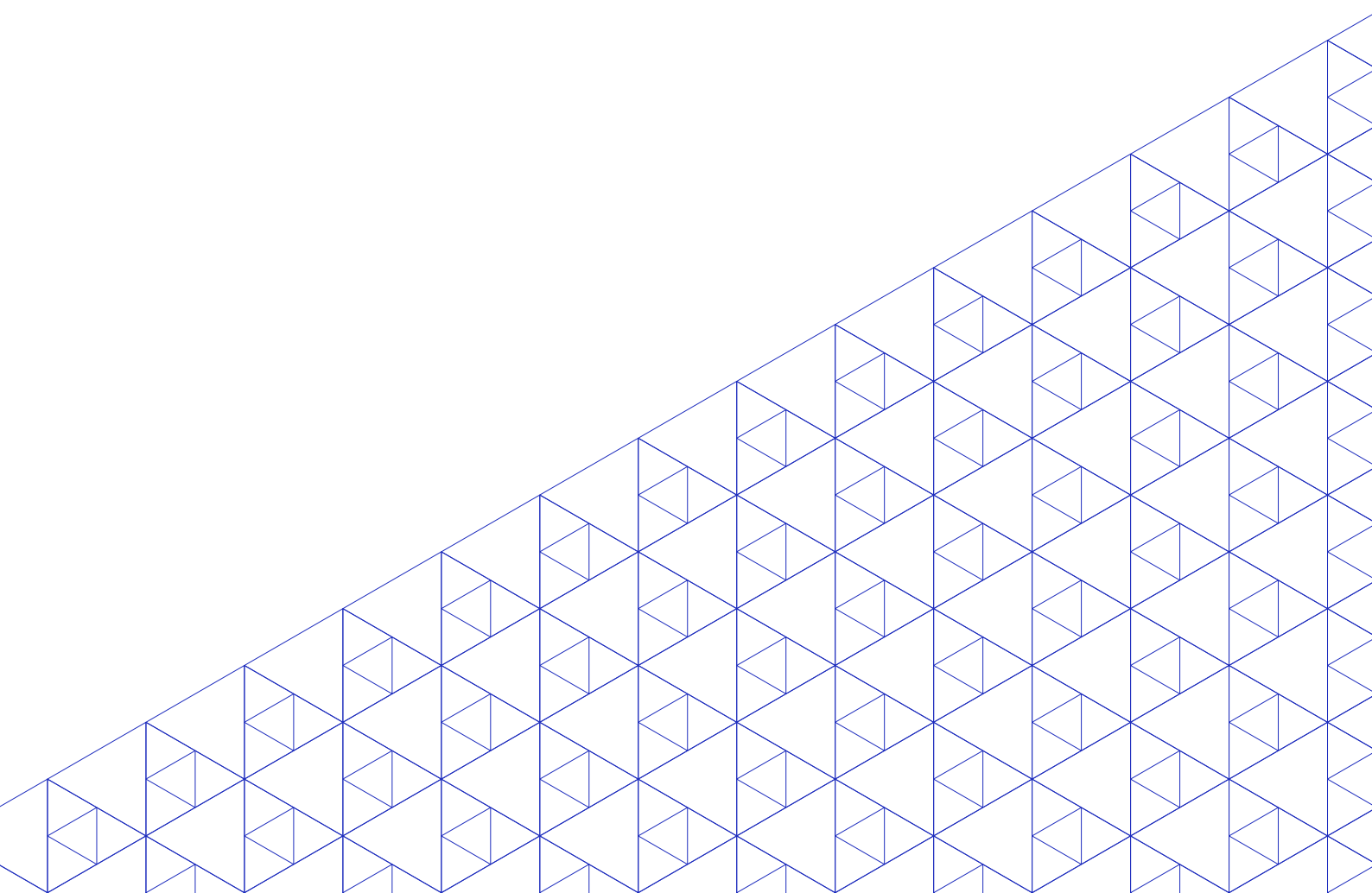




► Mobile internet, skills and structural transformation in Rwanda

Authors / Bernardo Caldarola, Marco Grazzi, Martina Occelli, Marco Sanfilippo





This is an open access work distributed under the Creative Commons Attribution 3.0 IGO License (<http://creativecommons.org/licenses/by/3.0/igo>). Users can reuse, share, adapt and build upon the original work, even for commercial purposes, as detailed in the License. The ILO must be clearly credited as the owner of the original work. The use of the emblem of the ILO is not permitted in connection with users' work.

Translations – In case of a translation of this work, the following disclaimer must be added along with the attribution: *This translation was not created by the International Labour Office (ILO) and should not be considered an official ILO translation. The ILO is not responsible for the content or accuracy of this translation.*

Adaptations – In case of an adaptation of this work, the following disclaimer must be added along with the attribution: *This is an adaptation of an original work by the International Labour Office (ILO). Responsibility for the views and opinions expressed in the adaptation rests solely with the author or authors of the adaptation and are not endorsed by the ILO.*

All queries on rights and licensing should be addressed to ILO Publications (Rights and Licensing), CH-1211 Geneva 22, Switzerland, or by email to rights@ilo.org.

ISBN: 9789220369470 (print)
ISBN: 9789220369487 (web-pdf)
ISBN: 9789220369494 (epub)
ISBN: 9789220369500 (mobi)
ISBN: 9789220369517 (html)
ISSN: 2708-3446

DOI URL: <https://doi.org/10.54394/XSTK4695>

The designations employed in ILO publications, which are in conformity with United Nations practice, and the presentation of material therein do not imply the expression of any opinion whatsoever on the part of the International Labour Office concerning the legal status of any country, area or territory or of its authorities, or concerning the delimitation of its frontiers.

The responsibility for opinions expressed in signed articles, studies and other contributions rests solely with their authors, and publication does not constitute an endorsement by the International Labour Office of the opinions expressed in them.

Reference to names of firms and commercial products and processes does not imply their endorsement by the International Labour Office, and any failure to mention a particular firm, commercial product or process is not a sign of disapproval.

ILO Working Papers summarize the results of ILO research in progress, and seek to stimulate discussion of a range of issues related to the world of work. Comments on this ILO Working Paper are welcome and can be sent to RESEARCH@ilo.org, liepmann@ilo.org.

Authorization for publication: Richard Samans, Director RESEARCH

ILO Working Papers can be found at: www.ilo.org/global/publications/working-papers

Suggested citation:

Caldarola, B., Grazzi, M., Occelli, M., Sanfilippo, M. 2022. *Mobile internet, skills and structural transformation in Rwanda*, ILO Working Paper 60 (Geneva, ILO).

Abstract

This paper examines the relationship between mobile internet, employment and structural transformation in Rwanda. Thanks to its ability to enable access to a wide range of ICT technologies, internet coverage has the potential to affect the dynamics and the composition of employment significantly. To demonstrate this, we have 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 to 2019. Our results show that an increase in mobile internet coverage affects the labour market in two ways. First, by increasing employment opportunities. Second, by contributing to changes in the composition of the labour market. Education, migration and shifts in demand are all instrumental in explaining our findings.

About the authors

Bernardo Caldarola is a PhD candidate at the Science Policy Research Unit, University of Sussex, Brighton, United Kingdom. His doctoral research looks at the drivers of structural transformation in sub-Saharan Africa, with a particular focus on the role of informal firms and workers. His research bridges several academic fields, such as development economics, economic geography and complexity economics. He has also collaborated with several international organizations and research centres, including the former UK Department for International Development, the Institute of Development Studies and the International Fund for Agricultural Development.

Marco Grazi is professor of Economic Policy at Università Cattolica del Sacro Cuore, Milan. His research focuses on the role of firms in shaping aggregate industry and country dynamics. To this end, his works embrace a number of fields, such as international trade, firm growth and its relation to innovation. He is currently working on the role of the latest wave of innovations (robot, artificial intelligence (AI) and digital technologies) in shaping firm and employment dynamics. He earned his PhD at the Scuola Superiore Sant'Anna in Pisa in 2006. Over the years he has been a visiting scholar at several institutions, namely Wharton School, University of Pennsylvania, Swiss Federal Institute of Technology Lausanne (EPFL), University of Cambridge and University of Notre Dame, Indiana.

Martina Ocelli is a Postdoctoral Associate in the Department of Global Development at the Cornell University and a Research Affiliate with the Department of Excellence in Economics and Management in the Era of Data Science (EMbeDS) at the Sant'Anna School of Advanced Studies in Pisa. She holds a PhD in Agricultural Economics from the Scuola Superiore Sant'Anna and a joint MSc in Development Economics from the Pavia University and Hohenheim University. Her work advocates for farmers' engagement with the scientific debate and she is currently focusing on the impact of new technologies on the agricultural sector.

Marco Sanfilippo is an Associate Professor at the University of Turin and a Research Affiliate at the Collegio Carlo Alberto. He holds a PhD in Development Economics from the University of Florence. His work explores the drivers of private sector development in low-income countries using firm- and individual-level data. He has published extensively on these topics and consults regularly with national and international organizations.

