-0.000385
		(0.00467)		(0.00496)	(0.000284)
Conflict PC		-72.36		1,352	-556.3*
		(4,280)		(3,942)	(325.0)
Lightning PC					-6.379***
					(0.973)
Lightning PC^2					17.24***
					(4.607)
Constant	8.806***	9.213***			0.249***
	(0.00858)	(0.357)			(0.0440)
Observations	11,258	10,989	11,220	10,989	10,989
R-squared	0.588	0.635			0.680
Household FE	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Mean DV	8.810	8.810	8.810	8.810	
Quantification			3.131	3.616	
Sargan (p value)			0.412	0.275	
Fstat					27.56

Note: The dependent variable measures the log of sales per worker of NFEs. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. The F-stat in column 5 reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution, after unlogging the dependent variable. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

the 'no anticipation' and 'no parallel trends' assumptions, and indicate that the effect of 3G adoption on NFE performance remains stable for both lower and higher treatment thresholds.

It is worth mentioning that column 5 in Table 4.5 shows only two positive coefficients among the control variables; the ones on the gender (female) and age of the NFE owner. The coefficient on gender (female) is strongly significant and negative, with a large point estimate,[15] indicating the presence of gender-based inequality structures, with female-owned NFEs showing systematically lower sales per worker. This result is consistent with the findings of Nagler and Naudé (2017), who identify as a possible cause the presence of time constraints for women, who are more often also involved in household duties.

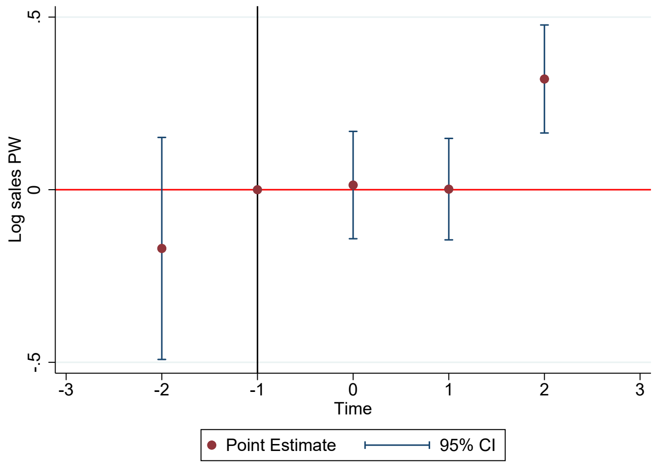

Figure 4.10: Event study: sales per worker. The event study design uses the first year in which a household hits 0.2 of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. The coefficients reported in the figure come from a model based on Equation (4.4), at the NFE level, including household, industry, and year fixed effects, incorporating the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95 per cent confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

This chapter examines various mechanisms that could possibly lead to a positive impact of internet adoption on NFE sales per worker. It investigates whether internet

[15] Although the coefficient is omitted in the following Table (4.7), the coefficient on gender (female) remains significant and negative.

adoption has an effect on access to credit, capital,[16] costs,[17] sales, or labour demand. Table 4.6 shows that mobile internet adoption does not have a statistically significant impact on capital, costs, nor credit (columns 1 to 5 in Table 4.6, corroborated by Figure C.2). Instead, internet adoption has a direct significant negative impact on the number of employees, and a significant positive impact on sales. The poor effect of internet adoption on credit is probably due to the limited adoption of mobile money in Nigeria: in 2014 only 2.3 per cent of the Nigerian population had access to a mobile money account. This is much lower than in other African countries, like Kenya, where in the same year 58 per cent of the population had access to mobile money (Demirgüç-Kunt et al., 2018).

We next focus on different industries. Results indicate that the positive effect of internet adoption on sales per worker is statistically significant only in trade activities (Table 4.7, column 9). These results are consistent with Houngbonon et al. (2022), who suggest that the concentration of benefits in trade services could be due to lower operating costs. An additional reason could be that mobile internet technologies are more suitable for retail and wholesale activities, due to the higher relevance of person-to-person interactions, when compared for instance to the manufacturing industry. However, it is also notable that, due to the small number of NFEs in some of the industries, the IV does not perform equally well for all of them. Concentration of effects in services is confirmed by the event study design presented in Figure C.3, which indicates that in the aggregate and FIRE services there is also a positive effect of mobile internet adoption on NFE performance, measured as sales per worker. However, as the FIRE industry is also the smallest one in the sample, these results should be interpreted cautiously.

We next examine whether internet access has an impact on structural change also through the change in the composition of industries. Table 4.8 illustrates the results

[16] Capital quantifies the current value of physical capital stock, including all tools, equipment, buildings, land, vehicles, inputs, supplies, and merchandise (goods for sale).

[17] Costs include (in the last month before each interview): salaries and wages, purchase of goods, transport, insurance, rent, interest on loans, raw materials, and other costs.

Table 4.6: 2SLS results, mechanisms

	Dependent variable:							
	Credit	Capital	Capital PW	Costs	Costs PW	Sales	Labour	3G Adop.
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	First stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3G Adoption	0.0421	-0.152	0.414	0.131	0.471	1.540**	-2.614*	
	(0.138)	(1.098)	(1.136)	(1.825)	(1.742)	(0.645)	(1.562)	
Lightning PC								-6.379***
								(0.973)
Lightning PC^2								17.24***
								(4.607)
Constant								0.249***
								(0.0440)
Observations	10,983	10,989	10,989	10,989	10,989	10,989	10,989	10,989
R-squared								0.680
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Household FE	YES	YES	YES	YES	YES	YES	YES	YES
Industry FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Mean DV	0.0424	10.17	9.330	7.902	7.139	9.671	2.649	
Quantification						1.910	-0.483	
Sargan (p value)	0.757	0.161	0.159	0.538	0.464	0.345	0.0846	
Fstat								27.56

Note: The dependent variables measure: the amount of credit received by the NFE (in thousand Nairas); the amount of capital stock owned by the NFE (in thousand Nairas), and its per-worker equivalent; the amount of costs sustained by the NFE in the last month (in thousand Nairas), and its per-worker equivalent; the amount of sales made the NFE in the last month (in thousand Nairas); and the number of labourers hired by the NFE. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

for the probability that a household h enters industry j in $t+1$, conditional on not being active in that industry in the previous time period t. We find that households with higher mobile adoption are less likely to enter the manufacturing industry, suggesting that they may be induced to specialise in industries – such as trade – where there are higher benefits from accessing internet. If new working opportunities stemming from the diffusion of mobile internet are more likely to arise in service activities, this could fuel the process of tertiarisation of the informal sector. Due to the availability of only two time periods for the transition of households, imposed by the way in which transition has been defined, it has not been possible to conduct an event study in this case. Moreover, the definition of household transition imposes to have a balanced panel structure, which explains the lower number of observations – corresponding to households which own at least one NFE both in 2012 and 2015.

Table 4.7: 2SLS results, (log) sales per worker, by industry

	Dependent variable:													
	Log sales per worker													
Sample:	Manuf.	1 stage	Constr.	1 stage	Transp.	1 stage	Serv.	1 stage	Trade	1 stage	Hosp.	1 stage	FIRE	1 stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
3G adoption	1.588		0.978		2.099		1.657		1.840*		8.446		-4.480	
	(2.168)		(5.526)		(2.697)		(1.009)		(1.034)		(7.544)		(5.215)	
Lightning density PC		-5.744***		-8.030**		-9.756***		-6.567***		-6.818***		-6.067*		-12.08**
		(1.647)		(3.681)		(3.251)		(1.097)		(1.158)		(3.531)		(5.944)
Lightning density PC^2		15.04***		22.86**		36.73*		17.63***		18.17***		32.72		25.88
		(5.373)		(11.11)		(20.82)		(5.218)		(5.257)		(22.67)		(17.42)
Constant		0.238***		1.034*		0.172		0.271***		0.289***		1.148		-0.380
		(0.0659)		(0.569)		(0.171)		(0.0573)		(0.0515)		(1.336)		(1.270)
Observations	1,849	1,849	250	250	425	425	7,452	7,452	5,856	5,856	388	388	122	122
R-squared		0.647		0.707		0.738		0.701		0.694		0.841		0.825
Controls	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Household FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES	YES
Mean DV	8.334		9.051		9.390		8.876		8.950		8.397		9.608	
Quantification	2.031		0.865		3.734		2.213		2.762		2428		-0.516	
Sargan (p value)	0.979		0.615		0.716		0.249		0.296		0.470		0.217	
Fstat		6.561		2.700		5.424		23.46		22.45		1.987		

Note: The dependent variables measures the log of sales per worker of NFEs. Odd numbered columns indicate the industrial sub-sample used in the 2SLS estimation, followed by their respective first-stage regression. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

Table 4.8: 2SLS results, household transition

	Transition to:			*Dep. var.:*
	Services	Manuf.	FIRE	3G Adoption
	2SLS	2SLS	2SLS	First stage
	(1)	(2)	(3)	(4)
3G Adoption	0.122	-0.540*	0.0179	
	(0.249)	(0.297)	(0.106)	
Lightning PC				-6.856***
				(1.203)
Lightning PC^2				19.18***
				(6.241)
Constant	-0.0946**	0.0328	0.0151	0.246***
	(0.0371)	(0.0418)	(0.0135)	(0.0478)
Observations	4,761	4,761	4,761	4,761
R-squared				0.641
Number of HH	1,590	1,590	1,590	
Controls	YES	YES	YES	YES
Household FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Fstat				25.16

Note: The dependent variables measure the household transition toward services, manufacturing, and FIRE services. A household transitions towards an industry if they start an NFE in that industry in $t+1$, conditional on not being active in the same industry in t. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: ratio of female members (female); members' mean age; number of household members; female-headed household; household is active in agriculture; percentage of household members in working age who can read;percentage of household members in working age who have ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

4.4.2 Inclusion

Results for sales and number of employees (columns 6 and 7 in Table 4.6) are useful to discuss if the structural changes induced by access to mobile internet are also inclusive. These indicate that the increase in NFE sales per worker is likely to be driven by the simultaneous effect of higher sales (for instance, allowing NFEs to reach a larger market) and lower labour requirements (a possible substitution effect of mobile internet rather than complementary to labour). Both results are confirmed by the event study approach (Figure 4.11), which indicates that these effects are more likely to occur after two time periods following mobile 3G internet adoption.

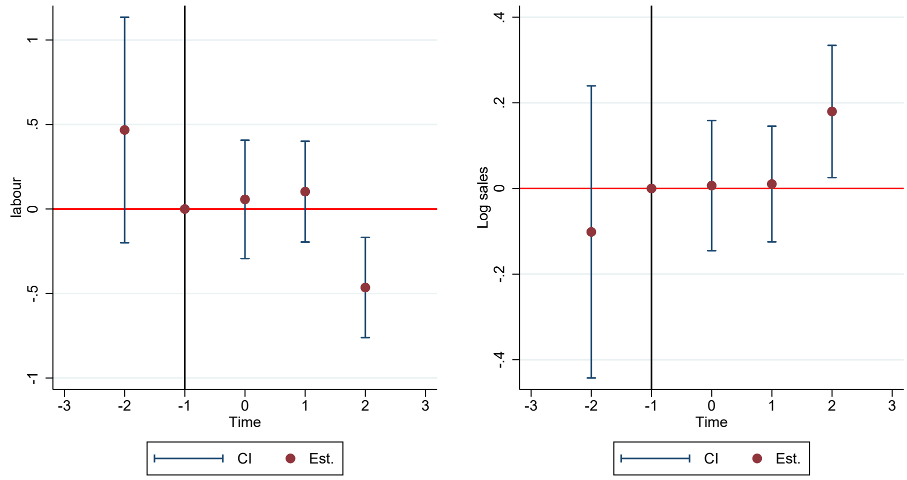

Figure 4.11: Event study: labour and sales. The event study design uses the first year in which a household hits 0.2 of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. The outcome variable used in the event study specification is indicated on the vertical axis. The coefficients reported in the figure come from a model based on Equation 4.4, including household, industry, and year fixed effects. The model uses NFEs as units of observation, and incorporates the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

In order to provide more detail on the labour substitution effect, the study examines whether NFEs reduced the number of employees sourced from the household or whether they hired labour. Table 4.9 indicates that there is a difference between within-household and external labour: the former decreases and the latter grows, indicating that, after adopting mobile internet, NFEs are more likely to free up household labour, in favour of employees from outside the household. One possible explanation is that, after adopting the new technology, firms retain only those household members who have the skills to master the new technology, and seek more of those outside. If it is the case that skilled labour is sought outside the household, the adoption of mobile internet may lead to labour market inequalities, favouring highly skilled workers over unskilled ones, but the data includes no information on the skills of the workers employed by the NFEs with which to test this. Another possible explanation is that household members find a more remunerative job outside the household; Bahia et al. (2020) show that internet adoption increases labour market participation of waged workers, while Strazzeri (2021) documents an increase in migration.

An attempt is made to test the latter hypothesis by estimating the effect of mobile internet adoption on the share of household members employed outside the household. Columns 1 and 2 in Table 4.10 indicate that mobile internet adoption by the household has a positive effect on the share of household members employed outside the household; this applies to NFE-owning households (column 2) and to all households too, irrespective of whether they own an NFE or not (column 1). The effect of the adoption of mobile internet is also positive if we consider any type of employment for all households (column 3), but not for NFE-owning households. Results suggest that the diffusion of mobile internet has been successful in reducing unemployment overall, but there is no increase in total employment for households that own an NFE; one possible explanation for the non-significant effect of 3G adoption in column 4 is that households that own an NFE are more likely to be in full employment already before adopting fast mobile internet. These results therefore suggest that while households are able to free household labour from their NFEs as a result of adopting the new technology, the excess labour can indeed be absorbed outside the household, although

Table 4.9: 2SLS results, household and non-household labour in NFEs

	Dependent variable:	
	Household labour	Non-household labour
	2SLS	2SLS
	(1)	(2)
3G Adoption	-6.338***	1.911*
	(2.010)	(1.014)
Observations	10,986	10,986
Controls	YES	YES
Household FE	YES	YES
Industry FE	YES	YES
Year FE	YES	YES
Mean DV	2.404	0.246

Note: The dependent variables measure the number of NFE workers from within (column 1) and from outside (column 2) the household. All estimations are run using a 2SLS estimator. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: owner's gender (female) and age; number of household members; female-headed household; household is active in agriculture; owner can read; owner has ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution.Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

Table 4.10: 2SLS results, labour within and outside the household

	Employment share:				*First stage:*	
	Outside		Overall		3G Adop.	
	All	NFE	All	NFE	All	NFE
	(1)	(2)	(3)	(4)	(5)	(6)
3G Adoption	0.209**	0.185*	0.415*	0.113		
	(0.107)	(0.110)	(0.249)	(0.295)		
Agriculture	-0.0190**	-0.00520	0.133***	0.0963***		
	(0.00825)	(0.0125)	(0.0194)	(0.0307)		
NFE	-0.0308***		0.152***			
	(0.00643)		(0.0148)			
Lightning PC					-3.666***	-6.782***
					(0.258)	(0.464)
Lightning PC^2					5.911***	18.80***
					(1.079)	(1.978)
Constant					0.190***	0.257***
					(0.0211)	(0.0372)
Observations	12,617	6,644	12,617	6,644	12,617	6,644
R-squared					0.617	0.661
Controls	YES	YES	YES	YES	YES	YES
Household FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
Sargan (p value)	0.26	0.07	0.16	0.39		
Fstat					217.1	139.5

Note: The dependent variables measure, respectively: the ratio of household members employed outside the the household over the total number of household members, across all households (column 1) and only in NFE-owning households (column 2); the ratio of household members employed both inside and outside the the household over the total number of household members, across all households (column 3) and only in NFE-owning households (column 4). All estimations are run using 2SLS. Column 5 reports the first stage results for models 1 and 3; column 6 reports the first stage results for models 2 and 4. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: ratio of female household members (female); household members' mean age; number of household members; female-headed household; household is active in agriculture; percentage of household members in working age who can read; percentage of household members in working age who have ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

this will not lead to an increase in the share of employed working age individuals. This finding partly confirms the presence of a labour-substitution effect: within NFEs, household labour is substituted both by the new technology, and by external labour. Within households, what is observed is more of a reallocation effect – the household labour that is pushed outside NFEs is re-absorbed outside the household.

These results suggest that, although mobile internet may increase the number of employees in a location, as also suggested in the literature, this is likely to come from the entry of new firms, or creation of employment in the formal sector, rather than from an increase in labour in existing firms in the informal sector.

The study then tests whether mobile internet leads to the entry of new firms into the informal sector. Table 4.11 (columns 1 and 2) shows that mobile internet adoption does not predict NFE ownership at the household level (or new entry of households that did not own an NFE in the previous year (Table 4.11, columns 4 and 5). This result suggests that, within the informal sector, only incumbent firms reap the benefits of mobile internet roll-out, increasing their sales.

At the same time, the adoption of mobile internet does not prevent the exit of households from the NFE market (columns 6 and 7), suggesting that incumbents who benefit from mobile internet adoption increase their market shares in the informal sector.

The insignificant effect of mobile internet adoption on firm entry contrasts with the evidence presented by Houngbonon et al. (2022) who document increased firm entry resulting from the diffusion of land line internet: however their focus is on formal firms.

Finally, an event study is conducted to explain NFE ownership, as entry is based on only two time periods. Figure 4.12 shows that mobile internet adoption has no positive effect on NFE ownership, and is even negative in the first year after treatment; in subsequent years, however, the effect becomes positive and weakly significant (90% confidence interval), indicating that households may take some time to find the resources or develop the skills needed to enter the NFE market, or they may decide to exit and take advantage of better paid opportunities. The fact that the internet has no effect on costs (Table 4.6) may also help explain why entry is not made easier by the adoption of mobile internet.

4.5 Conclusions

This chapter has attempted to establish a linkage between mobile internet adoption, structural change – measured in terms of NFE performance and industrial dynamics – and inclusion – measured by the creation of labour and entrepreneurial opportunities.

Table 4.11: 2SLS and OLS results, NFE ownership. entry, and exit

	Dependent variable:							
	Owner.		3G AdoP.	Entry		Exit		3G Adop.
	OLS FE	2SLS	First stage	OLS FE	2SLS	OLS FE	2SLS	First stage
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
3G Adoption	-0.0114	0.141		0.0775	-0.206	-0.126	-0.373	
	(0.0489)	(0.198)		(0.0912)	(0.362)	(0.0824)	(0.464)	
Lightning PC			-3.650***					-3.544***
			(0.548)					(0.980)
Lightning PC^2			5.909***					5.191***
			(1.331)					(1.451)
Constant	0.625***		0.189***	0.199*		0.0839		0.250***
	(0.0844)		(0.0241)	(0.110)		(0.101)		(0.0402)
Observations	13,297	13,297	13,297	8,912	8,912	8,912	8,912	8,912
R-squared	0.690		0.620	0.452		0.457		0.839
Controls	YES	YES	YES	YES	YES	YES	YES	YES
Household FE	YES	YES	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES	YES	YES
Sargan (p value)		0.28			0.68		0.50	
Fstat			29.69					6.537

Note: The dependent variables measure, respectively: whether households own an NFE (columns 1-2); whether households start an NFE at time $t+1$, conditional on not having an NFE at time t (columns 4-5); whether households leave the market at time $t+1$, conditional on having an NFE at time t (columns 6-7). All estimations are run using a 2SLS estimator. 3G adoption measures the percentage of 3G coverage in the household's LGA, weighted by the number of mobile phones per capita available in each household. All regressions include the following controls: ratio of female household members (female); household members' mean age; number of household members; female-headed household; household is active in agriculture; percentage of household members in working age can read; percentage of household members in working age who have ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Clustered standard errors in parenthesis: *** p<0.01, ** p<0.05, * p<0.1.

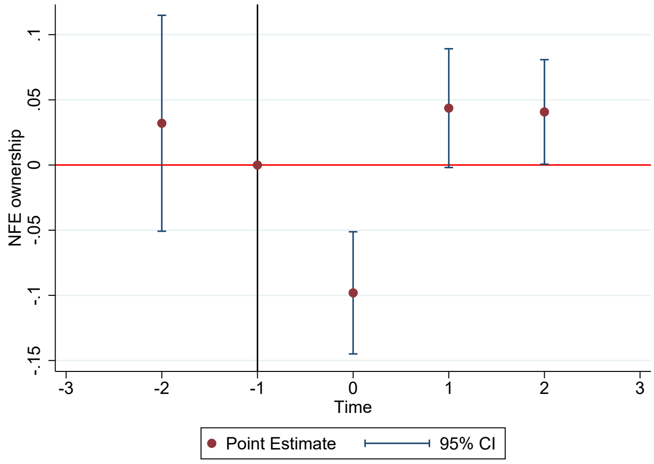

Figure 4.12: Event study: NFE ownership. The event study design uses the first year in which a household hits 0.2 of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. The outcome variable used in the event study specification measures whether households own an NFE. The coefficients reported in the figure come from a model based on Equation 4.4, including household, and year fixed effects. The model uses households as units of observation, and incorporates the following controls: percentage of female household members; mean age in the household; number of household members; female-headed household; household is active in agriculture; percentage of household members in working age who can read; percentage of household members in working age who have ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

The analysis builds on the Inclusive Structural Change framework (Ciarli et al., 2021b) and aims to identify the trade-offs and synergies that exist between innovation, structural change and inclusion in a context dominated by informal firms, in this case Nigeria.

First, the findings suggest that the adoption of a rapidly diffusing and ground-breaking technology, such as mobile internet, has a strong effect on the performance of Nigerian NFEs, and on their trajectory of structural change. Basing the analysis on a double identification strategy (IV approach exploiting geographical variation in lightning strikes, and an event study), the study finds that 3G adoption is associated with higher sales per worker, on average. However, benefits are likely to be concentrated in low-productivity services, particularly trade activities, with households moving away from manufacturing. Overall, the uptake of mobile internet is found to increase NFE performance, fuelling the process of tertiarisation of the economy.

The results on the effect of the internet on inclusive or exclusive structural change (Saha and Ciarli, 2018; Ciarli et al., 2021b) require a deeper discussion. The findings indicate that the improved NFE performance is also a result of lower labour requirements by firms; we investigated what happens to the labour shed by internet-adopting NFEs, finding that NFEs tend to free up household labour, and to retain (or increase) external labour, possibly because adopting firms will require new and higher levels of skills that are not available within the household. The study also examined what happens to excess household labour in NFEs by investigating the effect of mobile internet adoption on working opportunities outside household NFEs: results indicate that, overall, working opportunities outside the household increase for all households, whether or not they own an NFE. For NFE-owning households, however, mobile adoption does not increase the total employment rate (including labour employed within and outside the household), indicating that the household labour shed by NFEs does not negatively affect the employment rate in NFE owning households, who benefit from the creation of jobs outside the household. Household labour freed by the NFEs is therefore only displaced, from within the household to external activities. Coupled with the findings that, for households that do not own an NFE, total employment rates increase, and that NFEs hire more labour outside the household, it can be concluded that the overall effect of mobile internet adoption on labour outcomes has been positive in Nigeria.

Regarding the effect of mobile internet adoption on the creation of new entrepreneurial opportunities, the results portray a different scenario. In fact, mobile internet does not appear to encourage the entry of new firms, indicating that digitalisation in the informal sector may lead to concentration rather than creation of opportunities, as suggested by Altenburg et al. (2021). This finding can be explained by two results concerning the drivers of the increase in NFE performance. First, mobile adoption has a positive effect on sales, suggesting that the benefits are concentrated mainly in incumbent firms, which increase their market share. Second, mobile internet adoption does not affect the costs of operating an NFE: if costs are not reduced, barriers to entry will persist, thwarting entrepreneurship and new entries to the market, which are pivotal to making structural change an inclusive process. However, it cannot be ruled out that adopting households refrain from entering the market, given the availability of new job opportunities either in the formal or in the informal sector, but in any case outside the household (Bahia et al., 2020, 2021; Houngbonon et al., 2022). Moreover, as the effect of mobile internet on firm-performance is concentrated in services, households owning NFEs in other industries are not likely to reap any benefits, and will therefore be excluded by the process of structural transformation.

In summary, the analysis conducted in this chapter has shown that the adoption of 3G mobile internet does indeed have the potential to transform the economy and propel inclusive structural change in Nigeria, especially in terms of the creation of working opportunities. However, some challenges for inclusion persist if the benefits are only to be reaped by existing firms in the trade industry. Policies will play a fundamental role in unlocking the potential brought by the rapid (although late) diffusion of mobile internet in Nigeria and SSA, by making sure that workers displaced by household firms are occupied in industries that are compatible with their skills, in such a way as to enable the absorption of both skilled and unskilled labour. In fact, mobile internet has been shown to have a positive effect on the creation of new jobs, both inside and outside the household, and policies will be crucial to the alignment of incentives in a way that 'oils' the mechanisms of labour creation, reallocation, and upgrading described above. The inclusiveness of the transformation brought about by mobile internet will also depend on whether entry barriers are actively removed, as technological change will not reduce the cost of entry for new firms, but more likely will lead to concentration. The removal of such constraints will be essential if we are to tap into the potential of new entrants and to foster a virtuous and inclusive process of structural change.

Chapter 5

Mobile Internet, Skills and Structural Transformation in Rwanda[1]

5.1 Introduction

The diffusion of fast internet and complementary technologies has potentially disruptive effects on both the scale and composition of employment. Existing evidence from high-income countries shows that a higher complementarity between digital technologies and skills is leading to polarisation in the labour market (see, among others, Autor et al. 2003; Goos et al. 2014; Autor 2015; Buera et al. 2021). Turning to low-income economies, there is a general awareness of the relevance of technological change in either fuelling a catching-up process (see, among the others, Fagerberg and Verspagen 2021 and the literature reviewed by Vivarelli 2021) or in causing de-industrialisation (for its effect in combination with globalisation, see Rodrik 2016). There is, however, little evidence relating to the impact on the labour market of Infor-

[1]This chapter has been prepared as part of the "Skills and transitions" research project by the Research Department of the International Labour Organisation (ILO), and it has been co-authored with Marco Grazzi, Martina Occelli, and Marco Sanfilippo. A previous version of this chapter has been published on the ILO Working Paper Series: https://www.ilo.org/global/publications/working-papers/WCMS_843046/lang--en/index.htm.

mation and Communications Technology (ICT) applications made available through internet access. This in spite of the fact that ICT is one of the most transformative technologies, spreading rapidly across low- and middle-income countries.

Over the past two decades, low- and middle-income countries have experienced a substantial boost in the diffusion of broadband connectivity. In Sub-Saharan Africa, 30 per cent of the population now has access to the internet and mobile phone subscriptions stand at over 90 per cent: in both cases, the figures have more than doubled since 2010.[2] In a context in which hard infrastructure, such as fixed telephone lines and cables, is rarely available, mobile phones are the most common means by which Africans access the internet (Manacorda and Tesei, 2020).

In this chapter, we look at the expansion of the mobile internet network in Rwanda and analyse its implications in terms of structural transformation, focusing mainly on changes in the labour market.

The case of Rwanda is particularly interesting for the purposes of this study. The role of the ICT industry is deeply embedded in the national development strategies. The country's industrial policy until 2020, grounded in its Vision 2020 strategy (Ministry of Trade and Industry, Government of Rwanda, 2011),[3] aimed to diversify the economy and explicitly promote the transition towards a knowledge-based economy in which science and technology education, and ICT skills are actively encouraged. Since its inception in 2015, the Government's Smart Rwanda Master Plan has highlighted the objective of building a knowledge-based society, founded on the digital transformation of seven key areas, namely governance, education, health, finance, women and youth empowerment, trade and industry, and agriculture. This is combined with a strategy to ensure universal access to broadband connectivity. This package of policies led to the fact that by 2018, 4G mobile coverage had already reached over 96 per cent of

[2]World Bank, "World Bank Development Indicators", accessed 8 November 2021.

[3]While the Vision 2020 national development strategy has been updated with the more recent Vision 2050 strategy in December 2020 (Ministry of Finance and Economic Planning, Republic of Rwanda, 2020), throughout the study we will refer to policies that were in place over the period under consideration (between 2012 and 2019), in order to capture the synergy between Rwanda's policies and the contemporary developments of technical and structural change in the country.

Rwanda, with 47.7 per cent of the total population able to access the internet.[4]

The rapid roll out of internet has the potential to trigger changes in the structural transformation process and labour markets which are difficult to assess *a priori*. On the one hand, the exponential growth of ICT-related activities, such as mobile internet, and the drastic cost reduction of transmitting information have dramatically expanded the value of services in Africa (Fagerberg and Verspagen, 2021). On the other hand, the expansion of services at low levels of GDP per capita (Owusu et al., 2021), coupled with a stagnant industrialisation (Rodrik, 2016), makes the question of whether mobile internet can play a positive role in economic development all but trivial. Understanding the mechanisms at work between mobile internet, employment, and structural transformation is therefore an empirical question that we aim to explore in the context of Rwanda.

More specifically, in our analysis we link the roll out of mobile internet in Rwanda to a number of outcomes related to changes in the size and composition of employment in the country. This includes a shift towards more highly skilled occupations and/or activities characterised by higher value added per worker across industries.

The analysis is based on the collection and harmonisation of data from two main sources. The first source is the Global System for Mobile Communications Association (GSMA), which provides information on the coverage of different mobile technologies (2G, 3G and 4G) over time and across locations within the country. The second is individual-level data from population censuses and labour force surveys. The harmonisation of these two sources allows us to obtain consistent indicators of labour market participation covering a sufficiently long time span, which ranges from a baseline year with no internet coverage (2002) to the most recent year (2019). In the study we use districts, the second administrative level in Rwanda, as the unit of analysis.

Our analysis exploits the staggered – across districts and time – roll out of the 3G

[4]Ministry of ICT and Innovation (MINICT), "Digital Transformation Directorate General: Mandate" accessed 22 March 2022.

network and employs an econometric specification with district and time fixed effects that links changes in the coverage of mobile internet to changes in employment in each district over time. Given that the decision on where and when to introduce mobile technologies is unlikely to be 'as good as random', we base our analysis on an instrumental variable approach that – following existing literature (Manacorda and Tesei, 2020; Guriev et al., 2021) – exploits the geographic variation in the incidence of lighting strikes as a factor influencing the distribution of the mobile network within the country.

Our results show that improvements in the coverage of 3G mobile internet technologies affect the composition of the labour market in two distinct ways: (1) through an increase in the share of employed individuals, among whom are both skilled and unskilled workers, with the former increasing at a faster rate, given their relatively small initial size; (2) through a sectoral shift of employment towards services and, within the service sector, to some high value-added and skill-intensive industries. Results are robust to a battery of additional checks, including changes in the specification and the adoption of an event study approach. To rationalise some of these findings, we run additional analyses showing that improvements in mobile internet coverage are also related to: (1) an increase in the number of years of schooling in the younger population; and (2) an increased supply of workers in treated locations due to increasing shares of migrants. We also find evidence on demand-side mechanisms at play. Using administrative data on Rwandan formal firms (in all industries) we show evidence of agglomeration of more productive ones in locations with higher 3G coverage.

The remainder of the chapter is structured as follows: Section 5.2 presents the theoretical framework of the study, Section 5.3 introduces all the data used in the analysis; Section 5.4 discusses the empirical specification and the identification strategy based on a 2SLS estimator; Section 5.5 reports the main results and a set of robustness checks; Section 5.6 concludes.

5.2 Background literature

Structural transformation refers to the reallocation of production factors from low to high productive employment. Traditionally, structural transformation has consisted in the shift of economic activity from agriculture to manufacturing, and then to services (Kuznets, 1973). Technological change plays a significant role in driving the process of structural transformation, allowing the emergence of new, and more productive industries (Saviotti and Pyka, 2004). Trajectories of reallocation of labour and value added have differed across the world (Bah, 2011). Moreover, and despite the expectations of higher returns for countries far away from the technological frontier, not always has technical change led to catching up (Verspagen, 1991; Cirera and Maloney, 2017).

In Sub-Saharan Africa the shift from subsistence and informal activities to high-productivity occupations has been rather slow, with labour moving directly from agriculture to low-productivity services (McMillan et al., 2017a; Baccini et al., 2021). This process has been accompanied by a contraction of technologically dynamic and high productive industries, such as manufacturing (Rodrik, 2016). Causes are numerous: among others, lack of adequate policy (Mkandawire, 2014), human capital deficits (Stiglitz et al., 2013), insufficient infrastructures (Oqubay and Ohno, 2019), globalisation and labour-saving technical change (McMillan et al., 2014).

At the same time, the exponential growth of ICT-related activities has dramatically expanded the value of services (Fagerberg and Verspagen, 2021; Hsieh and Rossi-Hansberg, 2019), challenging the historical argument that an expansion of the service sector hampers long-run economic growth (Baumol, 1967). On the one hand, the continuous rise of Africa's service sector could offer new development opportunities, supported by technical progress (de Vries et al., 2015; Newfarmer et al., 2018). On the other hand, however, technical change may also favour workers with higher skills over those with lower skills, with technology replacing the latter (Autor et al., 2003; Akerman et al., 2015; Buera et al., 2021).

It has been argued that the current wave of technological change, characterised by the diffusion of ICT technologies, can represent a window of opportunity for African countries to leapfrog towards modern services (Fagerberg et al., 2021; Kaplinsky and Kraemer-Mbula, 2022). Nonetheless, the literature looking at the impact of technological change on structural transformation and employment in African contexts is scant. In order to establish whether the diffusion of digital technologies, such as mobile internet, will deliver a virtuous structural transformation, it is crucial to assess whether it will lead to the creation of employment, who will benefit from it, and in which sectors of the economy.

For many Africans, mobile internet has represented the first opportunity to connect to the internet. The diffusion of mobile internet coverage, along with other ICT and digital applications that it enables, can be considered a General Purpose Technology (GPT) (Cardona et al., 2013): given its disruptive nature, it has the potential to transform labour markets in different ways, which are difficult to assess *a priori* (Ciarli et al., 2021a). Kaplinsky and Kraemer-Mbula (2022) have noted that mobile phones do not depend on a centralised grid, they are cheap, can be shared by more than one user and, focusing on a distinguishing feature of GPTs, they have an impact across a large number of different economic activities, including farming (on this last application see, for instance, Mehrabi et al. 2021). Nonetheless, the wide range of innovations that might be enabled by internet connectivity is likely to exert contrasting effects on employment dynamics: studies anticipating the overall impact on the labour market are far from being unanimous.

Internet connectivity can be the engine of both labour-augmenting and labour-saving technological change. Greater connectivity affects labour-biased productivity directly, and can support human capital accumulation by increasing training opportunities (both on the job and in educational settings). Evidence summarised by Hjort and Tian (2021) shows that improved access to the internet has been found to enhance firms' productivity (India), workers' wages (Brasil) or both (China).[5] Mobile technologies

[5]There is evidence on the capacity of mobile internet diffusion to increase employment opportu-

can, however, be biased towards skilled workers (including those performing non-routine tasks), who can benefit disproportionately from better connectivity. This has the potential to increase labour market inequality. However, evidence on this is more nuanced. Bahia et al. (2021) find that, in Tanzania, it is mainly the better educated workers who take advantage of improvements in mobile connectivity. Hjort and Poulsen (2019), on the other hand, show that the arrival of fast internet in Africa has benefitted both poorly and more highly educated workers, although the latter have gained the greater benefit. This is possibly related to a demand side effect: fast internet coverage seems to promote both the entry and the performance of more productive and technologically intensive firms. Moreover, internet expansion unlocks the potential for firms to benefit from internet-enabled services, such as mobile money and e-commerce (Hjort and Tian, 2021). Mobile money, which requires internet connectivity for its underpinning infrastructure, has been found to stimulate demand both by increasing consumption and supply and by fostering enterprise development (for a review of the evidence see Suri et al. 2021). Electronic commerce, on the other hand, provides firms with the opportunity to expand into new markets at relatively low cost. Furthermore, evidence from African countries shows that the arrival of fast internet has promoted firms' export, with associated benefits for local employment (Hjort and Poulsen, 2019).

We aim to add to this debate, investigating links between mobile internet, employment and structural transformation in the context of a low-income country which extensively promotes mobile internet through policy.

nities. For instance, a recent paper by Bahia et al. (2020) on Nigeria reports significant employment uptake following the roll-out of mobile internet at the subnational level.

5.3 Data

5.3.1 Mobile internet

We collect information on mobile coverage in each of the 30 districts of Rwanda, drawing upon data made available by the GSMA in partnership with Collins Bartholomew.

The original data consist of a raster of 1 km × 1 km cells, with a layer of information for each technology (2G, 3G, 4G). While 2G (GSM) supports voice calls and messaging, the main technologies of interest in our study are 3G and 4G, which support the use of mobile broadband internet services. In each layer, cells take the value of 1 if the area is covered by a mobile signal, and 0 otherwise. In order to identify the share of the population with access to mobile internet at the district level, this information is combined with a population density grid, available at the same resolution and obtained from NASA's Socioeconomic Data and Applications Center.[6] In every district, the share of population with access to mobile internet is given by the sum of the population living in cells covered by mobile internet divided by the total population.[7]

Mobile internet technologies were introduced into Rwanda at the end of the 2000s. According to the GSMA data, before 2009 only the 2G technology was available. After 2009, 3G internet technologies started to be introduced in a staggered manner across districts and over time (see Figure 5.1). In contrast, the diffusion of the 4G network has been sudden. Developed in partnership with the South Korean firm, KT, the roll-out of the network began in 2015, reaching almost universal coverage within a couple of years. However, the number of subscriptions to the 4G network is still lagging behind those of other technologies.[8]

[6] Centre for International Earth Science Information Network CIESIN Columbia University (2018), "Population dynamics", accessed 25 February 2022.

[7] The data cover only up to 2018, and we assume no major changes in the following year. Results remain robust when dropping 2019.

[8] A recent report of the Rwanda Utilities Regulatory Authority (RURA) shows that currently only 1.2% of the total number of mobile broadband subscriptions use 4G technologies (see Rwanda Utilities Regulatory Authority (2021): Table 13).

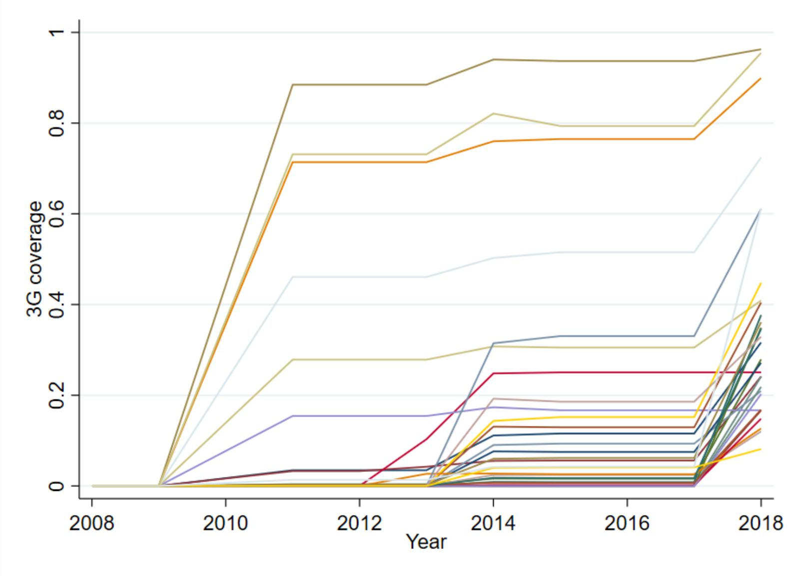

Figure 5.1: District-level mobile internet 3G coverage diffusion (2008–18). Source: GSMA data. Lines represent the mobile internet 3G coverage diffusion in each district.