Table of contents

Abstract	01
About the authors	01
<hr/>	
▶ Introduction	05
<hr/>	
▶ 1 Data	08
1.1 Mobile internet	08
1.2 Individual-level data	09
<hr/>	
▶ 2 Empirical specification	16
<hr/>	
▶ 3 Results and discussions	18
3.1 Robustness checks	22
3.2 Mechanisms and extensions	24
<hr/>	
▶ Conclusion	29
<hr/>	
Annex	30
References	39
Acknowledgements	41

List of Figures

Figure 1. District-level mobile internet 3G coverage diffusion (2008–18)	09
Figure 2. Share of skilled and unskilled workers over time among the working-age population, average across districts (selected years, 2002–19)	11
Figure 3. Percentage of skilled workers at district level (2002, 2012, 2019)	12
Figure 4. Share of people working in skilled occupations and mobile internet 3G coverage (2012 and 2019)	13
Figure 5. Industrial employment distribution, average across districts (selected years, 2002–19)	14
Figure 6. Relationship between mobile internet 3G coverage and the share of people with tertiary education (2012 and 2019)	15
Figure 7. Results by industries within the services sector	22
Figure 8. Event study	24
Figure 9. Share of migrants in each district in relation to the degree of 3G internet coverage (2012 and 2019)	26
Figure 10. Distribution of firms' productivity for two levels of diffusion of mobile internet 3G coverage (2011 and 2019)	28
Figure A.1. Geographical distribution of the average number of years of education (2002 and 2019)	30
Figure A.2. Mobile phone ownership in relation to the increase in 3G coverage (2017 and 2019)	31
Figure A.3. Having an internet connection at home in relation to the increase in 3G coverage (2017 and 2019)	32
Figure A.4. Direct and indirect exports as share of sales, for companies in the manufacturing and services sectors (2011 and 2019)	33
Figure A.5. Share of unskilled production workers (full-time employed) in manufacturing in relation to the ISIC code of the company (Rev. 3.1) (2006 and 2019)	34

List of Tables

Table 1. Share of occupations, average across districts (selected years, 2002–19)	11
Table 2. OLS and 2SLS results, employment	19
Table 3. 2SLS results, by type of occupation and sector of employment	21
Table 4. 2SLS results, by education	25
Table 5. 2SLS results, migrant workers	27
Table A.1. Share of industries, average across districts	35
Table A.2. Summary statistics	36
Table A.3. Robustness, province-specific time trends	37
Table A.4. Robustness, initial conditions	38
Table A.5. Robustness, accounting for 4G coverage	38

▶ Introduction

Technological change has been at the root of structural transformation, and has had an undeniable impact on employment dynamics. In this respect, the pervasiveness of high-speed internet and complementary technologies has potentially disruptive effects on the composition of employment. Existing evidence from developed countries has shown that a higher complementarity between digital technologies and skills is leading to polarization in the labour market (see, among others, Autor and Dorn 2013, Goos et al. 2014, Autor 2015, Buera et al. 2021). Turning to developing economies, there is a general awareness of the relevance of technological change in either fuelling a catching-up process (see, among the others, the literature reviewed by Vivarelli 2021) or in causing deindustrialization (for its effect in combination with globalization, see Rodrik 2016). There is, however, little evidence relating to the impact on the labour market of information and communications technology (ICT) applications made available through internet access. This is in spite of the fact that some of the most transformative technologies are spreading rapidly across the developing world.

The diffusion of the internet is perhaps the most notable case. Over the past two decades, developing 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 2010.¹ 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 paper, we look at the expansion of the mobile internet network in Rwanda and analyse its implications in terms of changes in the labour market for workers.

The case of Rwanda is particularly interesting for the purposes of this study. The role of the ICT sector is deeply embedded in national development strategies. The country's industrial policy, grounded in its Vision 2020 strategy (MINICOM 2011), aims to diversify the economy and explicitly promotes 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 sectors, 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. By 2018, 4G mobile coverage had already reached over 96 per cent of the territory, with 47.7 per cent of the total population able to access the internet.²

The rapid rollout of the internet has the potential to transform the labour markets in different ways, which are difficult to assess a priori. The diffusion of internet coverage, along with the ICT and digital applications that it enables, can be considered a general purpose technology (GPT). A perspective on the process which addresses the peculiarities of developing countries is provided by Kaplinsky and Kraemer-Mbula (2022). They noted that mobile phones do not depend on a centralized grid, they are cheap, can be shared by more than one user and, focusing on a distinguishing feature of GPT, 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). In this respect, a warning is due: the wide range of innovations that might be enabled by internet connectivity is likely to exert contrasting effects on employment dynamics, so that anticipating the overall impact can be a difficult task. On the one hand, process innovation is typically supposed to be associated with labour-saving effects, possibly mitigated by a "compensation effect" when lower prices stimulate demand, thus generating a need for additional workers. On the other hand, the positive effects of product

¹ World Bank, "World Bank Development Indicators", accessed 8 November 2021.

² Ministry of ICT and Innovation (MINICT), "Digital Transformation Directorate General: Mandate" accessed 22 March 2022.

innovation on employment appear to be less in dispute, having been shown empirically for a set of sub-Saharan countries, by Avenyo et al. (2019), among others.

In fact, internet connectivity is the gateway to many innovations and can be the engine of both labour-augmenting and labour-saving technological change. Greater connectivity affects labour productivity directly, in a labour-biased way, and can aid human capital accumulation, by increasing training opportunities (both on the job and in educational settings). Evidence summarized by Hjort and Tian (2021) shows that improved access to the internet has been found to enhance firms' productivity (India), workers' wages (Brazil) or both (China).³ Mobile technologies 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 the United Republic of 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 the evidence the authors provide relating to the demand side. In fact, in their analysis fast internet coverage promotes both the entry and the performance of more productive and technologically intensive firms. Undeniably, internet expansion unlocks the potential for firms to benefit from internet-enabled technologies, 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 in a number of developing countries (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. Evidence from several African countries shows that the arrival of fast internet has promoted firms' exports, with associated benefits for local employment (Hjort and Poulsen 2019).

Understanding whether similar mechanisms also apply to the case of Rwanda is therefore an empirical question that we aim to explore in this work. More specifically, in our analysis we link the rollout of mobile internet in Rwanda to a number of outcomes related to changes in the size and composition of employed individuals in the country. This includes a shift towards more highly skilled occupations and/or modern and higher value-added activities across sectors.

The analysis is based on the collection and harmonization 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 harmonization 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 – rollout of the 3G 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 lightning strikes as a factor influencing the distribution of the mobile network within the country.

³ There is evidence on the capacity of mobile internet diffusion to increase employment opportunities. For instance, a recent paper by Bahia et al. (2020) on Nigeria reports significant employment uptake following the rollout of mobile internet at the sub-national level.

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 being more affected, given the relatively small initial size of this group; (2) through a sectoral shift of employment towards services and, within that 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 rationalize 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 some initial evidence on demand-side mechanisms: using firm-level data from the World Bank Enterprise Survey (WBES), we show descriptively a prevalence of more productive firms, especially in the services sector, in locations with higher 3G coverage.

The remainder of the paper is structured as follows: section 1 introduces all the data used in the analysis; section 2 discusses the empirical specification and the identification strategy based on a 2SLS estimator; section 3 reports the main results and a set of robustness checks; and section 4 concludes.

▶ 1 Data

1.1 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.

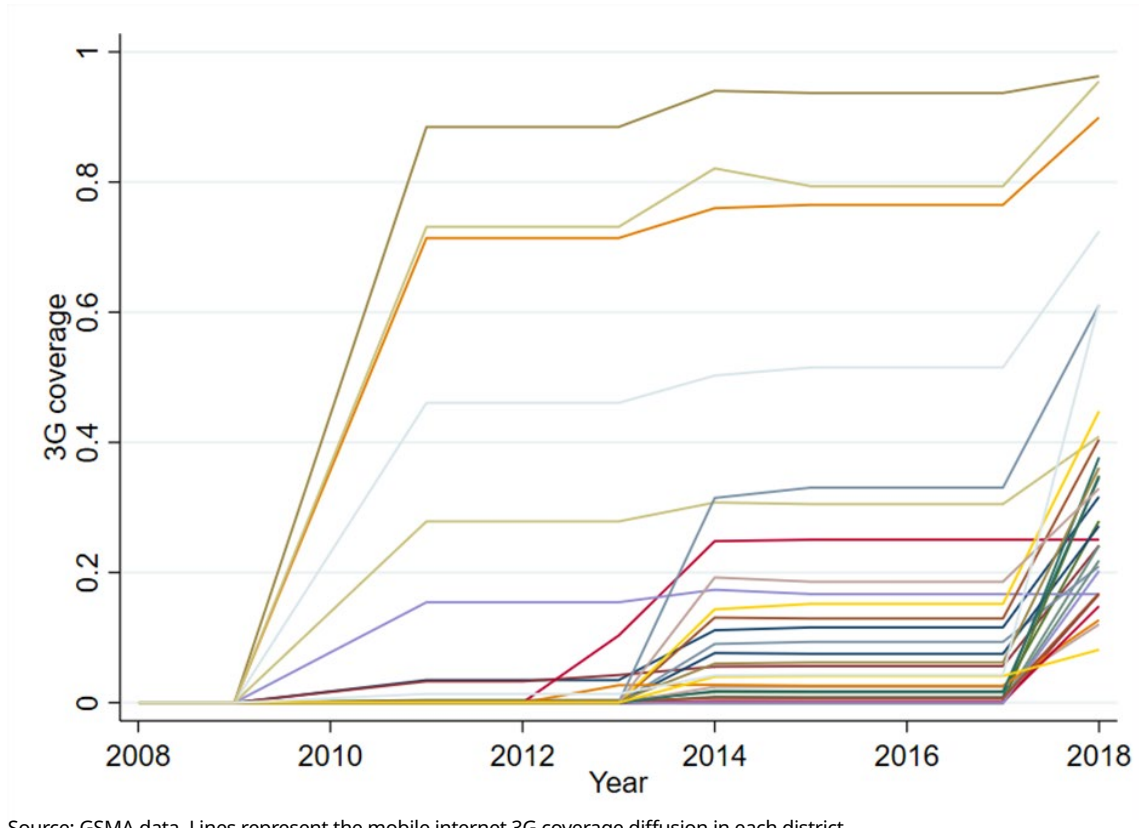
The original data consist of a raster of 1 km × 1 km cells, with a layer of information for each technology (2G, 3G, 4G). While 2G (GSM) supports voice calls and messaging, the main technologies of interest in our study are 3G and 4G, which support the use of mobile broadband internet services. In each 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.⁴ 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.⁵

Mobile internet technologies were introduced into Rwanda at the end of the 2000s. According to the GSMA data, before 2009 only 2G technology was available. After 2009, 3G internet technologies started to be introduced in a staggered manner across districts and over time (see figure 1). In contrast, the diffusion of the 4G network has been sudden. Developed in partnership with the South Korean firm, KT, the rollout 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,⁶ possibly due to costs.

⁴ NASA SEDAC, "Population dynamics", accessed 25 February 2022.

⁵ As the data cover only up to 2018, we assume no major changes in the following year.

⁶ A recent report of the Rwanda Utilities Regulatory Authority (RURA) shows that currently only 1.2 per cent of the total number of mobile broadband subscriptions use 4G technologies (see RURA 2021: Table 13).

▶ **Figure 1. District-level mobile internet 3G coverage diffusion (2008–18)**

Source: GSMA data. Lines represent the mobile internet 3G coverage diffusion in each district.

1.2 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 Census⁷ (from IPUMS International) with three waves (2017, 2018 and 2019) of the nationally representative Rwanda Labour Force Survey (RLFS)⁸ (from the National Institute of Statistics of Rwanda). We aggregate the individual-level information to obtain a district-level⁹ panel dataset on a sample restricted to the working-age population (which we define as covering individuals aged 15 to 64 years old).

⁷ 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 per cent 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.

⁸ The RLFS has been implemented in 2017, 2018 and 2019 to monitor the trend in employment and labour underutilization 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).

⁹ 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.

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,¹⁰ 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) (ILO, 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 harmonization 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).¹¹

Table 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.¹² 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 the country (figure 2) and by district (figure 3), with significant concentration in the urban districts, such as the capital, Kigali.

¹⁰ See the [19th International Conference of Labour Statisticians](#) (ILO 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).

¹¹ 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.

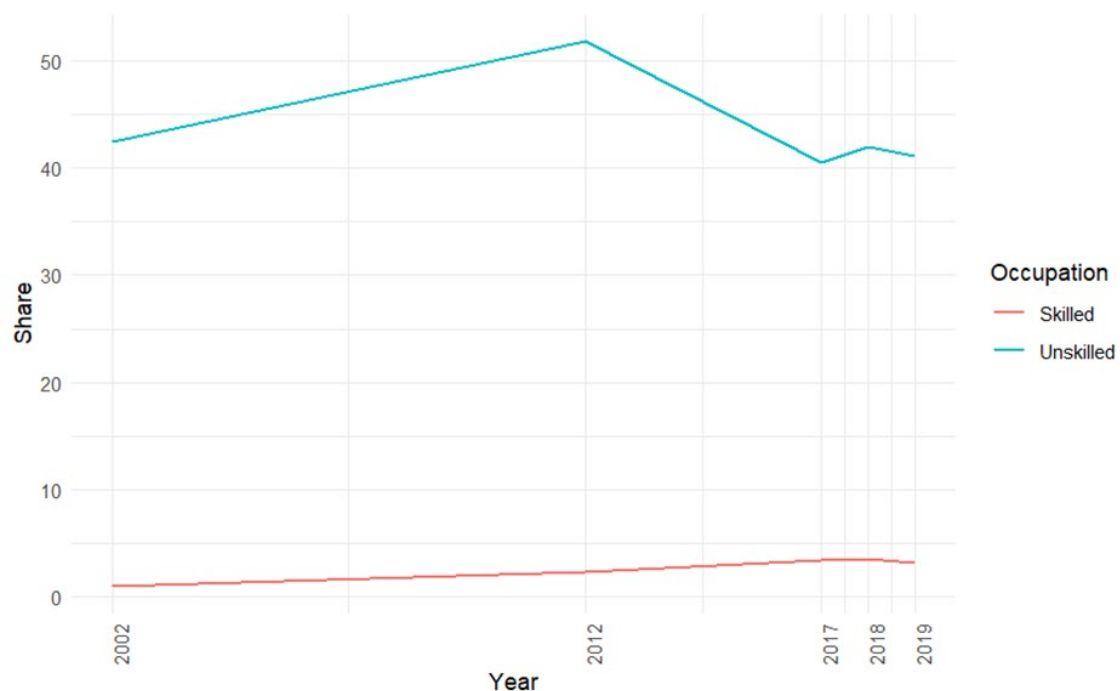
¹² For the full list of inclusions and exclusions of disaggregated occupations between ISCO 88 and ISCO 08, see CEDEFOP (2014, p. 26)

► **Table 1. Share of occupations, average across districts (selected years, 2002–19)**

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

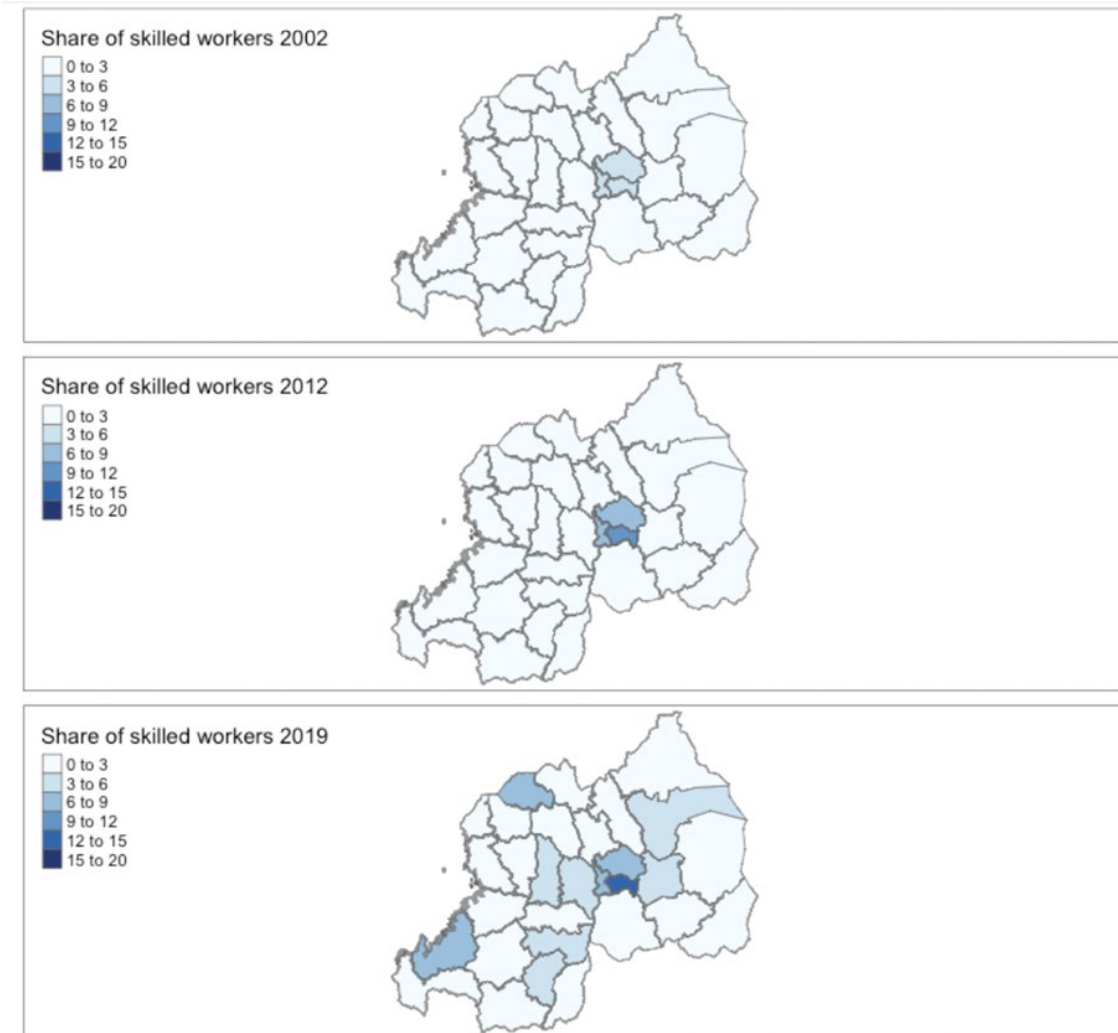
Note: * Not in employment refers to those individuals not currently working and to those working in subsistence farming.

Source: Authors' elaboration on national census and RLFS data.

► **Figure 2. Share of skilled and unskilled workers over time among the working-age population, average across districts (selected years, 2002–19)**

Note: Shares on the y-axis indicate the percentage of the working-age population employed in either skilled (red) or unskilled (blue) occupations.

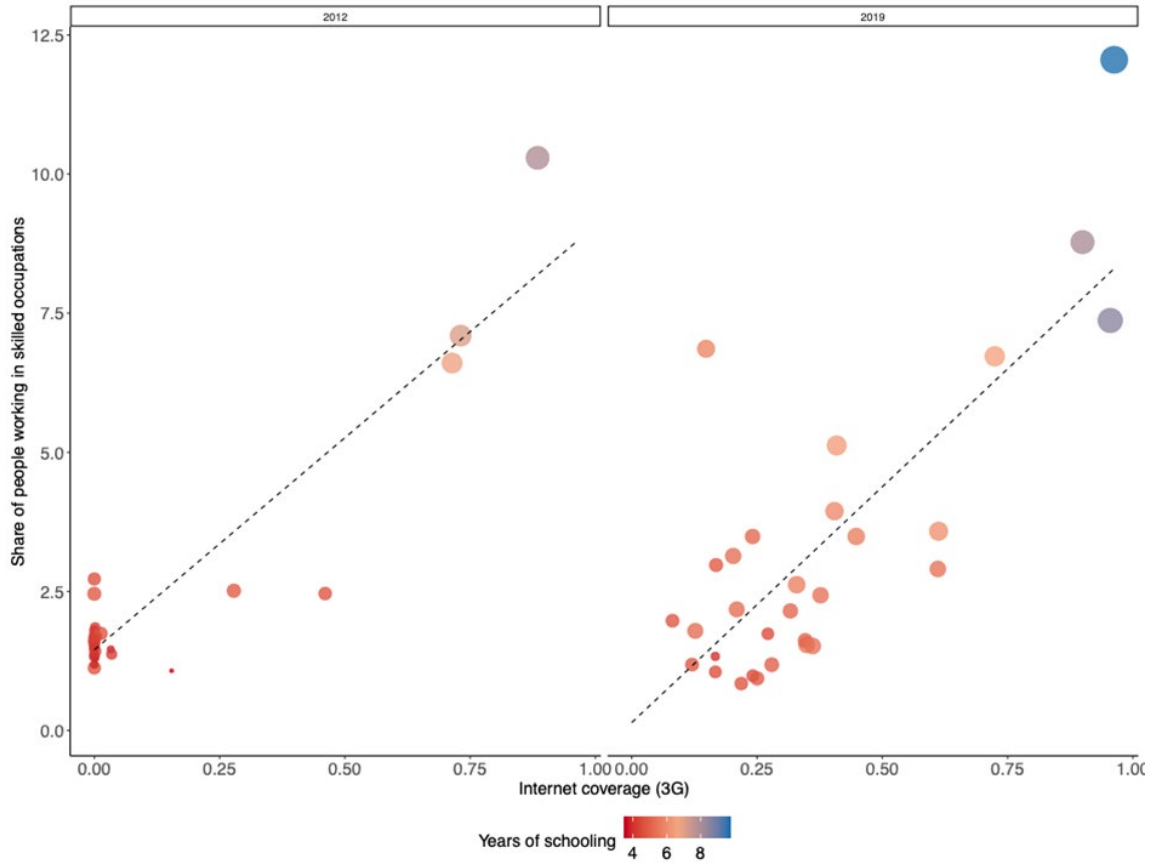
Source: Authors' elaboration on national census and RLFS data.