5.3.2 Individual-level data

To build our indicators of labour-market participation, we combine the two most recent waves (2002 and 2012) of the Rwanda National Population and Housing Census[9] (from IPUMS International) with three waves (2017, 2018 and 2019) of the nationally representative Rwanda Labour Force Survey (RLFS)[10] (from the National Institute of Statistics of Rwanda). We aggregate the individual-level information to obtain a district-level[11] panel dataset on a sample restricted to the working-age

[9]The third (2002) and fourth (2012) waves of the Rwanda National Population and Housing Census include demographic and socio-economic information about the total population of the country. The national territory is divided into enumeration areas (EAs), each the size of a village, typically including 150–200 households; for each EA, households are listed and 10% of randomly selected households in the EA are interviewed. In the third wave, 843,392 individuals were interviewed and in 2012 interviewees totalled 1,038,369.

[10]The RLFS has been implemented in 2017, 2018 and 2019 to monitor the trend in employment and labour underutilisation at the national, province and district level (National Institute of Statistics of Rwanda, 2018). Samples in each year are constructed using a two-stage sampling procedure: during the first stage, a stratified sample of EAs from the latest population census is drawn with probabilities proportional to size. During the second stage, a fixed number of households is selected with equal probability within each sample EA. Finally, all qualifying household members in the sample households are selected for survey interviewing: 77,719 (2017), 76,670 (2018) and 81,778 (2019).

[11]The district, the second administrative division of Rwanda (ADM2), is the lower level of geographic disaggregation at which we can combine the information of the censuses and the RLFS. Other administrative units are the province (ADM1) and the sector (ADM3). In 2006, Rwanda implemented a reform of its administrative boundaries: 12 provinces were replaced with 5 larger provinces and the number of districts dropped from 106 to 30. In our dataset, districts and provinces in 2002 have been collapsed in such a way as to reflect the administrative boundaries introduced by the 2006 reform.

population (which we define as covering individuals aged 15 to 64 years old).

Note that, following changes that occurred in the international labour statistics standard, which narrowed the definition of employment to those working for pay or profit,[12] throughout the sample we consider subsistence farmers as not being in employment.

Despite differences in scope, the combination of these two data sources is made possible by the presence of a wide range of comparable and geographically detailed demographic and socio-economic information. Both data sources provide individual and household sampling weights which allow creating representative figures at the district level. Based on this information, we compute indicators related to (i) occupations, (ii) industries and (iii) education.

Occupations: As we are keen to capture the dynamics of skilled occupations over time, we adopt the ISCO division of occupations into skill levels (ISCO-2008) (International Labour Office, 2012). Our data include a 3-digit ISCO 88 code for each employed individual in 2002 and a 4-digit ISCO 08 occupation for all individuals in subsequent years. Unfortunately, the break between the two classifications means that a one-to-one harmonisation exercise cannot be performed between classifications. However, ISCO major occupation groups (at the 1-digit level of the ISCO classification) have remained unchanged; this allows grouping based on skilled and unskilled occupations, according to the ISCO skill groups, to be created. Skilled occupations consist of skill levels 3 (professionals) and 4 (managers, technicians and associate professionals); unskilled workers are those belonging to skill levels 2 (clerical support, services and sale, skilled agricultural, craft and related trades, plant and machine operators) and 1 (elementary occupations).[13]

[12]See the 19th International Conference of Labour Statisticians (International Labour Organization, 2013) The Rwanda National Institute of Statistics has integrated these changes and, since 2017, the definition of employment no longer includes subsistence workers, leading to a consistent drop in the share of agricultural employment. For a discussion on some practical implications of this change in definition, see also Gaddis et al. (2020).

[13]We exclude armed forces and subsistence agricultural producers from the sample. The former only appear in 2019; the latter have been considered unemployed since 2017, creating an inconsistency in the longitudinal dimension of the dataset.

Table 5.1 shows the average, across districts, of the labour shares in each ISCO major occupation group between 2002 and 2019. The most striking figures are those related to elementary occupations, increasing from 2 per cent to 24 per cent between 2002 and 2019, and skilled agricultural workers, decreasing from 37 per cent in 2002 to 3 per cent in 2012. Rather than reflecting an abrupt shift in Rwanda's employment structure, this change is likely to be due to a reclassification of many agricultural occupations in the census (2002 and 2012) from skilled to elementary occupations (2017–19). Nevertheless, as these two groups both belong to the unskilled occupations group, the reclassification of agricultural occupations is not a major concern for our analysis.

It is worth noting that, while major occupational groups have remained unchanged, the more disaggregated occupations attributed to each group have changed. For instance, agricultural managers used to be considered as part of the Managers ISCO 88 major group (skill group 4) but have been moved to Skilled agricultural workers (skill group 2) under the new ISCO 08 classification.[14] As a result, jobs that used to be considered skilled under ISCO 88 are now considered unskilled under ISCO 08. Despite this, skilled occupations still exhibit a slow but steady upward trend both on average across districts (Figure 5.2), with significant concentration in the urban districts, such as the capital, Kigali.

Table 5.1: Share of occupations, average across districts (2002-–19)

Occupation	Skill level	2002	2012	2017	2018	2019
Crafts	Unskilled	1.4	3.2	3.27	3.63	3.59
Elementary	Unskilled	1.96	2.83	25.43	25.3	24.48
Services	Unskilled	1.51	5.07	7.7	8.3	8.5
Clerks	Unskilled	0.29	0.2	0.33	0.33	0.35
Machinery	Unskilled	0.34	0.96	0.97	1.11	1.14
Agriculture	Unskilled	37	39.48	2.76	3.3	3.08
Technicians	Skilled	0.28	0.63	0.56	0.49	0.56
Managers	Skilled	0.08	0.21	0.49	0.46	0.48
Professionals	Skilled	0.68	1.45	2.45	2.64	2.22
Not in employment*		56.45	45.97	56.05	54.44	55.61

Note: *Not in employment refers to those individuals not currently working and to those working in subsistence farming. Source: Authors' elaboration on national census and RLFS data.

[14] For the full list of inclusions and exclusions of disaggregated occupations between ISCO 88 and ISCO 08, see CEDEFOP (2014).

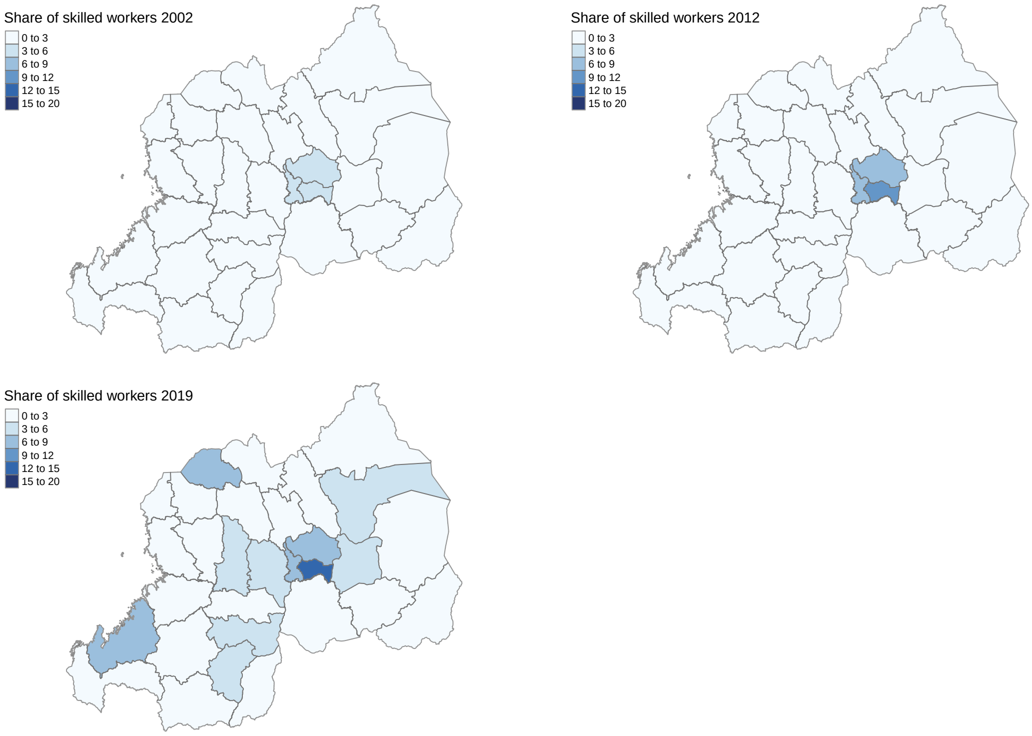

Figure 5.2: Percentage of skilled workers at district level (2002, 2012, 2019). Source: Authors' elaboration on national census and RLFS data.

The correlation between the share of people employed in skilled occupations and the rise of the country's mobile internet coverage is given in Figure 5.3. The figure incorporates information on education and shows that areas with higher internet coverage are those in which highly skilled workers, in terms of both the content of their occupation and their level of education, are employed.

Industries: Both the Rwanda population census and the labour force survey provide information on industries of employment, following the International Standard Industrial Classification (ISIC), revision 3.1. We create variables measuring the employment share of each of the ISIC industries. Table D.1 in the Appendix shows the labour shares across industries (ISIC major groups) over the time span under analysis (2002–19) averaged across districts. The descriptive evidence on employment by industry (Figure 5.4) indicates that, while employment in agriculture decreases, services expand over time. Growth in the tertiary sector has been driven mainly by the growth of trade activities, but with a rising trend in skilled services too, such as finance and health (Table D.1). Employment in manufacturing also increased,

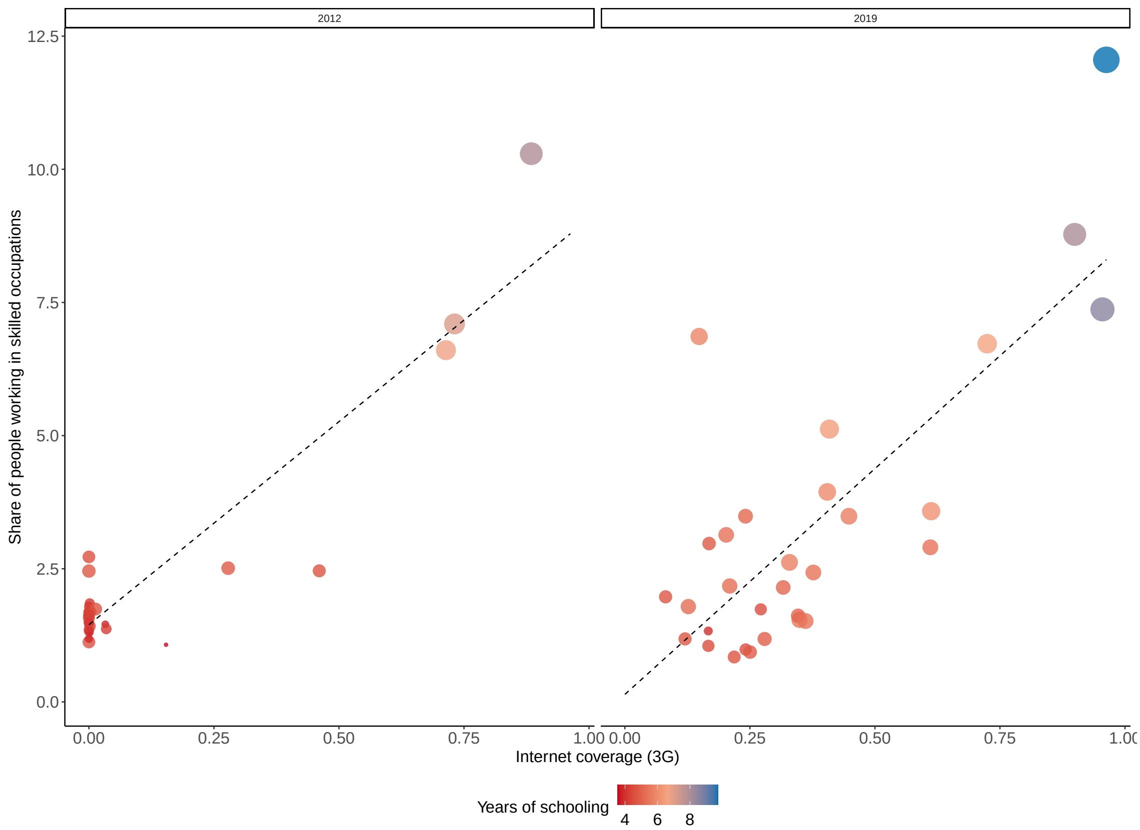

Figure 5.3: Share of people working in skilled occupations and mobile internet 3G coverage (2012 and 2019). The colour of the dots indicates the average schooling years of the working age population in each district, from low (red) to high (blue). Source: Authors' elaboration on national census and RLFS data.

although its share remains relatively small.

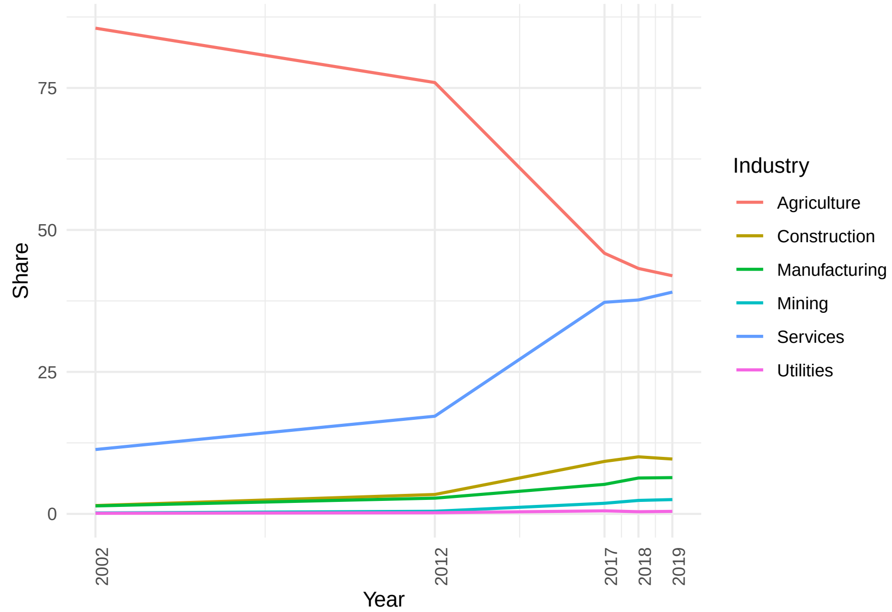

Figure 5.4: Shares on the Y axis indicate the percentage of the working-age population in employment by industry. In each year, shares sum to 100. Source: Authors' elaboration on national census and RLFS data.

Education: Education-related questions in the census and in the labour force surveys are not harmonised. We have selected two questions from the census which present the same formulation in the RLFS. The first is the level of education. Although the question is framed identically in the two questionnaires, the way in which responses are classified does not match. Therefore, we have harmonised answers into a categorical variable, including: (i) no education, (ii) primary (which in the census covers less than primary and primary respondents), (iii) secondary (which in the labour force survey includes lower secondary and upper secondary degree) and (iv) tertiary (which is identical in the two questionnaires). The second variable measures the number of years of education. This information has been collected directly only in the census; for the labour force survey, years of schooling have been elicited using the 2019 wave, which is the only one in which they were reported as a continuous variable. Hence, average years for each educational level are computed for 2019, and then attributed to individuals in 2017 and 2018 based on their educational level.[15]

[15]For example, if a respondent declares that they have achieved the primary diploma in 2017, we compute the number of years of education they have received by averaging out the years of education of a person with a primary diploma in 2019. As the education system has not undergone

For the entire period covered (2002–19), the enrollment age in elementary school is set at six years old. It should be noticed, however, that the so-called basic education, granted for free in Rwandese public schools for nine years (i.e. elementary and lower secondary education-up to grade 9), was extended to grade 12 in 2012. This resulted in a higher enrollment in upper secondary classes, with a jump from 21 per cent in 2011 to 30 per cent in 2017 (Neumann et al., 2012). The need to collapse upper and lower secondary education for all the years in the categorical response prevents the study from capturing this shift. Average years of education have increased between 2002 and 2019 across districts (see Figure D.1 in the Appendix). Furthermore, if we compare the share of people with tertiary education and the increase in mobile internet diffusion at the district level in the time window 2012–19 (Figure 5.3), we observe a positive correlation.

Additional variables: We use the RLFS data to cross-check the information available on mobile internet coverage. Specifically, we use questions asking respondents about (a) the presence of mobile phones (Figure D.2 in the Appendix) and (b) internet connections at home (Figure D.3 in the Appendix). In both cases, we find a correlation with the mobile internet data and the self-reported data from the RLFS, with the jump being driven by urban districts.[16]

5.4 Empirical specification

In our empirical analysis, we are interested in understanding how changes in the spatial and temporal variation of mobile phone coverage are correlated to changes in the composition of the labour force in Rwanda. Our empirical specification follows the existing literature (Manacorda and Tesei, 2020; Guriev et al., 2021) and links the roll-out of mobile internet coverage to the outcomes of interest, as follows:

any changes over these three years, we consider this assumption to be realistic. Note also that we consider only individuals who have completed the education level.

[16]We define an urban district as a district where at least 60% of the population lives in urban areas. In the years analysed, these are Gasabo, Nyarugenge and Kicukiro.

$$Y_{it} = \beta 3G_{it} + \gamma X'_{it} + \theta_i + \delta_t + \epsilon_{it} \tag{5.1}$$

where Y_{it} is one of the variables defining the labour market in district i at time t.

We will present results on the basis of three different sets of outcomes. First, the size of employment, using the share of persons in employment in the working-age population. Second, the distribution of workers by skill level. This analysis is based on the information drawn from the occupations classified as discussed in Section 5.3.2. Third, the distribution of workers across industries. This classification mimics the pattern of structural transformation of the country, by looking at whether increases in coverage of the mobile network correlate with shifts of workers from less to more modern activities across industries.

Following the discussion in Section 5.3, our variable of interest is $3G_{it}$, which measures the share of a district's i population covered by the 3G signal in any given year t. In our baseline specification, we use 3G expansion, as this technology was the first to allow users to browse and create online content. The expansion of 4G technology was sudden and quickly reached almost universal coverage in the country while still being the least widely adopted by users: these characteristics do not allow enough variation in the data to exploit. In contrast, the timing of the introduction of the 3G technologies is ideal to combine with labour force data. While the technology was formally introduced in the early 2000s, the roll-out covered only a few districts before 2012, and even those had minimal coverage (see Figure 5.1). After 2012, coverage expanded to other districts, but still not uniformly.

X'_{it} is a vector of district-specific controls. These include characteristics drawn from the survey data, i.e. the average age of the population and the percentage of female population on total. We also add variables that account for the geographic characteristics of the district.[17]

[17]These variables include the stability of malaria; terrain's ruggedness; the suitability of the

Finally, in all regressions we include a coefficient measuring the share of a district's population covered by the 2G network. This is added to ensure that our coefficient correctly identifies the contribution of the upgrade to 3G coverage, and not merely the expansion of the network. If a location is covered by the 3G network, it is in fact also covered by 2G. Hence, controlling for 2G should isolate the net contribution of the 3G technologies (a similar strategy is adopted by Bahia et al. 2021).

We include district (θ_i) and wave (δ_t) fixed effects. This reduces our identification to one that explores the changes over time in the outcomes of interest within each district which are (conditionally) correlated with the corresponding changes in the roll-out of the mobile broadband network. All the regressions are weighted using the districts' total population. Standard errors are clustered by district, which is the level of the treatment.

The estimation sample consists of a balanced panel covering the 30 Rwandan districts over the 5 waves of the combined censuses and national labour force surveys. Summary statistics of the variables of interest are reported in Table D.2 in the Appendix.

Identification strategy: Equation (5.1) will be will be correctly identified under very restrictive conditions, i.e. that the roll-out of the broadband network is not influenced by existing pre-trends (which however can be controlled for, as shown by (Rambachan and Roth, 2021), so that the treatment is 'as good as random', at least after conditioning for district fixed effects and time varying controls. These assumptions can arguably be questioned under different circumstances. Not only can initial conditions influence the decision to prioritise investments in connectivity, but the same could be said for some (omitted) variables that we cannot precisely account for in our analysis. In what follows, we try to address both of these issues while being aware that – absent an experimental setting – causal interpretation of the findings could be hard to achieve in our case.

---

terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the variable are measured at the district level and were made available by Alesina et al. (2021). As most of these variables are time invariant, we have interacted them with a time trend as done by Manacorda and Tesei (2020).

In order to deal with endogeneity, we employ an instrumental variable (IV) approach based on a two-stage least squares estimator (2SLS). To do this, we use an instrument previously adopted in other papers that consider the roll-out of the mobile network in contexts which, similar to ours, exploit subnational level information (Guriev et al., 2021; Manacorda and Tesei, 2020; Mensah, 2021). The instrument exploits differential intensities in lightning strikes across districts to explain differences in the coverage of the mobile network. The rationale for the use of such an instrument is that mobile phone infrastructure is affected by frequent electrostatic discharges caused by storms (Manacorda and Tesei, 2020). Hence, the more frequently an area is affected by lightning strikes, the more costly it becomes to construct such infrastructure (Guriev et al., 2021). With respect to the exclusion restriction – i.e. that lightning strikes do not affect our outcomes of interest directly – Andersen et al. (2012) provide evidence that the effect of lightning strikes on aggregate economic outcomes only occurs through its effect on investments in ICT technologies. They show in fact that the density of lightning strikes is a time-stationary process which has started showing a negative correlation with labour productivity growth of US states only after the 1990s, due to the effect of power spikes and dips caused by cloud-to-ground strikes on ICT user cost.

To build our instrument, we use lightning strike density data provided by the World Wide Lightning Location Network (WWLLN) Global Lightning Climatology and Timeseries. The raw data come in a raster of 5-arcminute cells (around 8 km $\times$ 8 km at Rwanda's latitude), with a unique layer measuring the number of daily strikes per square kilometre. The measure is taken every month and it is currently available for the period between 2010 and 2020. To capture a district's exposure to lightning strikes, we have averaged the lightning strike density over the period covered by the data in each cell[18] and aggregated cell values by district, taking their mean. The resulting measure of daily lightning strikes per square km in every district is then

[18] Although the definition of the instrument adopted is the best in terms of first-stage statistics, our results remain unaffected by changes in the construction of the instrument. In particular, we have experimented with (a) using initial values of lightning instead of their average over the period and (b) removing the population from the denominator.

converted into daily lightning strikes per inhabitant[19] by multiplying the measure by each district's area and dividing it by its population. The resulting time-invariant measure of daily lightning strikes per capita at the district level is then interacted with a time trend, following Guriev et al. (2021).

5.5 Results and discussion

In this section, we discuss the findings of our empirical analysis. Each regression relates one of the labour market outcomes to the expansion of broadband internet coverage within each district over time. The unit of observation is the district, which is also the level at which standard errors are clustered. We organise the discussion of the main results into three different sets of outcomes: employment, occupations, and industries.

Employment: Table 5.2 reports a first set of results linking mobile internet coverage to jobs, measured as the share of employment among the working-age population. Column 1 provides the unconditional ordinary least squares (OLS) estimates, while column 2 introduces district and year fixed effects, along with all the controls. In both cases, the coefficient of 3G coverage is positive and statistically significant, indicating a positive correlation with employment. The coefficient of the 2G coverage does not correlate significantly with the outcome, meaning that, if anything, the relationship between broadband internet and employment has mainly to do with the introduction of technologies that allow the internet to be accessed from mobile phones. Columns 3 and 4 report the first and the second stage of the IV estimate, respectively. The coefficient of the first stage regression (column 3) displays a negative coefficient that is highly statistically significant. This proves the validity of the instrument showing that those districts that are more likely to be affected by frequent lightning strikes have lower mobile network coverage. The F-statistic reported at the end of column 4 is well above 10, which further confirms the strength of the instrument adopted.

[19] As the size of the variable is small, to give a better interpretation of the coefficient of the first-stage regression we have computed it for 1,000 inhabitants.

The coefficient of interest in column 4 remains positive and is highly statistically significant.

Compared to the OLS estimation, the coefficient of the 2SLS estimation is larger. The size and the direction of the bias are similar to (if not smaller than) the results reported by Manacorda and Tesei (2020). There are a few possible reasons to expect a downward bias of the OLS coefficient. In addition to the possibility of a measurement error, which will bias the OLS coefficient to zero, and the presence of omitted variables, one explanation is that the districts most strongly influenced by the instrument are those with higher potential for employment, i.e. those starting from a position of lower employment levels.

As such, the economic interpretation of the coefficient is relevant. A move from the sample's 25th percentile of the distribution of mobile internet coverage to its 75th percentile is associated with an 11 percentage point increase in the share of employment, which is a 23.3 per cent increase from the sample average.

Occupations: The first two columns of Table 5.3 report findings covering the relationship between 3G mobile internet coverage and variables measuring the skill content of occupations. We find that the spread of mobile internet is positively related to a growth in both skilled and unskilled types of occupations. Although the size of the coefficients is higher for the unskilled, the quantification exercise shows that mobile internet matters relatively more for highly skilled employment, a finding that is consistent with related evidence from African countries (Hjort and Poulsen, 2019). A move from the 25th to the 75th percentile of the distribution of mobile coverage does, in fact, contribute to raising skilled employment by about 75 per cent, compared to 20 per cent for the unskilled.

Industries: Next, we check whether the roll out of mobile internet matters for the process of structural transformation occurring across industries at the district level. Over the past 20 years, Rwanda has experienced a process of structural transformation that is common among African countries, i.e. one that sees reduction

Table 5.2: OLS and 2SLS results, employment

	Dependent variable:			
	Employment			
	OLS	OLS	2SLS (1S)	2SLS (2S)
	(1)	(2)	(3)	(4)
3G	0.0632**	0.0475*		0.341***
	(0.0251)	(0.0265)		(0.0712)
2G		0.00247		-0.109
		(0.117)		(0.113)
Age		-0.0135		-0.0112
		(0.00848)		(0.0103)
Female		0.248		-0.456
		(0.249)		(0.274)
Malaria × t		4.28e-06		-0.00899
		(0.00557)		(0.00550)
Ruggedness × t		-0.000106***		-0.000132***
		(3.23e-05)		(4.21e-05)
Agricultural suitability × t		-0.0259		-0.0112
		(0.0207)		(0.0283)
Distance to coast × t		0.0393		-0.0567
		(0.0909)		(0.0857)
Distance to railway × t		-0.0608***		-0.0453*
		(0.0202)		(0.0252)
Distance to border × t		0.00132		-0.00639
		(0.00496)		(0.00526)
Lightning strikes (1,000 pop.) × t			-4.030***	
			(1.155)	
Constant	0.450***	1.019	-11.71	
	(0.00595)	(1.789)	(7.430)	
Observations	150	150	150	150
R-squared	0.062	0.684	0.863	
District FE	NO	YES	YES	YES
Year FE	NO	YES	YES	YES
Mean DV	0.463	0.463		0.463
Quantification	0.0200	0.0150		0.108
District Controls	NO	YES	YES	YES
F-stat				15.81

Note: The dependent variable measures the share of employed individuals among the working-age population. 3G and 2G measures the percentage of the population covered by the respective mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted using a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

in agricultural employment in tandem with the growth of available jobs in the service sector, rather than in manufacturing (see Rodrik 2016; Baccini et al. 2021). An interesting aspect of Rwanda's structural transformation is the focus on some of the service industries with higher potential in terms of jobs and value-added generation (Newfarmer et al., 2018). This includes the tourism industry, as well as financial and business services activities. The latter are also specifically targeted by the country's industrial policy's provisions. Understanding whether this process can be linked in some way to the roll-out of mobile internet coverage would therefore be of relevance. Results of this exercise, reported in columns 3 to 5 of Table 5.3, show that this does indeed seem to be the case. Districts that improved their internet connectivity are also those experiencing an increase in services-related employment. Expecting heterogeneity across services, we checked for the presence of specific patterns at the industry level. Results are plotted in Figure 5.5. While a positive coefficient is generally found for most of the industries within the service sector, those that are statistically different from zero include both highly skilled activities, such as finance and health, and low-skilled ones, such as private services to households. One possible reading of this result is that this particular trajectory of structural change (tertiarisation towards both high and low skills services) benefits the expansion of both high and low skilled occupations, creating job opportunities across different skill levels, and therefore benefiting both segments of the working population. However, with particular reference to the expansion of high-skills services, a potential bottleneck could be represented by the lack of highly skilled labour force suitable for occupations in these sectors Behuria and Goodfellow (2019).

5.5.1 Robustness checks

Alternative specifications: We first check the robustness of our results to alternative specifications. First, we run our analysis introducing province-specific time trends. There are five provinces in Rwanda, which were established in 2006. The introduction of such additional fixed effects, as shown in Table D.3 in the Appendix,

Table 5.3: 2SLS results, by type of occupation and industry of employment

	Dependent variable:				
	Skilled	Unskilled	Agriculture	Manuf.	Tertiary
	(1)	(2)	(3)	(4)	(5)
3G	0.0654***	0.276***	0.0957	0.0408	0.259**
	(0.0207)	(0.0741)	(0.196)	(0.0521)	(0.125)
Observations	150	150	150	150	150
District FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
District Controls	YES	YES	YES	YES	YES
Mean DV	0.0274	0.436	0.585	0.0442	0.285
Quantification	0.0207	0.0873	0.0303	0.0129	0.0820
F-stat	15.81	15.81	15.81	15.81	15.81

Note: The dependent variables measure, respectively, the share of skilled workers among the working-age population (Skilled); the share of unskilled workers among the working-age population (Unskilled); and the share of agricultural, manufacturing, and services (tertiary) in the district's total employment. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

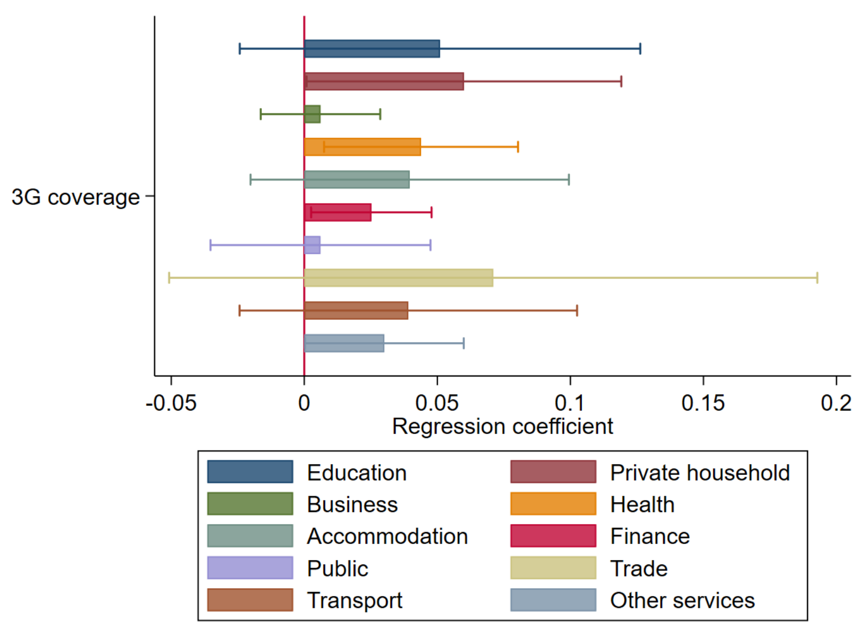

Figure 5.5: The graph reports the coefficient of the variable 3G as estimated from different regressions using the employment share of each service-related industry in the district's total employment as dependent variables. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of the female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator.

does not affect our estimates. Second, in order to deal with pre-trends more effectively, we run an exercise in which interaction terms between time trends and initial values of the outcome variables are included as additional regressors. This should help to alleviate the concern that districts with, for instance, high initial levels of skilled or agricultural employment prior to the roll-out of the 3G network, may experience different trajectories. Results, reported in Table D.4 in the Appendix, show that initial values interacted with time dummies do not alter either the size or the direction of the initial findings.

4G coverage: As discussed in Section 5.4, a potential issue of concern for our analysis is that most of our results are driven (or strengthened) by the introduction of 4G technology, which mainly occurred during the second half of the 2010s. The policy leading to the almost universal roll-out of 4G coverage, and the lack of individual-level data for the years during which this happened, pose a potential threat to our identification strategy. To understand whether the introduction of the latest mobile internet technology affects our results, we have replicated our analysis adding 4G coverage as an additional control. If the effects are explained by differential coverage of 4G, we should observe our coefficient of interest (3G coverage) weakening or losing statistical power. Nevertheless, as shown in Table D.5 in the Appendix, we find that the 3G coefficient explains all of the changes in labour market participation and composition, whereas the 4G coefficient is generally not statistically significant (an exception being the specification on the manufacturing sector, for which the 4G variable reports a negative and, weakly, significant coefficient).

Event study approach: As a final exercise, we take advantage of the panel structure of our data, which allows us to follow all the districts over different time periods, and of the staggered introduction of the treatment to estimate our relationships using an event study approach. Event studies are particularly useful when treatment is not randomised, but outcomes and trajectories before and after treatment, as well as across treated and control units, can be compared. For the purposes of this exercise we define the treatment as a binary variable, i.e. a dummy taking the value of 1 once a district achieves a certain coverage and 0 otherwise. More specifically, we use a value of 11 per cent coverage as a threshold. This value seems the most appropriate, given that it is both the overall sample median as well as the sample mean in 2012 (the first year in which we can observe 3G coverage in our districts).[20] We estimate the event study based on Equation (5.1), i.e. conditioning on the observables and district and year fixed effects and replacing the treatment with a number of lags and leads (a maximum of three on both terms), measuring the distance between each observation and the time at which a district was treated. Figure 5.6 provides a summary of the results. First, and importantly for identification purposes, on a visual inspection there is no evidence of pre-trends potentially affecting the estimation results. Second, the direction of the result is in line with those reported in the previous section. Third, most of the results show that the impact of granting access to mobile technologies is likely to improve over time, which is an important addition to our initial findings, offering some evidence on the dynamic impacts of mobile technologies.

Alternative individual data: We verify whether the relations estimated so far can be replicated using alternative sources of information on individuals' participation to the labour market. For these purposes, we use information from Rwanda's Demographic and Health Surveys (DHS). DHS collects nationally representative data on several socio-economic variables at both the household and individual levels. Relevant to this exercise is the fact that DHS surveys also include a module on employment, which records information about the employment status of each individual aged 15-49, as well as their main occupation. Based on occupational data we can identify whether an individual performs a skilled or an unskilled activity,[21] and if they are employed in the agricultural sector or in modern activities.[22]

[20] Results remain robust to alternative definitions of this threshold, including 20% and 50%.

[21] Following International Labour Office (2012) and Hjort and Poulsen (2019) an individual is classified as skilled if employed in one of the following occupations: clerical, skilled manual, sales, services and professional. Unskilled occupations are unskilled manual, domestic and (formal and informal) agriculture.

[22] Unfortunately, DHS do not include industry level information on employment, hence, following existing evidence workers in agricultural related occupations are assigned to the agricultural sector, while all the others to other industries (Diao et al., 2017).

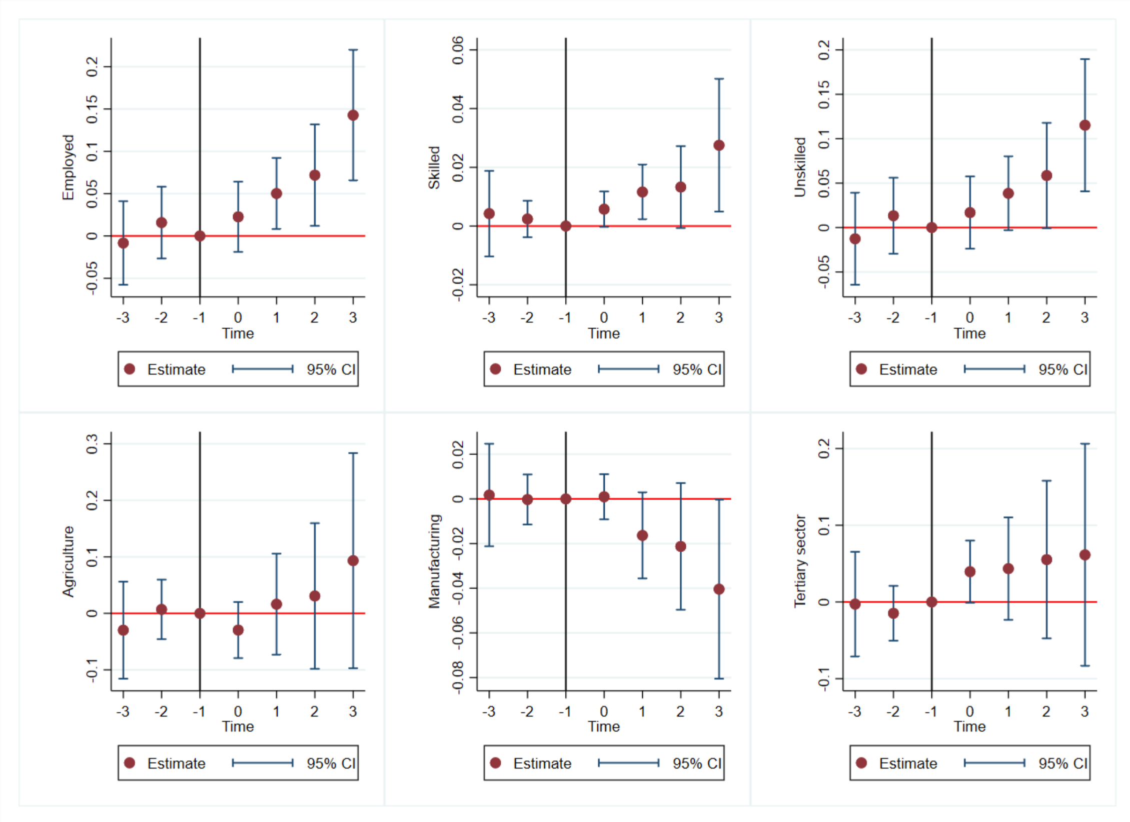

Figure 5.6: The event study design uses the first year in which a district hits 11% coverage of its population by the 3G network as treatment, corresponding to time 0 in the horizontal axis. The coefficients reported in the figure come from a model based on Equation (5.1), including district and year fixed effects, incorporating the following district-specific controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. Standard errors are clustered at the district level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

There are four waves of DHS currently available for Rwanda, run in 2005, 2010, 2014 and 2019. Together, they cover about 61 thousand individuals, equally distributed across districts and over time.[23] As the DHS is representative at the regional level (and within each region, at the urban/rural level), we do not aggregate the data at the district level, but treat them at the individual level adjusting our regressions using sample weights. We run regressions in which a given outcome, measured at the individual level, changes in response to changes in the 3G coverage in the district in which individuals are interviewed. All regressions include the 2G coefficient, as well as individual specific controls (their age and gender), district and year fixed effects. Results are summarised in Table D.6 in the Appendix, and are in line with our main results. They confirm that there is a positive relation between mobile internet coverage and employment. They also show that individuals are more likely to be employed in more skilled occupations and in non-agricultural activities following the roll-out of the mobile internet.

5.5.2 Mechanisms and extensions

In this section, we intend to extend our results by exploring some of the potential mechanisms at play in the relationship between mobile internet and changes in employment composition. We look more closely into three specific dimensions. The first is related to education levels of the working-age population. The second looks at migration. Finally, we conduct a preliminary analysis of possible demand-side factors, i.e. whether and how internet coverage has affected firms' characteristics.

Education: In Table 5.4, we replicate our results using indicators of educational attainments as outcome variables to understand whether the introduction of new technologies might have affected the educational choices of individuals. For this analysis, we have modified our sample in such a way as to consider only the cohort of individuals that were in their school age (i.e. 5 to 25 years of age) at the time of the

[23] By design, DHS targets women as the main respondent. As such, our sample includes about 69 per cent of women. The proper application of sampling weights, which we apply in our analysis following the DHS recommendations, takes this unbalance into account.

survey. This is done to avoid pooling both sets of new entrants to the labour market: on the one hand, the youngsters, whose educational choices might be directly affected by the current availability of internet connectivity; on the other hand, incumbents, whose levels of education are not affected by recent changes in mobile technologies. Results show that the diffusion of mobile internet has a positive effect on educational attainment: in fact, the former runs in parallel with a reduction in the share of pupils with primary or no education (column 1), and with a corresponding increase in the share of those with secondary or tertiary education (columns 2 and 3). More generally, an increase in mobile internet leads to an overall increase in the number of years of education, as reported in column 4.