► **Figure 3. 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 4. 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.

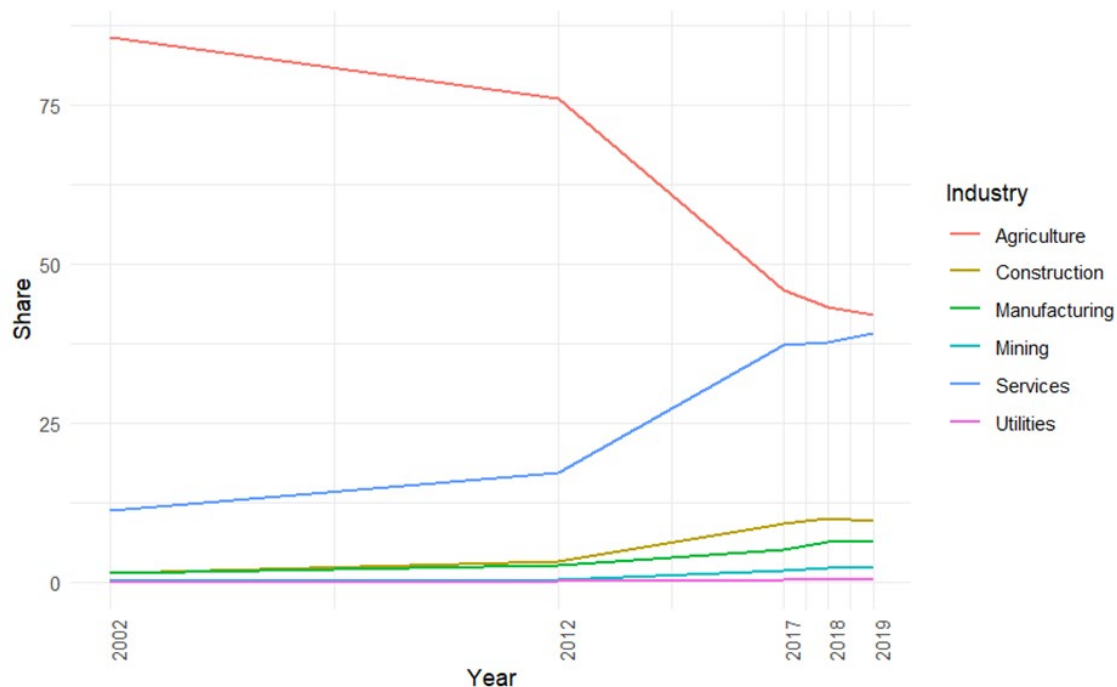
► Figure 4. Share of people working in skilled occupations and mobile internet 3G coverage (2012 and 2019)



Note: Darker colour and larger size indicate a higher degree of education associated with the skilled workers.

Source: Authors' elaboration on national census and RLFS data.

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 major sectors and industries. Table A1 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) 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 A1). Employment in manufacturing also increased, although its share remains relatively small.

► **Figure 5. Industrial employment distribution, average across districts (selected years, 2002–19)**

Note: Shares on the y-axis indicate the percentage of the working-age population in employment by industry. In each year, shares sum to 100.

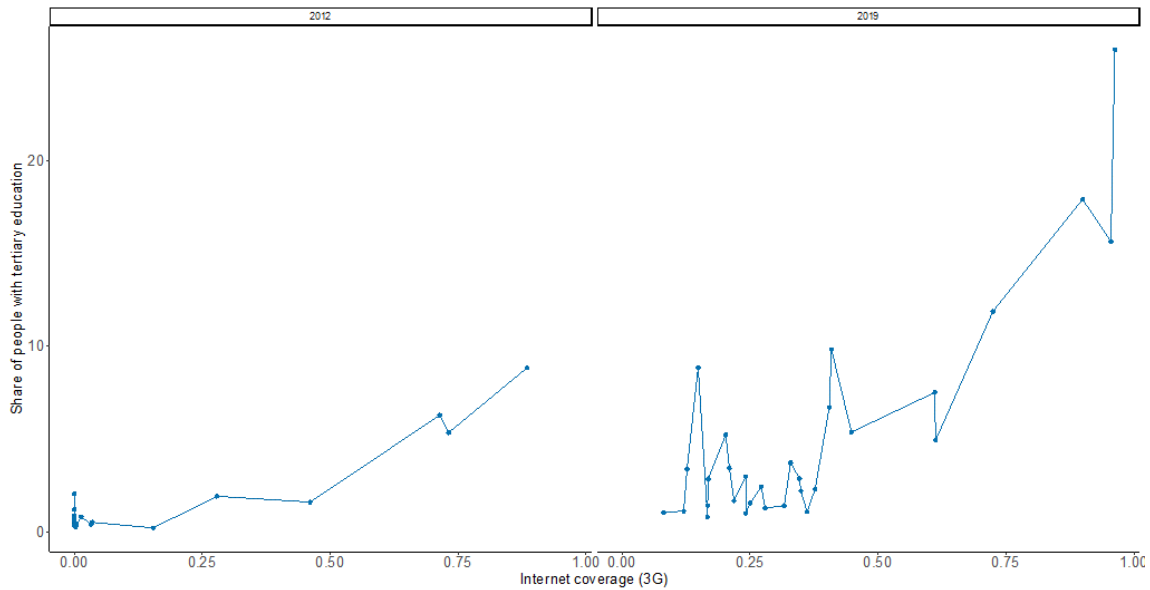
Source: Authors' elaboration on national census and RLFS data.

Education: Education-related questions in the census and in the labour force surveys are not harmonized. 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 harmonized 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.¹³ For the entire period covered (2002–19), the enrolment 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 enrolment in upper secondary classes, with a jump from 21 per cent in 2011 to 30 per cent in 2017 (Nkurunziza 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.

¹³ 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 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.

Average years of education have increased between 2002 and 2019 across districts (see figure A.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 6), we observe a positive correlation.

► **Figure 6. Relationship between mobile internet 3G coverage and the share of people with tertiary education (2012 and 2019)**



Note: The graph illustrates the relationship between the percentage of mobile internet coverage in Rwanda and the share of people with tertiary education, in 2012 and 2019. A higher degree of fast mobile internet coverage relates positively with a higher percentage of individuals who have completed tertiary education, with the relationship being stronger in 2019.

Source: Authors’ elaboration on national census, RLFS and GSMA data, Rwanda.

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 A.2 in the Appendix) and (b) internet connections at home (figure A.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.¹⁴

¹⁴ We define an urban district as a district where at least 60 per cent of the population lives in urban areas. In the years analysed, these are Gasabo, Nyarugenge and Kicukiro.

▶ 2 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 (Guriev et al. 2021, Manacorda and Tesei 2020) and links the rollout of mobile phone coverage to the outcomes of interest, as follows:

$$Y_{it} = \beta 3G_{it} + \gamma X'_{it} + \theta_i + \delta_t + \varepsilon_{it} \quad (1)$$

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 1.2. Third, the distribution of workers across sectors. 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 sectors.

Following the discussion in section 1, 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 be exploited. In contrast, the timing of the introduction of the 3G technologies is ideal to be combined with labour force data. While the technology was formally introduced in the early 2000s, the rollout covered only a few districts before 2012, and even those had very limited coverage (see figure 1). After 2012, coverage expanded to other districts as well, 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. We also add variables that account for the geographic characteristics of the district¹⁵ (as in Manacorda and Tesei 2020).

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 also 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 rollout 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.

¹⁵ These are mainly time invariant controls that, as in Manacorda and Tesei (2020), we introduce interacted with a time trend. The variables are the following: the natural logarithm of the geographic distance from the district centroid to the closest point on the national border; the natural logarithm of the geographic distance from the district centroid to the closest point on the coastline; the natural logarithm of the geographic distance from the district centroid to the closest point on a colonial railway; the mean stability of malaria transmission in a district; the mean agricultural suitability of the district; terrain ruggedness. All these variables come from Alesina et al. (2021, p. 16).

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 A.2 in the Appendix.

Identification strategy: Equation (1) will be correctly identified under very restrictive conditions, i.e. that the rollout of the broadband network is not influenced by existing pre-trends, 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 prioritize 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.

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 rollout of the mobile network in contexts which, similar to ours, exploit sub-national 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).

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-arc-minute cells (around 8 km × 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 cell¹⁶ and aggregated cell values by district, taking their mean. The resulting measure of daily lightning strikes per square km in every district is then converted into daily lightning strikes per inhabitant¹⁷ 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).

¹⁶ 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.

¹⁷ 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.

▶ 3 Results and discussions

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 organize the discussion of the main results into three different sets of outcomes: employment, occupations and sectors.

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

► Table 2. OLS and 2SLS results, employment

Variables	(1) OLS	(2) OLS	(3) First-stage 2SLS	(4) Second-stage 2SLS
3G	0.0632** (0.0251)	0.0475* (0.0265)		0.341*** (0.0710)
2G		0.00247 (0.117)		-0.109 (0.112)
Age		-0.0135 (0.00848)		-0.0112 (0.0103)
Female		0.248 (0.249)		-0.456 (0.273)
Malaria*t		4.28e-06 (0.00557)		-0.00899 (0.00548)
Ruggedness*t		-0.000106*** (3.23e-05)		-0.000132*** (4.19e-05)
Agricultural suitability*t		-0.0259 (0.0207)		-0.0112 (0.0281)
Distance to coast*t		0.0393 (0.0909)		-0.0567 (0.0853)
Distance to railway*t		-0.0608*** (0.0202)		-0.0453* (0.0251)
Distance to border*t		0.00132 (0.00496)		-0.00639 (0.00524)
Lightning strikes (1,000 pop.)*t			-4.030*** (1.155)	
Constant	0.450*** (0.00595)	1.019 (1.789)	-11.71 (7.430)	
Observations	150	150	150	150
R-squared	0.062	0.684	0.863	-0.401
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
F-stat				15.93

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.

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).¹⁸ 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 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.

Sectors: Next, we check whether the rollout of mobile internet correlates with the process of structural transformation occurring across sectors 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 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 rollout of mobile internet coverage would therefore be of relevance. Results of this exercise, reported in columns 3 to 5 of table 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 7. While a positive coefficient is generally found for most of the industries within the services 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.

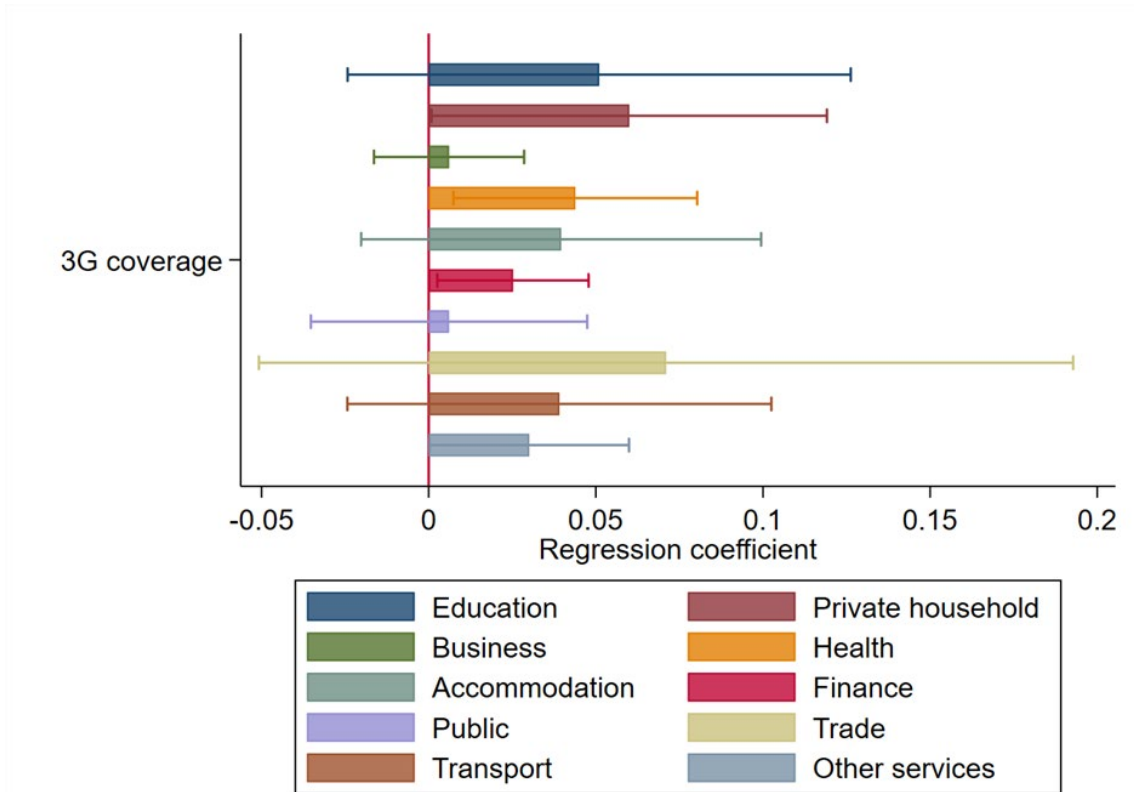
¹⁸ See their table A.III, reporting OLS coefficients up to 10 times smaller than the 2SLS.

► **Table 3. 2SLS results, by type of occupation and sector of employment**

	(1)	(2)	(3)	(4)	(5)
Variables	Skilled	Unskilled	Agriculture	Manufacturing	Tertiary sector
3G	0.0654*** (0.0206)	0.276*** (0.0739)	0.0957 (0.195)	0.0408 (0.0519)	0.259** (0.124)
Observations	150	150	150	150	150
R-squared	0.198	-0.356	0.080	0.016	0.073
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	15.93	15.93	15.93	15.93	15.93

Notes: 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 7. Results by industries within the services sector



Note: 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.

3.1 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 A.3 in the Appendix, 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 rollout of the 3G network, may experience different trajectories. Results, reported in table A.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.

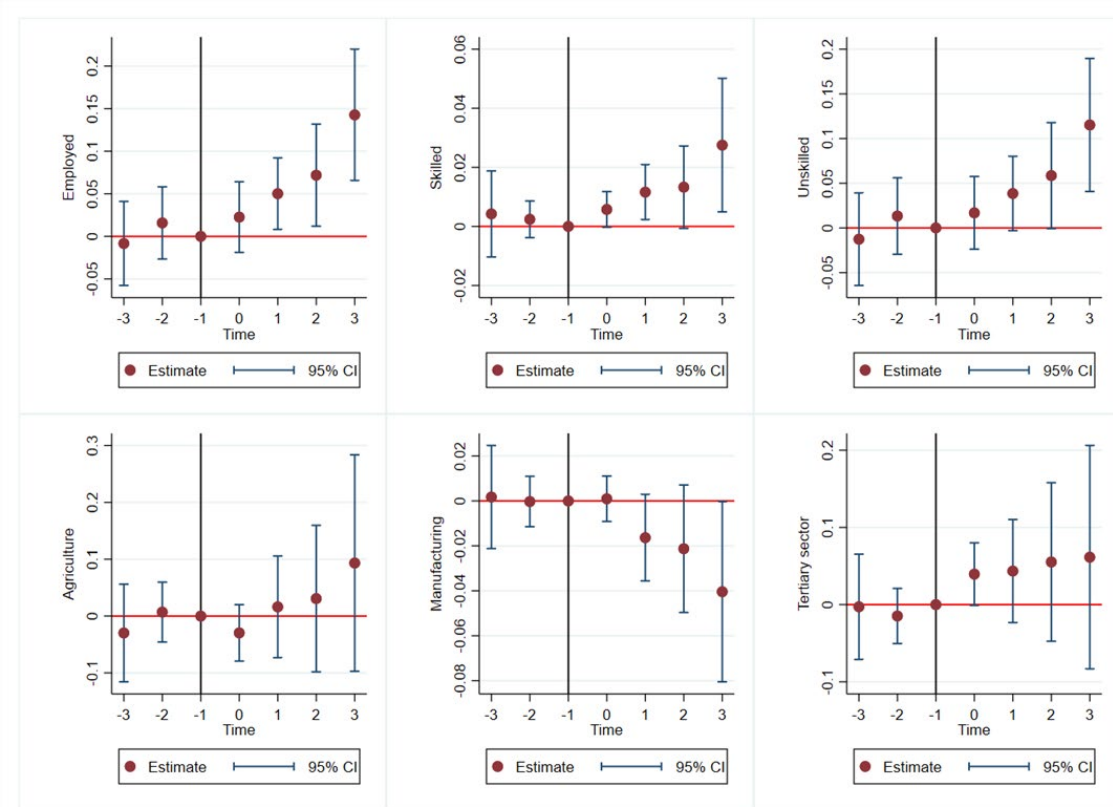
4G coverage: As discussed in section 1, a potential issue of concern for our analysis is that most of our results are driven (or strengthened) by the introduction of 4G technology, which mainly occurred during the second half of the 2010s. The policy leading to the almost universal rollout of 4G coverage, 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 A.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 randomized, but outcomes and trajectories before and after treatment, as well as across treated and control units, can be compared. A few important methodological issues have recently been discussed in relation to the design of event studies (for a review of the main issues, see Borusyak et al. 2021). One potential issue of concern, in our setting, is due to the continuous definition of the treatment (the share of a district's population being covered by 3G technologies) as well as the possibility that the same district may be treated more than once, given that the coverage of mobile internet increases over time. In order to avoid some of the potential biases arising from these issues, 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).¹⁹ We estimate the event study based on equation 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 8 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 potentially long-lasting impacts of mobile technologies.

¹⁹ Results remain robust to alternative definitions of this threshold, including 20 per cent and 50 per cent.

► Figure 8. Event study



Note: The event study design uses the first year in which a district hits 11 per cent 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 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 per cent confidence interval (CI). The graphs have been created using the STATA command eventdd.