Table 5.4: 2SLS results, by education

	Education:			
	Primary/no (1)	Secondary (2)	Tertiary (3)	Years (4)
3G	-0.457***	0.373***	0.0838***	0.0150***
	(0.0845)	(0.0821)	(0.0218)	(0.00519)
Observations	150	150	150	150
District FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
District Controls	YES	YES	YES	YES
Mean DV	0.856	0.134	0.00985	0.0299
Quantification	-0.145	0.118	0.0265	0.00476
F-stat	15.81	15.81	15.81	15.81

Note: The dependent variables measure, respectively, the population share of individuals with tertiary, secondary and primary (or no) education, and the number of years of education. The sample of individuals used for this exercise is restricted to those in the cohort aged 5–25 years old. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of the female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Migration: Changes in the distribution of economic activity are considered an important pull factor for internal migration in low- and middle-income countries.

Provided that improved mobile internet access generates differential gains across districts, one could expect a larger inflow of migrants into treated locations in comparison to other areas. Descriptive evidence seems to support this hypothesis (Figure D.4 in the Appendix). Districts with higher levels of internet coverage are also those with a higher share of migrants.

We can test this hypothesis more formally by replicating our main specification using a different set of outcomes related to migration. Information on migration can be obtained from the data by using a question that asks individuals about their previous place of residence and the timing of their move to their current district. Note that this question was not available in the 2017 and 2018 editions of the RLFS: those two waves are therefore excluded from this exercise.

We combine the information on migration with the employment status of workers to generate a variable that measures the share of employed migrants among the working-age population. Results of the 2SLS estimation using this variable as the dependent variable are reported in Table 5.5. As we can distinguish the origin of each individual, a migrant is defined according to whether they have relocated from a different district (column 1) or a different province (column 2) to the place that they are residing at the time of the interview. Results show that higher levels of mobile internet coverage make a location more attractive to migrant workers. Also, the specific definition of migrant applied to the variable does not make a significant difference to the results. In further analysis, we also find that this effect seems to be driven by migrants being employed in skilled occupations and in modern industries (both manufacturing and services). These additional results are reported in Table D.7 in the Appendix, and are based on a more restrictive definition of migration, i.e. individuals coming from a different province.[24]

Demand-side mechanisms: We use firm level data to investigate potential demand-

[24]Note that while the effect is positive on migrants ending up in both skilled and unskilled employment: a simple quantification based on evaluating a shift from the 25th to the 75th percentile of the 3G coverage distribution shows that the estimated improvement on skilled migrant employment is 128% higher than the actual average for skilled professions, as compared to 48% for unskilled ones.

Table 5.5: 2SLS results, migrant workers

	Migrants from:	
	Other districts	Other provinces
	(1)	(2)
3G	0.264**	0.190**
	(0.109)	(0.0836)
Observations	90	90
R-squared	0.618	0.659
District FE	YES	YES
Year FE	YES	YES
District Controls	YES	YES
Mean DV	0.161	0.113
Quantification	0.0836	0.0602
F-stat	10.66	10.66

Note: The dependent variables measure, respectively, the share of migrant workers relocated from other districts (column 1) or from other provinces (column 2) of Rwanda. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

side mechanisms that might corroborate the relationships that have been so far identified by looking at the supply-side only. More specifically, we are interested in understanding whether firms took advantage from the roll out of broadband internet across the country and over time.

For these purposes, we use information on firms registered in the following two national tax databases: (1) the Corporate Income Tax (CIT) database, recording revenues, expenditures and other financial indicators; and (2) the Pay As You Earn (PAYE) database, which includes information on employment. These databases can be matched using anonymised firm identifiers.[25] Data are available for the period 2008–2016, and cover all industries. For the purposes of this exercise we only use information on those firms for which we can correctly match CIT and PAYE data. This results in a sample of about 10,400 unique firms and 40 thousand observations. For these firms we calculate two indicators of productivity. First, a simple indicator of labour productivity, which we can measure as the log of total sales (from CIT) over the number of employees (from PAYE). Next, we also estimate Total Factor Productivity (TFP), though this is only possible for a smaller number of firms. We do this applying the approach by Levinsohn and Petrin (2003) (LP) that uses the costs of materials as a proxy for unobservable productivity shocks to correct for simultaneity bias. We also address potential collinearity in the first stage due to simultaneity bias in the labour coefficient by adopting the correction suggested by Ackerberg et al. (2015). We use total sales as a measure of output, the stock of assets at the beginning of period to measure capital, the wage bill for employment and the cost of goods sold to measure intermediate inputs.[26] Table D.8 in the Appendix reports production function coefficients at the industry level.

Table 5.6 reports the results of a regression in which the outcome of interest, that is a firm's TFP (column 1) or labour productivity (2),[27] is explained by the indicator

[25] Data are confidential, and they were obtained by the Rwanda Revenue Authority.

[26] Productivity is estimated using the `prodest` command in Stata (Rovigatti and Mollisi, 2018).

[27] The estimation samples for TFP and labour productivity are different, the latter being larger. Estimates based on TFP (labour productivity) are based on 1,322 (4,046) unique firms observed in the period 2010-2016 (2008-2016). Trade services is the most represented industry in both samples,

of district level coverage of the 3G network, as follows:

$$Y_{it} = \beta 3G_{dt} + \gamma 2G_{dt} + \theta_i + \delta_t + \epsilon_{it} \tag{5.2}$$

Where Y_{it} is a placeholder for the two outcome measures (TFP and sales per worker). All regressions include the 2G coverage (as before) as well as firm and year fixed effects. By including firm f.e. we can identify within-firm changes in productivity that correlates to changes in mobile internet coverage. Standard errors are clustered at the district level.

Table 5.6: OLS results, firm productivity

	Productivity:	
	TFP	LP
	(1)	(2)
3G	0.364**	1.763**
	(0.163)	(0.761)
Constant	1.540**	15.70***
	(0.588)	(1.094)
Observations	4,670	18,549
R-squared	0.951	0.678
Firm FE	YES	YES
Year FE	YES	YES

Note: The dependent variables measure, respectively, Total Factor Productivity (column 1) and Labour Productivity (column 2) of Rwandan firms, using Rwandan CIT-PAYE data. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the 2G mobile coverage, in addition to firm and year fixed effects. Standard errors are clustered at the district level *** p<0.01, ** p<0.05, * p<0.1.

Results show that there is a positive relation between mobile internet coverage and firm productivity: this is suggestive of demand side mechanisms at play, including more employment opportunities, possibly with higher skill content. However, it must be noted that the evidence provided by estimating Equation 5.2 is purely

accounting for 58% and 33% of the total, respectively.

correlational. An analysis of the impact of mobile internet access on firm size and productivity requires granular data on mobile coverage and information on internet adoption by firms. This represents a promising avenue for further research.

5.6 Conclusion

As numerous low- and middle-income economies are intensively exploiting the diffusion of high-speed internet technologies, the aim of this study is to investigate the effects of fast mobile internet coverage on the labour market and on structural transformation. Using Rwanda as a case study and exploiting the staggered diffusion of 3G coverage in the country during the past decade, we find that increases in mobile internet coverage positively affect the size of employment and its composition.

The increase in employment is observed in both skilled and unskilled types of occupations, although the quantification exercise shows that mobile internet is relatively more important for the former, also due to their initial lower shares. Finally, districts that improved their internet connectivity are also those experiencing an increase of employment in services, especially in high-value-added industries, such as finance and health. The estimations are robust to different econometric specifications (IV and event study) as well as to a battery of robustness checks. In trying to rationalise some of these findings, we also show that supply-side factors are activated by mobile internet coverage by means of (a) a higher intake of education by the cohorts currently of school age and (b) an increase in the share of migrant workers. On the demand side, we show that firms take advantage of higher 3G coverage by raising their productivity.

The findings of this study can be used to raise a few considerations on the process of structural transformation of Rwanda. The evidence on industrial reallocation of labour towards services is also informative for the debate on the tertiarisation of African economies, and on the role that technological change can play.

Through its development strategies, Rwanda has actively pursued the goal of transforming into a knowledge based economy – a commitment that was renewed in the recent Vision 2050 national development strategy (Ministry of Finance and Economic Planning, Republic of Rwanda, 2020). With the target of granting universal access to internet set at 80 per cent of the population in 2050, the Rwandan government aims at capitalising on this technology by continuing to invest in connectivity.

Our results look somehow consistent with this strategy. The direction of structural change that we find to be associated to the rapid diffusion of mobile internet is indeed pushing towards more skilled occupations and high-value added activities within the services. However, we also show evidence of large internet-driven employment growth rates in low-productivity services (e.g. trade) and this might indicate that low-productivity services are likely to keep providing a large share of the new employment opportunities in the future.

This phenomenon hints at increasing polarisation of the labour market that, with due distinction, could mimic what has already been reported in different institutional contexts (see among the many others Autor et al., 2003). As the interaction with the ongoing tertiarisation of the economy might further enhance polarisation in Rwanda and in countries experiencing a similar transition, further research and evidence-based policy are needed to govern such processes.

Chapter 6

Conclusions

The overall goal of this thesis was to contribute to the structural change literature by focusing on changes in employment shares, including the informal sector in the analysis of the structural transformation of the SSA countries analysed. The employment-based definition of structural change has been integrated with an analysis of the impact of technological diffusion and adoption on the micro-determinants of structural change, and on the size and skills composition of regional labour markets. This has allowed to make considerations on the distributional dimension of structural changes driven by the diffusion of a General Purpose Technology.

More in detail, the empirical chapters of this thesis have aimed to disentangle the process of structural transformation at the macro, subnational, and micro levels in SSA countries; the relative contribution of the formal and informal sectors to structural transformations, and how the two sectors are related; the role of technological change in driving such structural transformations at the micro-level in the informal sector; and whether the impacts of technology-led structural transformations differ for different workers and industries.

The thesis has built and expanded on the evidence that African countries have de-industrialised prematurely (Rodrik, 2016), with labour moving to low-productivity industries (McMillan et al., 2014, 2017b), contributing to the growth and persistence of the informal sector – the most significant employer across SSA countries. Against this backdrop, the thesis has explored the potentially active contribution of informality to structural change in SSA (Chapters 2 and 3). It has also analysed explicitly the role played by technical change in driving structural change and in achieving inclusionary or exclusionary outcomes (Chapters 4 and 5). Depending on the research question and data availability, the analysis has relied on trade data, firm census data, and household data from Ghana, Nigeria, and Rwanda. The empirical analysis combined micro data with subnational data on mobile internet coverage, lightning strikes, nighttime lights, and geography.

The empirical analysis has been guided by theoretical insights from the traditional (Lewis, 1954; Hirschman, 1958) and modern (Rodrik, 2018; McMillan et al., 2017a) structuralist, neo-Schumpeterian (Perez, 1986; Autor et al., 2003; Ciarli et al., 2021b), and economic complexity literature (Hidalgo et al., 2007; Tacchella et al., 2012), and by their intersections and differences. In particular, the structuralist understanding of structural change as labour shifts across industries has been complemented by the measurement of industry complexity to assess the extent to which transformations in the structure of employment are accompanied by an increase in the workers' capabilities (Chapters 2 and 3). The measure of industry complexity has been built using employment specialisation, rather than export, to include in the analysis all the (formal and informal) activities that contribute to the production of goods and services in a country, including those that are not traded. The expansion of non-tradable services is one of the most significant features of the current trends of structural change in SSA countries.

The structuralist and economic complexity contributions have then been complemented with theories and stylised facts from the neo-Schumpeterian literature on the role of technical change and technological capabilities in triggering structural transformations. The thesis has focused on the role of one of the fastest diffusing technologies in SSA: mobile internet.

Finally, building on the Inclusive Structural Change framework (Saha and Ciarli, 2018; Ciarli et al., 2021b), which combines contributions from the structuralist and neo-Schumpeterian literature, the thesis has studied the distributional impacts of technological and structural changes. In particular, the analysis examined whether such changes lead to more or less employment and entrepreneurial opportunities and for which workers, households, firms, and industries (Chapters 4 and 5).

The novel combination of different theoretical backgrounds has offered fresh empirical evidence on the underlying mechanisms of structural change in SSA and how they can be sustained and are inclusive. This thesis, therefore, offers a contribution that tackles a high-priority objective in the development policy agenda by advancing conceptually and empirically on aspects that are comparatively less explored in the development policy circles: the role of informality and technical change, and how these can be steered to achieve sustainable and inclusive development.

The following subsections discuss the main findings of the thesis concerning the contribution of informality to structural change and the effect of technological change on structural transformation and inclusion in SSA countries. Each subsection provides some policy implications.

6.1 The contribution of informality to structural change

Chapter 2 has compared the trajectories of structural change in Ghana's trade, formal, and informal sectors between 2003-13. It has analysed the changes in the composition of finely disaggregated industries across the three sectors in terms of exported value (trade sector) and employment (formal and informal sectors). Industry-level changes have been assessed against an employment-based measure of industry complexity. This measure is used as a proxy for the degree and intensity of capabilities required to specialise in each industry, measured with the employment specialisation patterns of Ghanaian districts. The analysis has documented whether the trade and production structure in the three sectors has been shifting towards more or less complex industries.

The results have shown that, while Ghanaian trade and formal sectors have shifted towards more complex industries, employment in the informal sector has moved towards less complex industries, with overall minimal changes in its employment composition. Despite the increased complexity, trade specialisation has moved towards natural resources, like oil, which has led to limited employment creation, both in upstream and downstream related industries. The most encouraging results are those from the formal sector, where formal manufacturing employment has shifted to more complex manufacturing industries, although shrinking compared to mining. Moreover, the analysis has singled out the transformation of the informal sector over a ten-year period in Ghana, shedding light on how its dynamics differ from other sectors of the economy.

The chapter also raises important methodological implications, arguing that relying on the analysis of trade specialisation patterns might overlook essential components of structural change. The chapter's results have shown that the patterns of export specialisation do not provide a reasonable estimation of the production structure of low- and middle-income countries like Ghana and, therefore, of the potential capabilities available in the country. Using an employment-based measure of industrial complexity, which has been named the Industry Complexity Index, has allowed to include many non-tradable activities in the analysis of industrial capabilities. Despite their prevalence in SSA, the complexity-based approaches limited to trade-based indices have not considered such activities.

Chapter 3 has examined the linkages between the formal and informal sectors to study the possible channels through which informality can contribute to structural change. The formal and informal industrial co-location patterns at the subnational level in Ghana have been analysed to test the hypotheses that informal industries contribute actively to Ghana's structural change by contributing to the aggregate industrial variety and interacting with formal industries. Such hypotheses were tested by estimating the formal and informal employment concentration in Ghanaian districts and the concentration of formal and formal industries in the same districts. The co-location of industry pairs has been explained by the relatedness of their capabilities, input-output linkages, and complexity differentials.

First, descriptive results have shown that, on average, formal and informal industries follow different co-location patterns. For instance, the retail industry may co-locate with some manufacturing industries when both the former and the latter are informal, but not when they are both formal. This result indicates that co-location patterns follow drivers beyond the mere industrial nature of activities. Moreover, informal industries agglomerate more than formal ones. Industries with similar industry complexity and complementary capabilities in the informal sector are more likely to concentrate in the same districts. Conversely, formal industries display an enclave-like pattern, characterised by no relatedness in terms of capabilities and heterogeneous levels of complexity, and therefore little linkages with co-located industries. Input-output linkages, instead, have a positive, non-linear effect on the co-location between industries within the informal and formal sectors. They also explain the co-location of formal and informal industries. The input-output relationship is the only factor driving informal industries to co-locate with formal ones (among the three channels tested here). Informal-formal co-location is also more likely to occur across different macro-industries (1-digit ISIC groups) rather than within.

Overall, the findings have highlighted that informal industries co-locate with related industries across Ghanaian districts, unlike formal ones, which show few links and are concentrated only in a few Ghanaian districts. While informal industries appear to be vertically integrated with formal ones, their different levels of complexity and the differences in terms of underlying capabilities may represent an obstacle for labour to move from informal to formal activities.

Additionally, these results have offered a methodological contribution to measuring complexity. The extant evidence on complexity has found that more complex activities tend to cluster together (Hidalgo, 2018; Balland et al., 2020). Instead, these results – particularly those on the drivers of formal and informal-formal co-location – have shown that industrial co-location may lead to clusters that are rather heterogeneous in terms of their overall complexity.

6.1.1 Policy implications

Overall, the results on the contribution of the informal sector to structural change in Ghana in Chapters 2 and 3 call for policies that improve the integration of informal activities in the process of structural transformation, while leveraging and acknowledging their specific industrial specialisation, capabilities, and agglomeration patterns. On the one hand, results have identified favourable conditions for related diversification and labour reallocation in the informal sector – as suggested by the relatedness of informal co-location patterns. On the other hand, they have revealed that a mismatch between informal and formal sectors may make it more difficult for the less complex components in the informal sector to shift to more complex formal activities. These findings provide a rationale for industrial policies to improve the coordination of "resources and abilities that are hidden, scattered, or badly utilised" (Hirschman, 1958, p. 5). For instance, this could be pursued by promoting policies leading to the formation of a local production system, characterised by high technological linkages among local industries (Hirschman, 1977; Andreoni, 2019), and adopting a systemic approach that acknowledges the centrality of informality in creating employment opportunities (Kraemer-Mbula and Wunsch-Vincent, 2016) and facilitates the relocation of the informal workers who have moved to less complex industries.

The trajectory of informal industries in Ghana and their co-location patterns with formal activities highlight the importance of accumulating capabilities in the informal sector. Such capabilities need to be related to the overall productive structure, which determines the opportunities for diversification into related, more productive industries. The results of Chapters 2 and 3 have shown that informal industries represent a potentially large reserve of capabilities that could be leveraged to support structural transformation; given the heterogeneity of their industrial composition, the relatedness of their capabilities, and their tendency to agglomerate. Nevertheless, results have also shown that the Ghanaian informal sector has not followed the evolution of the formal and trade sectors toward more complex industries. This trajectory indicates a potential mismatch between the capabilities available in the informal sector and those required by the emerging industries in the formal and trade sectors. The mismatch is confirmed by the results of Chapter 3, which show that informal industries are not always related to formal ones, hampering opportunities for informal labour to upgrade by moving towards formal industries.

Given its relevance in terms of employment and entrepreneurial opportunities, the informal sector needs to be at the core of structural change, and the mismatch in capabilities with the formal sector should be overcome to avoid widening the capability gap. Policies will have to create incentives for capital investments and training to support informal firms. These are, particularly in the African context, as important as formal innovation activities such as Research and Development (Paus et al., 2022). Moreover, based on the evidence that informal capabilities tend to agglomerate geographically, policies should exploit those to promote diversification and upgrading strategies within the informal sector linked (technologically and in terms of capabilities) with the aggregate productive structure. Without addressing these issues, informal firms and employment will remain at the margins of the productive ecosystem in SSA, as the mismatch between the available stock of capabilities and those required by emerging modern sectors may thwart the contribution of informality to structural change via the reallocation of labour towards more complex industries.

Finally, industrial policies in Ghana could promote strategies of economic diversification that build upon and improve the composition of capabilities in domestic employment, taking into account the high shares of informality. The development of new and related capabilities will be particularly relevant in the perspective of increasing regional integration between African countries, as suggested by the recent implementation of the African Continental Free Trade Area (ACFTA), promoted by the African Union.[1] The treaty aims at establishing a free-trade area among participant countries, to "attain sustainable and inclusive socio-economic development, gender equality and structural transformation of the State Parties", and to "promote industrial development through diversification and regional value chain development, agricultural development and food security".[2] With its aim of establishing regional value chains in the continent, the ACFTA propositions may push African countries to reconsider their diversification and specialisation strategies, and therefore their process of accumulation of capabilities, which will have to take into account the presence of a large informal sector to deliver inclusive economic growth and development (UNCTAD, 2021). The employment-based complexity approach proposed in Chapters 2 and 3 can offer a valuable tool to identify industries with higher transformative potential and to orient African strategic diversification in alignment with the available stock of capabilities in each country, including those of the parts of the population running informal activities.

6.2 Technological change and inclusive structural change

Chapter 4 has focused on estimating the impact of technological change on the structural transformation of informal activities at the micro-level, and the inclusive or exclusive nature of such transformation. The chapter has used data on Nigerian households, combined with mobile internet coverage maps at the level of local government areas (LGA), and relied on the Inclusive Structural Change framework

[1] https://au.int/en/cfta.

[2] Full text of the agreement establishing the African Continental Free Trade Agreement: https://au.int/sites/default/files/treaties/36437-treaty-consolidated_text_on_cfta_-_en.pdf.

(Saha and Ciarli, 2018; Ciarli et al., 2021a). The analysis has looked at the effect of mobile internet adoption on two measures of structural change in the informal sector: the labour productivity of existing households non-farming enterprises (NFEs) within industries and the probability that households open new activities in new industries. Next, the chapter has analysed the distributional impacts of internet-driven structural change by estimating its effect on the entry of new firms and household labour: respectively, entrepreneurial opportunities and job creation/destruction.

Concerning the first measure of structural change, findings have revealed that mobile adoption has had a sizeable effect on the performance of NFEs, as demonstrated by the growth of their sales per worker for households adopting mobile internet. However, only firms in low-productivity services such as wholesale and retail trade were likely to benefit from the adoption of mobile internet. Instead, with respect to the second measure of structural change, findings have shown that households adopting mobile internet were less likely to move to manufacturing industries, confirming the premature de-industrialisation trend (Rodrik, 2016).

To investigate the distributional impacts, the analysis has first decomposed labour productivity improvements in sales and employment effects. It was found that the positive effect of internet mobility on labour productivity is due to both higher sales and lower labour inputs. Next, the effect on employment was broken down into within- and outside-household labour: the results indicate that only household labour shrinks while non-household labour increases. This effect may be due to the need to source more skilled labour, because household members move to better-paid jobs in the formal sector, or a combination of both these mechanisms. Investigating the impact of mobile internet adoption on the employment rates in households that own an NFE, it was found that household members who no longer work for the NFE following mobile internet adoption are entirely re-absorbed by working opportunities outside the household. The total effect of mobile internet adoption on the employment rates of household members is therefore null: mobile adoption, which leads to higher labour productivity by reducing the number of workers from the household, has led to a process of labour reallocation from inside to outside the household. Regarding the employment rates of households that do not own an NFE, the results revealed an increase in the employment rates in households that do not own NFEs due to mobile internet adoption. Based on these findings, we can conclude that the overall effect of mobile internet adoption on employment creation in the Nigerian informal sector has been positive.

Second, the chapter investigated the impact of mobile internet adoption on the entry of new NFEs. Results showed that mobile internet adoption does not lead to the entry of new firms. As noted above, this is partly explained by the increase in the sales of incumbent firms.

In summary, mobile adoption sustains structural change among Nigerian NFEs in terms of increasing productivity, but not in relation to shifting activities to more complex and value added industries such as manufacturing. This process of structural change is also inclusive in terms of creating more working opportunities, but is exclusive in terms of reducing opportunities for new firms to enter.

Using a similar approach, Chapter 5 investigated the impact of the staggered diffusion of mobile internet in Rwandan districts on their structural transformation at the regional level, measured by the skill and industry composition of Rwandan labour markets. The analysis has combined GSMA network coverage maps with individual-level information from national population censuses and labour force surveys, creating a district-level dataset of Rwanda that covers the period 2002–2019. In particular, the chapter has studied the effects of mobile internet diffusion on the creation of skilled and unskilled employment across different industries to assess whether internet diffusion has benefited one group more than the other, leading to a process of skill-biased structural change.

The findings indicate that aggregate employment in Rwanda has grown considerably with the diffusion of fast mobile internet. Both skilled and unskilled jobs have grown, with unskilled jobs growing more in absolute terms and skilled jobs expanding faster.

On the supply side, the effect of mobile internet on job creation is associated with a larger intake of migrant workers. This result indicates that areas where mobile internet has been diffusing more broadly may attract migrant workers. This is especially true for highly educated migrant workers. Additionally, the growth of skilled jobs is associated with higher educational attainments of individuals in their schooling age, suggesting that students who will soon enter the labour market may find themselves better equipped to take up jobs requiring higher and more specific skills.

The study found that also in Rwanda, mobile internet is associated with a higher total factor and labour productivity of Rwandan firms. Coupled with the result that employment has grown in the Rwandan districts with higher mobile internet coverage, the analysis has shown that firms have benefited from access to the new technology, catering to increasingly large markets and creating new employment opportunities. In relation to changes in the sectoral composition, results have shown that the sectoral allocation of labour that follows the diffusion of mobile internet is skewed towards services, with a contraction of the manufacturing sector. Within the service sector, low-productivity services – such as wholesale and retail trade, private household services, and hospitality – expand more in absolute terms, although some modern sectors – like finance and transport – have also grown considerably.

Overall, the results have indicated that the diffusion of mobile internet in Rwanda is also likely to strengthen the trend of premature de-industrialisation, leading to structural change towards services where skilled and unskilled workers are likely to benefit from job creation.

6.2.1 Policy implications

There is widespread optimism about the adoption of ICT in low- and middle-income countries, given its potential to bootstrap into the adoption of new, related technologies and to leapfrog the productive structure away from traditionally low-productivity activities (Hjort and Poulsen, 2019; Fagerberg et al., 2021; Kaplinsky and Kraemer-Mbula, 2022). However, it has been argued that without the appropriate conditions, the diffusion of ICT in low- and middle-income countries may lead to exclusionary outcomes (Altenburg et al., 2021). The results of Chapters 4 and 5 have contributed in this direction. They can be used, in combination with the Inclusive Structural Change framework (Saha and Ciarli, 2018; Ciarli et al., 2021a), to formulate policy recommendations aiming at ensuring the inclusiveness of the internet-driven structural change in SSA.

First, the evidence presented in Chapters 4 and 5 has indicated that mobile internet may steer structural change towards services, both in the informal sector (Nigeria) and in the overall labour market (Rwanda). This is consistent with the general observation that services are growing both in terms of employment and value added across SSA countries (Owusu et al., 2021). These results have shown that the diffusion of the internet partly explains the tertiarisation trend. On the one hand, low- and middle-income countries may build on this tertiarisation to leapfrog their productive structure towards knowledge-intensive industries (Fagerberg et al., 2021; Kaplinsky and Kraemer-Mbula, 2022). To reap the benefits of technological diffusion, it will be necessary that access to ICTs is not hampered by pre-existing barriers, such as the lack of appropriate skills. Digital literacy programmes can play a crucial role in levelling the playing field.

On the other hand, the tertiarisation of the economy presents at least two challenges. The first is the exclusion of manufacturing workers and entrepreneurs, which may find themselves poorly equipped in terms of the necessary skills to take advantage of the emerging opportunities in services. The growth of low-productivity services poses the second challenge. While these services are often labour-intensive, providing a source of income for African households (Behuria and Goodfellow, 2019), the growth of services, such as retail and wholesale trade, may reinforce the trend of labour reallocation towards below-average productivity industries, reinforcing the process of growth-reducing structural change described by McMillan et al. (2014). Once more, industrial policy can play a pivotal role in incentivising investments in those sectors that, at the same time, ensure high levels of productivity while also creating opportunities to absorb labour from below-average productivity industries.

Second, Chapter 4 has shown that the relationship between technical change and structural change can entail a trade-off. Whereas the adoption of mobile internet by Nigerian informal firms positively affects firms' performance and labour creation, existing entry barriers may be reinforced by the growth of incumbents' market share, as entry costs do not decrease. Besides this exclusionary mechanism, preventing the entry of new actors may slow down the rate of diffusion of the new technology, thwarting the introduction of innovations made possible by the adoption of the internet, and slowing down the process of structural change. For this reason, policies should ensure that market barriers are removed in such a way as to allow new entrepreneurs to reach the market. Moreover, unlocking the job-creation potential of the internet will also depend on the presence of mechanisms to allow the seamless reallocation of 'excess' workers from more productive firms to new activities (whether formal or informal). This will require an extension of traditional policy tools available to formal firms (such as grants and skills-upgrading programmes) to include also informal firms (Kraemer-Mbula and Wunsch-Vincent, 2016).

Finally, the experience of Rwanda described in Chapter 5 shows that while both skilled and unskilled jobs have grown, skilled jobs grow faster than unskilled ones. In the long run, this mechanism could lead to a polarisation of the labour market. The analysis could not look at the distribution of wages, as this information is not available in the data. Nevertheless, these results raise a flag on the possibility that technological change in low- and middle-income settings only favours workers with high skills, also based on the experience of other countries (Autor et al., 2003; Buera et al., 2021). In this respect, the Rwandan government has already put in place policies to guarantee access to ICT literacy and skills in both past and current national development strategies (Ministry of Trade and Industry, Government of Rwanda, 2011; Ministry of Finance and Economic Planning, Republic of Rwanda, 2020), training its current and future workforce. The availability of internet infrastructure may be frustrated by the lack of relevant skills to apply the new technology to production and innovation. At the same time, however, the equal distribution of benefits will depend upon the accessibility of new technologies. While Rwanda – a relatively small country – has reached universal mobile broadband signal coverage in 2020 (also thanks to targeted, intentional policies steering the roll-out of mobile internet), the situation in larger countries like Nigeria appears to be different, with mobile internet coverage concentrated in urban districts. The persistence of the digital divide could therefore represent an obstacle to the diffusion of ICTs and their opportunities, reinforcing pre-existing inequality structures between regions/subnational areas, and increasing within-country inequalities. Policy mechanisms should be in place to deliver outcomes that the market alone may not be able to deliver.

6.3 Limitations of the analysis and avenues for further research

This thesis has relied on a broad set of data sources, including household surveys, firm censuses, and geospatial data. While disaggregated data is becoming increasingly available in SSA, many limitations persist. Some of these are examined below, followed by a discussion on some promising avenues for further research.

Informality. Chapters 2 to 4 devote a great deal of attention towards the study of the informal sector in SSA. However, informality is a concept that lacks a univocal definition. Formal registration of firms is a broadly used criterion, also adopted in this thesis. Nevertheless, there are aspects of informality that persist among formal firms (for instance, keeping informal accounts) and elements of formality in unregistered firms (such as the participation in formal supply chains). The analysis conducted in this thesis cannot capture these nuances. Moreover, the reduced (although growing) data availability on informality makes the study of informal activities a challenging endeavour. Chapters 2 and 4 rely on household data – namely, on household farming and non-farming activities – to estimate, respectively, the change in the industrial composition of informal economic activities and the impact of mobile internet adoption on informal firms. However, as the data used is self-reported, potential measurement errors cannot be ruled out (from disclosure to transcription).

Moreover, we choose to rely on household-run activities to identify the informal sector, both in Chapters 2 and 4, derived from household surveys. However, while such data is comprehensive and nationally representative, informality may go beyond the farming and non-farming enterprises declared by households. Finally, Chapter 2 compares structural changes in the informal sector (proxied by household farming and non-farming activities) with structural changes in the formal and trade sectors. While Chapter 3 establishes a connection between the formal and informal sectors, this thesis does not measure the impact of a change in export specialisation (measured in terms of value exported) on the composition of the formal and informal sectors (measured using changes in relative shares of industry-level employment), limiting the scope of the analysis of the relationship between the three sectors. Additional information on firms – such as value added – would allow a better comparison of the formal and informal sectors with the trade sector (Chapter 2) and a more refined measure of firm productivity (Chapter 4).

Furthermore, the quantitative nature of the methodology adopted in the analysis conducted in the four empirical chapters is limited by its focus on aggregates (Chapter 2 and 3) and averages (Chapter 4 and 5), overlooking upgrading and innovation occurring at the margin within the informal sector. The observation of these micro-level, potentially scattered dynamics requires a more qualitative approach (for some examples, see de Beer and Tumaine 2020; Dias and Fernandez 2020; Kraemer-Mbula and Monaco 2020), and could be extremely informative on the future evolution of informality in Sub-Saharan Africa. It therefore constitutes an important and promising avenue for further research.

Capabilities. The analysis of capabilities also features prominently in all of the empirical analyses of this thesis. Chapters 2 and 3 use a measure of industry complexity to estimate the degree of capability intensity to specialise in a given industry. Following the extant literature on economic complexity (Hidalgo and Hausmann, 2009; Tacchella et al., 2012; Freire, 2021), the measurement of industry complexity relies on the assumption that more complex industries require more capabilities. However, a better understanding of the capabilities at the micro-level (workers, entrepreneurs, and firms) is required to investigate how both individuals and firms upgrade in the industry space, how they accumulate capabilities, master new technologies, and innovate. Information on workers' skills, firm-level technologies, and innovation practices and outputs could be useful for testing such upgrading mechanisms at the micro level. Moreover, industry capabilities, and therefore the Industry Complexity Index, change over time as the relative sophistication of industries also changes along with technical advances. Longitudinal, comprehensive data on firms and their employment would also offer a more refined, time-varying complexity measure. However, such data is hard to obtain and is available only for a limited set of countries (to the best of my knowledge, the only SSA country with a publicly available census of firms is Ghana). Finally, the measurement of skills implemented in Chapter 5 relies on a rather coarse definition of skilled and unskilled jobs, following the definition by the International Labour Office (2012). A more fine-grained measurement of skills at the level of workers, including, for instance, unpacking the variety of medium skills, could help offer a more detailed picture of the labour polarisation trends driven by technological change. Information on skills is not available for the Nigerian NFEs analysed in Chapter 4, therefore impeding inference on the skill level of labourers that leave/join NFEs and other productive activities.

Technological change. Chapters 4 and 5 focus, respectively, on the impact of mobile internet adoption and diffusion on structural change at the micro (Nigeria) and subnational (Rwanda) level. However, the GSMA data on mobile internet coverage used in both chapters is only available at the district level, making the measurement of mobile internet adoption in Nigeria likely less precise when compared to lower levels of geographical disaggregation. The original data comes at a much lower geographical unit (8 km × 8 km cells). However, its use is pay-walled by the data provider, and its cost was higher than the funding allowed to complete this doctoral thesis. Further refinement of the analysis conducted in Chapter 4 will be pursued by seeking research funding to allow the acquisition of the disaggregated data. This limitation is also relevant to Chapter 5, which uses the (30) Rwandan districts (Admin-2) as a level of analysis, imposing a constraint on the number of observations. A more disaggregated (Admin-3) level of analysis could make the estimation more robust. Moreover, both chapters use mobile internet as a proxy for ICT diffusion. Whereas the internet has many applications, functioning as a gateway to access many other ICTs, these cannot be measured using the data available. The effect of mobile adoption on structural change estimated in Chapter 4 takes only partially into account the possibility that firms and workers could benefit from the diffusion of the technology even if they do not adopt it. Conversely, Chapter 5 only considers the diffusion of mobile internet, as details on adoption are not available at the micro-level. Finally, the analysis of Chapter 5 is limited to the employment effects of technological change, which may hide growing wage inequalities within regional labour markets, flagged by the fast rate of growth of skilled occupations.

Building on the results gathered in this thesis, combined with the three sets of limitations above, three promising avenues for further research can be identified.

First, it has been documented how SSA countries have exhibited heterogeneous trajectories of structural change (Bah, 2011; McMillan et al., 2017b). Chapter 2 has shown that national-level trends of structural change may hide a remarkable level of heterogeneity, as indicated by the different trajectories observed for the Ghanaian informal, formal, and trade sectors. International comparisons of structural change trends across the three sectors, assessed against industrial complexity – measured using employment specialisation across SSA subnational areas – could, at the same time, provide a more encompassing picture of the structural change trends in SSA while using a measure of industry complexity that builds on a broader set of data.

Second, Chapters 2 and 3 have shown that the informal and formal sectors are diverging in terms of productive capabilities. Future research should investigate if this mismatch is due to an upgrade of informal workers and firms towards the formal sector, or because the informal sector has been downgrading in terms of productive capabilities. Related to this, micro-level determinants of worker/firm upgrading, or constraints leading to downgrading, and their interaction with technological change require further investigation.

Third, recent advances have suggested that jobs may respond to technological change not only based on their skills, but also on the routine intensity of their tasks (Goos et al., 2014, 2021; Cirillo et al., 2021). More granular information on the nature of jobs created/replaced by technological change could provide a much more detailed picture of the dynamics of inclusion/exclusion that follow the diffusion of breakthrough, general purpose technologies. Connected to this point is also the importance of other forms of technological change. The measures of internet adoption and diffusion used in Chapters 4 and 5 are employed as a proxy for the diffusion of ICTs, which, however, are not measured directly. An analysis of the impact of internet diffusion would benefit from singling out its effect on further ICT applications and the emergence of ICT related industries. Examples of ICT-related technologies are automation and robotisation technologies, additive manufacturing, and Artificial Intelligence. These technologies bear the promise of a strong transformative potential: the effect of their diffusion and adoption in SSA remains largely unexplored so far and represents a very promising avenue for further research.

Bibliography

Acemoglu, D. and Guerrieri, V. (2008). Capital deepening and nonbalanced economic growth. *Journal of Political Economy*, 116(3).

Acemoglu, D., Johnson, S., and Robinson, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *American Economic Review*, 91(5):1369–1401.

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451.

African Center for Economic Transformation (2014). Growth with Depth. Africa Transformation Report, ACET, Accra, Ghana.

African Center for Economic Transformation (2017). Agriculture Powering Africa's Economic Transformation. Africa Transformation Report, ACET, Accra, Ghana.

African Development Bank (2019). African Economic Outlook 2019. African Economic Outlook, African Development Bank, Cote d'Ivoire.

Aghion, P., Akcigit, U., Bergeaud, A., Blundell, R., and Hemous, D. (2019). Innovation and Top Income Inequality. *The Review of Economic Studies*, 86(1):1–45.

Aghion, P. and Howitt, P. (1997). *Endogenous Growth Theory*, volume 1. The MIT Press, 1 edition.

Aker, J. C. and Mbiti, I. M. (2010). Mobile phones and economic development in Africa. *Journal of Economic Perspectives*, 24(3):207–32.

Akerman, A., Gaarder, I., and Mogstad, M. (2015). The skill complementarity of broadband internet. *The Quarterly Journal of Economics*, 130(4):1781–1824.

Alcorta, L., Foster-McGregor, N., Verspagen, B., and Szirmai, A. (2021). *New Perspectives on Structural Change: Causes and Consequences of Structural Change in the Global Economy*. Oxford University Press, Oxford.

Alesina, A., Hohmann, S., Michalopoulos, S., and Papaioannou, E. (2021). Intergenerational mobility in Africa. *Econometrica*, 89(1):1–35.

Alesina, A., Michalopoulos, S., and Papaioannou, E. (2016). Ethnic inequality. *Journal of Political Economy*, 124(2):428–488.

Altenburg, T., Brandi, C., Pegels, A., Stamm, A., and Vrolijk, K. (2021). COVID-19 – Turning point in the process of industrialization? *Inclusive and Sustainable Industrial Development Working Paper Series*, WP 8.

Andersen, T. B., Bentzen, J., Dalgaard, C.-J., and Selaya, P. (2012). Lightning, IT diffusion, and economic growth across U.S. States. *Review of Economics and Statistics*, 94(4):903–924.

Andreoni, A. (2019). A Generalized Linkage Approach to Local Production Systems Development in the Era of Global Value Chains, with Special Reference to Africa. In Kanbur, R., Noman, A., and Stiglitz, J. E., editors, *The Quality of Growth in Africa*, pages 264–295. Columbia University Press.

Angrist, J. (2004). Treatment effect heterogeneity in theory and practice. *The Economic Journal*, 114:C52–C83.

Arimah, B. C. (2001). Nature and Determinants of the Linkages between Informal and Formal Sector Enterprises in Nigeria. *African Development Review*, 13(1):114–144.

Arndt, C., McKay, A., and Tarp, F., editors (2016a). *Growth and Povery in Sub-Saharan Africa*. Oxford University Press.

Arndt, C., McKay, A., and Tarp, F. (2016b). Two cheers on growth and poverty in Sub-Saharan Africa. WIDER Policy Brief 2016/1, UNU-WIDER, Helsinki, Finland.

Arrow, K. J. (1962). The economic implications of learning by doing. *The Review of Economic Studies*, 29(3).

Autor, D. (2022). The labor market impacts of technological change: From unbridled enthusiasm to qualified optimism to vast uncertainty. Working Paper 30074, National Bureau of Economic Research.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2020). The fall of the labor share and the rise of superstar firms. *The Quarterly Journal of Economics*, 135(2):645–709.

Autor, D. H. (2015). Why Are There Still So Many Jobs? The History and Future of Workplace Automation. *Journal of Economic Perspectives*, 29(3):3–30.

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of technological change: An empirical exploration. *The Quarterly Journal of Economics*, 118(4):1279–1333.

Avenyo, E. K. (2018). Informal sector innovation in Ghana: Data set and descriptive analysis. *UNU-Merit Working Paper Series*, 2018-030.

Avenyo, E. K., Konte, M., and Mohnen, P. (2019). The employment impact of product innovations in sub-Saharan Africa: Firm-level evidence. *Research Policy*, 48(9):103806.

Avenyo, E. K., Konte, M., and Mohnen, P. (2020). Product innovation and informal market competition in sub-Saharan Africa. *Journal of Evolutionary Economics*.

Baccini, L., Fiorini, M., Hoekman, B., and Sanfilippo, M. (2021). Services and Economic Development in Africa. Igc working paper, International Growth Centre.

Bah, E.-h. M. (2011). Structural Transformation Paths Across Countries. *Emerging Markets Finance and Trade*, 47(sup2):5–19.

Bahar, D., Rosenow, S., Stein, E., and Wagner, R. (2019). Export take-offs and acceleration: Unpacking cross-sector linkages in the evolution of comparative advantage. *World Development*, 117:48–60.

Bahia, K., Castells, P., Cruz, G., Masaki, T., Pedrós, X., Pfutze, T., Rodríguez-Castelán, C., and Winkler, H. (2020). The welfare effects of mobile broadband internet. Evidence from Nigeria. Policy Research Working Paper 9230, World Bank.

Bahia, K., Castells, P., Cruz, G., Masaki, T., Rodríguez-Castelán, C., and Sanfelice, V. (2021). Mobile Broadband Internet, Poverty and Labor Outcomes in Tanzania. IZA Discussion Paper Series 14720, IZA - Institute of Labour Economics, Bonn, Germany.

Balassa, B. (1965). Trade Liberalisation and "Revealed" Comparative Advantage. *The Manchester School*, 33(2):99–123.

Baldwin, R. (2013). Trade and industrialization after globalization's second unbundling: How building and joining a supply chain are different and why it matters. In *Globalization in an age of crisis: Multilateral economic cooperation in the twenty-first century*, pages 165–212. University of Chicago Press.

Balland, P.-A., Jara-Figueroa, C., Petralia, S. G., Steijn, M. P. A., Rigby, D. L., and Hidalgo, C. A. (2020). Complex economic activities concentratein large cities. *Nature Human Behaviour*, 4:248–254.

Balland, P.-A. and Rigby, D. (2017). The geography of complex knowledge. *Economic Geography*, 93(1):1–23.

Barabási, A. L. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barrero, J. M., Bloom, N., and Davis, S. J. (2021). Internet access and its implications for productivity, inequality, and resilience. Working Paper 29102, National Bureau of Economic Research.

Baumol, W. J. (1967). Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis. *The American Economic Review*, 57(3):415–426.

Behuria, P. and Goodfellow, T. (2019). Leapfrogging Manufacturing? Rwanda's Attempt to Build a Services-Led 'Developmental State'. *The European Journal of Development Research*, 31(3):581–603.

Bell, M. and Pavitt, K. (1992). Accumulating Technological Capability in Developing Countries. *The World Bank Economic Review*, 6(1):257–281.

Benjamin, N. C. and Mbaye, A. A. (2020). The Failure of Structural Transformation in Francophone Africa and the Rise of the Informal Sector. In Mbaye, A. A., Golub, S. S., and Gueye, F., editors, *Formal and Informal Enterprises in Francophone Africa: Moving Toward a Vibrant Private Sector*, pages 137–154. International Development Research Centre, Ottawa.

Bhorat, H., Kanbur, R., Rooney, C., and Steenkamp, F. (2020). Sub-Saharan Africa's Manufacturing Sector: Building Complexity. In Ravi Kanbur, A. N. and Stiglitz, J. E., editors, *The Quality of Growth in Africa*, pages 235–264. Columbia University Press.

Bloch, R. and Owusu, G. (2012). Linkages in ghana's gold mining industry: Challenging the enclave thesis. *Resources Policy*, 37(4):434–442. Making the Most of Commodities: The Determinants of Linkages in Africa.

Bosch, M. and Maloney, W. F. (2010). Comparative analysis of labor market dynamics using Markov processes: An application to informality. *Labour Economics*, 17:621–631.

Boschma, R. (1999). The rise of clusters of innovative industries in belgium during the industrial epoch. *Research Policy*, 28:853–871.

Boschma, R. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1):61–74.

Boschma, R. (2021). The geographical dimension of structural change. In *New*

Perspectives on Structural Change: Causes and Consequences of Structural Change in the Global Economy, pages 172–187. Oxford University Press.

Boschma, R. and Frenken, K. (2009). Technological relatedness and regional branching. In Bathelt, H., Feldman, M., and Kogler, D., editors, *Dynamic Geographies of Knowledge Creation and Innovation*. Routledge, Taylor and Francis.

Boschma, R. and Frenken, K. (2011). The emerging empirics of evolutionary economic geography. *Journal of Economic Geography*, 11(2):295–307.

Boschma, R. and Iammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic Geography*, 85(3):289–311.

Bourguignon, F. (2004). The Poverty-growth-inequality triangle. Working Paper 125, Indian Council for Research on International Economic Relations (ICRIER), New Delhi.

Boyack, K. W. and Klavans, R. (2010). Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately? *Journal of the American Society for Information Science and Technology*, 61(12):2389–2404.

Buera, F. J., Kaboski, J. P., Rogerson, R., and Vizcaino, J. I. (2021). Skill-Biased Structural Change. *The Review of Economic Studies*, 89(2):592–625.

Bustos, S., Gomez, C., Hausmann, R., and Hidalgo, C. A. (2012). The Dynamics of Nestedness Predicts the Evolution of Industrial Ecosystems. *PLoS ONE*, 7(11):1–8.

Cadot, O., Carrère, C., and Strauss-Kahn, V. (2011). Export Diversification: What's behind the Hump? *The Review of Economics and Statistics*, 93(2):590–605.

Caldarelli, G., Cristelli, M., Gabrielli, A., Pietronero, L., Scala, A., and Tacchella, A. (2012). A Network Analysis of Countries' Export Flows: Firm Grounds for the Building Blocks of the Economy. *PLoS ONE*, 7(10): e47278(10):e47278.

Caldarola, B., Grazzi, M., Occelli, M., and Sanfilippo, M. (2022). Mobile internet, skills and structural transformation in Rwanda. ILO Working Paper 64, International Labour Organisation, Geneva.

Cardona, M., Kretschmer, T., and Strobel, T. (2013). ICT and productivity: conclusions from the empirical literature. *Information Economics and Policy*, 25(3):109–125.

CEDEFOP (2014). Coping with changes in international classifications of sectors and occupations: Application in skills forecasting. Cedefop research paper 43, CEDEFOP, Luxembourg.

Centre for International Earth Science Information Network CIESIN Columbia University (2018). Gridded population of the world, version 4 (gpwv4): Population density, revision 11.

Chataway, J., Hanlin, R., and Kaplinsky, R. (2014). Inclusive innovation: an architecture for policy development. *Innovation and Development*, 4(1):33–54.

Chen, M. A. (2012). The informal economy: Definitions and theories and policies. In *WIEGO Working Paper*, volume 1.

Chenery, H. (1979). *Structural Change and Development Policy*. World Bank, Washington, DC.

Chenery, H. B. (1960). Patterns of industrial growth. *The American Economic Review*, 50(4):624–654.

Chenery, H. B. (1975). The structuralist approach to development policy. *The American Economic Review: Papers and Proceedings of the Eighty-seventh Annual Meeting of the American Economic Association*, 65(2):310–316.

Chenery, H. B., Robinson, S., Syrquin, M., and Feder, S. (1986). *Industrialization and growth*. Oxford University Press.

Christiaensen, L. and Todo, Y. (2014). Poverty Reduction During the Rural–Urban

Transformation – The Role of the Missing Middle. *World Development*, 63:43–58. Economic Transformation in Africa.

Christiaensen, L., Weerdt, J. D., and Todo, Y. (2013). Urbanization and poverty reduction: The role of rural diversification and secondary towns. *Agricultural Economics*, 44(4-5):435–447.

Ciarli, T., Kenney, M., Massini, S., and Piscitello, L. (2021a). Digital technologies, innovation, and skills: Emerging trajectories and challenges. *Research Policy*, 50(7):104289.

Ciarli, T., Kofol, C., and Menon, C. (2015). Business as unusual. An explanation of the increase of private economic activity in high-conflict areas in Afghanistan. LSE Research Online Documents on Economics, London School of Economics.

Ciarli, T., Savona, M., and Thorpe, J. (2021b). Innovation for Inclusive Structural Change. In Lee, J.-D., Lee, K., Meissner, D., Radosevic, S., and Vonortas, N., editors, *The Challenges of Technology and Economic Catch-Up in Emerging Economies*. Oxford University Press.

Cicerone, G., McCann, P., and Venhorst, V. A. (2020). Promoting regional growth and innovation: Relatedness, revealed comparative advantage and the product space. *Journal of Economic Geography*, 20(1):293–316.

Cimoli, M. and Dosi, G. (1995). Technological paradigms, patterns of learning and development: An introductory roadmap. *Journal of Evolutionary Economics*, 5(3):243–268.

Cirera, X. and Maloney, W. F. (2017). *Resolving the Innovation Paradox*. World Bank Group, Washington, DC.

Cirillo, V., Evangelista, R., Guarascio, D., and Sostero, M. (2021). Digitalization, routineness and employment: An exploration on Italian task-based data. *Research Policy*, 50(7):104079.

Clark, C. (1940). *The Conditions of Economic Progress*. Macmillan, London.

Clarke, D. and Tapia-Schythe, K. (2021). Implementing the panel event study. *The Stata Journal: Promoting communications on statistics and Stata*, 21(4):853–884.

Collier, P. and Dercon, S. (2014). African agriculture in 50 years: smallholders in a rapidly changing world? *World development*, 63:92–101.

Collier, P. and Duponchel, M. (2013). The economic legacy of civil war: Firm-level evidence from Sierra Leone. *Journal of Conflict Resolution*, 57(1):65–88.

Cristelli, M., Gabrielli, A., Tacchella, A., Caldarelli, G., and Pietronero, L. (2013). Measuring the intangibles: A metrics for the economic complexity of countries and products. *PLoS ONE*, 8(8).

Cristelli, M., Tacchella, A., Cader, M., Roster, K., and Pietronero, L. (2017). On the predictability of growth. Policy Research Working Paper 8117, World Bank, Washington, DC.

de Beer, J. and Tumaine, N. (2020). Innovation in refugee economies: supporting intermediaries that embrace informality. In Charmes, J., editor, *Research Handbook on Development and the Informal Economy*. Edward Elgar Publishing, Cheltenham, UK.

de Brauw, A., Mueller, V., and Lee, H. L. (2014). The Role of Rural–Urban Migration in the Structural Transformation of Sub-Saharan Africa. *World Development*, 63:33–42. Economic Transformation in Africa.

de Vries, G., Timmer, M., and de Vries, K. (2015). Structural Transformation in Africa: Static Gains, Dynamic Losses. *The Journal of Development Studies*, 51(6):674–688.

Delgado, M., Porter, M. E., and Stern, S. (2016). Defining clusters of related industries. *Working Paper 20375, NBER Working Papers Series.*

Demirgüç-Kunt, A., Klapper, L., Singer, D., Ansar, S., and Hess, J. (2018). The global findex database 2017: Measuring financial inclusion and the fintech revolution. Technical report, World Bank, Washington DC.

DeStefano, T., Kneller, R., and Timmis, J. (2018). Broadband infrastructure, ICT use and firm performance: Evidence for UK firms. *Journal of Economic Behavior & Organization*, 155(C):110–139.

Di Meglio, G., Gallego, J., Maroto, A., and Savona, M. (2018). Services in Developing Economies: The Deindustrialization Debate in Perspective. *Development and Change*, 49(6):1495–1525.

Diao, X., Harttgen, K., and McMillan, M. (2017). The Changing Structure of Africa's Economies. *The World Bank Economic Review*, 31(2):412–433.

Diao, X., Kweka, J., and McMillan, M. (2018a). Small firms, structural change and labor productivity growth in Africa: Evidence from Tanzania. *World Development*, 105:400–415.

Diao, X., Kweka, J., McMillan, M., and Qureshi, Z. (2019a). Economic Transformation in Africa from the Bottom Up: New Evidence from Tanzania. *The World Bank Economic Review*, 34:S58–S62.

Diao, X., Magalhaes, E., and McMillan, M. (2018b). Understanding the Role of Rural Non-Farm Enterprises in Africa's Economic Transformation: Evidence from Tanzania. *The Journal of Development Studies*, 54(5):833–855.

Diao, X. and McMillan, M. (2018). Toward an Understanding of Economic Growth in Africa: A Reinterpretation of the Lewis Model. *World Development*, 109:511–522.

Diao, X., McMillan, M., and Rodrik, D. (2019b). The Recent Growth Boom in Developing Economies: A Structural-Change Perspective. In Nissanke, M. and Ocampo, J. A., editors, *The Palgrave Handbook of Development Economics*, pages 281–334. Springer International Publishing.

Dias, S. M. and Fernandez, L. (2020). Formalisation from the ground: the case of waste pickers' cooperatives. In Charmes, J., editor, *Research Handbook on Development and the Informal Economy*. Edward Elgar Publishing.

Dinh, H. T., Palmade, V., Chandra, V., and Cossar, F. (2012). Manufacturing in

Africa: Targeted Policies to Enhance Private Investment and Create Jobs. Africa development forum, World Bank.

Diodato, D., Neffke, F., and O'Clery, N. (2018). Why do industries coagglomerate? How Marshallian externalities differ by industry and have evolved over time. *Journal of Urban Economics*, 106:1–26.

Djido, A. I. and Shiferaw, B. A. (2018). Patterns of labor productivity and income diversification – empirical evidence from uganda and nigeria. *World Development*, 105:416–427.

Dorosh, P. and Thurlow, J. (2014). Can Cities or Towns Drive African Development? Economywide Analysis for Ethiopia and Uganda. *World Development*, 63:113–123. Economic Transformation in Africa.

Dosi, G. (1988). Sources, procedures, and microeconomic effects of innovation. *The Journal of Economic Literature*, 26:1120–1171.

Duarte, M. and Restuccia, D. (2010). The Role of the Structural Transformation in Aggregate Productivity*. *The Quarterly Journal of Economics*, 125(1):129–173.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What Causes Industry Agglomeration? Evidence from Coagglomeration Patterns. *American Economic Review*, 100(3):1195–1213.

Elvidge, C. D., Baugh, K., Zhizhin, M., Hsu, F. C., and Ghosh, T. (2017). VIIRS night-time lights. *International Journal of Remote Sensing*, 38:5860–5879.

Elvidge, C. D., Baugh, K. E., Anderson, S. J., Sutton, P. C., and Ghosh, T. (2012). The Night Light Development Index (NLDI): A spatially explicit measure of human development from satellite data. *Social Geography*, 7(1):23–35.

Erten, B. and Ocampo, J. A. (2013). Super Cycles of Commodity Prices Since the Mid-Nineteenth Century. *World Development*, 44:14–30.

Fagerberg, J., Kraemer-Mbula, E., and Lorenz, E. (2021). An Evolutionary Analysis of Transformative Change in LDCs: the cases of Kenya and Rwanda. TIK Working Papers on Innovation Studies 20210623, TIK Centre for technology, innovation and culture, Oslo, Norway.

Fagerberg, J. and Verspagen, B. (2021). Technological revolutions, structural change & catching-up. In Alcorta, L., Foster-McGregor, N., Verspagen, B., and Szirmai, A., editors, *New perspectives on structural change: causes and consequences of structural change in the global economy*, chapter 7, pages 131–155. Oxford University Press.

Felipe, J., Kumar, U., Abdon, A., and Bacate, M. (2012). Product complexity and economic development. *Structural Change and Economic Dynamics*, 23(1):36–68.

Fischer, A. G. B. (1938). Production, Primary, Secondary and Tertiary. *Econ. Records*, 15(1):24–38.

Fourastié, J. (1949). *Le Grand Espoir du XXe Siècle: progrès technique, progrès économique, progrès social.* Presses Universitaires de France, Paris.

Freeman, C. and Perez, C. (1988). Structural crises of adjustment, business cycles and investment behaviour. In Dosi, G., editor, *Technical change and economic theory*, pages 38–66. Pinter Publishers, London, NY.

Freire, C. (2021). Economic Complexity Perspectives on Structural Change. In *New Perspectives on Structural Change*, pages 188–214. Oxford University Press.

Frenken, K., Van Oort, F., and Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5):685–697.

Fu, W., Diez, J. R., and Schiller, D. (2013). Interactive learning, informal networks and innovation: Evidence from electronics firm survey in the Pearl River Delta, China. *Research Policy*, 42(3):635–646.

Fu, X. (2020). *Innovation under the Radar: The Nature and Sources of Innovation in Africa.* Cambridge University Press.

Fu, X., Mohnen, P., and Zanello, G. (2018). Innovation and productivity in formal and informal firms in Ghana. *Technological Forecasting & Social Change*.

Fu, X., Pietrobelli, C., and Soete, L. (2011). The role of foreign technology and indigenous innovation in the emerging economies: technological change and catching-up. *World development*, 39(7):1204–1212.

Gaddis, I., Oseni, G., Palacios-Lopez, A., and Pieters, J. (2020). Who Is Employed? Evidence from Sub-Saharan Africa on Redefining Employment. Policy Research Working Paper 9370, World Bank, Washington, DC.

Galetti, J. R. B., Tessarin, M. S., and Morceiro, P. C. (2021). Skill relatedness, structural change and heterogeneous regions: evidence from a developing country. *Papers in Regional Science*.

Gerschenkron, A. (1962). *Economic backwardness in historical perspective*. Harvard University Press, Cambridge, MA.

Ggombe, K. and Newfarmer, R. (2017). Rwanda: From devastation to services-first transformation. *WIDER Working Paper 2017/84*, 84:36.

Ghana Statistical Service (2015). Summary Report: Integrated Business Establishment Survey. Technical report, Ghana Statistical Service, Accra.

Ghana Statistical Service (2016). Ghana Labour Force Survey Report. Technical report, Ghana Statistical Service.

Giuliani, E. (2007). The selective nature of knowledge networks in clusters: Evidence from the wine industry. *Journal of Economic Geography*, 7(2):139–168.

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., and Shleifer, A. (1992). Growth in cities. *Journal of Political Economy*, 100:1126–1152.

Gollin, D., Jedwab, R., and Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth*, 21(1):35–70.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014). The agricultural productivity gap. *The Quarterly Journal of Economics*, 129(2):939–993.

Gollin, D., Parente, S. L., and Rogerson, R. (2007). The food problem and the evolution of international income levels. *Journal of Monetary Economics*, 54(4):1230–1255.

Gomez-Lievano, A. (2018). Methods and Concepts in Economic Complexity. *arXiv preprint*, arXiv:1809.10781.

Gomez-Lievano, A. and Patterson-Lomba, O. (2019). The drivers of urban economic complexity and their connection to urban economic performance. *arXive preprint*, arXiv:1812.02842.

Goos, M., Manning, A., and Salomons, A. (2014). Explaining Job Polarization: Routine-Biased Technological Change and Offshoring. *American Economic Review*, 104(8):2509–2526.

Goos, M., Rademakers, E., and Röttger, R. (2021). Routine-Biased technical change: Individual-Level evidence from a plant closure. *Research Policy*, 50(7):104002.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: Theory. *Econometrica*, 52(3):681–700.

GSMA (2015). Digital inclusion and the role of mobile in Nigeria. Technical report, GSM Association.

Gupta, A., Ponticelli, J., and Tesei, A. (2020). Technology adoption and productivity: the role of mobile phones in agriculture. Working Paper 27192, National Bureau of Economic Research.

Guriev, S., Melnikov, N., and Zhuravskaya, E. (2021). 3G Internet and Confidence in Government. *The Quarterly Journal of Economics*, 136(4):2533–2613.

Haggblade, S., Hazell, P., and Reardon, T. (2010). The rural non-farm economy: Prospects for growth and poverty reduction. *World Development*, 38(10):1429–1441.

Hanlin, R. and Kaplinsky, R. (2016). South-south trade in capital goods - the market-driven diffusion of appropriate technology. *The European Journal of Development Research*, 28(3):361–378.

Haraguchi, N., Cheng, C. F. C., and Smeets, E. (2017). The Importance of Manufacturing in Economic Development: Has This Changed? *World Development*, 93:293–315.

Harding, T. and Venables, A. J. (2016). The implications of natural resource exports for nonresource trade. *IMF Economic Review*, 64(2):268–302.

Harris, J. R. and Todaro, M. P. (1970). Migration, unemployment and development: a two-sector analysis. *The American economic review*, 60(1):126–142.

Hart, K. (1973). Informal Income Opportunities and Urban Employment in Ghana. *The Journal of Modern African Studies*, 11(1).

Hausmann, R. and Chauvin, J. (2015). Diversification in Rwanda: Moving to the Adjacent Possible. *CID Working Paper*.

Hausmann, R. and Hidalgo, C. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences of the United States of America*, 106(26):10570–10575.

Hausmann, R., Hidalgo, C., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A., and Yildirim, M. (2014). *The Atlas of Economic Complexity*. MIT Press, Cambridge MA.

Hausmann, R., Hwang, J., and Rodrik, D. (2007). What you export matters. *Journal of Economic Growth*, 12(1):1–25.

Hausmann, R. and Klinger, B. (2006). Structural transformation and patterns of comparative advantage in the product space. CID Working Paper RWP06-041, Centre for International Development, Harvard University.

Hausmann, R., Klinger, B., and Lawrence, R. (2008). Examining beneficiation.

CID Working Paper Series 162, Centre for International Development, Harvard University, Cambridge, MA.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2014). Growth and structural transformation. In *Handbook of economic growth*, volume 2, pages 855–941. Elsevier.

Herrendorf, B., Rogerson, R., and Ákos Valentinyi (2022). New evidence on sectoral labour productivity: implications for industrialization and development. Working Paper 29834, NBER.

Hidalgo, C. (2018). *Why information grows: The evolution of order, from atoms to economies*. Basic Books.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., and Zhu, S. (2018). The principle of relatedness. In *Unifying Themes in Complex Systems IX*, pages 451–457. Springer International Publishing.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity. *Proceedings of the National Academy of Sciences*, 106(26):10570–10575.

Hidalgo, C. A., Klinger, B., Barabasi, A.-L., and Hausmann, R. (2007). The Product Space Conditions the Development of Nations. *Science*, 317(5837):482–487.

Hirschman, A. O. (1958). *The Strategy of Economic Development*. Yale University Press, New Haven.

Hirschman, A. O. (1977). *The Essential Hirschman*. Princeton University Press.

Hjort, J. and Poulsen, J. (2019). The Arrival of Fast Internet and Employment in Africa. *American Economic Review*, 109(3):1032–1079.

Hjort, J. and Tian, L. (2021). The Economic Impact of Internet Connectivity in Developing Countries. *SSRN Journal*.

Houngbonon, G. V., Mensah, J. T., and Traore, N. (2022). The Impact of Internet

Accesson Innovation and Entrepreneurship in Africa. Policy Research Working Paper 9945, World Bank, Washington, DC.

Hsieh, C.-T. and Rossi-Hansberg, E. (2019). The industrial revolution in services. Working Paper 25968, National Bureau of Economic Research.

Hume, N. and Terazono, E. (2021). Markets weigh prospect of new commodities supercycle. *Financial Times (online).*

Iammarino, S. (2005). An evolutionary integrated view of regional systems of innovation: Concepts, measures and historical perspectives. *European Planning Studies*, 13(4):497–519.

Im, K. S., Pesaran, M. H., and Shin, Y. (2003). Testing for unit roots in heterogeneous panels. *Journal of econometrics*, 115(1):53–74.

Imbs, J. and Wacziarg, R. (2003). Stages of Diversification. *American Economic Review*, 93(1):63–86.

International Fund for Agricultural Development (2019). Creating Opportunities for Rural Youth. Rural Development Report 2019, International Fund for Agricultural Development, Rome.

International Labour Office (2012). International Standard Classification of Occupations: ISCO-08. techreport, ILO, Geneva.

International Labour Organization (2013). 19th international conference of labour statisticians. 2–11 october 2013. Technical report, ILO, Geneva.

International Labour Organization (2018). Women and men in the informal economy: a statistical picture. Third edition, ILO, Geneva.

International Monetary Fund (2019). Growth Slowdown, Precarious Recovery. World Economic Outlook, 2019, International Monetary Fund. OCLC: 1101275363.

Jacobs, J. (1969). *The economy of cities.* Vintage.

Jayne, T. S., Chamberlin, J., and Benfica, R. (2018). Africa's Unfolding Economic Transformation. *The Journal of Development Studies*, 54(5):777–787.

Kaldor, N. (1966). *Causes of the Slow rate of Economic Growth in the United Kingdom*. Cambridge University Press, Cambridge.

Kaldor, N. (1970). The case for regional policies. *Scottish journal of political economy*, 17(3):337–348.

Kanbur, R., Akbar, N., and Stiglitz, J. E., editors (2019). *The Quality of Growth in Africa*. Columbia University Press.

Kaplan, J. O. and Lau, K. H.-K. (2021). *The WWLLN Global Lightning Climatology and timeseries (WGLC)*.

Kaplinsky, R. and Kraemer-Mbula, E. (2022). Innovation and uneven development: The challenge for low- and middle-income economies. *Research Policy*, 51(2).

Keeble, D. and Wilkinson, F. (2000). *High-Technology Clusters, Networking and Collective Learning in Europe*. Ashgate, Aldershot, England.

Kraemer-Mbula, E. and Monaco, L. (2020). Informality and innovation: An avenue towards bottom-up industrialisation for Africa? In Charmes, J., editor, *Research Handbook on Development and the Informal Economy*, pages 363–386. Edward Elgar Publishing.

Kraemer-Mbula, E. and Wunsch-Vincent, S. (2016). Introduction. In Kraemer-Mbula, E. and Wunsch-Vincent, S., editors, *The Informal Economy in Developing Nations: Hidden Engine of Innovation?*, pages 1–12. Cambridge University Press.

Krueger, J. J. (2008). Productivity and structural change: a review of the literature. *Journal of Economic Surveys*, 22(2):330–363.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy*, 99(3).

Kruse, H., Mensah, E., Sen, K., and de Vries, G. (2021). A manufacturing renaissance?

Industrialization trends in the developing world. UNU WIDER Working Paper 2021/08, UNU WIDER, Maastricht, The Netherlands.

Kuznets, S. (1966). *Modern Economic Growth: Rate, Structure, and Spread.* Yale University Press, New Haven and London.

Kuznets, S. (1973). Modern economic growth: Findings and reflections. *The American Economic Review*, 63(3):247–258.

La Porta, R. and Shleifer, A. (2014). Informality and Development. *Journal of Economic Perspectives*, 28(3):109–126.

Lazzeretti, L., Capone, F., and Boix, R. (2012). Reasons for Clustering of Creative Industries in Italy and Spain. *European Planning Studies*, 20(8):1243–1262.

Lenzen, M., Kanemoto, K., Moran, D., and Geschke, A. (2012). Mapping the structure of the world economy. *Environmental Science & Technology*, 46(15):8374—-8381.

Leontief, W. (1953). Domestic production and foreign trade; the American capital position reexamined. *Proceedings of the American Philosophical Society*, 97:66–92.

Levinsohn, J. and Petrin, A. (2003). Estimating Production Functions Using Inputs to Control for Unobservables. *Review of Economic Studies*, 70(2):317–341.

Lewis, A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(8):139–191.

Lewis, A. (1979). The dual economy revisited. *The Manchester School*, 47(3):211–229.

Liao, S., Kim, I. S., Miyano, S., and Zhu, F. (2020). *concordance: Product Concordance.* R package version 2.0.0.

Lin, J., Cader, M., and Pietronero, L. (2020). What African Industrial Development Can Learn from East Asian Successes: The Role of Complexity and Economic Fitness. EMCompass Note 88, The World Bank Group, Washington, DC.

Lin, J. Y. (2012). From Flying Geese To Leading Dragons: New Opportunities and

Strategies for Structural Transformation in Developing Countries. Policy Research Working Paper 5702, World Bank, Washington, DC.

Lin, J. Y., Krueger, A., and Rodrik, D. (2011). New structural economics: A framework for rethinking development. *The World Bank Research Observer*, 26(2):193–223.

Lorentz, A., Ciarli, T., Savona, M., and Valente, M. (2019). Structural Transformations and Cumulative Causation: Towards an Evolutionary Micro-foundation of the Kaldorian Growth Model. *Working Papers of BETA 2019-15, Bureau d'Economie Théorique et Appliquée, UDS Strasbourg.*

Lundvall, B. Å. and Lema, R. (2014). Growth and structural change in Africa: Development strategies for the learning economy. *African Journal of Science, Technology, Innovation and Development*, 6(5):455–466.

Maloney, W. F. (2004). Informality Revisited. *World Development*, 32(7):1159–1178.

Manacorda, M. and Tesei, A. (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *ECTA*, 88(2):533–567.

Marshall, A. (1920). *Principles of Economics.* Macmillan.

Matsuyama, K. (2009). Structural Change in an Interdependent World: A Global View of Manufacturing Decline. *Journal of the European Economic Association*, 7(2-3):478–486.

McCormick, D., Manga, E., Upadhyaya, R., Kamau, P., Wamalwa, H., and Ngigi, S. (2020). Informality and development in Africa. In Charmes, J., editor, *Research Handbook on Development and the Informal Economy*, pages 79–115. Edward Elgar Publishing.

McMillan, M. and Rodrik, D. (2011). Globalization, structural change and productivity growth. Working Paper 17143, NBER.

McMillan, M., Rodrik, D., and Sepulveda, C. (2017a). Structural Change, Fundamentals and Growth: A Framework and Case Studies. NBER Working Paper 23378, NBER.

McMillan, M., Rodrik, D., and Íñigo Verduzco-Gallo (2014). Globalization, Structural Change, and Productivity Growth, with an Update on Africa. *World Development*, 63:11–32.

McMillan, M. and Zeufack, A. (2022). Labor Productivity Growth and Industrialization in Africa. *Journal of Economic Perspectives*, 36(1):3–32.

McMillan, M. S., Rodrik, D., and Sepúlveda, C., editors (2017b). *Structural change, fundamentals, and growth: A framework and case studies*. International Food Policy Research Institute (IFPRI), Washington, D.C.

Meagher, K. (2013). Unlocking the Informal Economy: A Literature Review on Linkages Between Formal and Informal Economies in Developing Countries. *Wiego Working Paper 27*.

Mealy, P., Farmer, J. D., and Teytelboym, A. (2019). Interpreting economic complexity. *Science Advances*, 5(1).

Mehrabi, Z., McDowell, M., Ricciardi, V., Levers, C., Martinez, J., Mehrabi, N., Wittman, H., Ramankutty, N., and Jarvis, A. (2021). The global divide in data-driven farming. *Nature Sustainability*, 4(2):154–160.

Mensah, J. T. (2021). Mobile Phones and Local Economic Development: A Global Evidence. *SSRN Journal*.

Ministry of Finance and Economic Planning, Republic of Rwanda (2020). Vision 2050. Technical report, MINECOFIN.

Ministry of Trade and Industry, Government of Rwanda (2011). National Industrial Policy of Rwanda. Technical report, MINICOM.

Mkandawire, T. (2014). Can Africa turn from recovery to development? *Current History*, 113(763):171–177.

Moreno, A. I., Pieters, J., and Erumban, A. A. (2012). Subcontracting and the size and composition of the informal sector: Evidence from indian manufacturing. *IZA Discussion Paper*, 6785.

Morrison, A. and Rabellotti, R. (2009). Knowledge and information networks in an Italian wine cluster. *European Planning Studies*, 17(7):983–1006.

Mukim, M. (2015). Coagglomeration of formal and informal industry: evidence from India. *Journal of Economic Geography*, 15(2):329–351.

Myrdal, G. (1957). *Economic theory and under-developed regions*. Duckworth London.

Nagler, P. and Naudé, W. (2017). Non-farm entrepreneurship in rural sub-Saharan Africa: New empirical evidence. *Food Policy*, 67:175–191.

Nakano, Y., Tsusaka, T. W., Aida, T., and Pede, V. O. (2018). Is farmer-to-farmer extension effective? The impact of training on technology adoption and rice farming productivity in Tanzania. *World Development*, 105:336–351.

National Institute of Statistics of Rwanda (2018). Labour force survey 2017. Technical report, NISR.

Naudé, N. S. O. W. (2015). Non-Farm Enterprise and Productivity and Spatial Autocorrelation in Rural Africa: Evidence from Ethiopia and Nigeria. *IZA Discussion Paper*, 8295.

Neffke, F., Hartog, M., Boschma, R., and Henning, M. (2018). Agents of structural change: The role of firms and entrepreneurs in regional diversification. *Economic Geography*, 94(1):23–48.

Neffke, F. and Henning, M. (2008). Revealed Relatedness: Mapping Industry Space. Papers in Evolutionary Economic Geography 08.19, Utrecht University.

Neffke, F. and Henning, M. (2013). Skill Relatedness and Firm Diversification. *Strategic Management Journal*, 34:297–316.

Neffke, F., Henning, M., Boschma, R., Lundquist, K.-J., and Olander, L.-O. (2011).

The dynamics of agglomeration externalities along the life cycle of industries. *Regional Studies*, 45(1):49–65.

Nelder, J. A. and Wedderburn, R. W. M. (1972). Generalized linear models. *Journal of the Royal Statistical Society. Series A (General)*, 135(3):370–384.

Nelson, R. R. and Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. The Belknap Press of Harvard University Press, Cambridge, Massachussets and London, UK.

Neumann, D., Nkurunziza, J., Broekhuis, A., and Hooimeijer, P. (2012). Free Education in Rwanda: Just One Step towards Reducing Gender and Sibling Inequalities. *Education Research International*, 2012:396019.

Newfarmer, R. S., Tarp, F., and Page, J., editors (2018). *Industries without smokestacks: industrialization in Africa reconsidered*. WIDER studies in development economics. Oxford University Press, Oxford, United Kingdom. OCLC: on1030900452.

Newman, C., Page, J. M., Rand, J., Shimeles, A., Söderbom, M., and Tarp, F. (2016). *Made in Africa: Learning to compete in industry*. Brookings Institution Press, Washington, D.C.

Newman, M. E. J. (2018). *Networks*. Oxford University Press, Oxford, United Kingdom ; New York, NY, United States of America, second edition edition.

Newman, M. E. J. and Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 69(2 2):1–16.

OECD (2015). Innovation Policies for Inclusive Development. Scaling Up Inclusive Innovations. Technical report, OECD, Paris.

Oqubay, A. and Ohno, K. (2019). *How nations learn: Technological learning, industrial policy, and catch-up*. Oxford University Press.

Osei, R. D. and Jedwab, R. (2017). Structural change in a poor African country: New historical evidence from Ghana. In McMillan, Margaret S.; Rodrik, D. and Sepúlveda, C., editors, *Structural change, fundamentals, and growth: A framework and case studies*, pages 161–196. International Food Policy Research Institute (IFPRI), Washington, D.C.

Owusu, S., Szirmai, A., and Foster-McGregor, N. (2021). The rise of services in the global economy. *New Perspectives on Structural Change: Causes and Consequences of Structural Change in the Global Economy*, page 270.

Page, J. (2012). Can Africa Industrialise? *Journal of African Economies*, 21(suppl 2):ii86–ii124.

Papke, L. E. and Woolridge, J. M. (1996). Econometric Methods and for Fractional Response Variables with an Application to 401 (K) Plan Participation Rates. *Journal of Applied Econometrics*, 11:619–632.

Paunov, C. (2013). *Innovation and Inclusive Development*. OECD Publishing.

Paus, E., Robinson, M., and Tregenna, F. (2022). Firm innovation in Africa and Latin America: Heterogeneity and country context. *Industrial and Corporate Change*, 31(2):338–357.

Perez, C. (1986). New Technologies and Development. In *Small Countries Confronting the Technological Revolut*. Pinter, London.

Perry, G. E., Maloney, W. F., Arias, O. S., Fajnzylber, P., Mason, A. D., and Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion*. The World Bank.

Pesaran, M. H., Shin, Y., and Smith, R. J. (2001). Bounds testing approaches to the analysis of level relationships. *Journal of Applied Econometrics*, 16(3):289–326.

Petralia, S., Balland, P.-A., and Morrison, A. (2017). Climbing the ladder of technological development. *Research Policy*, 46(5):956–969.

Pietronero, L., Cristelli, M., Gabrielli, A., Mazzilli, D., Pugliese, E., Tacchella, A., and Zaccaria, A. (2017). Economic complexity: "buttarla in caciara" vs a constructive approach. *arXiv*, 1709.05272.

Piketty, T. (2018). *Capital in the twenty-first century.* Harvard University Press.

Prebisch, R. (1950). *The Economic Development of Latin America and its Principal Problem, Economic Commission for Latin America.* Economic Commission for Latin America, United Nations Department of Economic Affairs, Lake Success, New York.

Pugliese, E., Cimini, G., Patelli, A., Zaccaria, A., Pietronero, L., and Gabrielli, A. (2019). Unfolding the innovation system for the development of countries: coevolution of science, technology and production. *Scientific Reports*, 9(1):16440.

Raleigh, C., Linke, A., Hegre, H., and Karlsen, J. (2010). Introducing ACLED: An Armed Conflict Location and Event Dataset: Special Data Feature. *Journal of Peace Research*, 47(5):651–660.

Ramana, M. and Agyapong, P. (2016). Thinking big? Ghana, small reactors, and nuclear power. *Energy Research & Social Science*, 21:101–113.

Rambachan, A. and Roth, J. (2021). An honest approach to parallel trends. Technical report, Working Paper.

Ravallion, M. (2004). Pro-poor growth: A primer. Working paper, SSRN.

Rodrik, D. (2008). The real exchange rate and economic growth. *Brookings Papers on Economic Activity*, 39(2 (Fall)):365–439.

Rodrik, D. (2013). Unconditional Convergence in Manufacturing. *The Quarterly Journal of Economics*, 128(1):165–204.

Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, 21(1):1–33.

Rodrik, D. (2018). An African Growth Miracle? *Journal of African Economies*, 27(1):10–27.

Romer, P. M. (1986). Increasing returns and long-run growth. *The Journal of Political Economy*, 94(5):1002–1037.

Rosenberg, N. (1979). Technological Interdependence in the American Economy: History of Technology. *Technology and Culture*, 20(1):25–50.

Rosenberg, N. and Frischtak, C. R. (1983). Long waves and economic growth: A critical appraisal. *The American Economic Review*, 73(2):146–151.

Rostow, W. (1959). The Stages of Economic Growth. *The Economic History Review*, 12(1):1–16.

Rovigatti, G. and Mollisi, V. (2018). Theory and practice of total-factor productivity estimation: The control function approach using stata. *The Stata Journal*, 18(3):618–662.