3.2 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 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 of school age (i.e. 5 to 25 years of age) at the time of the 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 4. 2SLS results, by education**

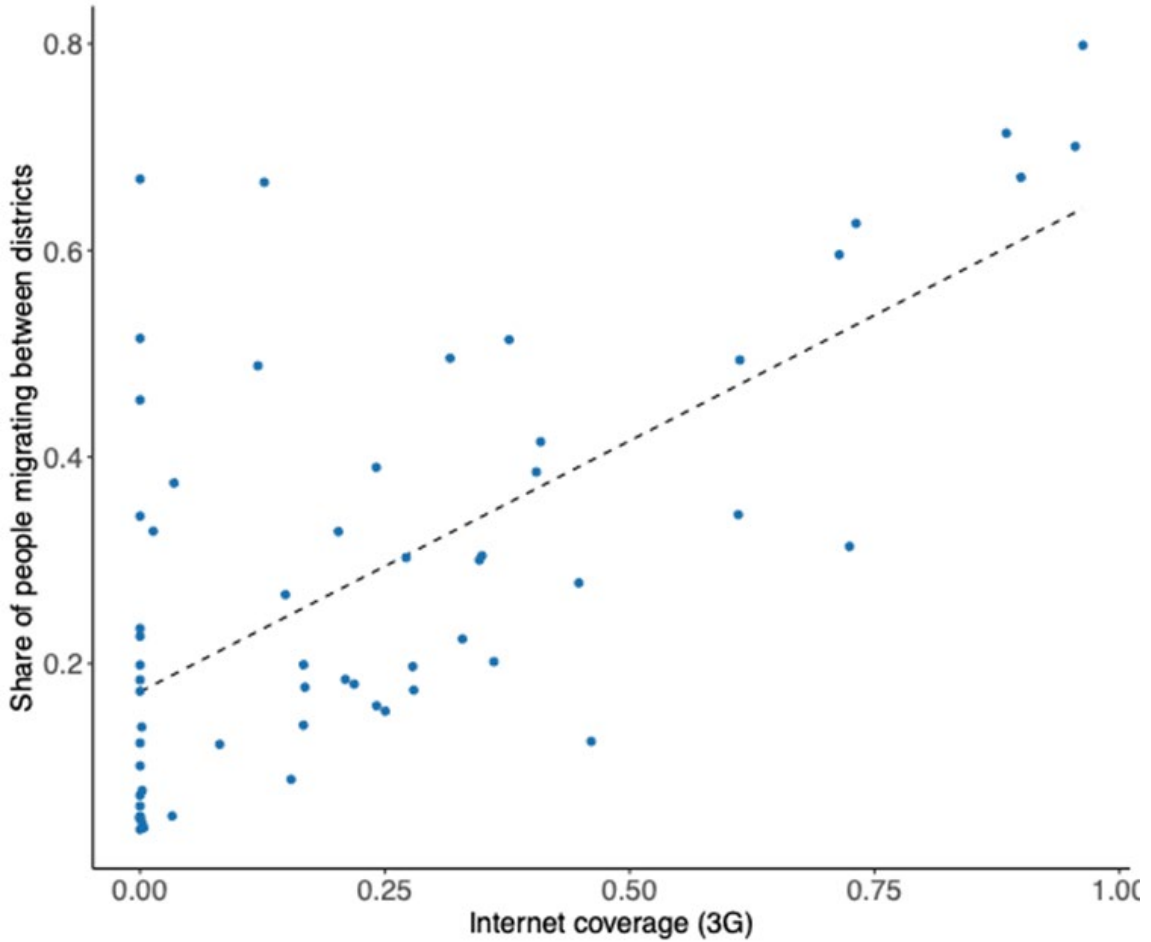
Variables	(1) Primary or no education	(2) Secondary	(3) Tertiary	(4) Years of schooling
3G	-0.457*** (0.0842)	0.373*** (0.0818)	0.0838*** (0.0217)	0.0150*** (0.00517)
Observations	150	150	150	150
R-squared	-0.468	-0.624	0.129	0.067
District FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
District controls	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	15.93	15.93	15.93	15.93

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 developing 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 9). 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.

► Figure 9. Share of migrants in each district in relation to the degree of 3G internet coverage (2012 and 2019)



Note: The graph reports the share of people migrating between districts and the increase in fast internet coverage (3G).

Source: Authors' elaboration on national census and RLFS data.

We combine the information on migrations 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. As we can distinguish the origin of a migrant, 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 sectors (both manufacturing and services).²⁰

²⁰ These additional results, which are not reported for reasons of space but are available upon request, are based on a more restrictive definition of migration, i.e. individuals coming from a different province.

► Table 5. 2SLS results, migrant workers

Variables	(1) Migrants from other districts	(2) Migrants from other provinces
3G	0.264** (0.108)	0.190** (0.0831)
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	10.80	10.80

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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

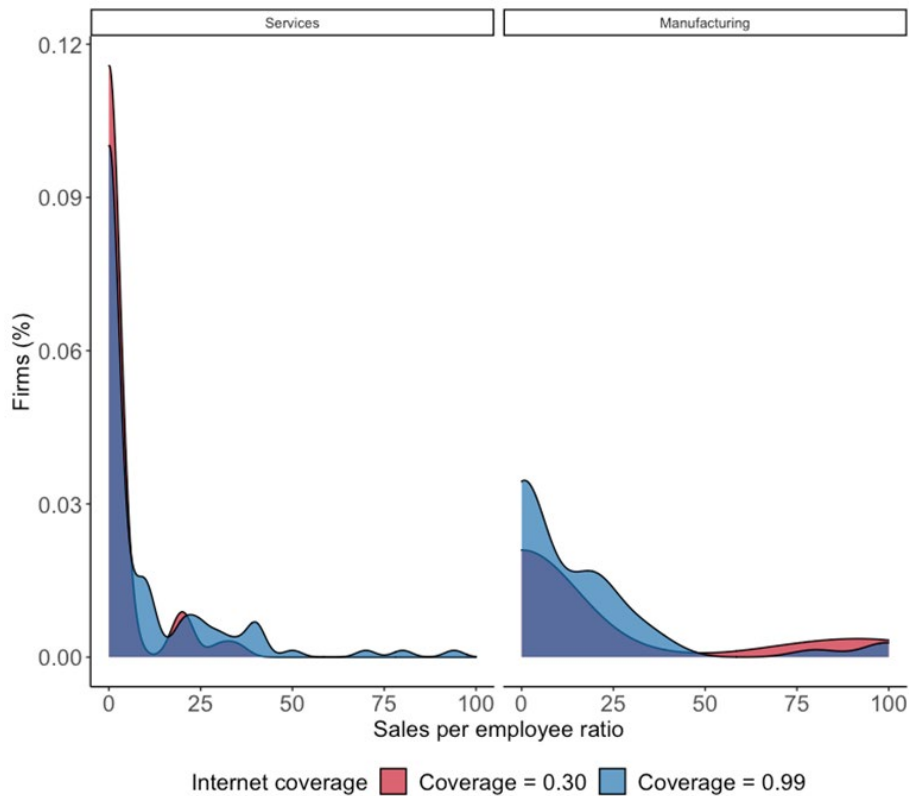
Demand-side mechanisms: The WBES data allow us to scope descriptively possible demand-side mechanisms which might corroborate relationships that have already been identified by looking at supply-side dynamics. We are interested in understanding whether enterprises took advantage of the rollout of fast internet, possibly moving towards operations and routines which require the utilization of a mobile internet connection. Note that the descriptive nature of these exercises does not exclude the emergence of patterns driven by co-founding factors; however, we find the evidence informative of demand-side mechanisms that have contributed to the Rwandese structural transformation.

First, we focus on exports, as these are activities which signal the ability of a company to interact in a global market and often require a reliable internet connection for their operation. We calculated exports as a percentage of the total sales of the enterprise. The share of direct and indirect exports with respect to the total sales increased between 2011²¹ and 2019 (see figure A.4 in the Appendix). The relation holds for the manufacturing as well as the services sector. However, we observe that the average does not seem to increase. What changes is the right tail of the distribution, which is thicker. Second, we observe from the supply-side evidence that the share of people with tertiary education and the share of workers in highly skilled jobs tend to increase with more widespread 3G coverage. Descriptive evidence suggests that this could be linked to lower demand from firms for unskilled labour (see figure A.5 in the Appendix): considering only the manufacturing sector, we see that in all ISIC industries the share of unskilled productive workers decreased between 2011 and 2019.

²¹ We do not utilize 2006 as the baseline year for this exercise, as there are no data available for that year.

Finally, we merge the mobile internet coverage data with the WBES dataset to relate mobile internet diffusion with a proxy for productivity: sales per employee. As the WBES reports geographical data only at the provincial level for the area of Kigali, the Southern province and the Western province, we computed the mean value of the mobile internet coverage across districts²² within these three provinces for each of the WBES years, 2011 and 2019.²³ Further, we select two values of mobile internet coverage which are representative of low coverage (diffusion in 30 per cent of the territory) and high coverage (diffusion in 90 per cent of the territory). We find that higher diffusion of mobile internet coverage seems to correlate with better performances in terms of productivity (figure 10). The positive relationship is driven mainly by the service sector, while for manufacturing firms differences in internet coverage do not seem to have a clear-cut relationship with productivity.

► **Figure 10. Distribution of firms’ productivity for two levels of diffusion of mobile internet 3G coverage (2011 and 2019)**



Note: The graph shows the relationship between firms’ productivity and internet coverage. Internet coverage is here indicated with two cut-off values, which are representative of low internet coverage (i.e. internet coverage = 0.30, meaning that fast internet diffusion is present in 30 per cent of the territory) and high internet coverage (i.e. internet coverage = 0.99, meaning that fast internet diffusion is present in 99 per cent of the territory). The two panels of the graph group the firms by sector, dividing manufacturing (on the right) from services (on the left) firms.

Source: Authors’ elaboration on WBES data, Rwanda.

²² Taking the mean across districts in each year is, inevitably, averaging out differences between mobile internet coverage diffusion within each province: in case of Kigali, these differences are almost negligible (e.g. ±0.09 in 2019) while for the Southern and Western provinces differences are more considerable (e.g. ±0.15 in 2019 for the Southern and ±0.27 in 2019 for the Western province). Therefore, we select just two values of internet coverage and interpret this evidence cautiously.

²³ We do not include 2006, as the mobile internet coverage in that year was null or negligible in the three provinces under consideration.