Rwanda Utilities Regulatory Authority (2021). Statistics report for telecom, media and broadcasting sector as of the first quarter of the year 2021. Technical report, RURA.

Sachs, J. D. and Warner, A. (2001). The curse of natural resources. *European Economic Review*, 45(4-6):827–838.

Saha, A. and Ciarli, T. (2018). Innovation, Structural Change, and Inclusion. A Cross Country PVAR Analysis. Working Paper Series 01/2018, SPRU, Brighton, UK.

Saviotti, P. P. and Pyka, A. (2004). Economic development by the creation of new sectors. *Journal of Evolutionary Economics*, 14(1):1–35.

Savona, M. (2021). Revisiting high development theory to explain upgrading prospects in business services global value chains. *The European Journal of Development Research*, 33(2):206–226.

Saxenian, A. (1996). *Regional Advantage: Culture and Competition in Silicon Valley and Route 128*. Harvard University Press.

Sbardella, A., Perruchas, F., Napolitano, L., Barbieri, N., and Consoli, D. (2018a). Green Technology Fitness. *Entropy*, 20(776).

Sbardella, A., Pugliese, E., and Pietronero, L. (2017). Economic development and wage inequality: A complex system analysis. *PLOS ONE*, 12(9):1–26.

Sbardella, A., Pugliese, E., Zaccaria, A., and Scaramozzino, P. (2018b). The role of complex analysis in modelling economic growth. *Entropy*, 20(11):883.

Schumpeter, J. (1943). *Capitalism, Socialism and Democracy.* Harper, New York.

Schumpeter, J. A. (1934). *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest, and the Business Cycle.* Harvard University Press, Cambridge, MA.

Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24(4):265–269.

Stiglitz, J. E., Lin, J. Y., Monga, C., and Patel, E. (2013). Industrial policy in the African context. Policy Research Working Paper 6633, World Bank.

Stock, J. and Watson, M. (1993). A simple estimator of cointegrating vectors in higher order integrated systems. *Econometrica*, 61(4):783–820.

Stock, J. and Yogo, M. (2005). Testing for Weak Instruments in Linear IV Regression. In Andrews, D. W. K. and Stock, J. H., editors, *Identification and Inference for Econometric Models*, pages 80–109. Cambridge University Press, Cambridge, UK.

Strazzeri, M. (2021). Internet usage and migration decisions: Evidence from Nigerian micro data. PhD Thesis.

Su, D. and Yao, Y. (2017). Manufacturing as the key engine of economic growth for middle-income economies. *Journal of the Asia Pacific Economy*, 22(1):47–70.

Suri, T., Aker, J., Batista, C., Callen, M., Ghani, T., Jack, W., Klapper, L., Riley, E., Schaner, S., and Sukhtankar, S. (2021). Mobile money. *VoxDevLit*, 2(1).

Sutton, J. and Kpentey, B. (2012). *An enterprise map of Ghana*, volume 2. International Growth Centre, London Publishing Partnership.

Syrquin, M. (2010). Kuznets and pasinetti on the study of structural transformation: Never the twain shall meet? *Structural Change and Economic Dynamics*, 21(4):248–257.

Szirmai, A. (2012). Industrialisation as an engine of growth in developing countries, 1950–2005. *Structural Change and Economic Dynamics*, 23(4):406–420.

Tacchella, A., Cristelli, M., Caldarelli, G., Gabrielli, A., and Pietronero, L. (2012). A New Metrics for Countries' Fitness and Products' Complexity. *Scientific Reports*, 2(723):1–7.

Tacchella, A., Mazzilli, D., and Pietronero, L. (2018). A dynamical systems approach to gross domestic product forecasting. *Nature Physics*, 14(8):861–865.

The Growth Lab at Harvard University (2019). *Growth Projections and Complexity Rankings, V2 [Data set].* The Growth Lab at Harvard University.

The Growth Lab at Harvard University (2022). The Atlas of Economic Complexity. Website, The Growth Lab at Harvard University.

Thirlwall, A. (1979). The balance of payments contraint as an explanation of internationalgrowth rate differences. *Banca Nazionale del Lavoro*, 32:45–53.

Timmer, M. P., Miroudot, S., and de Vries, G. J. (2018). Functional specialisation in trade. *Journal of Economic Geography*, 19(1):1–30.

UNCTAD (2021). Reaping the Potential Benefits of the African Continental Free Trade Area for Inclusive Growth. Economic Development in Africa Report, United Nations Conference on Trade and Development, Geneva.

United Nations Statistical Division (2002). *International standard industrial classification of all economic activities (ISIC), rev. 3.1.* United Nations Publications.

Valensisi, G. and Davis, J. (2011). Least developed countries and the green transition:

Towards a renewed political economy agenda. Maastricht School of Management Working Papers 27, Maastricht School of Management.

Vargas, M., Bottai, C., Kozlowski, D., and Pintar, N. (2020). *economiccomplexity: Computational Methods for Economic Complexity*. R package version 1.1.

Verspagen, B. (1991). A new empirical approach to catching up or falling behind. *Structural Change and Economic Dynamics*, 2(2):359–380.

Vivarelli, M. (2021). Globalization, Structural Change, and Technological Change in Emerging Economies. In Foster-McGregor, N., Alcorta, L., Szirmai, A., and Verspagen, B., editors, *New Perspectives on Structural Change*. Oxford University Press.

Waltman, L., van Eck, N. J., and Noyons, E. C. M. (2010). A unified approach to mapping and clustering of bibliometric networks. *CoRR*, abs/1006.1032.

Weidmann, N. B. and Schutte, S. (2017). Using night light emissions for the prediction of local wealth. *Journal of Peace Research*, 54(2):125–140.

World Bank (2022). World Development Indicators.

Young, A. (2012). The African Growth Miracle. *Journal of Political Economy*, 120(4):696–739.

Appendix A

Appendix to Chapter 2

A.1 Treemaps

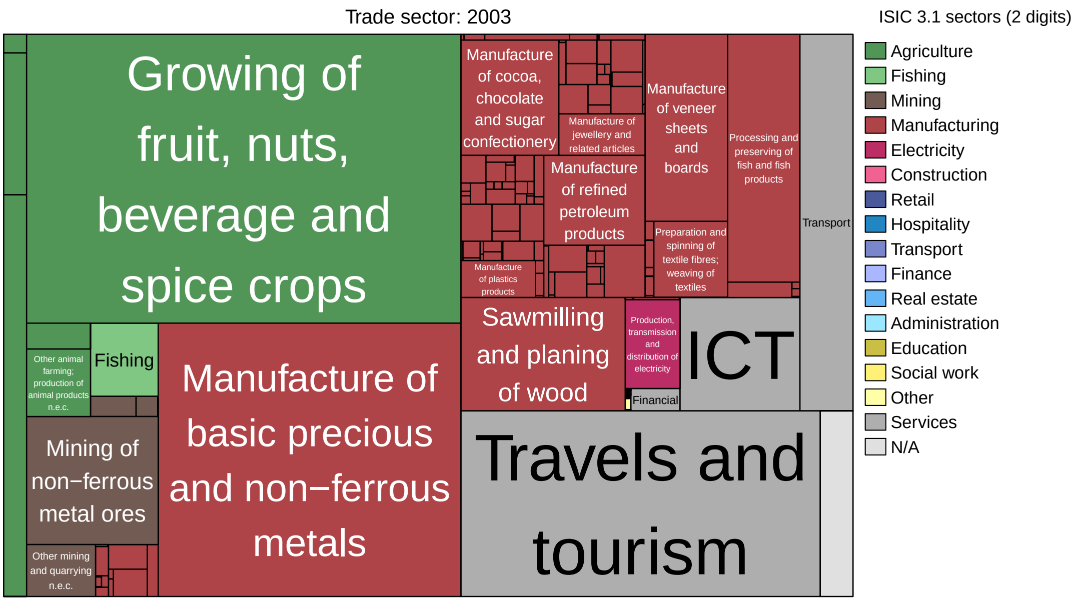

Figure A.1: Exported goods and services by Ghana – ISIC industries (2003).

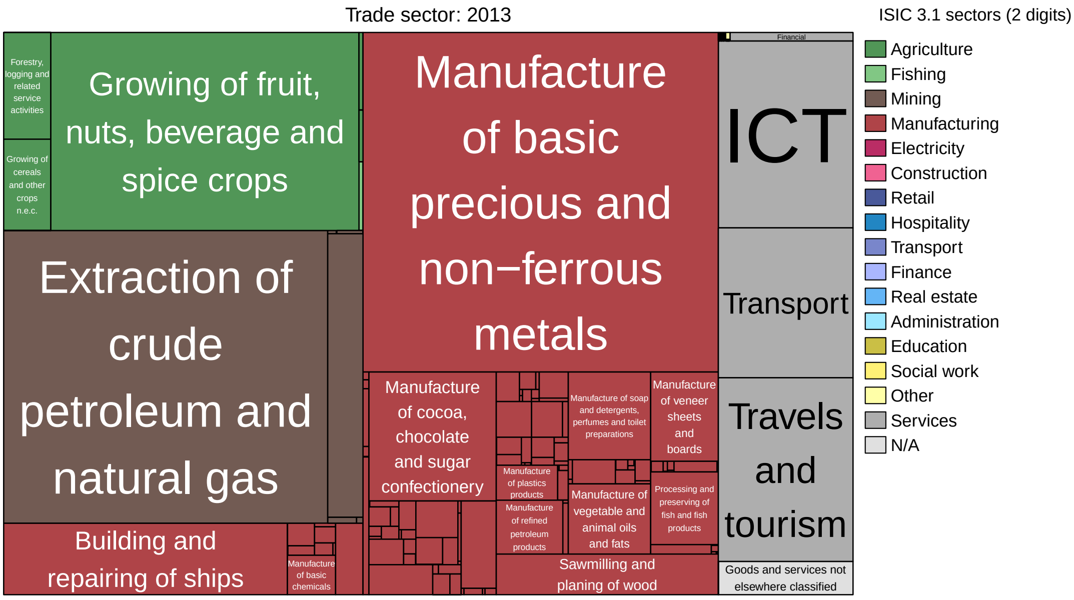

Figure A.2: Exported goods and services by Ghana – ISIC industries (2013).

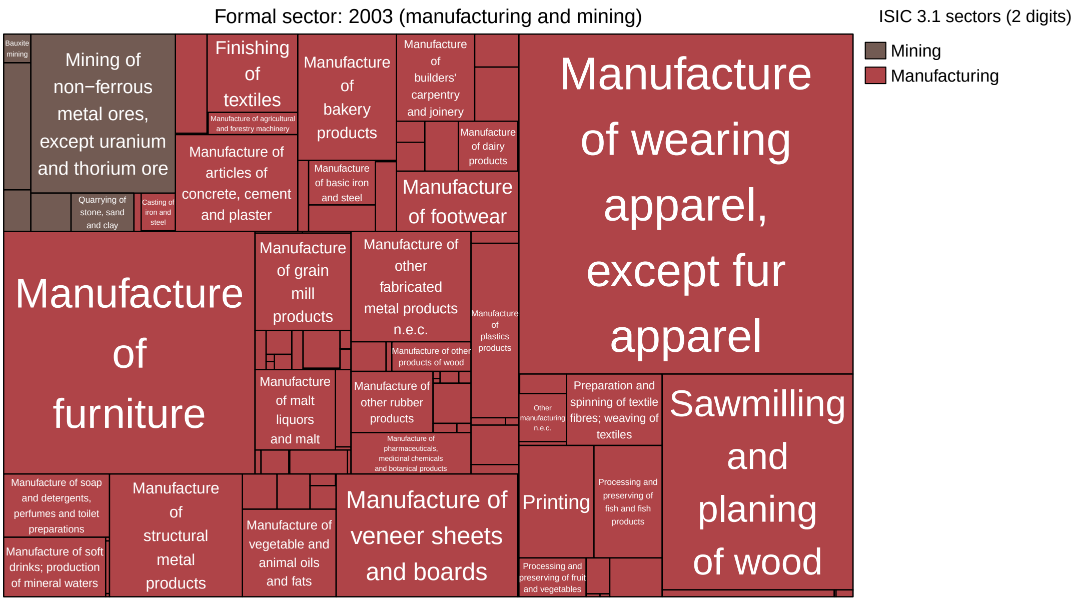

Figure A.3: Proportion of formal employment in Ghana by ISIC industries – manufacturing and mining (2003).

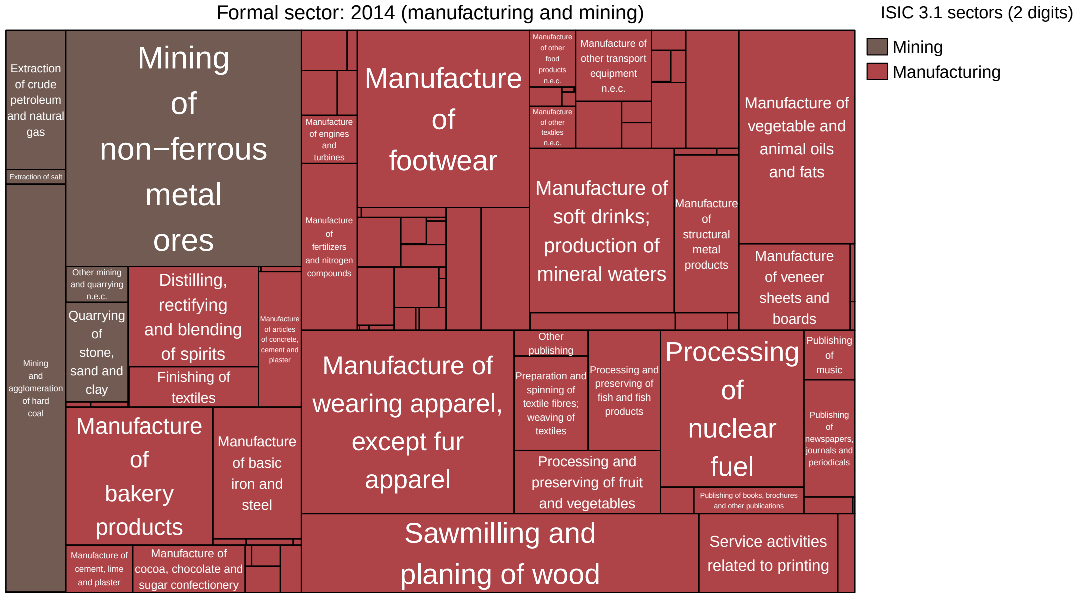

Figure A.4: Proportion of formal employment in Ghana – manufacturing and mining (2014).

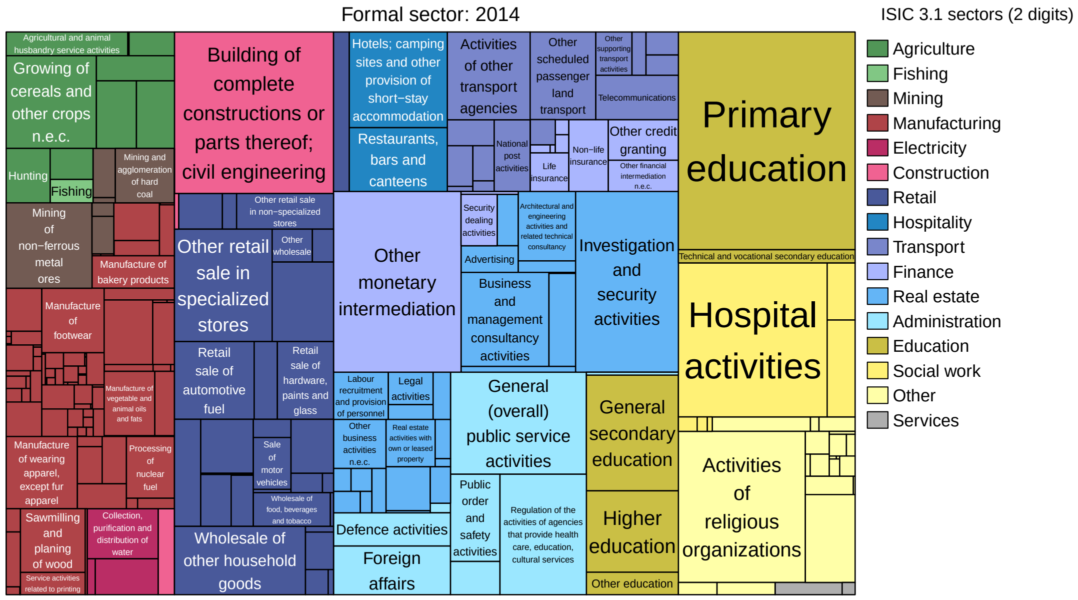

Figure A.5: Proportion of formal employment in Ghana by ISIC industries (2014).

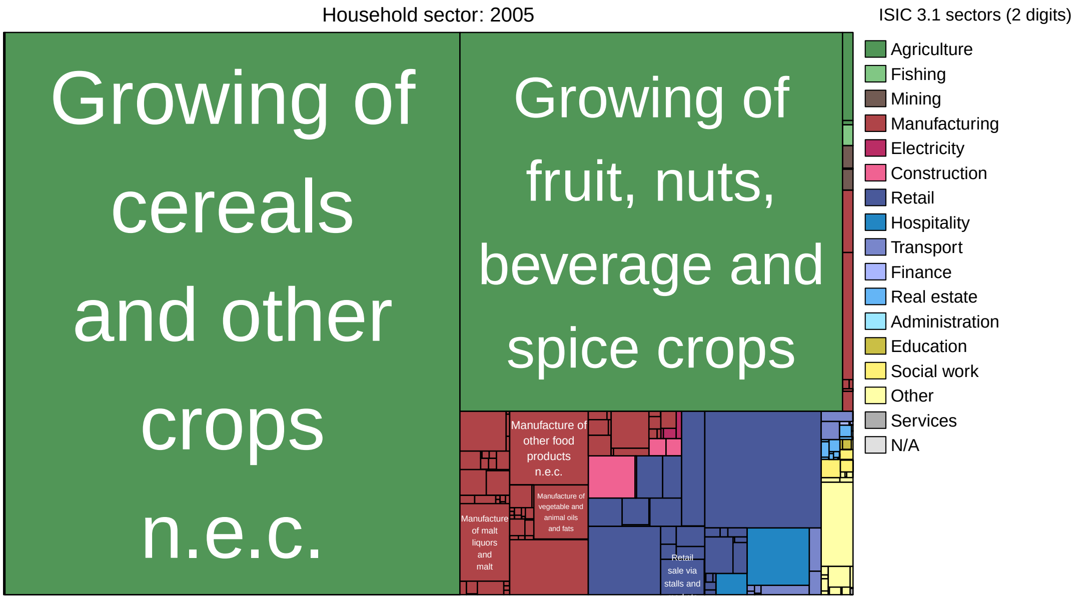

Figure A.6: Proportion of employment in household enterprises by ISIC industries – GLSS (2005).

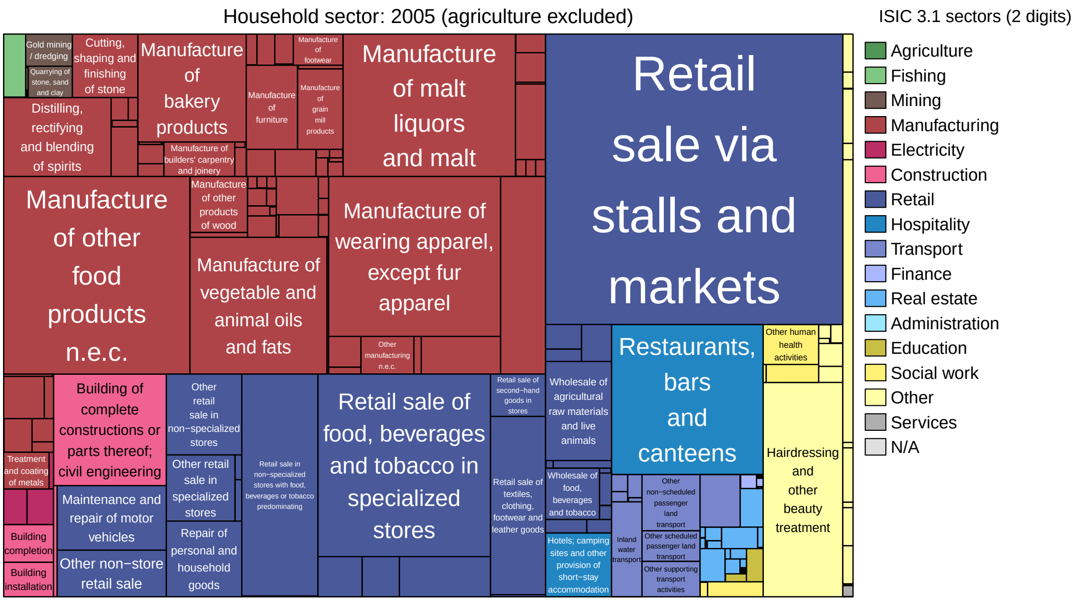

Figure A.7: Proportion of employment in household enterprises by ISIC industries (agriculture excluded) – GLSS (2005).

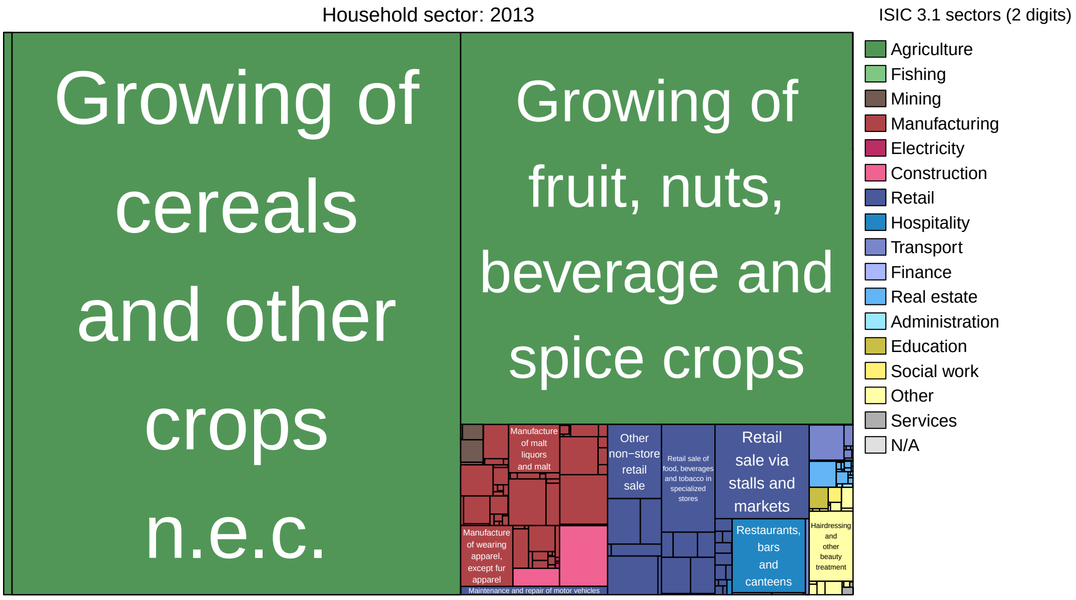

Figure A.8: Proportion of employment in household enterprises by ISIC industries – GLSS (2013).

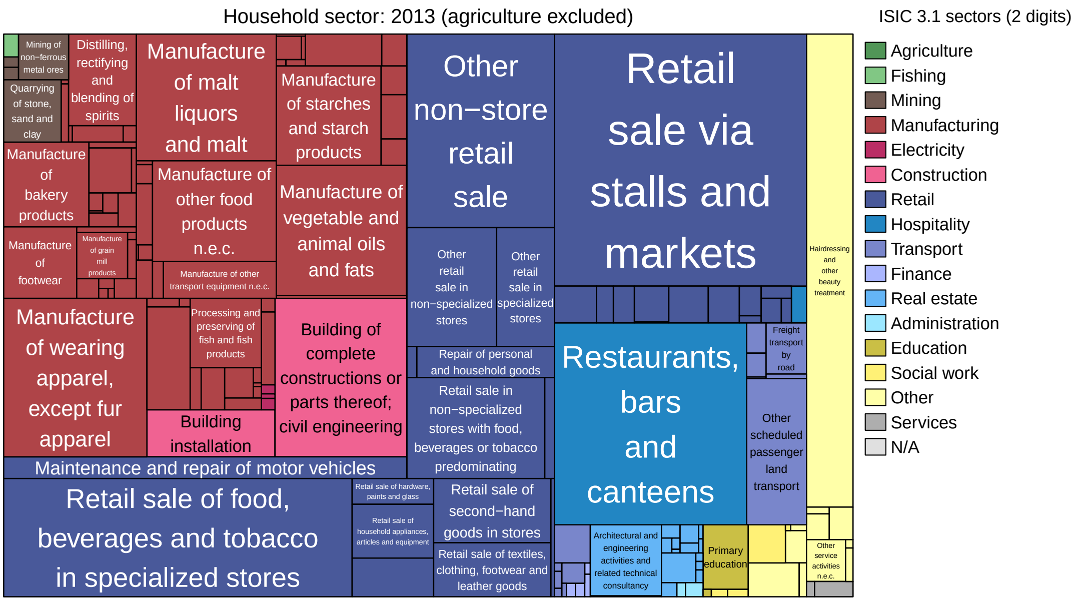

Figure A.9: Proportion of employment in household non-farming enterprises by ISIC industries (agriculture excluded) – GLSS (2013).

A.2 Tradable services – correspondence table

Table A.1: Correspondence between Harmonised System tradable services and ISIC industries

Tradable services label (HS)	ISIC corresponding industries
ICT	4-digit ISIC codes: 3000, 3130, 3210, 3220, 3230, 3312, 3313, 5150, 7123, 6420, 7210, 7221, 7229, 7230, 7240, 7250, 7290
Financial	All 4-digit ISIC industries in “Financial intermediation” (1-digit label)
Transport	All 4-digit ISIC industries in “Transport, storage and communication” (1-digit label)
Travels and tourism	4-digit ISIC code: 6304

Note: Please refer to Appendix B for the full ISIC labels

A.3 Validating the complexity and fitness framework

To provide empirical support for the complexity and fitness framework used for our analysis, we compute the two-way correlation between fitness and a widely used proxy for economic development at the subnational level: nighttime luminosity (Elvidge et al., 2012; Alesina et al., 2016; Weidmann and Schutte, 2017). Building on evidence that measures of economic complexity (using trade data) are good predictors of national growth (Cristelli et al., 2017; Tacchella et al., 2018), we assume that the "fittest" districts are specialised in more sophisticated industries, which require more capabilities and lead to higher economic output.

Figure A.10 plots the correlation between district fitness, calculated as described in Section 2.3.2, and district average night time luminosity. Nightlights satellite imagery was made available by the Earth Observation Group and was collected as part of the VIIRS time series (Elvidge et al., 2017) and was used to produce a high quality global Nighttime Light map. The data comes in a raster of 15 arc seconds resolution (approx. 500m at the Equator). Superimposing Ghana's districts coordinates on the raster for 2014 (the year in which the IBES data were collected), the nightlight data were averaged by district to create a measure of average nighttime luminosity.

The positive slope of the fitted line between fitness and average nightlights (Figure A.10) shows that there is a positive relation between economic fitness and economic development in Ghana's districts, as predicted at the country level. To ensure that the results are not driven by higher population density in districts engaged in more diversified economic activities, Table A.2 presents the results of an Ordinary Least Squares regression of fitness on nightlight luminosity, without (column 1) and with (column 2) the log of the district population density as covariate. Neither the results nor the R-squared change substantially with the introduction of population density. The coefficient of the log of nightlights remains positive and statistically significant, although the point estimate is slightly smaller after controlling for population. Of

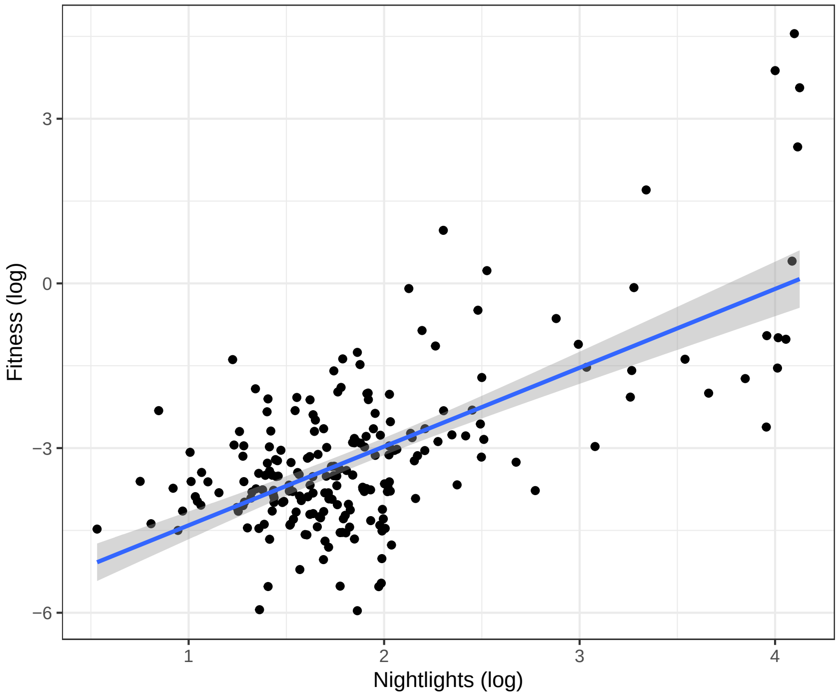

Figure A.10: Correlation between fitness and nighttime luminosity.

course, these results do not provide evidence that the ICI, our measure of industry complexity, causes higher economic development in a district (measured by nightlight intensity).

Table A.2: Correlation between district fitness and nighttime luminosity

	Dependent variable:	
	Fitness (log)	
	(1)	(2)
Avg. Nightlights (log)	1.436***	1.351***
	(0.188)	(0.184)
Pop. density (log)		0.056
		(0.051)
Constant	−5.843***	−5.936***
	(0.328)	(0.360)
Observations	216	216
R^2	0.428	0.431
Adjusted R^2	0.425	0.426
Residual Std. Error	1.158 (df = 214)	1.157 (df = 213)
F Statistic	159.893*** (df = 1; 214)	80.632*** (df = 2; 213)

Note: All coefficients are estimated using OLS. Observations are Ghanaian districts. Heteroskedasticity robust s.e. in parentheses: *p<0.1; **p<0.05; ***p<0.01.

A.4 Final dataset

Table A.3: Final dataset

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
0111	Growing of cereals and other crops n.e.c.	Agriculture	1.93	0.90	0.00	1.75	53.51	53.09	-6.91
0112	Growing of vegetables, horticultural specialties and nursery products	Agriculture	0.35	0.07	0.00	0.54	0.19	0.00	-4.97
0113	Growing of fruit, nuts, beverage and spice crops	Agriculture	26.28	12.78	0.00	0.57	30.32	32.34	-6.81
0121	Farming of cattle, sheep, goats, horses, asses, mules and hinnies; dairy farming	Agriculture	0.09	0.00	0.00	0.01	0.02	0.00	-7.38
0122	Other animal farming; production of animal products n.e.c.	Agriculture	0.90	0.04	0.00	0.28	0.01	0.00	-7.38
0130	Growing of crops combined with farming of animals (mixed farming)	Agriculture	0.00	0.00	0.00	0.41	0.00	0.00	-5.13
0140	Agricultural and animal husbandry service activities	Agriculture	0.00	0.00	0.00	0.60	0.00	0.00	-7.59
0150	Hunting	Agriculture	0.00	0.00	0.00	0.50	0.00	0.00	-6.08
0200	Forestry, logging and related service activities	Agriculture	0.68	1.05	0.00	0.23	0.13	0.99	-7.08
0501	Fishing	Fishing	1.04	0.06	0.00	0.21	0.05	0.01	-5.25
1010	Mining and agglomeration of hard coal	Mining	0.00	0.00	0.00	0.67	0.00	0.01	-4.54
1020	Mining and agglomeration of lignite	Mining	0.00	0.00	0.00	NA	0.00	0.00	NA
1030	Extraction and agglomeration of peat	Mining	0.00	0.00	0.00	NA	0.00	0.00	NA
1110	Extraction of crude petroleum and natural gas	Mining	0.19	19.86	0.00	0.23	0.00	0.00	-4.22
1200	Mining of uranium and thorium ores	Mining	0.00	0.00	0.00	0.00	0.00	0.00	4.70
1310	Mining of iron ores	Mining	0.00	0.00	0.00	0.00	0.00	0.00	2.43
1320	Mining of non-ferrous metal ores	Mining	3.53	2.10	4.80	1.53	0.00	0.07	-5.81
1322	Bauxite mining	Mining	0.00	0.00	0.16	NA	0.00	0.00	NA
1323	Manganese mining	Mining	0.00	0.00	0.24	NA	0.00	0.00	NA
1410	Quarrying of stone, sand and clay	Mining	0.00	0.01	0.50	0.17	0.05	0.11	-6.04
1421	Mining of chemical and fertilizer minerals	Mining	0.00	0.02	0.00	NA	0.00	0.00	NA
1422	Extraction of salt	Mining	0.09	0.01	0.72	0.02	0.00	0.00	-4.88
1429	Other mining and quarrying n.e.c.	Mining	0.75	0.04	0.32	0.06	0.00	0.00	-3.92
1511	Production, processing and preserving of meat and meat products	Manufacturing	0.01	0.01	0.18	0.02	0.06	0.03	-6.99
1512	Processing and preserving of fish and fish products	Manufacturing	3.75	1.03	1.61	0.24	0.29	0.15	-6.03
1513	Processing and preserving of fruit and vegetables	Manufacturing	0.21	0.12	0.55	0.26	0.00	0.01	-5.98
1514	Manufacture of vegetable and animal oils and fats	Manufacturing	0.45	1.21	1.72	0.68	0.62	0.50	-7.08
1520	Manufacture of dairy products	Manufacturing	0.14	0.03	0.62	0.05	0.02	0.01	-4.34
1531	Manufacture of grain mill products	Manufacturing	0.18	0.15	1.95	0.06	0.12	0.07	-8.64
1532	Manufacture of starches and starch products	Manufacturing	0.17	0.22	0.04	0.01	0.03	0.31	-8.02
1533	Manufacture of prepared animal feeds	Manufacturing	0.00	0.00	0.05	0.02	0.01	0.00	-5.35
1541	Manufacture of bakery products	Manufacturing	0.02	0.01	2.61	0.56	0.38	0.21	-6.57

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
1542	Manufacture of sugar	Manufacturing	0.02	0.00	0.00	0.00	0.00	0.00	-2.05
1543	Manufacture of cocoa, chocolate and sugar confectionery	Manufacturing	2.36	3.44	0.50	0.15	0.00	0.01	-5.44
1544	Manufacture of macaroni, noodles, couscous and similar farinaceous products	Manufacturing	0.02	0.04	0.00	NA	0.00	0.00	-1.51
1549	Manufacture of other food products n.e.c.	Manufacturing	0.16	0.23	0.22	0.07	1.21	0.37	-5.82
1551	Distilling, rectifying and blending of spirits	Manufacturing	0.01	0.13	0.66	0.36	0.28	0.18	-6.89
1552	Manufacture of wines	Manufacturing	0.02	0.01	0.00	NA	0.03	0.00	NA
1553	Manufacture of malt liquors and malt	Manufacturing	0.01	0.00	1.35	0.05	0.81	0.52	-7.11
1554	Manufacture of soft drinks; production of mineral waters	Manufacturing	0.03	0.02	1.27	0.66	0.00	0.02	-6.85
1600	Manufacture of tobacco products	Manufacturing	0.05	0.01	0.13	NA	0.00	0.00	NA
1711	Preparation and spinning of textile fibres; weaving of textiles	Manufacturing	1.17	0.10	1.43	0.19	0.10	0.13	-6.91
1712	Finishing of textiles	Manufacturing	0.00	0.00	1.49	0.14	0.01	0.02	-4.79
1721	Manufacture of made-up textile articles, except apparel	Manufacturing	0.03	0.02	0.05	0.00	0.01	0.01	-5.96
1722	Manufacture of carpets and rugs	Manufacturing	0.00	0.00	0.00	0.00	0.00	0.01	1.10
1723	Manufacture of cordage, rope, twine and netting	Manufacturing	0.00	0.00	0.12	0.02	0.01	0.00	-3.82
1729	Manufacture of other textiles n.e.c.	Manufacturing	0.01	0.00	0.01	0.05	0.00	0.01	-5.12
1730	Manufacture of knitted and crocheted fabrics and articles	Manufacturing	0.14	0.00	0.01	0.00	0.04	0.01	-6.57
1810	Manufacture of wearing apparel, except fur apparel	Manufacturing	0.05	0.03	23.77	1.08	0.90	0.67	-7.73
1820	Dressing and dyeing of fur; manufacture of articles of fur	Manufacturing	0.00	0.00	0.01	0.00	0.01	0.00	-4.99
1911	Tanning and dressing of leather	Manufacturing	0.02	0.00	0.00	NA	0.01	0.00	NA
1912	Manufacture of luggage, handbags and the like	Manufacturing	0.01	0.01	0.06	0.01	0.00	0.00	-6.39
1920	Manufacture of footwear	Manufacturing	0.04	0.04	1.55	0.84	0.05	0.15	-4.95
2010	Sawmilling and planing of wood	Manufacturing	3.90	1.89	8.62	0.86	0.03	0.01	-7.03
2021	Manufacture of veneer sheets and boards	Manufacturing	3.23	1.26	4.66	0.26	0.01	0.00	-4.31
2022	Manufacture of builders' carpentry and joinery	Manufacturing	0.03	0.01	1.43	0.01	0.08	0.06	-8.16
2023	Manufacture of wooden containers	Manufacturing	0.00	0.00	0.19	NA	0.01	0.01	-6.39
2029	Manufacture of other products of wood	Manufacturing	0.18	0.09	0.49	NA	0.11	0.00	NA
2101	Manufacture of pulp, paper and paperboard	Manufacturing	0.02	0.01	0.02	0.00	0.00	0.00	-1.99
2102	Manufacture of corrugated paper and paperboard	Manufacturing	0.06	0.04	0.17	0.03	0.00	0.00	-1.16
2109	Manufacture of other articles of paper and paperboard	Manufacturing	0.01	0.00	0.29	NA	0.01	0.00	NA
2211	Publishing of books, brochures and other publications	Manufacturing	0.01	0.00	0.01	0.08	0.01	0.00	-4.24
2212	Publishing of newspapers, journals and periodicals	Manufacturing	0.00	0.00	0.43	0.16	0.00	0.00	0.11
2213	Publishing of music	Manufacturing	0.00	0.00	0.01	0.07	0.00	0.00	-3.98
2219	Other publishing	Manufacturing	0.00	0.00	0.04	0.05	0.01	0.00	-6.19
2221	Printing	Manufacturing	0.00	0.00	1.77	NA	0.00	0.00	NA
2222	Service activities related to printing	Manufacturing	0.00	0.00	0.25	0.30	0.02	0.01	-4.38
2230	Reproduction of recorded media	Manufacturing	0.00	0.00	0.00	0.00	0.00	0.00	-4.87
2310	Manufacture of coke oven products	Manufacturing	0.00	0.00	0.00	0.00	0.01	0.00	-3.22
2320	Manufacture of refined petroleum products	Manufacturing	1.91	0.75	0.40	NA	0.00	0.00	NA
2330	Processing of nuclear fuel	Manufacturing	0.00	0.00	0.00	0.62	0.00	0.03	-5.20

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
2411	Manufacture of basic chemicals	Manufacturing	0.19	0.40	0.16	NA	0.00	0.00	NA
2412	Manufacture of fertilizers and nitrogen compounds	Manufacturing	0.10	0.19	0.01	0.26	0.00	0.00	-2.39
2413	Manufacture of plastics in primary forms and of synthetic rubber	Manufacturing	0.04	0.01	0.12	0.02	0.00	0.00	-3.27
2421	Manufacture of pesticides and other agrochemical products	Manufacturing	0.01	0.02	0.03	0.02	0.00	0.00	-2.41
2422	Manufacture of paints, varnishes and similar coatings	Manufacturing	0.04	0.10	0.32	0.06	0.00	0.00	0.58
2423	Manufacture of pharmaceuticals, medicinal chemicals and botanical products	Manufacturing	0.13	0.05	1.04	NA	0.01	0.00	NA
2424	Manufacture of soap and detergents, perfumes and toilet preparations	Manufacturing	0.21	1.51	1.40	0.17	0.05	0.07	-6.28
2429	Manufacture of other chemical products n.e.c.	Manufacturing	0.08	0.16	0.02	0.02	0.00	0.00	-4.49
2430	Manufacture of man-made fibres	Manufacturing	0.01	0.01	0.25	NA	0.00	0.00	-1.08
2511	Manufacture of rubber tyres and tubes	Manufacturing	0.05	0.06	0.00	0.06	0.00	0.00	-1.13
2519	Manufacture of other rubber products	Manufacturing	0.01	0.06	1.05	0.00	0.01	0.00	-2.72
2520	Manufacture of plastics products	Manufacturing	0.58	0.45	1.75	0.02	0.01	0.02	-6.80
2610	Manufacture of glass and glass products	Manufacturing	0.09	0.32	0.02	0.02	0.01	0.00	-6.37
2691	Manufacture of non-structural non-refractory ceramic ware	Manufacturing	0.02	0.00	0.14	0.01	0.07	0.03	-6.47
2692	Manufacture of refractory ceramic products	Manufacturing	0.00	0.00	0.09	0.01	0.02	0.00	-5.69
2693	Manufacture of structural non-refractory clay and ceramic products	Manufacturing	0.02	0.01	0.25	NA	0.01	0.00	NA
2694	Manufacture of cement, lime and plaster	Manufacturing	0.06	0.17	0.31	0.09	0.00	0.00	-3.42
2695	Manufacture of articles of concrete, cement and plaster	Manufacturing	0.01	0.05	2.48	0.16	0.04	0.01	-6.07
2696	Cutting, shaping and finishing of stone	Manufacturing	0.00	0.00	0.00	NA	0.14	0.00	-3.09
2699	Manufacture of other non-metallic mineral products n.e.c.	Manufacturing	0.00	0.00	0.04	0.00	0.00	0.00	-4.34
2710	Manufacture of basic iron and steel	Manufacturing	0.12	0.41	0.62	0.32	0.00	0.01	-4.77
2720	Manufacture of basic precious and non-ferrous metals	Manufacturing	17.34	25.27	0.37	NA	0.00	0.00	NA
2731	Casting of iron and steel	Manufacturing	0.00	0.00	0.27	NA	0.00	0.00	NA
2732	Casting of non-ferrous metals	Manufacturing	0.00	0.00	0.00	0.01	0.00	0.00	-1.65
2811	Manufacture of structural metal products	Manufacturing	0.02	0.03	3.41	0.28	0.01	0.02	-6.46
2812	Manufacture of tanks, reservoirs and containers of metal	Manufacturing	0.01	0.01	0.24	0.07	0.00	0.00	-6.89
2813	Manufacture of steam generators, except central heating hot water boilers	Manufacturing	0.00	0.00	0.00	0.00	0.00	0.00	-3.52
2891	Forging, pressing, stamping and roll-forming of metal; powder metallurgy	Manufacturing	0.00	0.00	0.00	0.00	0.00	0.00	-6.88
2892	Treatment and coating of metals	Manufacturing	0.00	0.00	0.02	0.04	0.05	0.03	-7.56
2893	Manufacture of cutlery, hand tools and general hardware	Manufacturing	0.15	0.09	0.35	0.03	0.02	0.01	-7.22
2899	Manufacture of other fabricated metal products n.e.c.	Manufacturing	0.24	0.26	2.77	0.16	0.01	0.02	-5.12
2911	Manufacture of engines and turbines	Manufacturing	0.18	0.03	0.00	0.07	0.00	0.02	-6.53
2912	Manufacture of pumps, compressors, taps and valves	Manufacturing	0.04	0.07	0.00	0.00	0.00	0.00	2.97
2913	Manufacture of bearings, gears, gearing and driving elements	Manufacturing	0.00	0.01	0.00	NA	0.00	0.00	-1.73

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
2914	Manufacture of ovens, furnaces and furnace burners	Manufacturing	0.00	0.00	0.03	0.00	0.00	0.00	-3.19
2915	Manufacture of lifting and handling equipment	Manufacturing	0.02	0.40	0.01	0.00	0.00	0.01	3.41
2919	Manufacture of other general-purpose machinery	Manufacturing	0.06	0.04	0.00	0.02	0.00	0.00	-3.43
2921	Manufacture of agricultural and forestry machinery	Manufacturing	0.03	0.02	0.42	0.01	0.01	0.00	-6.16
2922	Manufacture of machine tools	Manufacturing	0.07	0.07	0.09	0.03	0.00	0.00	-5.76
2923	Manufacture of machinery for metallurgy	Manufacturing	0.00	0.00	0.01	0.00	0.00	0.00	0.06
2924	Manufacture of machinery for mining, quarrying and construction	Manufacturing	0.08	0.64	0.04	0.00	0.00	0.00	-4.10
2925	Manufacture of machinery for food, beverage and tobacco processing	Manufacturing	0.04	0.05	0.29	0.06	0.01	0.00	-5.26
2926	Manufacture of machinery for textile, apparel and leather production	Manufacturing	0.01	0.00	0.00	0.00	0.00	0.00	-2.42
2927	Manufacture of weapons and ammunition	Manufacturing	0.00	0.00	0.04	0.00	0.00	0.00	-5.65
2929	Manufacture of other special-purpose machinery	Manufacturing	0.06	0.05	0.01	0.01	0.00	0.00	-4.24
2930	Manufacture of domestic appliances n.e.c.	Manufacturing	0.07	0.03	0.00	0.01	0.00	0.00	-5.02
3000	Manufacture of office, accounting and computing machinery	Manufacturing	0.03	0.02	0.00	0.00	0.00	0.00	-0.25
3110	Manufacture of electric motors, generators and transformers	Manufacturing	0.01	0.05	0.00	0.02	0.00	0.00	-0.42
3120	Manufacture of electricity distribution and control apparatus	Manufacturing	0.03	0.10	0.00	NA	0.00	0.00	NA
3130	Manufacture of insulated wire and cable	Manufacturing	0.00	0.01	0.09	0.02	0.00	0.00	-4.00
3140	Manufacture of accumulators, primary cells and primary batteries	Manufacturing	0.02	0.05	0.00	NA	0.00	0.00	-1.82
3150	Manufacture of electric lamps and lighting equipment	Manufacturing	0.02	0.01	0.00	0.04	0.00	0.00	-3.29
3190	Manufacture of other electrical equipment n.e.c.	Manufacturing	0.01	0.02	0.00	0.12	0.01	0.01	-3.93
3210	Manufacture of electronic valves and tubes and other electronic components	Manufacturing	0.00	0.00	0.00	NA	0.00	0.00	NA
3220	Manufacture of television, telephone and radio transmitters	Manufacturing	0.01	0.07	0.00	0.01	0.00	0.02	-7.71
3230	Manufacture of television and radio receivers	Manufacturing	0.05	0.06	0.06	0.03	0.01	0.03	-7.35
3311	Manufacture of medical and surgical equipment	Manufacturing	0.02	0.14	0.01	NA	0.00	0.00	NA
3312	Manufacture of instruments and appliances for measuring, checking, testing, navigating	Manufacturing	0.03	0.09	0.00	NA	0.00	0.00	NA
3313	Manufacture of industrial process control equipment	Manufacturing	0.00	0.00	0.00	NA	0.00	0.00	NA
3320	Manufacture of optical instruments and photographic equipment	Manufacturing	0.00	0.01	0.00	NA	0.00	0.00	NA
3330	Manufacture of watches and clocks	Manufacturing	0.03	0.00	0.00	0.01	0.00	0.01	-5.28
3410	Manufacture of motor vehicles	Manufacturing	0.09	0.06	0.00	0.00	0.02	0.00	-4.04
3420	Manufacture of bodies (coachwork) for motor vehicles	Manufacturing	0.10	0.07	0.31	0.01	0.01	0.00	-4.70
3430	Manufacture of parts and accessories for motor vehicles	Manufacturing	0.03	0.04	0.05	NA	0.00	0.00	NA
3511	Building and repairing of ships	Manufacturing	0.08	4.25	0.06	0.00	0.00	0.01	-6.29
3512	Building and repairing of pleasure and sporting boats	Manufacturing	0.00	0.00	0.00	NA	0.00	0.00	-5.07

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
3520	Manufacture of railway and tramway locomotives and rolling stock	Manufacturing	0.00	0.00	0.00	NA	0.00	0.00	NA
3530	Manufacture of aircraft and spacecraft	Manufacturing	0.20	0.07	0.00	NA	0.00	0.00	NA
3591	Manufacture of motorcycles	Manufacturing	0.10	0.01	0.00	NA	0.03	0.00	-2.16
3592	Manufacture of bicycles and invalid carriages	Manufacturing	0.03	0.01	0.01	NA	0.02	0.00	-1.51
3599	Manufacture of other transport equipment n.e.c.	Manufacturing	0.01	0.00	0.05	0.15	0.00	0.12	-7.80
3610	Manufacture of furniture	Manufacturing	0.21	0.03	12.75	0.00	0.14	0.01	-5.03
3691	Manufacture of jewellery and related articles	Manufacturing	0.75	0.03	0.13	NA	0.03	0.00	NA
3692	Manufacture of musical instruments	Manufacturing	0.03	0.00	0.03	0.01	0.00	0.01	-5.16
3693	Manufacture of sports goods	Manufacturing	0.00	0.00	0.00	NA	0.00	0.00	NA
3694	Manufacture of games and toys	Manufacturing	0.00	0.00	0.02	NA	0.00	0.00	4.02
3699	Other manufacturing n.e.c.	Manufacturing	0.03	0.01	0.48	NA	0.07	0.00	NA
3710	Recycling of metal waste and scrap	Manufacturing	0.00	0.00	0.00	0.02	0.00	0.00	-3.15
4010	Production, transmission and distribution of electricity	Electricity	1.02	0.00	0.00	0.47	0.03	0.00	-7.04
4020	Manufacture of gas; distribution of gaseous fuels through mains	Electricity	0.00	0.00	0.00	0.06	0.03	0.00	-5.45
4100	Collection, purification and distribution of water	Electricity	0.00	0.00	0.00	0.74	0.01	0.00	-6.68
4510	Site preparation	Construction	0.00	0.00	0.00	0.03	0.01	0.00	-5.75
4520	Building of complete constructions or parts thereof; civil engineering	Construction	0.00	0.00	0.00	5.38	0.41	0.62	-5.38
4530	Building installation	Construction	0.00	0.00	0.00	0.29	0.06	0.18	-6.31
4540	Building completion	Construction	0.00	0.00	0.00	NA	0.06	0.00	NA
5010	Sale of motor vehicles	Retail	0.00	0.00	0.00	0.41	0.01	0.02	-1.77
5020	Maintenance and repair of motor vehicles	Retail	0.00	0.00	0.00	0.33	0.23	0.26	-6.53
5030	Sale of motor vehicle parts and accessories	Retail	0.00	0.00	0.00	0.14	0.03	0.05	-5.73
5040	Sale, maintenance and repair of motorcycles and related parts and accessories	Retail	0.00	0.00	0.00	0.03	0.04	0.02	-8.38
5050	Retail sale of automotive fuel	Retail	0.00	0.00	0.00	1.30	0.00	0.03	-7.84
5110	Wholesale on a fee or contract basis	Retail	0.00	0.00	0.00	0.52	0.01	0.01	-6.23
5121	Wholesale of agricultural raw materials and live animals	Retail	0.00	0.00	0.00	0.48	0.22	0.04	-7.58
5122	Wholesale of food, beverages and tobacco	Retail	0.00	0.00	0.00	0.54	0.09	0.04	-6.64
5131	Wholesale of textiles, clothing and footwear	Retail	0.00	0.00	0.00	0.03	0.02	0.01	-5.12
5139	Wholesale of other household goods	Retail	0.00	0.00	0.00	1.87	0.00	0.03	-3.62
5141	Wholesale of solid, liquid and gaseous fuels and related products	Retail	0.00	0.00	0.00	0.12	0.01	0.01	-6.55
5142	Wholesale of metals and metal ores	Retail	0.00	0.00	0.00	0.02	0.00	0.00	-6.72
5143	Wholesale of construction materials	Retail	0.00	0.00	0.00	NA	0.00	0.00	NA
5149	Wholesale of other intermediate products, waste and scrap	Retail	0.00	0.00	0.00	0.03	0.01	0.02	-5.96
5151	Wholesale of computers, computer peripheral equipment and software	Retail	0.00	0.00	0.00	0.13	0.00	0.00	-2.86

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
5152	Wholesale of electronic and telecommunications parts and equipment	Retail	0.00	0.00	0.00	0.04	0.01	0.00	-5.32
5159	Wholesale of other machinery, equipment and supplies	Retail	0.00	0.00	0.00	0.22	0.00	0.00	-5.39
5190	Other wholesale	Retail	0.00	0.00	0.00	0.28	0.01	0.01	-3.25
5211	Retail sale in non-specialized stores with food, beverages or tobacco predominating	Retail	0.00	0.00	0.00	1.02	0.56	0.41	-8.55
5219	Other retail sale in non-specialized stores	Retail	0.00	0.00	0.00	0.72	0.20	0.31	-7.81
5220	Retail sale of food, beverages and tobacco in specialized stores	Retail	0.00	0.00	0.00	0.38	1.03	1.21	-7.92
0523	Other retail trade of new goods in specialized stores	Retail	0.00	0.00	0.00	NA	0.01	0.00	NA
5231	Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles	Retail	0.00	0.00	0.00	0.59	0.12	0.09	-7.73
5232	Retail sale of textiles, clothing, footwear and leather goods	Retail	0.00	0.00	0.00	0.33	0.32	0.19	-6.55
5233	Retail sale of household appliances, articles and equipment	Retail	0.00	0.00	0.00	0.59	0.05	0.13	-6.80
5234	Retail sale of hardware, paints and glass	Retail	0.00	0.00	0.00	0.91	0.06	0.06	-6.72
5239	Other retail sale in specialized stores	Retail	0.00	0.00	0.00	2.30	0.15	0.20	-6.64
5240	Retail sale of second-hand goods in stores	Retail	0.00	0.00	0.00	0.01	0.08	0.22	-6.10
5251	Retail sale via mail order houses	Retail	0.00	0.00	0.00	0.00	0.00	0.01	-4.85
5252	Retail sale via stalls and markets	Retail	0.00	0.00	0.00	0.24	2.84	1.87	-7.09
5259	Other non-store retail sale	Retail	0.00	0.00	0.00	0.11	0.17	0.84	-7.68
5260	Repair of personal and household goods	Retail	0.00	0.00	0.00	0.15	0.19	0.13	-7.93
5510	Hotels; camping sites and other provision of short-stay accommodation	Hospitality	0.00	0.00	0.00	1.97	0.14	0.02	-5.98
5520	Restaurants, bars and canteens	Hospitality	0.00	0.00	0.00	1.31	0.75	1.14	-8.45
6010	Transport via railways	Transport	0.00	0.00	0.00	0.13	0.00	0.00	-3.04
6021	Other scheduled passenger land transport	Transport	0.00	0.00	0.00	1.20	0.06	0.26	-6.56
6022	Other non-scheduled passenger land transport	Transport	0.00	0.00	0.00	NA	0.11	0.00	NA
6023	Freight transport by road	Transport	0.00	0.00	0.00	0.16	0.00	0.06	-4.76
6030	Transport via pipelines	Transport	0.00	0.00	0.00	0.00	0.00	0.00	-2.04
6110	Sea and coastal water transport	Transport	0.00	0.00	0.00	0.16	0.00	0.00	1.47
6120	Inland water transport	Transport	0.00	0.00	0.00	0.05	0.09	0.00	-4.07
6210	Scheduled air transport	Transport	0.00	0.00	0.00	0.13	0.00	0.00	-1.85
6301	Cargo handling	Transport	0.00	0.00	0.00	0.22	0.01	0.01	-5.15
6302	Storage and warehousing	Transport	0.00	0.00	0.00	0.14	0.00	0.00	-6.64
6303	Other supporting transport activities	Transport	0.00	0.00	0.00	0.33	0.07	0.04	-1.41
6304	Activities of travel agencies and tour operators	Transport	0.00	0.00	0.00	0.25	0.01	0.00	-5.06
6309	Activities of other transport agencies	Transport	0.00	0.00	0.00	1.52	0.01	0.02	-4.87
6411	National post activities	Transport	0.00	0.00	0.00	0.54	0.00	0.01	-6.23
6412	Courier activities other than national post activities	Transport	0.00	0.00	0.00	0.03	0.00	0.00	-4.42
6420	Telecommunications	Transport	0.00	0.00	0.00	0.77	0.07	0.01	-3.52
6511	Central banking	Finance	0.00	0.00	0.00	0.04	0.00	0.00	-5.74

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
6519	Other monetary intermediation	Finance	0.00	0.00	0.00	4.82	0.00	0.00	-6.66
6591	Financial leasing	Finance	0.00	0.00	0.00	0.05	0.00	0.00	-5.50
6592	Other credit granting	Finance	0.00	0.00	0.00	0.60	0.00	0.00	-6.51
6599	Other financial intermediation n.e.c.	Finance	0.00	0.00	0.00	0.46	0.00	0.00	-5.60
6601	Life insurance	Finance	0.00	0.00	0.00	0.31	0.00	0.00	-5.33
6602	Pension funding	Finance	0.00	0.00	0.00	0.03	0.00	0.00	-5.91
6603	Non-life insurance	Finance	0.00	0.00	0.00	0.59	0.00	0.00	-7.06
6711	Administration of financial markets	Finance	0.00	0.00	0.00	0.01	0.00	0.00	-3.95
6712	Security dealing activities	Finance	0.00	0.00	0.00	0.39	0.00	0.00	-2.75
6719	Activities auxiliary to financial intermediation n.e.c.	Finance	0.00	0.00	0.00	NA	0.01	0.00	NA
6720	Activities auxiliary to insurance and pension funding	Finance	0.00	0.00	0.00	0.03	0.00	0.00	-5.05
7010	Real estate activities with own or leased property	Real estate	0.00	0.00	0.00	0.49	0.01	0.00	-3.72
7020	Real estate activities on a fee or contract basis	Real estate	0.00	0.00	0.00	0.13	0.00	0.00	-4.68
7111	Renting of land transport equipment	Real estate	0.00	0.00	0.00	0.13	0.00	0.00	-5.47
7121	Renting of agricultural machinery and equipment	Real estate	0.00	0.00	0.00	NA	0.01	0.00	NA
7130	Renting of personal and household goods n.e.c.	Real estate	0.00	0.00	0.00	0.03	0.00	0.01	-7.01
7210	Hardware consultancy	Real estate	0.00	0.00	0.00	0.14	0.00	0.00	-3.48
7221	Software publishing	Real estate	0.00	0.00	0.00	0.02	0.00	0.00	-4.06
7229	Other software consultancy and supply	Real estate	0.00	0.00	0.00	0.10	0.00	0.00	-4.47
7230	Data processing	Real estate	0.00	0.00	0.00	0.01	0.01	0.00	-5.86
7240	Database activities and online distribution of electronic content	Real estate	0.00	0.00	0.00	0.52	0.00	0.00	-7.16
7250	Maintenance and repair of office, accounting and computing machinery	Real estate	0.00	0.00	0.00	0.14	0.00	0.01	-4.36
7290	Other computer-related activities	Real estate	0.00	0.00	0.00	0.08	0.00	0.00	-4.59
7310	Research and experimental development on natural sciences and engineering (NSE)	Real estate	0.00	0.00	0.00	0.41	0.00	0.00	-5.87
7411	Legal activities	Real estate	0.00	0.00	0.00	0.30	0.00	0.00	-6.15
7412	Accounting, bookkeeping and auditing activities; tax consultancy	Real estate	0.00	0.00	0.00	0.23	0.00	0.00	-6.09
7413	Market research and public opinion polling	Real estate	0.00	0.00	0.00	0.17	0.00	0.00	-0.48
7414	Business and management consultancy activities	Real estate	0.00	0.00	0.00	1.73	0.00	0.01	-6.79
0742	Architectural, engineering and other technical activities	Real estate	0.00	0.00	0.00	NA	0.00	0.00	NA
7421	Architectural and engineering activities and related technical consultancy	Real estate	0.00	0.00	0.00	0.83	0.03	0.15	-4.49
7422	Technical testing and analysis	Real estate	0.00	0.00	0.00	0.15	0.00	0.00	-5.40
7430	Advertising	Real estate	0.00	0.00	0.00	0.32	0.00	0.00	1.37
7491	Labour recruitment and provision of personnel	Real estate	0.00	0.00	0.00	0.54	0.00	0.00	-4.64
7492	Investigation and security activities	Real estate	0.00	0.00	0.00	3.89	0.00	0.00	-3.34
7493	Building-cleaning and industrial-cleaning activities	Real estate	0.00	0.00	0.00	0.14	0.00	0.00	-4.39

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
7494	Photographic activities	Real estate	0.00	0.00	0.00	0.05	0.03	0.03	-7.26
7495	Packaging activities	Real estate	0.00	0.00	0.00	0.12	0.00	0.00	-2.07
7499	Other business activities n.e.c.	Real estate	0.00	0.00	0.00	0.54	0.03	0.01	-6.58
7511	General (overall) public service activities	Administration	0.00	0.00	0.00	2.90	0.00	0.00	-8.04
7512	Regulation of the activities of agencies that provide health care, education, cultural services	Administration	0.00	0.00	0.00	2.18	0.00	0.00	-7.82
7513	Regulation of and contribution to more efficient operation of business	Administration	0.00	0.00	0.00	0.34	0.00	0.00	-7.16
7514	Supporting service activities for the government as a whole	Administration	0.00	0.00	0.00	0.06	0.00	0.01	-6.69
7521	Foreign affairs	Administration	0.00	0.00	0.00	1.19	0.00	0.00	-6.72
7522	Defence activities	Administration	0.00	0.00	0.00	0.81	0.00	0.00	-4.68
7523	Public order and safety activities	Administration	0.00	0.00	0.00	0.90	0.00	0.00	-7.77
7530	Compulsory social security activities	Administration	0.00	0.00	0.00	0.04	0.00	0.00	-5.87
8010	Primary education	Education	0.00	0.00	0.00	8.05	0.02	0.09	-8.73
8021	General secondary education	Education	0.00	0.00	0.00	2.23	0.00	0.00	-8.50
8022	Technical and vocational secondary education	Education	0.00	0.00	0.00	0.49	0.00	0.00	-7.39
8030	Higher education	Education	0.00	0.00	0.00	1.57	0.00	0.00	-6.50
8090	Other education	Education	0.00	0.00	0.00	0.43	0.01	0.00	-7.00
8511	Hospital activities	Social work	0.00	0.00	0.00	4.80	0.00	0.00	-8.12
8512	Medical and dental practice activities	Social work	0.00	0.00	0.00	0.33	0.03	0.00	-7.38
8519	Other human health activities	Social work	0.00	0.00	0.00	0.57	0.07	0.04	-7.82
8520	Veterinary activities	Social work	0.00	0.00	0.00	0.02	0.00	0.00	-6.52
8531	Social work activities with accommodation	Social work	0.00	0.00	0.00	0.06	0.00	0.00	-6.98
8532	Social work activities without accommodation	Social work	0.00	0.00	0.00	0.03	0.03	0.00	-5.58
9000	Sewage and refuse disposal, sanitation and similar activities	Other	0.00	0.00	0.00	0.92	0.02	0.00	-6.34
0911	Activities of business, employers and professional organizations	Other	0.00	0.00	0.00	NA	0.00	0.00	NA
9111	Activities of business and employers organizations	Other	0.00	0.00	0.00	0.30	0.01	0.00	-5.97
9112	Activities of professional organizations	Other	0.00	0.00	0.00	0.11	0.00	0.00	-6.28
9120	Activities of trade unions	Other	0.00	0.00	0.00	0.11	0.00	0.00	-6.85
9191	Activities of religious organizations	Other	0.00	0.00	0.00	4.02	0.01	0.02	-8.78
9192	Activities of political organizations	Other	0.00	0.00	0.00	0.03	0.00	0.00	-7.20
9211	Motion picture and video production and distribution	Other	0.00	0.00	0.00	0.04	0.01	0.00	-6.36
9212	Motion picture projection	Other	0.00	0.00	0.00	0.00	0.01	0.00	-6.75
9213	Radio and television activities	Other	0.00	0.00	0.00	NA	0.00	0.00	NA
9214	Dramatic arts, music and other arts activities	Other	0.01	0.01	0.00	0.09	0.02	0.05	-6.49
9219	Other entertainment activities n.e.c.	Other	0.00	0.00	0.00	0.07	0.02	0.02	-7.22
9220	News agency activities	Other	0.00	0.00	0.00	0.04	0.00	0.00	-5.53
9232	Museums activities and preservation of historic sites and buildings	Other	0.00	0.00	0.00	0.00	0.00	0.00	-5.20

Table A.3: Final dataset *(continued)*

Code	Label (4-digit)	ISIC group	Trade 2003	Trade 2013	Formal 2003	Formal 2014	Household 2005	Household 2013	Log fitness
9233	Botanical and zoological gardens and nature reserves activities	Other	0.00	0.00	0.00	0.02	0.00	0.00	-6.10
9241	Sporting activities	Other	0.00	0.00	0.00	0.10	0.00	0.00	-5.70
9249	Other recreational activities	Other	0.00	0.00	0.00	0.18	0.01	0.02	-7.14
9301	Washing and (dry-)cleaning of textile and fur products	Other	0.00	0.00	0.00	0.15	0.01	0.01	-4.55
9302	Hairdressing and other beauty treatment	Other	0.00	0.00	0.00	0.08	0.56	0.65	-7.57
9303	Funeral and related activities	Other	0.00	0.00	0.00	0.02	0.01	0.01	-6.99
9309	Other service activities n.e.c.	Other	0.00	0.00	0.00	0.01	0.09	0.05	-6.33
9600	Undifferentiated goods-producing activities of private households for own use	Other	0.00	0.00	0.00	NA	0.01	0.00	NA
9500	Unspecified services	Services	0.00	0.00	0.00	0.03	0.00	0.02	-3.36
9501	Travels and tourism	Services	13.96	5.18	0.00	NA	0.00	0.00	-5.06
9502	Transport	Services	4.19	4.23	0.00	NA	0.00	0.00	NA
9503	ICT	Services	2.85	5.27	0.00	NA	0.00	0.00	-2.57
9504	Financial	Services	0.23	0.22	0.00	NA	0.00	0.00	-4.59
9700	Undifferentiated service-producing activities of private households for own use	Services	0.00	0.00	0.00	NA	0.00	0.00	NA
9900	Extraterritorial organizations and bodies	Services	0.00	0.00	0.00	0.17	0.00	0.00	-3.75
9999	Goods and services not elsewhere classified	N/A	1.28	0.94	0.00	NA	0.00	0.00	NA

Appendix B

Appendix to Chapter 3

B.1 Network representation

The adjacency matrix $\mathbf{II^{CL}}$ can be represented as a network of co-located industries, where nodes are industries and the weighted edges represent the probability for industries to co-locate (Figure B.1). In the figure, less relevant edges have been removed ($\phi < 0.5$, where ϕ is defined by Equation (B.1)) in order to provide a better visualisation given the high degree of connectedness of the network, and the high number of weak connections. The productive structure can be further investigated by looking at the 'communities' of industrial sectors – clusters of industries that tend to be more linked to each other and less with other communities. For this purpose, a divisive method based on edge betweenness is used. Edge betweenness quantifies the likelihood that a node will be on the shortest path between other nodes. Edge betweenness community detection the graph to be divided in such a way that the modularity – degree of connection – is maximised within clusters/communities, making it possible to analyse the group structure of the network (Newman and Girvan, 2004).

We now move on to a comparison of how different measures of specialisation affect the structure of the resulting adjacency matrices for different thresholds. The table

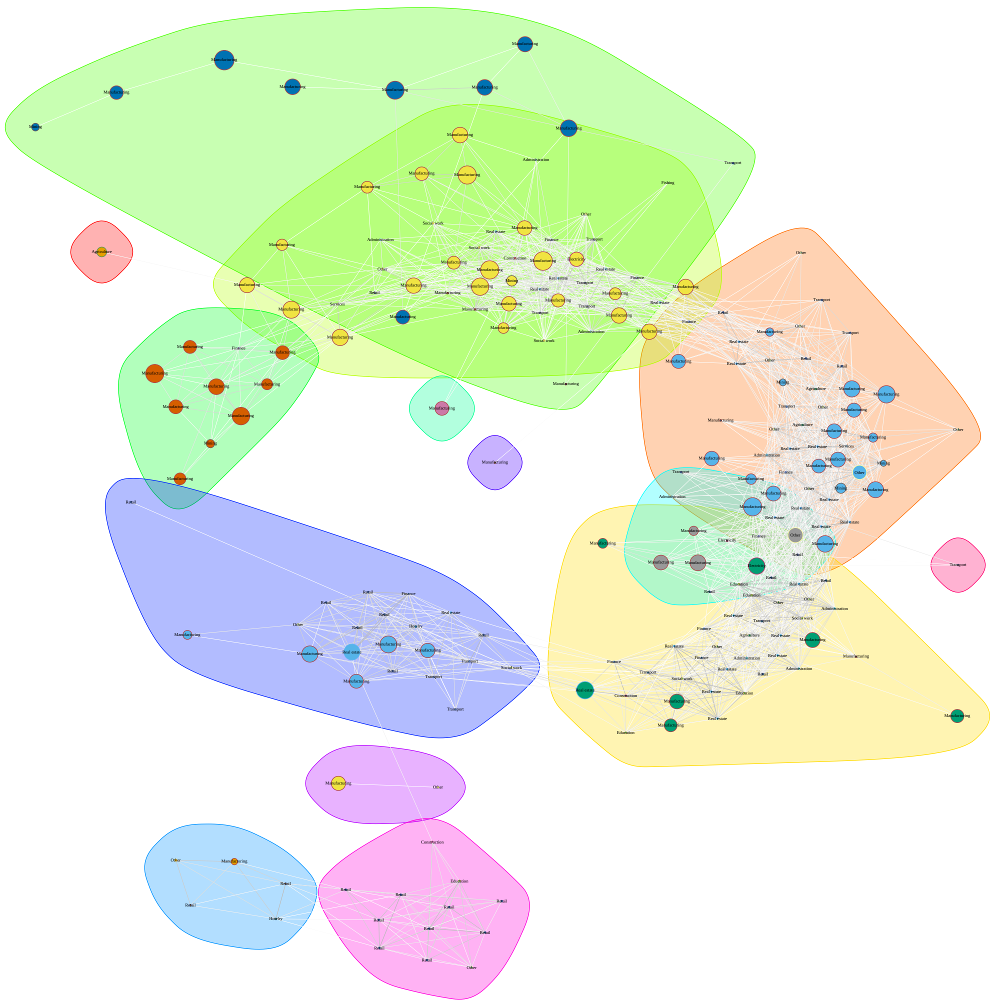

Figure B.1: Visual representation of the aggregate productive structure of Ghana based on domestic employment, with community structure.

Table B.1: Comparison of matrix densities

Industry	Spec. Index > 1	Spec. Index > 2	Spec. Index > 3
All	0.86	0.51	0.24
Formal	0.75	0.37	0.19
Informal	0.79	0.51	0.3

below shows that, for higher thresholds of the specialisation index, the production structure of Ghana appears to be less and less connected at all levels.

As not all connections are equally weighted – some industries being more strongly connected than others – the adjacency matrices require a more in-depth description in terms of the distribution of the nodes' degree centrality (the number of industries to which each industry is connected) with additional information on the strength of their connections, measured by the weight of each individual edge. These two measures can be combined in a unique measure – the node's strength – which can be calculated as the sum of all the weights of the edges held by the industry/node.

Figure B.2 uses a non-parametric density function to show the distribution of degree and strength of the nodes, and the distribution of the weight of the edges. For all levels of activity (formal, informal and aggregate), increasing the threshold of the specialisation index from 1 to 3 progressively transforms the degree distribution of the networks from left to right skewed, indicating that if we consider only activities in which districts are highly specialised, the productive structure becomes less and less connected, with only a few industries/nodes having a large number of connections to other nodes. This pattern becomes even more evident if we take into account the strength rather than the unweighted degree of the nodes: the former suggests that the productive structures of Ghana are similar to scale-free networks, a property often observed in real-world networks which have a large number of weakly connected nodes, and a small group of nodes working as hubs (Barabási and Albert, 1999). However, this property remains to be tested, and the theoretical considerations on the emergence of the scale free property fall beyond the scope of the analysis in this chapter.

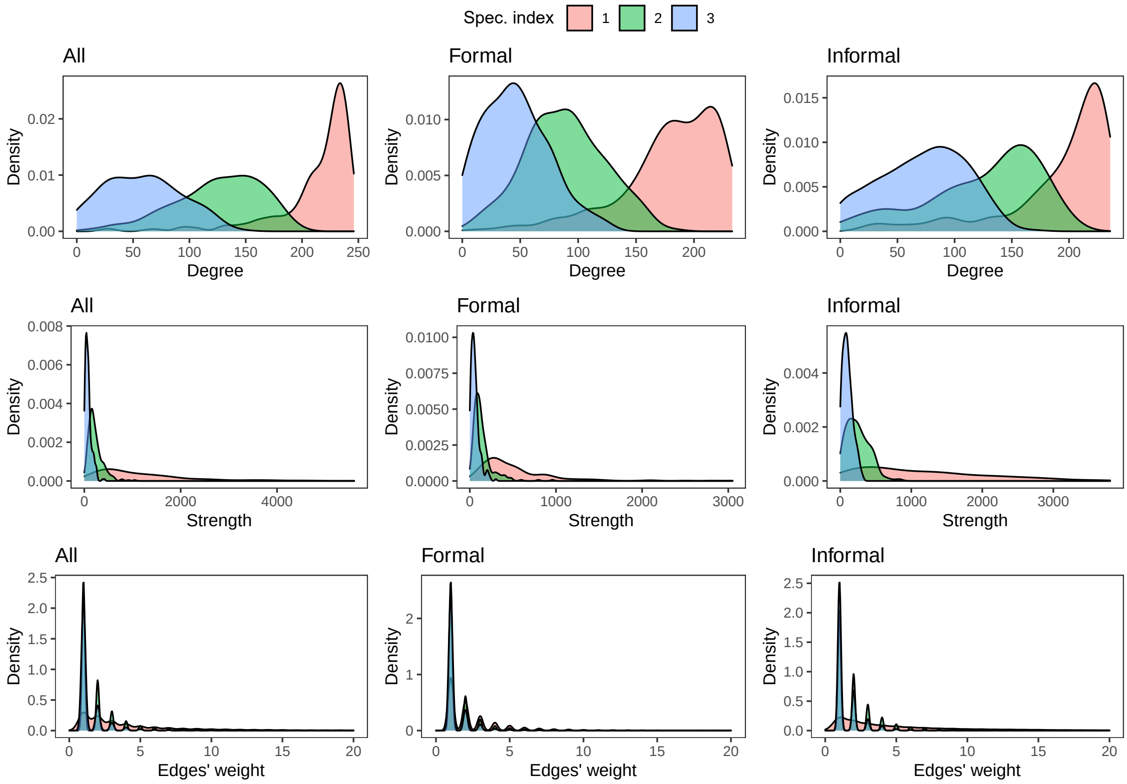

Figure B.2: Description of the adjacency matrices. The plots describe the overall (first column), formal (second column) and informal (third column) adjacency matrices, for different cutoffs of the specialisation index, as indicated by the legend. The measures used are node degree (number of edges), node strength (degree weighted by node's edges weights) and edges' weight. Each node represents and industry, while edges measure the absolute frequency of co-location between to industries.

B.1.1 Randomisation tests

In order to assess whether the adjacency matrix used to reproduce the Ghanaian industry space is capturing a non-random structure, the analysis relies on graph randomisation tests. Following Newman (2018), randomisation tests are performed by comparing the characteristics of the formal and informal productive structures to a null model using a degree-preserving randomisation of the adjacency matrices. Unlike the majority of random graphs, degree-preserving randomisation allows us to keep the degree distribution of a given node which would otherwise follow a Poisson distribution – quite different from the degree distributions observed in Figure B.2. This feature is particularly relevant to test graphs with extremely skewed degree distributions and with high density, as is the case for the Ghanaian industry space. One thousand random graphs were estimated by reshuffling the edges among the nodes with constant density, and subsequently the empirical measures of transitivity and average path length were compared with the distribution of the same measures in the random models, in such a way as to determine whether the observed graph was comparable to the random ones.

Figure B.3 shows that the high density of the Ghanaian productive structures is reflected in the empirical measures of average path length and transitivity across formal, informal, and all activities. The first is very close to one, meaning that the average distance between all pairs of nodes is slightly greater than one node. The second indicates that most nodes participate in closed triangles, with a value close to 1. In the randomization test, these two empirical measures were compared with a distribution of the same measures computed using the 1000 random graphs generated by the degree-preserving randomisation. As shown in Figure B.3, the empirical measures fall well beyond the distributions obtained by the random graphs, allowing us to exclude the hypothesis that the formal, informal, and aggregate productive structures of Ghana follow a random pattern of co-location.

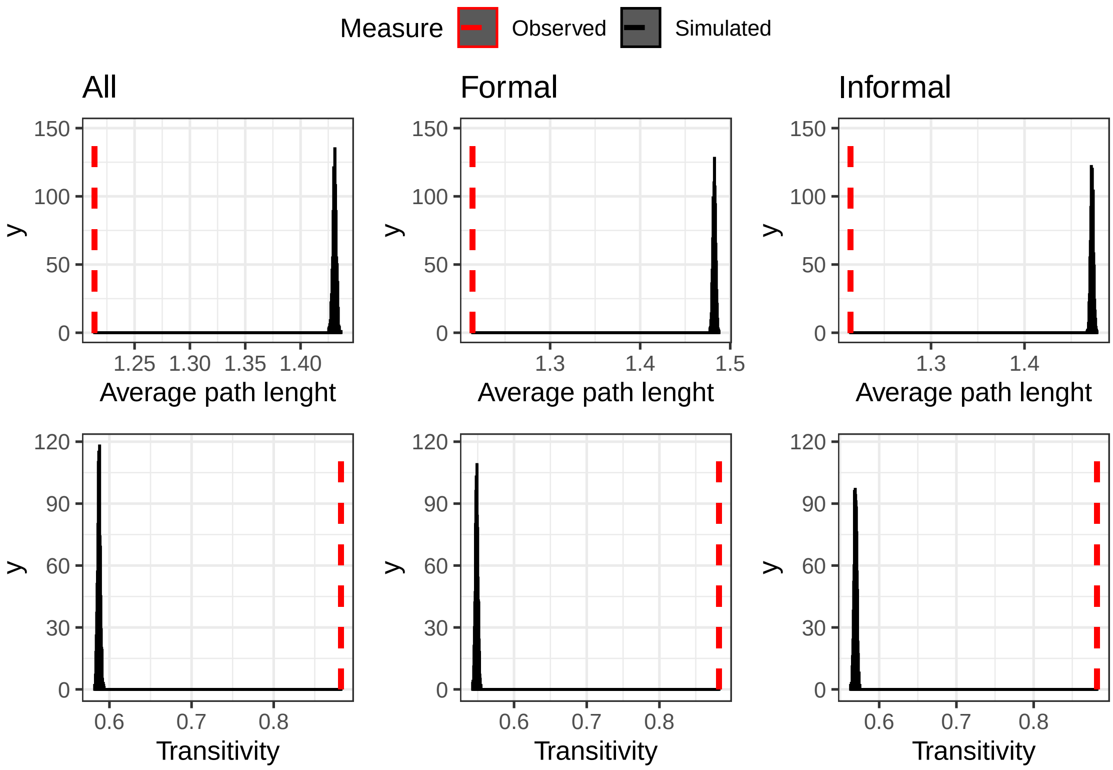

Figure B.3: Results of the randomisation tests. The plots report the distribution of two measures obtained from simulating 1000 random graphs using a degree-preserving randomisation technique based on the overall (first column), formal (second column) and informal (third column) incidence matrices. The black histograms indicate the simulated distributions of the measures indicated in the X axes. The red line pins down the empirical (observed) measure Such measures are average path length (the average geodesic distance between all possible pairs of nodes in the network, first row) and transitivity (a measure of the nodes' tendency to cluster together, second row).