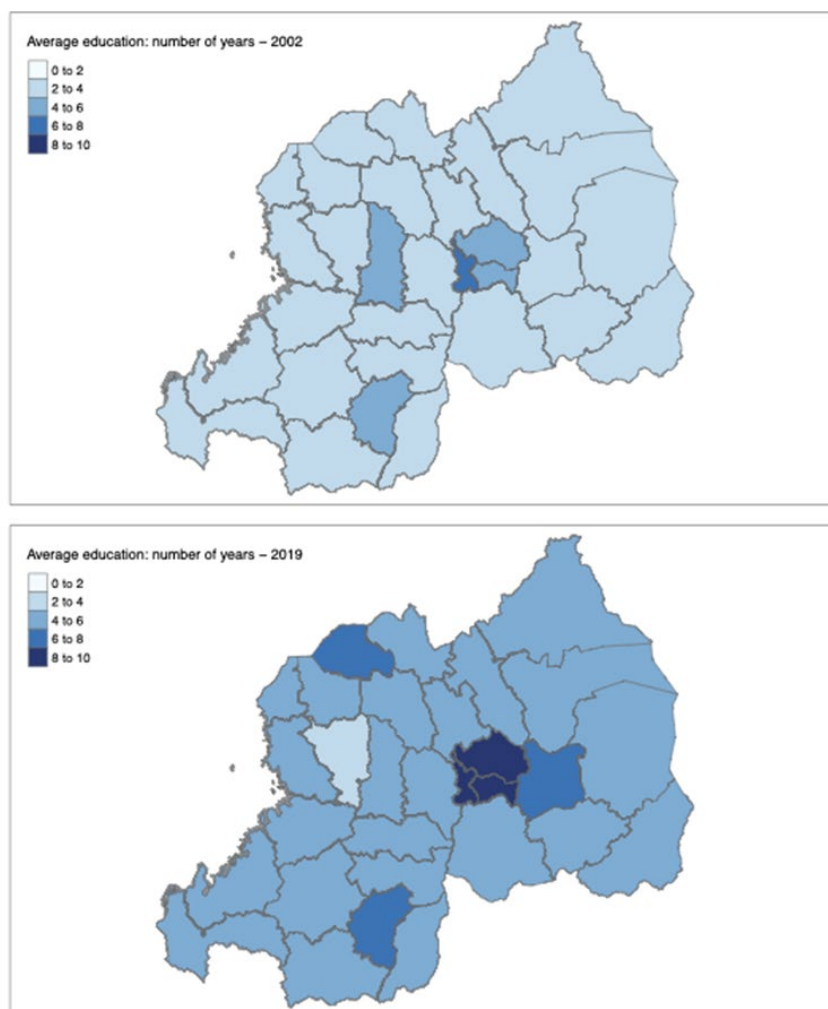
▶ Conclusion

As numerous developing economies are intensively exploiting the diffusion of high-speed internet technologies, the aim of this study is to investigate the effects of fast internet coverage on the labour market and on structural transformation.

Using Rwanda as a particular 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 seen in both skilled and unskilled types of occupations. Although the size of the coefficients is greater for the unskilled, the quantification exercise shows that mobile internet is relatively more important for skilled employment, given the initial lower share of the latter. Finally, districts that improved their internet connectivity are also those experiencing an increase of employment in services, especially in high-value-added sectors, such as finance and health. The estimations are robust to two different econometric specifications (IV and event study) as well as to a battery of robustness checks. In trying to rationalize 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, preliminary evidence seems to indicate an upgrade of firms in treated locations

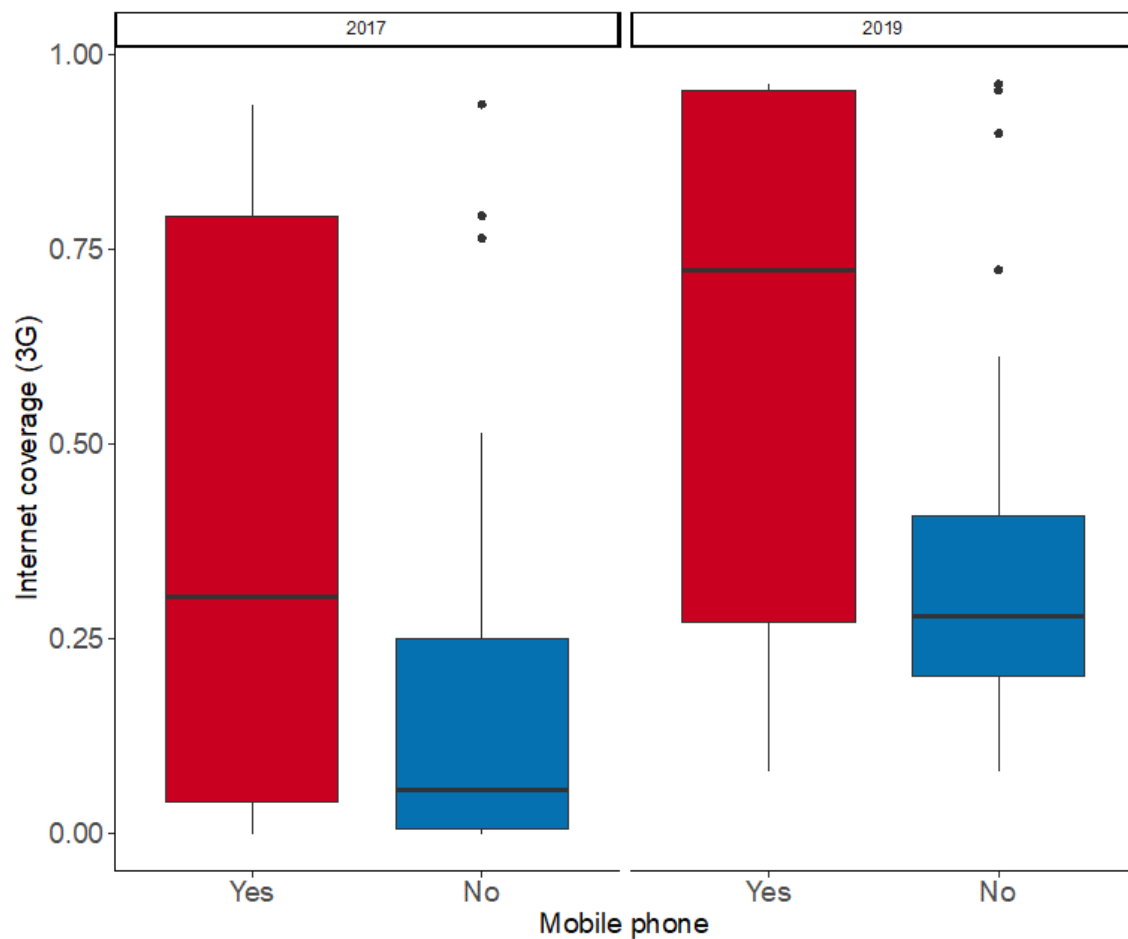
Annex

► Figure A.1. Geographical distribution of the average number of years of education (2002 and 2019)



Source: Authors' elaboration on RLFS data.

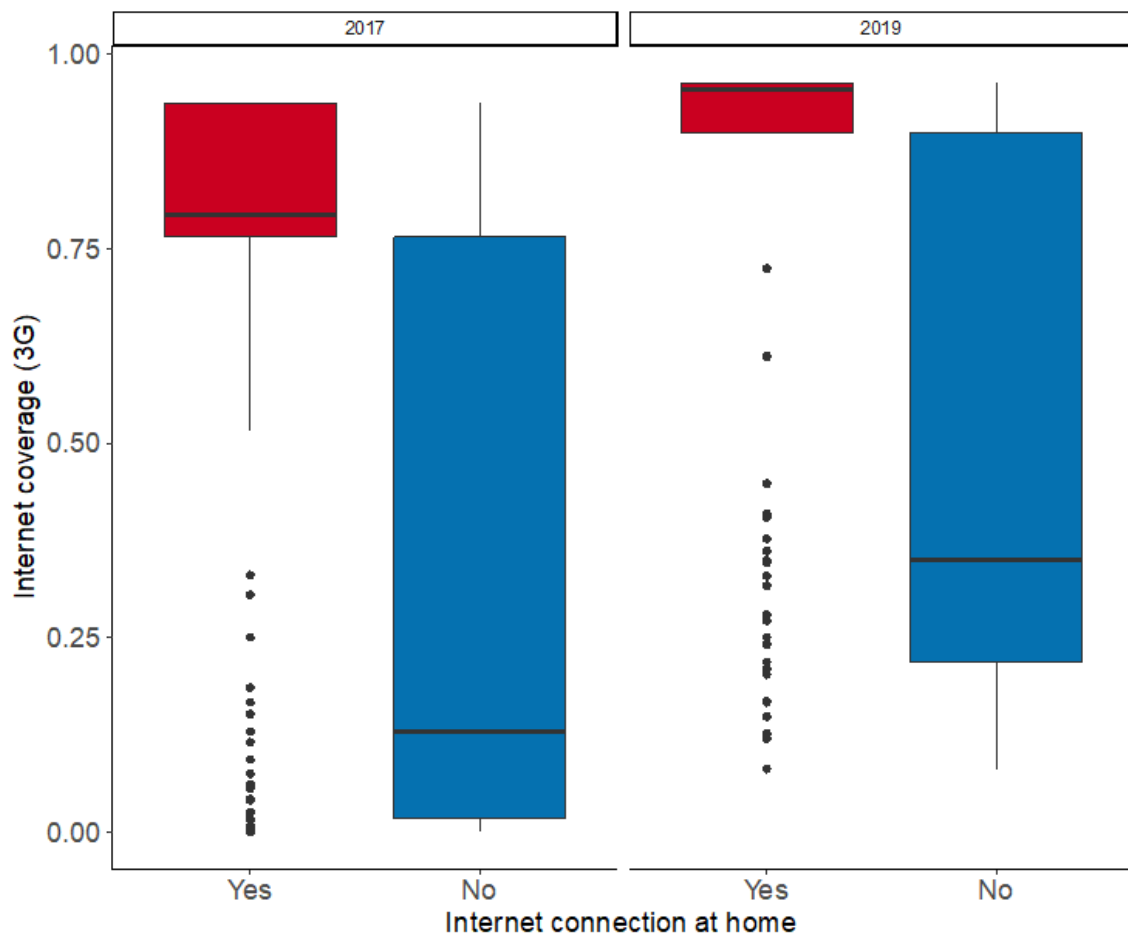
▶ Figure A.2. Mobile phone ownership in relation to the increase in 3G coverage (2017 and 2019)



Note: Red indicates urban districts, blue indicates rural districts.