B.2 Comparing standardisation measures

A comparison between two different measures is proposed in order to validate the choice of cosine similarity to standardise the co-location matrices. Below, the cosine similarity measure is compared to a measure of proximity based on the concept of minimum conditional probability as proposed by Hidalgo et al. (2007). The proximity ϕ_{ij} can be computed as the minimum conditional probability that districts will be specialised at the same time in industries i and j at a given specialisation index cutoff:

$$\phi_{ij} = min\{P(x_i|x_j); P(x_j|x_i)\} \tag{B.1}$$

where x_i and x_j take value 1 if a district is specialised in industry i or j and 0 otherwise.

Figure B.4 compares the proximity-based co-location and the cosine-based colocation across all the incidence matrices (aggregate, formal and informal); the measures of co-location computed with the two methods are strongly correlated, although the correlation looks like an heteroskedastic one, with higher degrees of variation for higher values of proximity and cosine similarity.

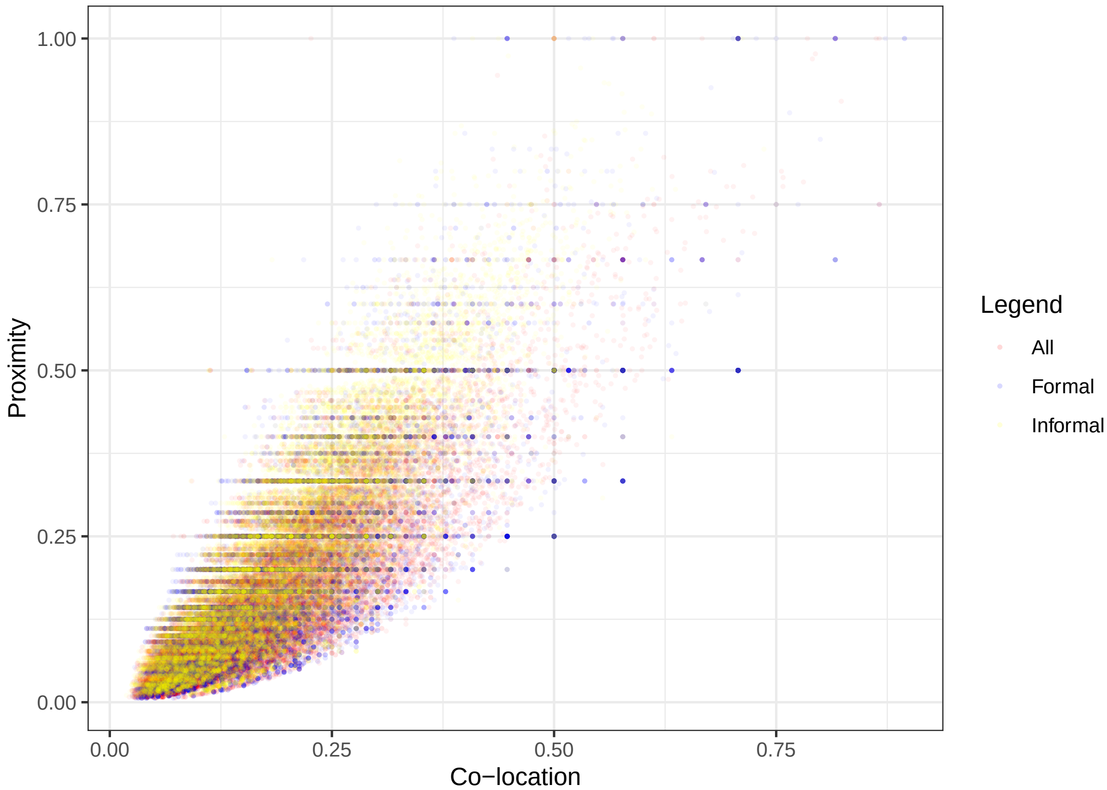

Figure B.4: Correlation between proximity and co-location. The plot shows the correlation between the cosine-transformed industrial co-location measure used in this study, and the proximity measure proposed by Hidalgo et al. (2007), where each dot represents a pair of industries. In both cases, the IBES data is used to compute the measures. The correlation is shown for the overall (red), formal (blue) and informal (yellow) co-location.

B.3 District fitness

Table B.2: Ghanaian districts ordered by fitness

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Tema Metropolis	1	1	2	4.550	109	5.314	110	1.086	101
Kma	2	3	1	3.877	97	-5.860	78	5.347	99
Accra Metropolis	3	2	3	3.565	132	2.554	134	-0.054	130
La Dade Kotopon	4	5	10	2.487	59	-7.661	54	-2.864	70
Sekondi Takoradi	5	8	5	1.703	79	-9.262	62	-1.728	90
Ho Municipal	6	10	13	0.967	57	-9.976	50	-3.025	65
Ledzokuku / Krowor	7	4	24	0.406	69	-7.196	62	-4.175	58
Tamale North Sub Metro	8	16	4	0.233	77	-11.008	55	-0.445	64
Ga West	9	22	14	-0.077	52	-12.524	36	-3.162	65
Wa Municipal	10	7	7	-0.094	67	-8.586	49	-2.419	67
Nsawam Adoagyiri	11	38	11	-0.488	47	-14.352	26	-2.921	52
Sagnerigu	12	15	38	-0.640	50	-10.562	31	-4.878	43
Sunyani Municipal	13	18	12	-0.858	64	-11.673	47	-2.956	74
Ga East	14	9	6	-0.952	69	-9.761	51	-2.175	73
Adenta	15	13	15	-0.989	44	-10.447	32	-3.275	52
Ashaiman	16	27	23	-1.017	50	-13.119	26	-3.983	62
New Juaben Municipal	17	12	16	-1.109	61	-10.164	40	-3.377	61
Bolgatanga Municipal	18	25	18	-1.140	61	-12.927	41	-3.762	57
Keta Municipal	19	114	22	-1.256	35	-17.926	19	-3.969	37
South Dayi	20	123	33	-1.375	26	-18.011	17	-4.700	20
Kpone Katamanso	21	19	9	-1.381	54	-11.697	27	-2.729	53
Krachi East	22	46	37	-1.388	30	-15.109	17	-4.867	26
Ada East	23	56	17	-1.478	36	-15.456	23	-3.590	21
Asokore Mampong Municipal	25	36	19	-1.542	59	-14.196	24	-3.864	61
Ga South	26	21	25	-1.587	53	-11.924	46	-4.207	58
Gonja Central	27	35	179	-1.594	33	-14.154	17	-7.398	21
Ningo Prampram	28	50	42	-1.714	53	-15.185	31	-5.140	48
La Nkwantanang Madina	29	17	20	-1.734	60	-11.306	31	-3.959	61
Akatsi South	30	52	45	-1.895	45	-15.210	21	-5.327	37
Upper Manya	31	130	103	-1.920	32	-18.238	16	-6.389	30
Bawku Municipal	32	26	46	-1.978	43	-13.020	22	-5.328	35
Agona East	33	23	178	-1.998	29	-12.724	18	-7.397	22
Awutu Senya East Municipal	34	64	29	-1.999	58	-16.056	28	-4.493	62
Hohoe Municipal	35	45	43	-2.008	54	-15.082	34	-5.157	40
Techiman Municipal	36	11	39	-2.020	62	-10.013	38	-4.952	59

Table B.2: Ghanaian districts ordered by fitness *(continued)*

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Cape Coast Metro	37	25	30	-2.071	52	-12.927	36	-4.612	69
Lawra	38	119	63	-2.078	27	-17.978	20	-5.769	18
Upper Denkyira East	39	51	61	-2.104	39	-15.197	23	-5.750	37
Denkyembour	40	24	154	-2.117	24	-12.860	13	-6.973	22
Prestea / Huni Valley	41	86	49	-2.121	35	-17.148	15	-5.496	28
Afigya Kwabre	42	109	26	-2.308	39	-17.797	14	-4.275	35
Pru	43	71	70	-2.317	42	-16.568	15	-5.817	38
Wassa Amenfi Central	44	60	139	-2.318	27	-15.649	13	-6.767	25
Shama	45	61	95	-2.320	39	-15.677	18	-6.275	35
Nkwanta South	46	39	57	-2.339	37	-14.739	26	-5.684	23
Yilo Krobo	47	30	40	-2.368	48	-13.658	25	-4.995	47
Berekum	48	6	47	-2.392	48	-8.456	23	-5.369	54
East Gonja	49	203	28	-2.490	31	-21.657	6	-4.346	30
Ahanta West	50	57	8	-2.520	38	-15.498	20	-2.525	40
Ketu South	51	37	62	-2.561	44	-14.305	29	-5.760	42
Ga Central Municipal	52	31	32	-2.617	45	-13.719	20	-4.696	47
Tarkwa Nsuaem	53	55	60	-2.646	25	-15.448	29	-5.737	21
Ajumako-Enyan-Esiam	54	120	44	-2.648	30	-17.980	14	-5.270	29
Bawku West	54	106	56	-2.648	33	-17.757	22	-5.599	26
Bongo	55	69	106	-2.690	32	-16.543	14	-6.442	23
Builsa North	56	87	109	-2.696	31	-17.156	16	-6.458	26
Kadjebi	57	121	54	-2.698	34	-17.993	18	-5.542	33
Asante Akim Central Municipal	58	28	69	-2.731	46	-13.620	29	-5.798	43
Akwapem South	59	20	110	-2.759	30	-11.886	18	-6.467	31
Agona West	60	63	41	-2.766	54	-15.845	27	-5.046	60
Ejisu Juaben	61	33	48	-2.778	39	-14.040	24	-5.479	41
South Tongu	62	107	65	-2.787	36	-17.792	17	-5.778	30
Akwapem North	63	48	127	-2.811	28	-15.169	20	-6.656	29
Yendi Municipal	64	74	83	-2.825	41	-16.814	22	-6.118	32
Gomoa East	65	34	81	-2.842	38	-14.107	22	-6.105	33
Nkoranza South	66	113	80	-2.876	41	-17.882	20	-6.096	33
Lower Manya	67	96	51	-2.881	39	-17.445	20	-5.506	51
Kwahu West	68	70	35	-2.898	46	-16.554	27	-4.790	48
Birim Municipal	69	29	52	-2.904	43	-13.635	32	-5.520	34
Sissala West	70	85	104	-2.910	35	-17.139	25	-6.432	24
Nzema East	71	47	74	-2.945	32	-15.163	22	-5.956	30
Sefwi Akontombra	72	59	75	-2.959	23	-15.604	17	-5.958	32
Suhum Municipal	73	72	31	-2.961	41	-16.570	23	-4.617	39
Kwabre	74	176	59	-2.971	38	-19.594	14	-5.734	38

Table B.2: Ghanaian districts ordered by fitness *(continued)*

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Jomoro	75	42	84	-2.977	39	-14.984	23	-6.128	35
West Mamprusi	76	94	114	-2.981	33	-17.328	21	-6.499	22
Ejura Sekye Dumase	77	149	27	-2.988	32	-18.804	15	-4.307	28
East Akim	78	141	69	-3.022	30	-18.552	11	-5.798	43
Asutifi North	79	65	97	-3.040	37	-16.066	21	-6.307	38
Asuogyaman	80	53	36	-3.043	22	-15.350	16	-4.806	48
Awutu Senya	81	79	98	-3.046	33	-17.029	14	-6.330	30
Atebubu Amantin	82	95	78	-3.077	38	-17.365	18	-6.043	37
Adansi North	83	92	90	-3.116	28	-17.301	14	-6.200	32
Birim North	84	131	68	-3.124	34	-18.278	20	-5.794	27
Ketu North	85	161	55	-3.133	32	-19.150	16	-5.552	29
Mfantsiman	86	90	82	-3.138	35	-17.266	17	-6.106	41
North Tongu	87	76	64	-3.149	34	-16.942	10	-5.776	39
Lambussie Karni	88	83	66	-3.155	34	-17.136	17	-5.781	25
Atwima Nwabiagya	89	77	58	-3.164	40	-16.975	19	-5.730	44
Atiwa	90	135	150	-3.185	28	-18.360	13	-6.930	26
Central Tongu	91	132	161	-3.210	34	-18.306	18	-7.045	28
Tain	92	156	53	-3.230	38	-18.949	13	-5.522	37
Komenda Edna Eguafo / Abirem	93	67	94	-3.232	33	-16.507	19	-6.259	33
Effutu	94	80	67	-3.257	40	-17.052	23	-5.791	40
Kwahu South	95	184	47	-3.263	35	-19.862	14	-5.369	28
Atwima Mponua	96	62	200	-3.275	20	-15.720	15	-7.960	15
Amansie West	97	75	21	-3.331	36	-16.817	20	-3.966	32
Pusiga	98	102	135	-3.335	32	-17.625	10	-6.727	25
Upper West Akim	99	126	88	-3.381	31	-18.110	13	-6.172	30
Mpohor	100	159	118	-3.408	14	-19.066	4	-6.559	22
Jaman South	101	84	77	-3.419	29	-17.138	18	-6.036	31
Asunafo North	102	32	107	-3.444	31	-13.998	19	-6.449	39
Asikuma / Odoben / Brakwa	103	98	85	-3.446	30	-17.476	19	-6.141	34
Dormaa Municipal	104	100	79	-3.461	38	-17.571	21	-6.052	37
Ellembelle	105	124	93	-3.482	26	-18.040	15	-6.249	22
Mampong Municipal	106	88	112	-3.491	43	-17.171	26	-6.484	41
Sissala East	106	164	142	-3.491	32	-19.185	13	-6.826	20
Kintampo North	107	129	73	-3.495	42	-18.218	17	-5.943	42
Adaklu	108	156	157	-3.496	20	-18.949	8	-6.989	14
Kasena Nankana East	109	43	148	-3.503	32	-15.051	24	-6.923	28
Offinso Municipal	110	40	123	-3.505	32	-14.788	17	-6.608	34
Ada West	111	146	86	-3.509	30	-18.631	9	-6.147	26
Assin North	112	44	120	-3.510	35	-15.078	16	-6.573	34

Table B.2: Ghanaian districts ordered by fitness *(continued)*

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Tano North	113	111	87	-3.513	31	-17.837	15	-6.151	31
Biakoye	114	150	113	-3.518	29	-18.812	13	-6.493	28
Jasikan	115	97	89	-3.519	30	-17.459	19	-6.180	21
Aowin	116	110	108	-3.607	28	-17.811	29	-6.453	21
Twifo Ati-Morkwa	117	133	114	-3.609	25	-18.320	12	-6.499	29
Sene West	118	127	156	-3.611	29	-18.124	13	-6.977	31
Suaman	119	142	162	-3.615	27	-18.585	13	-7.061	24
Saboba	119	158	111	-3.615	27	-18.990	15	-6.470	21
Gomoa West	120	78	72	-3.651	35	-16.987	22	-5.927	30
Bosomtwe /Atwima / Kwanwoma	121	93	99	-3.670	31	-17.322	19	-6.340	24
Bole	122	139	122	-3.674	35	-18.480	16	-6.597	28
Kwaebibirem	123	54	105	-3.675	26	-15.443	18	-6.440	27
Sunyani West	124	99	130	-3.685	38	-17.563	18	-6.690	31
Techiman North	125	188	96	-3.696	18	-20.245	8	-6.279	17
Sefwi Bibiani-Anhwiaso Bekwai	126	81	155	-3.713	31	-17.067	27	-6.974	29
Kwahu Afram Plains North	127	186	141	-3.732	30	-20.030	11	-6.816	34
Kpando Municipal	128	66	101	-3.737	36	-16.362	16	-6.383	37
Krachi West	129	187	149	-3.742	31	-20.243	14	-6.926	20
West Gonja	130	103	116	-3.746	34	-17.630	20	-6.537	30
Assin South	131	41	191	-3.760	19	-14.905	16	-7.578	15
Sekyere East	132	201	71	-3.761	29	-21.455	6	-5.908	32
Asante Akim North	133	101	144	-3.769	26	-17.601	12	-6.854	26
Atwima Kwanwoma	134	116	100	-3.774	29	-17.955	12	-6.349	32
Bekwai Municipal	135	73	131	-3.784	32	-16.706	21	-6.698	32
Ahafo Ano North	135	152	160	-3.784	29	-18.824	15	-7.026	29
Amansie Central	136	190	50	-3.791	22	-20.274	12	-5.503	20
Afigya Sekyere	136	193	110	-3.791	33	-20.512	11	-6.467	30
Mamprusi East	137	58	119	-3.794	31	-15.576	28	-6.568	30
Sawla/Tuna/Kalba	138	155	170	-3.800	27	-18.883	7	-7.158	22
Dormaa East	139	183	143	-3.813	33	-19.810	15	-6.844	25
Ayensuano	140	165	124	-3.815	27	-19.320	11	-6.610	25
West Akim	141	148	92	-3.816	38	-18.770	18	-6.238	40
Wenchi	142	115	75	-3.819	29	-17.930	14	-5.958	47
Dormaa West	143	185	110	-3.866	18	-19.940	14	-6.467	19
Akyem Mansa	144	89	180	-3.873	24	-17.222	16	-7.426	20
Nkwanta North	145	171	140	-3.885	35	-19.468	13	-6.782	35
Nandom	146	145	125	-3.887	34	-18.623	16	-6.617	31
Jirapa	147	49	169	-3.904	31	-15.184	16	-7.157	27
Kasena Nankana West	148	82	76	-3.908	26	-17.076	21	-5.982	28

Table B.2: Ghanaian districts ordered by fitness *(continued)*

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Kpandai	149	177	128	-3.912	30	-19.606	13	-6.658	28
Abura / Asebu / Kwamankese	150	138	117	-3.919	28	-18.436	16	-6.542	30
Asante Akim South	151	144	115	-3.926	29	-18.602	18	-6.503	28
Ho West	152	182	91	-3.933	23	-19.782	11	-6.225	27
Ahafo Ano South	153	147	137	-3.956	32	-18.664	17	-6.761	31
Bia West	154	125	189	-3.969	30	-18.101	21	-7.575	23
Builsa South	155	112	184	-3.972	22	-17.872	11	-7.492	18
Juabeso	156	108	163	-3.985	24	-17.795	18	-7.070	18
Garu Tempane	157	179	173	-3.990	25	-19.688	14	-7.272	21
Kwahu East	158	195	134	-3.991	29	-20.669	8	-6.722	27
Zabzugu	159	137	164	-4.024	31	-18.424	13	-7.100	22
Tano South	160	68	145	-4.031	34	-16.522	19	-6.857	34
Nadowli-Kaleo	161	151	121	-4.041	27	-18.820	16	-6.593	21
Sekyere Afram Plains North	162	175	165	-4.048	24	-19.587	9	-7.112	24
Adansi South	163	140	152	-4.083	25	-18.549	14	-6.958	32
Upper Denkyira West	164	163	166	-4.117	28	-19.179	13	-7.114	30
Kumbumgu	165	170	172	-4.127	21	-19.449	12	-7.184	15
Bia East	166	157	183	-4.147	20	-18.962	10	-7.454	18
Sekyere Central	167	122	196	-4.148	24	-18.002	11	-7.789	20
Daffiama Bussie	168	202	132	-4.151	20	-21.600	6	-6.706	16
Asutifi South	169	134	182	-4.156	25	-18.321	21	-7.446	20
Asunafo South	170	167	175	-4.164	28	-19.405	13	-7.314	23
Nanumba South	171	91	195	-4.193	27	-17.297	16	-7.787	19
Sefwi-Wiawso	172	173	102	-4.208	23	-19.491	16	-6.388	17
Tolon	173	153	158	-4.225	25	-18.852	12	-6.998	26
Shai Osu Doku	174	136	168	-4.250	16	-18.380	7	-7.155	20
Akatsi North	175	206	167	-4.271	20	-21.877	7	-7.154	21
Nanumba North	176	143	147	-4.289	27	-18.587	12	-6.908	27
Savelugu Nanton	177	105	185	-4.290	25	-17.672	16	-7.497	17
Offinso North	178	191	136	-4.295	29	-20.302	11	-6.758	32
Mion	179	181	197	-4.322	22	-19.757	8	-7.825	14
Wassa Amenfi West	180	172	126	-4.378	18	-19.480	11	-6.623	20
Krachi Nchumuru	181	199	192	-4.388	20	-21.175	8	-7.703	18
Fanteakwa	181	174	151	-4.388	26	-19.559	12	-6.933	25
Jaman North	182	118	153	-4.404	22	-17.965	11	-6.961	29
Karaga	183	104	202	-4.406	25	-17.651	15	-8.092	14
Bosome Freho	184	162	159	-4.435	21	-19.162	12	-7.020	16
Tatale	185	198	133	-4.439	24	-21.044	8	-6.712	22
Binduri	186	192	193	-4.454	24	-20.398	9	-7.760	16

Table B.2: Ghanaian districts ordered by fitness *(continued)*

District	Rank all	Rank for.	Rank inf.	Log Fitness	Diversity	Log Fitness for.	Diversity for.	Log Fitness inf.	Diversity inf.
Gushiegu	187	128	146	-4.463	23	-18.180	11	-6.898	16
Bodi	188	189	199	-4.464	21	-20.252	10	-7.956	18
Wa West	189	168	194	-4.476	19	-19.421	8	-7.782	17
Kwahu Afram Plains South	190	200	177	-4.500	19	-21.257	7	-7.358	22
Ekumfi	191	178	176	-4.510	14	-19.608	9	-7.349	13
Talensi	192	180	190	-4.537	21	-19.727	10	-7.577	17
Kintampo South	193	204	204	-4.540	20	-21.746	7	-8.216	12
Agotime Ziope	194	160	171	-4.544	21	-19.088	11	-7.171	22
Birim South	195	166	129	-4.575	20	-19.404	14	-6.673	20
Twifo Lower Denkyira	196	120	188	-4.581	17	-17.980	8	-7.551	17
Afadzato South	197	169	138	-4.656	20	-19.427	11	-6.763	21
Wassa Amenfi East	198	117	174	-4.661	17	-17.957	12	-7.301	21
North Dayi	199	207	181	-4.695	18	-21.914	5	-7.429	16
Chereponi	200	196	186	-4.766	20	-20.897	9	-7.537	15
Nkoranza North	201	197	187	-4.806	19	-20.926	9	-7.543	16
Nabdam	202	208	203	-5.013	19	-22.001	8	-8.106	13
Mamprugu Moagduri	203	154	205	-5.032	19	-18.870	7	-8.282	13
Wa East	204	114	201	-5.213	11	-17.926	10	-8.008	8
Bunkpurugu Yonyo	205	210	198	-5.461	14	-22.527	6	-7.864	14
Banda	206	209	206	-5.515	15	-22.019	6	-8.304	14
Sene East	207	211	208	-5.522	17	-23.598	4	-8.598	14
Sekyere Afram Plains	208	205	209	-5.524	14	-21.787	6	-8.623	10
Wassa East	209	194	207	-5.945	7	-20.657	10	-8.473	9
North Gonja	210	212	210	-5.963	8	-24.074	3	-9.436	7

B.4 Industrial complexity

Table B.3: 4-digit ISIC industries by complexity

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
1	1	-	1200	Mining of uranium and thorium ores	4.7	1	3.34	1	-	-
2	-	1	3694	Manufacture of games and toys	4.02	1	-	-	4.39	1
3	3	-	2915	Manufacture of lifting and handling equipment	3.41	2	0.6	2	-	-
4	3	1	2912	Manufacture of pumps, compressors, taps and valves	2.97	3	0.6	2	4.39	1
5	7	-	1310	Mining of iron ores	2.43	2	-9.17	2	-	-
6	49	13	6110	Sea and coastal water transport	1.47	3	-16.09	4	-4.4	4
7	15	60	7430	Advertising	1.37	4	-11.73	5	-7.98	12
8	2	12	1722	Manufacture of carpets and rugs	1.1	2	0.66	1	-3.64	2
9	5	2	2422	Manufacture of paints, varnishes and similar coatings, printing ink and mastics	0.58	3	-8.56	3	0.43	1
10	9	33	2212	Publishing of newspapers, journals and periodicals	0.11	2	-9.99	2	-6.44	8
11	10	4	2923	Manufacture of machinery for metallurgy	0.06	2	-11.21	1	-1.11	1
12	1	11	3000	Manufacture of office, accounting and computing machinery	-0.25	2	3.34	1	-3.47	1
13	32	6	3110	Manufacture of electric motors, generators and transformers	-0.42	4	-13.88	4	-2.54	4
14	19	15	7413	Market research and public opinion polling	-0.48	4	-12.21	4	-4.57	3
15	-	22	2430	Manufacture of man-made fibres	-1.08	1	-	-	-5.16	1
16	1	27	2511	Manufacture of rubber tyres and tubes; retreading and rebuilding of rubber tyres	-1.13	2	3.34	1	-6.26	4
17	1	18	2102	Manufacture of corrugated paper and paperboard and of containers of paper and paperboard	-1.16	3	3.34	1	-4.82	3
18	42	5	6303	Other supporting transport activities	-1.41	4	-15.28	4	-1.74	3
19	-	17	1544	Manufacture of macaroni, noodles, couscous and similar farinaceous products	-1.51	2	-	-	-4.81	2
19	-	21	3592	Manufacture of bicycles and invalid carriages	-1.51	2	-	-	-5.1	2
20	35	10	2732	Casting of non-ferrous metals	-1.65	4	-14.26	3	-3.28	1
21	-	16	2913	Manufacture of bearings, gears, gearing and driving elements	-1.73	4	-	-	-4.79	4
22	109	85	5010	Sale of motor vehicles	-1.77	4	-19.86	6	-8.78	20
23	-	28	3140	Manufacture of accumulators, primary cells and primary batteries	-1.82	2	-	-	-6.27	2
24	44	20	6210	Scheduled air transport	-1.85	6	-15.42	6	-4.89	4
25	31	3	2101	Manufacture of pulp, paper and paperboard	-1.99	4	-13.87	1	-0.97	2
26	33	7	6030	Transport via pipelines	-2.04	3	-13.91	2	-2.8	2
27	2	25	1542	Manufacture of sugar	-2.05	2	0.66	1	-6.08	2
28	59	14	7495	Packaging activities	-2.07	3	-16.7	3	-4.56	4
29	-	19	3591	Manufacture of motorcycles	-2.16	1	-	-	-4.86	1

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
30	60	63	2412	Manufacture of fertilizers and nitrogen compounds	-2.39	5	-16.81	5	-8.09	8
31	13	15	2421	Manufacture of pesticides and other agrochemical products	-2.41	6	-11.66	4	-4.57	2
32	16	16	2926	Manufacture of machinery for textile, apparel and leather production	-2.42	2	-11.74	1	-4.79	1
33	4	39	2519	Manufacture of other rubber products	-2.72	4	-7.25	1	-6.86	4
34	72	50	6712	Security dealing activities	-2.75	7	-17.67	4	-7.53	9
35	35	62	5151	Wholesale of computers, computer peripheral equipment and software	-2.86	4	-14.26	4	-8.06	11
36	20	18	6010	Transport via railways	-3.04	5	-12.32	4	-4.82	1
37	-	42	2696	Cutting, shaping and finishing of stone	-3.09	6	-	-	-7.09	6
38	27	41	3710	Recycling of metal waste and scrap	-3.15	9	-13.2	5	-7.04	7
39	2	37	2914	Manufacture of ovens, furnaces and furnace burners	-3.19	8	0.66	1	-6.77	8
40	1	31	2310	Manufacture of coke oven products	-3.22	4	3.34	1	-6.39	5
41	38	93	5190	Other wholesale	-3.25	7	-14.95	7	-9.01	23
42	11	52	2413	Manufacture of plastics in primary forms and of synthetic rubber	-3.27	7	-11.5	3	-7.63	10
43	76	9	3150	Manufacture of electric lamps and lighting equipment	-3.29	10	-18.04	10	-3.11	1
44	37	51	7492	Investigation and security activities	-3.34	6	-14.79	5	-7.59	11
45	68	-	9500	Unspecified services	-3.36	4	-17.33	4	-	-
46	65	30	2694	Manufacture of cement, lime and plaster	-3.42	8	-17.04	6	-6.35	8
47	3	46	2919	Manufacture of other general-purpose machinery	-3.43	6	0.6	2	-7.21	5
48	90	68	7210	Hardware consultancy	-3.48	4	-19.02	4	-8.25	14
49	42	29	2813	Manufacture of steam generators, except central heating hot water boilers	-3.52	8	-15.28	5	-6.31	4
49	119	113	6420	Telecommunications	-3.52	8	-20.23	9	-9.71	33
50	67	67	5139	Wholesale of other household goods	-3.62	9	-17.29	9	-8.24	17
51	93	42	7010	Real estate activities with own or leased property	-3.72	13	-19.21	15	-7.09	11
52	108	47	9900	Extraterritorial organizations and bodies	-3.75	5	-19.85	6	-7.27	6
53	18	49	1723	Manufacture of cordage, rope, twine and netting	-3.82	7	-12.18	2	-7.42	4
54	39	45	1429	Other mining and quarrying n.e.c.	-3.92	6	-15.07	3	-7.17	3
55	36	142	3190	Manufacture of other electrical equipment n.e.c.	-3.93	8	-14.77	3	-10.42	40
56	29	48	6711	Administration of financial markets	-3.95	6	-13.25	4	-7.38	2
57	43	99	2213	Publishing of music	-3.98	8	-15.29	5	-9.2	26
58	71	1	3130	Manufacture of insulated wire and cable	-4	10	-17.63	10	4.39	1
59	3	58	3410	Manufacture of motor vehicles	-4.04	5	0.6	2	-7.81	3
60	50	32	7221	Software publishing	-4.06	6	-16.12	4	-6.41	3
61	80	24	6120	Inland water transport	-4.07	9	-18.62	7	-6.05	3
62	51	26	2924	Manufacture of machinery for mining, quarrying and construction	-4.1	5	-16.17	3	-6.14	4

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
63	77	13	1110	Extraction of crude petroleum and natural gas	-4.22	8	-18.19	6	-4.4	5
64	8	56	2929	Manufacture of other special-purpose machinery	-4.24	10	-9.81	2	-7.72	8
64	23	90	2211	Publishing of books, brochures and other publications	-4.24	11	-12.89	8	-8.93	21
65	69	66	2021	Manufacture of veneer sheets; manufacture of plywood, laminboard, particle board and other panels and boards	-4.31	8	-17.44	7	-8.23	11
66	5	65	1520	Manufacture of dairy products	-4.34	8	-8.56	2	-8.17	10
66	84	57	2699	Manufacture of other non-metallic mineral products n.e.c.	-4.34	8	-18.72	3	-7.73	6
67	52	138	7250	Maintenance and repair of office, accounting and computing machinery	-4.36	8	-16.18	2	-10.3	45
68	61	87	2222	Service activities related to printing	-4.38	9	-16.83	11	-8.88	15
69	89	40	7493	Building-cleaning and industrial-cleaning activities	-4.39	11	-18.96	9	-6.97	11
70	88	36	6412	Courier activities other than national post activities	-4.42	15	-18.88	11	-6.61	8
71	58	92	7229	Other software consultancy and supply	-4.47	12	-16.64	8	-8.99	24
72	86	55	2429	Manufacture of other chemical products n.e.c.	-4.49	10	-18.86	4	-7.69	8
72	74	105	7421	Architectural and engineering activities and related technical consultancy	-4.49	15	-17.79	9	-9.42	30
73	124	59	1010	Mining and agglomeration of hard coal	-4.54	10	-20.43	10	-7.87	6
74	78	101	9301	Washing and (dry-)cleaning of textile and fur products	-4.55	18	-18.43	11	-9.27	25
75	30	74	7290	Other computer-related activities	-4.59	9	-13.73	5	-8.47	12
76	79	86	7491	Labour recruitment and provision of personnel	-4.64	7	-18.56	5	-8.87	8
77	130	57	7020	Real estate activities on a fee or contract basis	-4.68	9	-20.71	10	-7.73	11
77	156	61	7522	Defence activities	-4.68	16	-21.65	21	-8.04	6
78	4	73	3420	Manufacture of bodies (coachwork) for motor vehicles; manufacture of trailers and semi-trailers	-4.7	12	-7.25	1	-8.46	12
79	73	63	6023	Freight transport by road	-4.76	15	-17.77	10	-8.09	9
80	14	150	2710	Manufacture of basic iron and steel	-4.77	10	-11.71	4	-10.58	45
81	47	135	1712	Finishing of textiles	-4.79	9	-15.84	3	-10.23	37
82	17	78	5251	Retail sale via mail order houses	-4.85	15	-12.13	5	-8.61	14
83	22	69	2230	Reproduction of recorded media	-4.87	14	-12.57	2	-8.28	11
83	94	107	6309	Activities of other transport agencies	-4.87	12	-19.22	11	-9.53	16
84	113	71	1422	Extraction of salt	-4.88	7	-20.08	5	-8.35	4
85	104	109	1920	Manufacture of footwear	-4.95	11	-19.71	7	-9.56	20
86	139	102	0112	Growing of vegetables, horticultural specialties and nursery products	-4.97	8	-20.99	7	-9.29	15
87	1	79	1820	Dressing and dyeing of fur; manufacture of articles of fur	-4.99	14	3.34	1	-8.62	13
88	24	80	2930	Manufacture of domestic appliances n.e.c.	-5.02	16	-12.96	4	-8.63	18
89	8	75	3610	Manufacture of furniture	-5.03	10	-9.81	2	-8.5	11
90	142	23	6720	Activities auxiliary to insurance and pension funding	-5.05	8	-21.17	7	-5.57	4

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
91	116	86	6304	Activities of travel agencies and tour operators; tourist assistance activities n.e.c.	-5.06	13	-20.17	12	-8.87	13
92	-	93	3512	Building and repairing of pleasure and sporting boats	-5.07	9	-	-	-9.01	9
93	45	127	2899	Manufacture of other fabricated metal products n.e.c.	-5.12	23	-15.75	9	-10.07	37
93	63	103	1729	Manufacture of other textiles n.e.c.	-5.12	11	-16.97	4	-9.36	15
93	41	97	5131	Wholesale of textiles, clothing and footwear	-5.12	14	-15.19	5	-9.13	17
94	141	8	0130	Growing of crops combined with farming of animals (mixed farming)	-5.13	12	-21.16	13	-3	2
95	48	72	6301	Cargo handling	-5.15	11	-15.98	5	-8.4	10
96	1	88	3692	Manufacture of musical instruments	-5.16	13	3.34	1	-8.9	14
97	133	121	2330	Processing of nuclear fuel	-5.2	15	-20.77	11	-9.89	31
97	87	53	9232	Museums activities and preservation of historic sites and buildings	-5.2	14	-18.87	10	-7.66	5
98	99	94	0501	Fishing	-5.25	11	-19.48	4	-9.02	12
99	53	103	2925	Manufacture of machinery for food, beverage and tobacco processing	-5.26	14	-16.2	6	-9.36	23
100	25	77	3330	Manufacture of watches and clocks	-5.28	23	-13.07	6	-8.6	22
101	34	111	5152	Wholesale of electronic and telecommunications parts and equipment	-5.32	15	-14.07	3	-9.6	26
102	136	85	6601	Life insurance	-5.33	25	-20.89	22	-8.78	20
103	97	78	1533	Manufacture of prepared animal feeds	-5.35	15	-19.38	6	-8.61	10
104	161	144	4520	Building of complete constructions or parts thereof; civil engineering	-5.38	21	-21.72	22	-10.51	35
105	117	150	5159	Wholesale of other machinery, equipment and supplies	-5.39	16	-20.18	13	-10.58	34
106	145	34	7422	Technical testing and analysis	-5.4	16	-21.32	17	-6.49	4
107	121	84	1543	Manufacture of cocoa, chocolate and sugar confectionery	-5.44	8	-20.27	7	-8.73	4
108	111	82	4020	Manufacture of gas; distribution of gaseous fuels through mains	-5.45	14	-19.97	10	-8.68	11
109	106	127	7111	Renting of land transport equipment	-5.47	22	-19.73	12	-10.07	34
110	132	35	6591	Financial leasing	-5.5	16	-20.76	13	-6.5	6
111	101	97	9220	News agency activities	-5.53	14	-19.63	8	-9.13	8
112	153	38	8532	Social work activities without accommodation	-5.58	16	-21.5	16	-6.8	7
113	142	168	6599	Other financial intermediation n.e.c.	-5.6	23	-21.17	24	-10.89	57
114	6	104	2927	Manufacture of weapons and ammunition	-5.65	12	-9.01	1	-9.37	11
115	56	106	2692	Manufacture of refractory ceramic products	-5.69	25	-16.59	7	-9.46	26
116	99	129	9241	Sporting activities	-5.7	22	-19.48	15	-10.13	34
117	157	71	5030	Sale of motor vehicle parts and accessories	-5.73	18	-21.67	13	-8.35	13
118	160	64	6511	Central banking	-5.74	22	-21.71	21	-8.11	7
119	167	70	4510	Site preparation	-5.75	20	-21.9	13	-8.31	9

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
120	70	110	2922	Manufacture of machine tools	-5.76	16	-17.5	4	-9.59	22
121	158	76	1320	Mining of non-ferrous metal ores, except uranium and thorium ores	-5.81	16	-21.68	16	-8.51	8
122	55	118	1549	Manufacture of other food products n.e.c.	-5.82	18	-16.53	4	-9.85	22
123	151	89	7230	Data processing	-5.86	21	-21.46	13	-8.91	15
124	143	44	7530	Compulsory social security activities	-5.87	24	-21.28	20	-7.13	5
124	170	81	7310	Research and experimental development on natural sciences and engineering (NSE)	-5.87	17	-22.29	23	-8.67	12
125	96	98	6602	Pension funding	-5.91	24	-19.26	15	-9.16	12
126	140	100	5149	Wholesale of other intermediate products, waste and scrap	-5.96	18	-21.07	13	-9.22	12
126	40	122	1721	Manufacture of made-up textile articles, except apparel	-5.96	30	-15.18	4	-9.93	29
127	158	91	9111	Activities of business and employers organizations	-5.97	17	-21.68	20	-8.97	6
128	125	112	1513	Processing and preserving of fruit and vegetables	-5.98	19	-20.54	9	-9.7	16
128	148	148	5510	Hotels; camping sites and other provision of short-stay accommodation	-5.98	34	-21.37	33	-10.56	56
129	85	140	1512	Processing and preserving of fish and fish products	-6.03	18	-18.79	4	-10.36	24
130	165	100	1410	Quarrying of stone, sand and clay	-6.04	27	-21.81	17	-9.22	18
131	134	116	2695	Manufacture of articles of concrete, cement and plaster	-6.07	37	-20.79	27	-9.82	37
132	175	141	0150	Hunting, trapping and game propagation including related service activities	-6.08	27	-22.7	23	-10.41	36
133	155	43	7412	Accounting, bookkeeping and auditing activities; tax consultancy	-6.09	22	-21.64	27	-7.11	6
134	154	83	9233	Botanical and zoological gardens and nature reserves activities	-6.1	13	-21.53	9	-8.69	4
134	91	77	5240	Retail sale of second-hand goods in stores	-6.1	34	-19.11	10	-8.6	23
135	172	77	7411	Legal activities	-6.15	22	-22.39	28	-8.6	17
136	128	116	2921	Manufacture of agricultural and forestry machinery	-6.16	27	-20.61	8	-9.82	25
137	21	147	2219	Other publishing	-6.19	26	-12.44	5	-10.55	47
138	135	182	6411	National post activities	-6.23	31	-20.87	13	-11.26	70
138	181	145	5110	Wholesale on a fee or contract basis	-6.23	17	-23.02	15	-10.53	27
139	168	117	9112	Activities of professional organizations	-6.28	23	-21.96	21	-9.84	23
139	99	140	2424	Manufacture of soap and detergents, cleaning and polishing preparations, perfumes and toilet preparations	-6.28	25	-19.48	9	-10.36	33
140	1	131	3511	Building and repairing of ships	-6.29	29	3.34	1	-10.17	30
141	146	149	4530	Building installation	-6.31	27	-21.34	21	-10.57	43
142	100	134	9309	Other service activities n.e.c.	-6.33	31	-19.56	12	-10.21	31
143	149	132	9000	Sewage and refuse disposal, sanitation and similar activities	-6.34	26	-21.41	15	-10.19	21
144	95	137	9211	Motion picture and video production and distribution	-6.36	31	-19.24	11	-10.27	37
145	83	129	2610	Manufacture of glass and glass products	-6.37	40	-18.7	11	-10.13	40
146	-	127	2023	Manufacture of wooden containers	-6.39	29	-	-	-10.07	27

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
146	54	124	1912	Manufacture of luggage, handbags and the like, saddlery and harness	-6.39	29	-16.31	4	-9.97	29
147	173	142	2811	Manufacture of structural metal products	-6.46	40	-22.46	15	-10.42	46
148	108	125	2691	Manufacture of non-structural non-refractory ceramic ware	-6.47	33	-19.85	7	-9.98	29
149	123	165	9214	Dramatic arts, music and other arts activities	-6.49	34	-20.35	15	-10.8	55
150	177	108	8030	Higher education	-6.5	42	-22.77	45	-9.55	27
151	190	167	6592	Other credit granting	-6.51	32	-23.48	35	-10.85	55
152	183	81	8520	Veterinary activities	-6.52	40	-23.04	32	-8.67	18
153	127	114	5020	Maintenance and repair of motor vehicles	-6.53	52	-20.59	22	-9.73	44
153	122	156	2911	Manufacture of engines and turbines, except aircraft, vehicle and cycle engines	-6.53	29	-20.33	14	-10.65	45
154	182	136	5141	Wholesale of solid, liquid and gaseous fuels and related products	-6.55	38	-23.03	34	-10.24	36
154	107	126	5232	Retail sale of textiles, clothing, footwear and leather goods	-6.55	49	-19.79	7	-10.03	39
155	148	161	6021	Other scheduled passenger land transport	-6.56	39	-21.37	23	-10.73	51
156	28	180	1541	Manufacture of bakery products	-6.57	32	-13.24	5	-11.23	67
156	9	123	1730	Manufacture of knitted and crocheted fabrics and articles	-6.57	39	-9.99	2	-9.96	33
157	169	184	7499	Other business activities n.e.c.	-6.58	25	-22.22	22	-11.31	60
158	147	133	6302	Storage and warehousing	-6.64	26	-21.36	17	-10.2	20
158	131	178	5239	Other retail sale in specialized stores	-6.64	38	-20.74	20	-11.14	66
158	174	150	5122	Wholesale of food, beverages and tobacco	-6.64	46	-22.64	41	-10.58	46
159	191	162	6519	Other monetary intermediation	-6.66	41	-23.51	62	-10.74	47
160	187	152	4100	Collection, purification and distribution of water	-6.68	38	-23.28	41	-10.6	45
161	171	119	7514	Supporting service activities for the government as a whole	-6.69	38	-22.38	37	-9.86	23
162	193	149	5234	Retail sale of hardware, paints and glass	-6.72	59	-23.68	44	-10.57	66
162	194	160	7521	Foreign affairs	-6.72	35	-23.73	39	-10.72	45
162	115	143	5142	Wholesale of metals and metal ores	-6.72	24	-20.16	13	-10.5	24
163	52	153	9212	Motion picture projection	-6.75	35	-16.18	3	-10.61	36
164	201	115	7414	Business and management consultancy activities	-6.79	31	-24.43	44	-9.8	22
165	139	140	5233	Retail sale of household appliances, articles and equipment	-6.8	54	-20.99	23	-10.36	50
165	66	144	2520	Manufacture of plastics products	-6.8	37	-17.09	7	-10.51	40
166	186	166	0113	Growing of fruit, nuts, beverage and spice crops	-6.81	30	-23.2	21	-10.82	47
167	163	157	9120	Activities of trade unions	-6.85	37	-21.77	24	-10.66	32
167	166	194	1554	Manufacture of soft drinks; production of mineral waters	-6.85	44	-21.82	26	-11.47	85
168	26	162	2891	Forging, pressing, stamping and roll-forming of metal; powder metallurgy	-6.88	48	-13.12	2	-10.74	45
169	75	171	2812	Manufacture of tanks, reservoirs and containers of metal	-6.89	48	-17.86	8	-10.98	61
169	112	176	1551	Distilling, rectifying and blending of spirits; ethyl alcohol production from fermented materials	-6.89	32	-19.98	8	-11.11	46

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
170	82	160	1711	Preparation and spinning of textile fibres; weaving of textiles	-6.91	36	-18.65	4	-10.72	36
170	188	169	0111	Growing of cereals and other crops n.e.c.	-6.91	26	-23.34	23	-10.94	39
171	185	131	8531	Social work activities with accommodation	-6.98	50	-23.13	24	-10.17	34
172	102	164	1511	Production, processing and preserving of meat and meat products	-6.99	44	-19.65	17	-10.78	36
172	152	151	9303	Funeral and related activities	-6.99	46	-21.47	12	-10.59	44
173	189	120	8090	Other education	-7	46	-23.43	54	-9.88	39
174	98	154	7130	Renting of personal and household goods n.e.c.	-7.01	52	-19.45	11	-10.62	51
175	159	187	2010	Sawmilling and planing of wood	-7.03	35	-21.69	17	-11.34	76
176	198	96	4010	Production, transmission and distribution of electricity	-7.04	57	-24.07	73	-9.12	26
177	197	54	6603	Non-life insurance	-7.06	43	-23.94	63	-7.67	13
178	142	174	1514	Manufacture of vegetable and animal oils and fats	-7.08	42	-21.17	16	-11.07	47
178	179	128	0200	Forestry, logging and related service activities	-7.08	41	-22.83	38	-10.08	15
179	46	146	5252	Retail sale via stalls and markets	-7.09	39	-15.83	3	-10.54	36
180	12	172	1553	Manufacture of malt liquors and malt	-7.11	39	-11.52	3	-11.04	40
181	162	169	9249	Other recreational activities	-7.14	54	-21.73	29	-10.94	52
182	184	183	7240	Database activities and online distribution of electronic content	-7.16	52	-23.05	52	-11.29	65
182	196	95	7513	Regulation of and contribution to more efficient operation of business	-7.16	52	-23.83	56	-9.1	16
183	126	172	9192	Activities of political organizations	-7.2	46	-20.56	13	-11.04	45
184	120	173	9219	Other entertainment activities n.e.c.	-7.22	53	-20.26	16	-11.05	58
184	57	175	2893	Manufacture of cutlery, hand tools and general hardware	-7.22	55	-16.6	5	-11.1	59
185	105	159	7494	Photographic activities	-7.26	64	-19.72	15	-10.69	55
186	64	158	3230	Manufacture of television and radio receivers, sound or video recording or reproducing apparatus, and associated goods	-7.35	78	-17.01	6	-10.67	62
187	138	175	0121	Farming of cattle, sheep, goats, horses, asses, mules and hinnies; dairy farming	-7.38	55	-20.95	8	-11.1	51
187	178	181	0122	Other animal farming; production of animal products n.e.c.	-7.38	59	-22.79	25	-11.24	66
187	200	139	8512	Medical and dental practice activities	-7.38	47	-24.31	48	-10.32	33
188	199	131	8022	Technical and vocational secondary education	-7.39	64	-24.22	72	-10.17	34
189	92	192	2892	Treatment and coating of metals; general mechanical engineering on a fee or contract basis	-7.56	72	-19.15	6	-11.44	74
190	114	163	9302	Hairdressing and other beauty treatment	-7.57	90	-20.1	16	-10.75	62
191	195	189	5121	Wholesale of agricultural raw materials and live animals	-7.58	57	-23.77	55	-11.36	59
192	202	127	0140	Agricultural and animal husbandry service activities, except veterinary activities	-7.59	50	-24.46	53	-10.07	26
193	176	193	5259	Other non-store retail sale	-7.68	73	-22.72	28	-11.46	70

Table B.3: 4-digit ISIC industries by complexity *(continued)*

Rank	Rank for.	Rank inf.	Code	ISIC 4-digit ind.	ICI (Log)	Ubiquity	ICI for. (Log)	Ubiquity for.	ICI inf. (Log)	Ubiquity inf.
194	62	179	3220	Manufacture of television and radio transmitters and apparatus for line telephony and line telegraphy	-7.71	94	-16.94	9	-11.21	74
195	118	177	1810	Manufacture of wearing apparel, except fur apparel	-7.73	92	-20.21	11	-11.13	77
195	150	188	5231	Retail sale of pharmaceutical and medical goods, cosmetic and toilet articles	-7.73	99	-21.44	31	-11.35	81
196	205	130	7523	Public order and safety activities	-7.77	86	-25.27	112	-10.14	29
197	164	186	3599	Manufacture of other transport equipment n.e.c.	-7.8	105	-21.79	28	-11.33	89
198	129	196	5219	Other retail sale in non-specialized stores	-7.81	74	-20.69	23	-11.66	83
199	207	155	7512	Regulation of the activities of agencies that provide health care, education, cultural services and other social services, excluding social security	-7.82	68	-25.69	91	-10.64	30
199	203	170	8519	Other human health activities	-7.82	83	-24.55	87	-10.95	61
200	192	201	5050	Retail sale of automotive fuel	-7.84	95	-23.67	64	-12.12	108
201	157	191	5220	Retail sale of food, beverages and tobacco in specialized stores	-7.92	115	-21.67	36	-11.39	83
202	103	195	5260	Repair of personal and household goods	-7.93	74	-19.69	11	-11.51	72
203	137	197	1532	Manufacture of starches and starch products	-8.02	82	-20.92	8	-11.74	76
204	206	126	7511	General (overall) public service activities	-8.04	84	-25.3	113	-10.03	21
205	208	190	8511	Hospital activities	-8.12	118	-26.18	144	-11.37	66
206	81	198	2022	Manufacture of builders' carpentry and joinery	-8.16	91	-18.64	10	-11.98	81
207	144	202	5040	Sale, maintenance and repair of motorcycles and related parts and accessories	-8.38	109	-21.31	25	-12.15	96
208	110	199	5520	Restaurants, bars and canteens	-8.45	139	-19.91	19	-12.01	117
209	209	166	8021	General secondary education	-8.5	144	-26.35	168	-10.82	62
210	143	200	5211	Retail sale in non-specialized stores with food, beverages or tobacco predominating	-8.55	149	-21.28	29	-12.06	121
211	180	203	1531	Manufacture of grain mill products	-8.64	138	-22.99	26	-12.45	123
212	210	185	8010	Primary education	-8.73	184	-26.4	195	-11.32	85
213	204	204	9191	Activities of religious organizations	-8.78	160	-24.86	116	-12.57	148

Appendix C

Appendix to Chapter 4

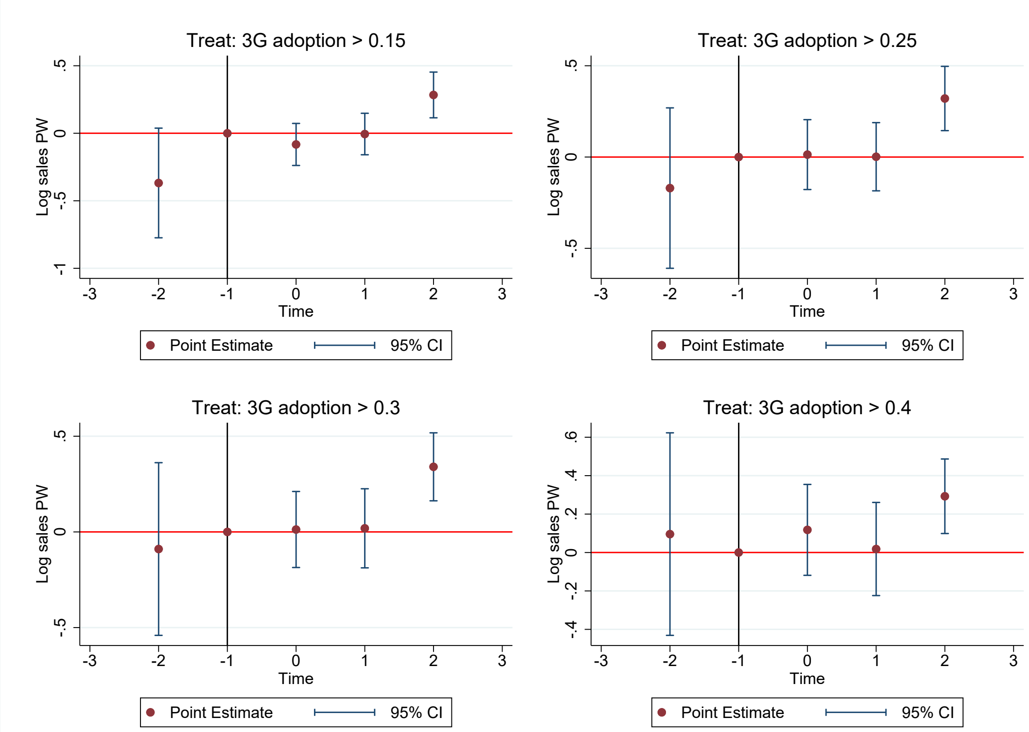

Figure C.1: Event study: robustness checks. Each event study design uses the first year in which a household hits a different threshold of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. Thresholds are indicated on the top of each graph. The outcome variable used in the event study specification is the log of the NFE sales per worker. The coefficients reported in the figure come from a model based on Equation (4.4), including household, industry, and year fixed effects. The model uses NFEs as units of observation, and incorporates the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

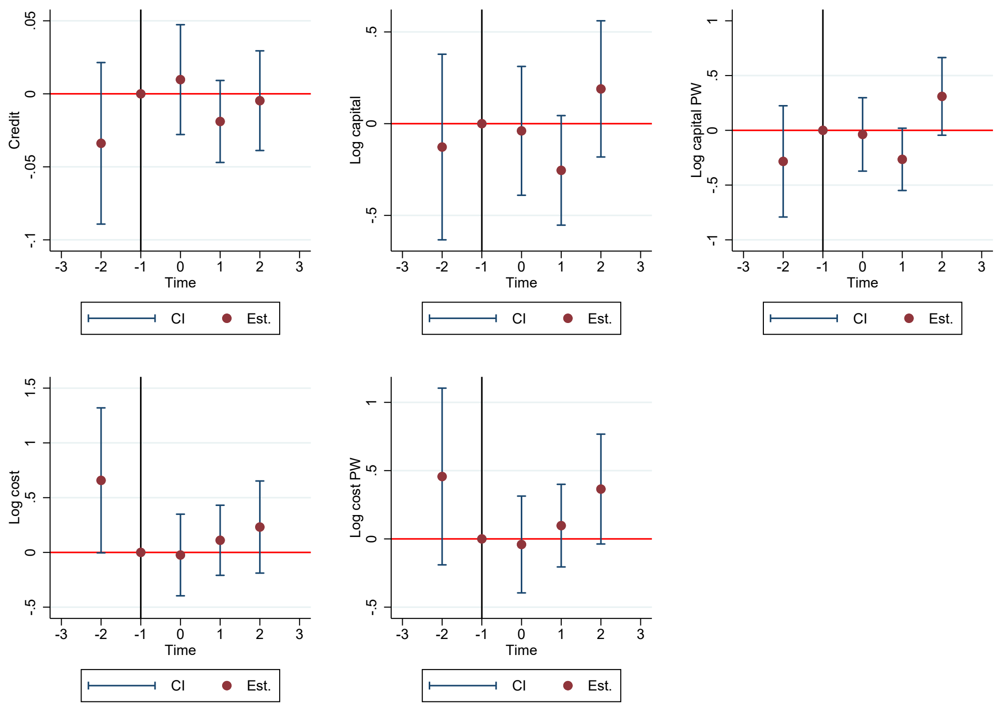

Figure C.2: Event study: mechanisms. The event study design uses the first year in which a household hits 0.2 of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. The outcome variable used in the event study specification is indicated on the vertical axis. The coefficients reported in the figure come from a model based on Equation (4.4), including household, industry, and year fixed effects. The model uses NFEs as units of observation, and incorporates the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

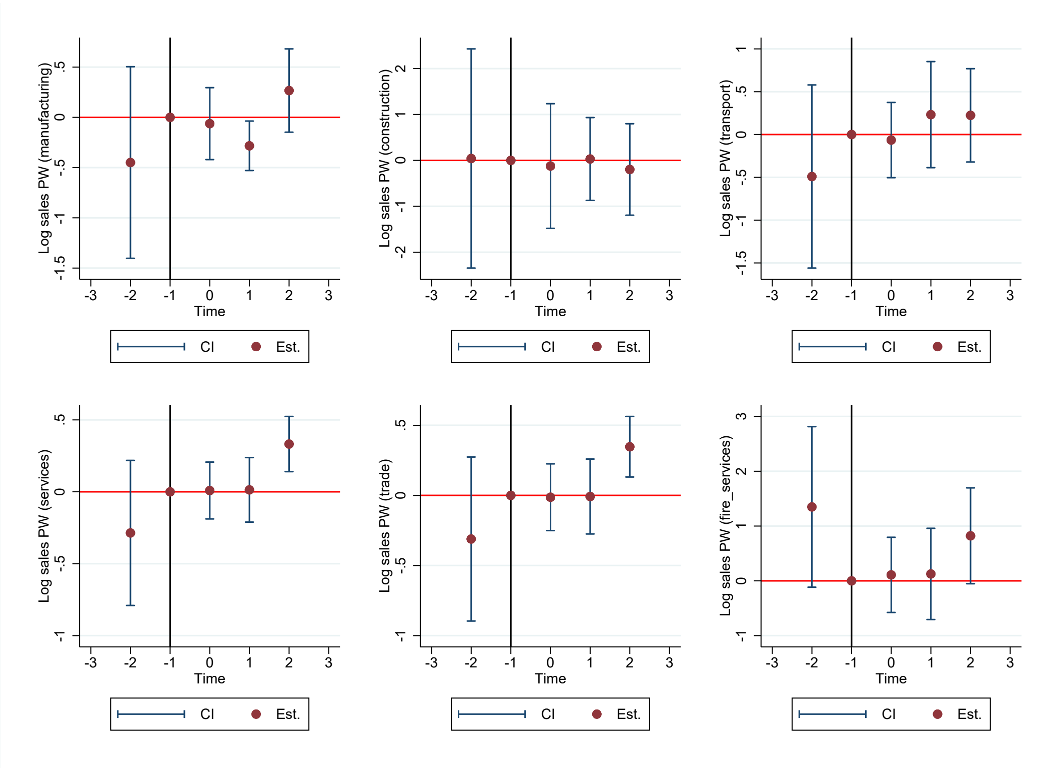

Figure C.3: Event study: sales per worker, by industry. The event study design uses the first year in which a household hits 0.2 of 3G mobile internet adoption as treatment, corresponding to time 0 in the horizontal axis. The outcome variable and the industry-level sample used in the event study specification are indicated on the vertical axis. The coefficients reported in the figure come from a model based on Equation (4.4), including household, industry, and year fixed effects. The model uses NFEs as units of observation, and incorporates the following controls: owner's gender (female); owner's age; number of household members; female-headed household; household is active in agriculture; owner can read; owner ever attended school; distance of the enumerator area from main road, town, market, and capital city (in km); number of conflicts per capita in the LGA. Standard errors are clustered at the enumerator area level. Regression coefficients are reported together with their 95% confidence interval (CI). The graphs have been created using the Stata command `eventdd`.

Appendix D

Appendix to Chapter 5

D.1 Figures

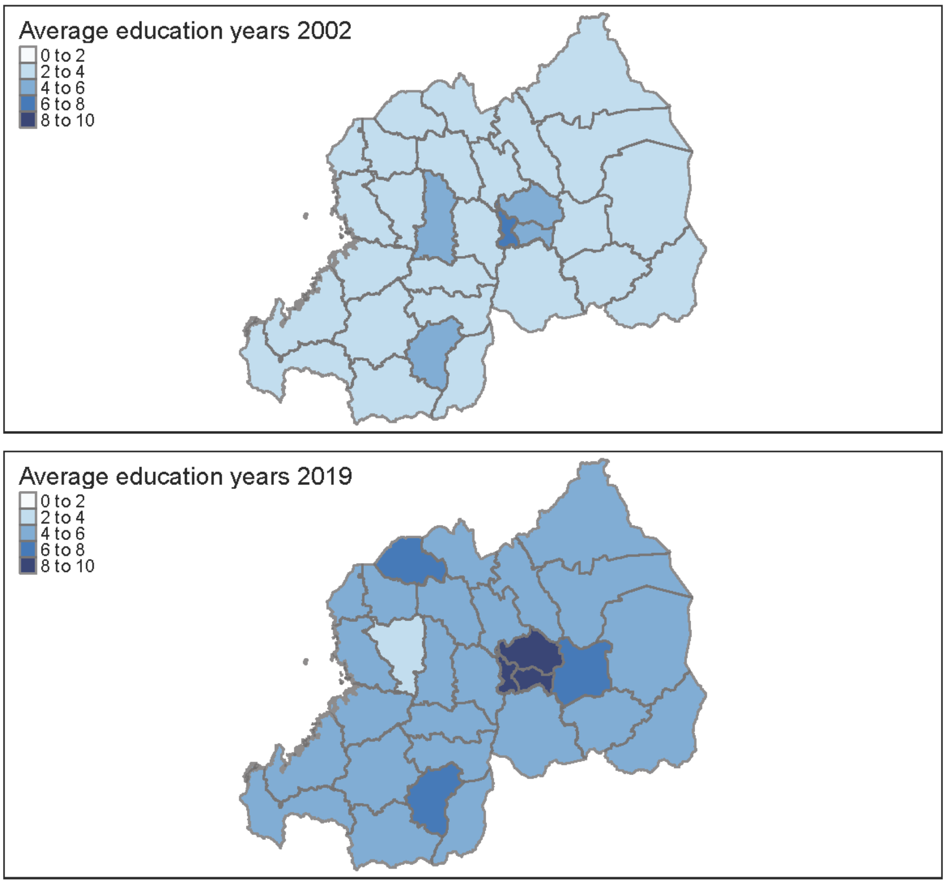

Figure D.1: Geographical distribution of the average number of years of education (2002 and 2019). Source: Authors' elaboration on RLFS data.

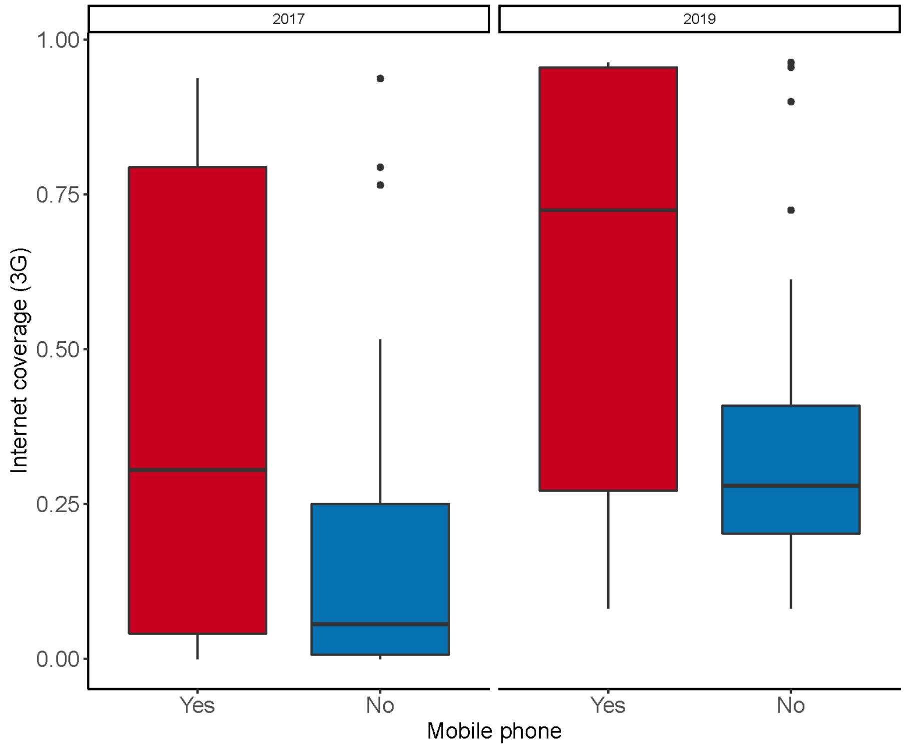

Figure D.2: Mobile phone ownership in relation to the increase in 3G coverage (2017 and 2019). Note: Red indicates urban districts, blue indicates rural districts. Source: Authors' elaboration on RLFS data.

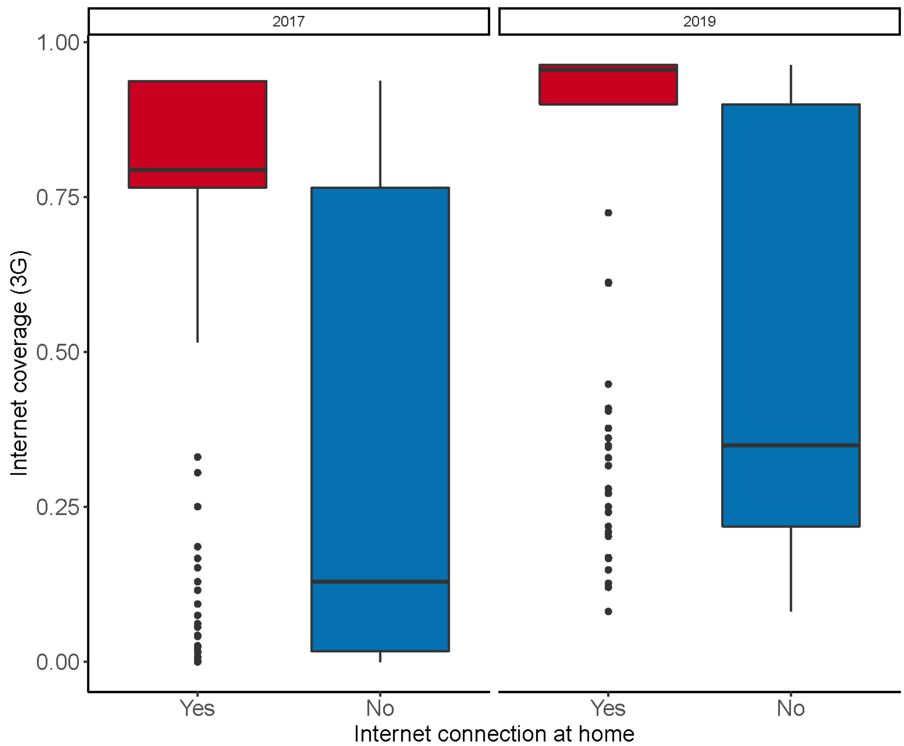

Figure D.3: Having an internet connection at home in relation to the increase in 3G coverage (2017 and 2019). Note: Red indicates urban districts, blue indicates rural districts. Source: Authors' elaboration on RLFS data.

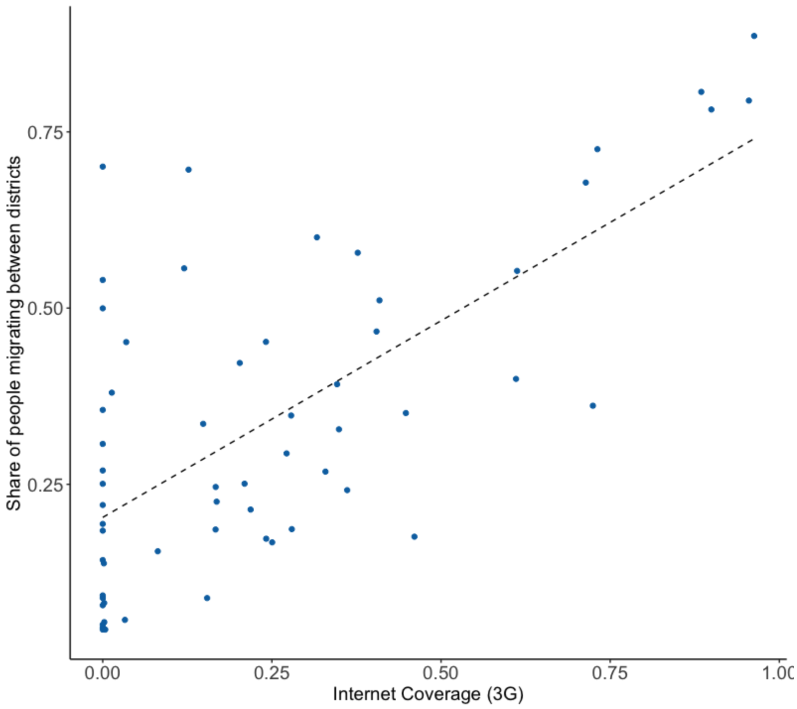

Figure D.4: The graph reports the share of people migrating between districts and the increase in fast internet coverage (3G) between 2017 and 2019. Source: Authors' elaboration on national census and RLFS data.

D.2 Tables

Table D.1: Share of industries, average across districts

Industry	2002	2012	2017	2018	2019
Manufacturing	1.4	2.77	5.2	6.32	6.39
Private work in households	2.72	2.99	6.74	6.89	6.65
Education	1.33	1.8	3.83	3.35	3.84
Construction	1.47	3.42	9.24	10.05	9.66
Accommodation	0.25	0.84	1.19	2.03	2.73
Trade	3.32	5.15	14.73	14.14	13.99
Other Services	0.79	1.15	2.18	2.56	2.63
Public	0.78	0.99	1.9	1.63	1.29
Business	0.31	1.1	0.62	0.7	0.67
Transport	1.23	2.09	3.97	4.21	5.08
Agriculture	85.53	75.95	45.89	43.22	41.93
Utilities	0.1	0.21	0.54	0.38	0.44
Finance	0.09	0.32	0.59	0.71	0.87
Health	0.51	0.76	1.51	1.45	1.29
Mining	0.16	0.46	1.87	2.37	2.52

Note: Authors' elaboration on national census and RLFS data.

Table D.2: Descriptive statistics

Statistic	N	Mean	St. Dev.	Min	Max
Occupations					
Skilled	150	0.027	0.024	0.004	0.131
Unskilled	150	0.436	0.064	0.276	0.619
Employed	150	0.463	0.067	0.290	0.636
Industries					
Agriculture	150	0.585	0.258	0.048	0.970
Manufacturing	150	0.044	0.028	0.003	0.170
Tertiary Sector	150	0.285	0.199	0.024	0.783
Education					
Primary or less	150	0.812	0.151	0.369	0.995
Secondary	150	0.154	0.112	0.005	0.386
Tertiary	150	0.034	0.050	0.000	0.260
Years of education	150	5.032	1.300	2.910	9.731
Demographic					
Age	150	32.072	1.366	27.780	34.748
Female	150	0.531	0.022	0.448	0.578
Migration (district)	90	0.272	0.198	0.032	0.798
Migration (province)	90	0.191	0.167	0.012	0.612
Internet					
2G	150	0.773	0.386	0.000	1.000
3G	150	0.203	0.263	0.000	0.963
4G	150	0.497	0.416	0.000	1.000
Geographic					
Malaria stability	150	0.528	0.668	0.000	2.338
Terrain ruggedness	150	223.686	68.723	107.334	382.963
Agricultural suitability	150	0.495	0.123	0.308	0.827
Distance from coast	150	1,076.845	43.879	979.015	1,149.148
Distance from railway	150	234.493	28.792	165.686	281.594
Distance from capital	150	58.820	30.770	6.447	139.695

Table D.3: Robustness, province-specific time trends

	Dependent variable:					
	Employed	Skilled	Unskilled	Agriculture	Manuf.	Tertiary
	(1)	(2)	(3)	(4)	(5)	(6)
3G	0.448**	0.0795**	0.369**	0.0243	0.0714	0.488***
	(0.179)	(0.0365)	(0.175)	(0.234)	(0.0777)	(0.171)
Observations	150	150	150	150	150	150
District FE	YES	YES	YES	YES	YES	YES
Province Trends	YES	YES	YES	YES	YES	YES
District Controls	YES	YES	YES	YES	YES	YES
Mean DV	0.463	0.0274	0.436	0.585	0.0442	0.285
Quantification	0.142	0.0252	0.117	0.00769	0.0226	0.155
F-stat	5.369	5.369	5.369	5.369	5.369	5.369

Note: The dependent variables measure, respectively, the share of skilled workers among the working-age population (Skilled); the share of unskilled workers among the working-age population (Unskilled) and the share of agricultural, manufacturing and services in the district's total employment. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table D.4: Robustness, initial conditions

	Dependent variable:					
	Employed	Skilled	Unskilled	Agriculture	Manuf.	Tertiary
	(1)	(2)	(3)	(4)	(5)	(6)
3G	0.337***	0.0620*	0.300***	-0.336	0.137*	0.653***
	(0.0713)	(0.0316)	(0.0865)	(0.227)	(0.0799)	(0.195)
Observations	150	150	150	150	150	150
District FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
District Controls	YES	YES	YES	YES	YES	YES
Initial DV	YES	YES	YES	YES	YES	YES

Note: The dependent variables measure, respectively, the share of skilled workers among the working-age population (Skilled); the share of unskilled workers among the working-age population (Unskilled) and the share of agricultural, manufacturing and services in the district's total employment. 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of a district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. In addition, all regressions include initial values of the dependent variable, interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table D.5: Robustness, accounting for 4G coverage

	Dependent variable:					
	Employed	Skilled	Unskilled	Agriculture	Manuf.	Tertiary
	(1)	(2)	(3)	(4)	(5)	(6)
3G	0.278***	0.0691***	0.209**	0.0214	0.0882	0.362*
	(0.0932)	(0.0242)	(0.0964)	(0.288)	(0.0661)	(0.178)
4G	0.156	-0.00910	0.166	0.183	-0.117*	-0.253
	(0.131)	(0.0241)	(0.138)	(0.326)	(0.0659)	(0.196)
Observations	150	150	150	150	150	150
District FE	YES	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES	YES
District Controls	YES	YES	YES	YES	YES	YES
Mean DV	0.463	0.0274	0.436	0.585	0.0442	0.285
Quantification	0.0879	0.0219	0.0660	0.00676	0.0279	0.115
F-stat	13.13	13.13	13.13	13.13	13.13	13.13

Note: The dependent variables measure, respectively, the share of skilled workers among the working-age population (Skilled); the share of unskilled workers among the working-age population (Unskilled) and the share of agricultural, manufacturing and services in the district's total employment. 3G and 4G measure the percentage of the population covered by the respective mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. In addition, all regressions include initial values of the dependent variable, interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table D.6: OLS results, individual DHS data

	Dependent variable:				
	Employed (1)	Skilled (2)	Unskilled (3)	Agriculture (4)	No Agriculture (5)
3G	0.0501*	0.115***	-0.0511	-0.104***	0.156***
	(0.0295)	(0.0321)	(0.0350)	(0.0393)	(0.0416)
2G	-0.0242	-0.0605**	0.0427	0.0519	-0.0746*
	(0.0395)	(0.0287)	(0.0496)	(0.0566)	(0.0415)
Female	0.00765	-0.0801***	0.0882***	0.146***	-0.138***
	(0.00599)	(0.00369)	(0.00649)	(0.00640)	(0.00476)
Age	0.00939***	0.00202***	0.00753***	0.0111***	-0.00164***
	(0.000201)	(0.000170)	(0.000239)	(0.000214)	(0.000197)
Constant	0.451***	0.0951***	0.343***	0.129**	0.319***
	(0.0363)	(0.0263)	(0.0461)	(0.0522)	(0.0389)
Observations	61,065	61,105	61,105	61,105	61,105
R-squared	0.127	0.107	0.136	0.235	0.180
District FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
Mean DV	0.808	0.147	0.658	0.558	0.249

Note: The dependent variables are dummies measuring, respectively, if an individual is employed (column 1), if he/she is employed in a skilled or an unskilled occupation (columns 2-3), if he/she is employed in or outside the agricultural sector (columns 4-5). 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include individual specific controls (their age and gender) as well as a variable measuring the 2G mobile technology coverage of the district's total population. All regressions are adjusted for the complex sample design of DHS using the approach recommended by the data provider (for more details, see the following link). Mean DV is the average value of the dependent variable in the estimation sample. Standard errors are computed after adjusting regressions for their sampling design. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table D.7: 2SLS results, migrant workers by occupation and industry

	Share of migrant workers:				
	Skilled	Unskilled	Agriculture	Manufacturing	Tertiary
	(1)	(2)	(3)	(4)	(5)
3G	0.0371**	0.155*	0.0760	0.0237*	0.0903***
	(0.0168)	(0.0824)	(0.0900)	(0.0135)	(0.0316)
Observations	90	90	90	90	90
R-squared	0.504	0.627	0.584	-0.034	0.730
District FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES
District Controls	YES	YES	YES	YES	YES
Mean DV	0.00908	0.103	0.0582	0.00434	0.0388
Quantification	0.0117	0.0491	0.0241	0.00751	0.0286
F-stat	10.66	10.66	10.66	10.66	10.66

Note: The dependent variables measure, respectively, the share of migrant workers relocated from other provinces in skilled (column 1) and unskilled (column 2) occupations, and in agriculture, manufacturing and tertiary sectors (columns 3 to 5). 3G measures the percentage of the population covered by the mobile technology in each district. All regressions include the following controls: the 2G mobile technology coverage of the district's total population; the average age of the district's population; the share of female population in the district's total population; the stability of malaria; terrain's ruggedness; the suitability of the terrain for agricultural use; the distance (in km) to the nearest coast; the distance (in km) to the closest colonial railway; the distance (in km) to the nearest border. All the geographic variables are interacted with a time trend. All regressions are estimated using a 2SLS estimator. The F-stat reports the results of the Kleibergen-Paap rk Wald F statistic. Mean DV is the average value of the dependent variable in the estimation sample. The quantification reports the estimated change in the mean of the dependent variable resulting from a shift in the variable of interest (3G) from the 25th to the 75th percentile of its distribution. Standard errors are clustered at the district level. *** p<0.01, ** p<0.05, * p<0.1.

Table D.8: Production function coefficients

Sector	Labour	Capital	Materials
Agriculture, forestry and fishing	0.1226	0.0336	0.8085
Mining and quarrying	0.3511	-0.1564	0.8169
Manufacturing	0.1691	0.0338	0.7223
Electricity, gas and air conditioning supply	0.0065	0.2676	0.4716
Water supply; sewerage, waste management	0.1609	0.005	0.684
Construction	0.246	0.0029	0.5292
Wholesale and retail trade; repair of motor vehicles	0.0808	-0.0095	0.8627
Transportation and storage	0.4445	0.1327	0.3315
Accommodation and food service activities	0.3331	-0.0269	0.694
Information and communication	0.2502	0.0265	0.6372
Financial and insurance activities	-0.08	0.0064	0.9529
Real estate activities	-0.1753	0.2195	0.6612
Professional, scientific and technical activities	0.178	0.0733	0.617
Administrative and support service activities	0.148	-5e-04	0.8361
Education	0.6176	0.0239	0.1754
Human health and social work activities	0.5411	0.0706	0.1795
Arts, entertainment and recreation	0.3231	0.1506	0.2562
Other service activities	0.1705	-0.0439	0.785

Note: The table reports coefficients of the production function estimated for each industry (ISIC rev. 4 “sections” following the methodology described in Section 5.5.2.