Source: Authors' elaboration on RLFS data.

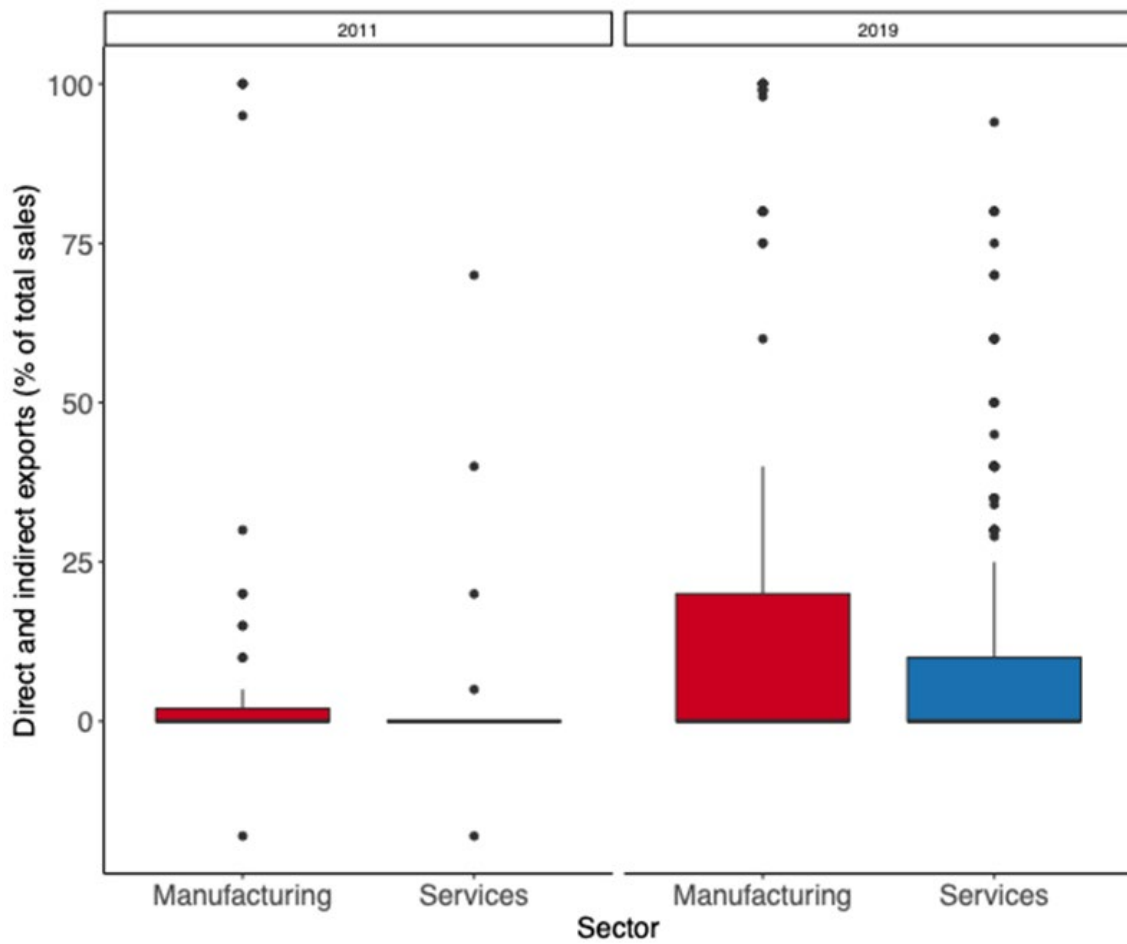
► Figure A.3. Having an internet connection at home in relation to the increase in 3G coverage (2017 and 2019)



Note: Red indicates urban districts, blue indicates rural districts.

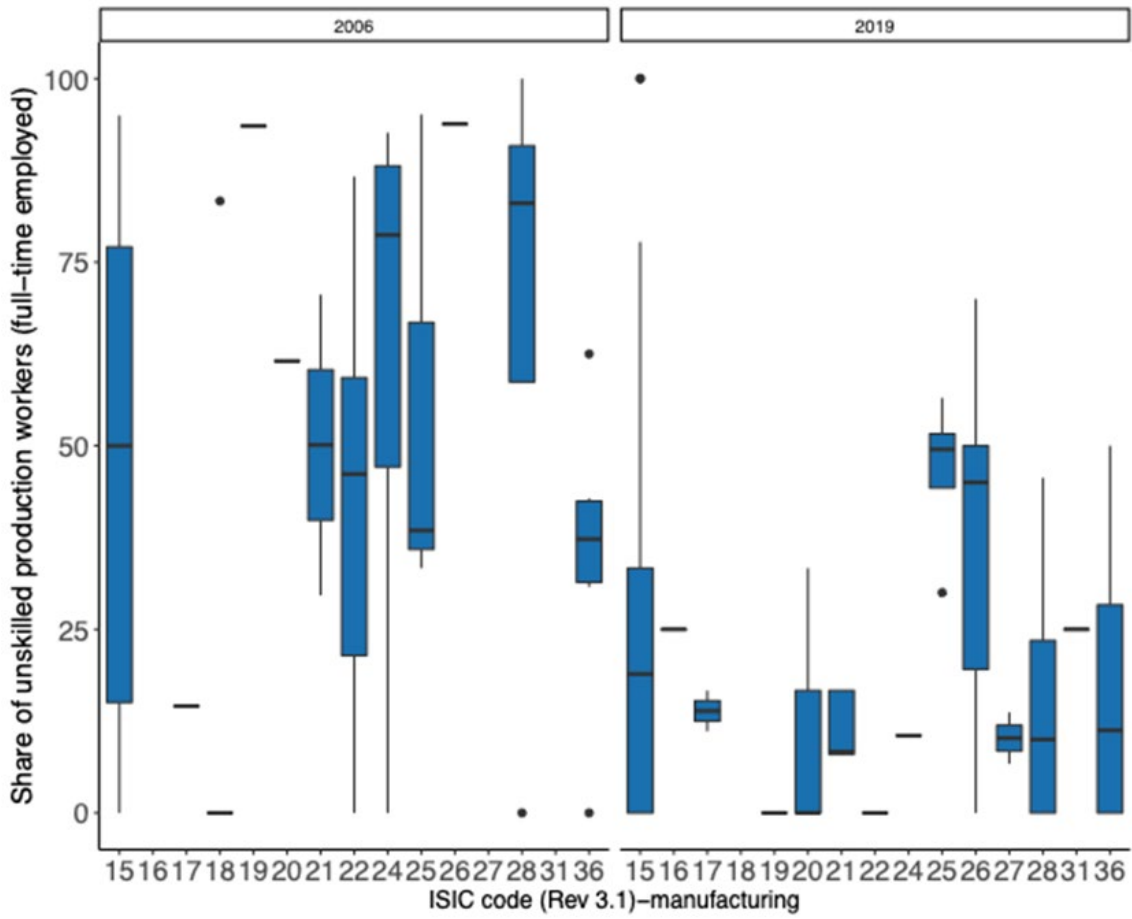
Source: Authors' elaboration on RLFS data.

► Figure A.4. Direct and indirect exports as share of sales, for companies in the manufacturing and services sectors (2011 and 2019)



Source: Authors' elaboration on RLFS data.

► Figure A.5. Share of unskilled production workers (full-time employed) in manufacturing in relation to the ISIC code of the company (Rev. 3.1) (2006 and 2019)



Source: Authors' elaboration on WBES data.

► **Table A.1. Share of industries, average across districts**

Industries	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

Source: Authors' elaboration on census and RLFS data.

► Table A.2. Summary statistics

Statistic	N	Mean	St. Dev.	Min.	Pctl(25)	Pctl(75)	Max.
Occupations							
Skilled	150	0.027	0.024	0.004	0.012	0.034	0.131
Unskilled	150	0.436	0.064	0.276	0.405	0.467	0.619
Employed	150	0.463	0.067	0.290	0.420	0.499	0.636
Industries							
Agriculture	150	0.585	0.258	0.048	0.396	0.839	0.970
Manufacturing	150	0.044	0.028	0.003	0.021	0.061	0.170
Services	150	0.285	0.199	0.024	0.105	0.400	0.783
Education							
Primary or less	150	0.812	0.151	0.369	0.708	0.961	0.995
Secondary	150	0.154	0.112	0.005	0.033	0.238	0.386
Tertiary	150	0.034	0.050	0.000	0.005	0.044	0.260
Years of education	150	5.032	1.300	2.910	4.112	5.688	9.731
Demographic							
Age	150	32.072	1.366	27.780	30.921	33.041	34.748
Female	150	0.531	0.022	0.448	0.520	0.546	0.578
Migration (district)	90	0.161	0.124	0.023	0.069	0.218	0.515
Migration (province)	90	0.112	0.107	0.008	0.028	0.180	0.472
Internet							
2G	150	0.773	0.386	0	0.8	1.0	1
3G	150	0.203	0.263	0	0	0.3	1
4G	150	0.497	0.416	0	0	0.9	1
Geographic							
Malaria stability	150	0.528	0.668	0.000	0.037	0.558	2.338
Terrain ruggedness	150	223.686	68.723	107.334	173.435	275.200	382.963
Agricultural suitability	150	0.495	0.123	0.308	0.418	0.563	0.827
Distance from coast	150	1,076.845	43.879	979.015	1,047.169	1,111.904	1,149.148
Distance from railway	150	234.493	28.792	165.686	210.827	259.344	281.594
Distance from capital	150	58.820	30.770	6.447	36.136	78.365	139.695

► Table A.3. Robustness, province-specific time trends

Variables	(1) Employed	(2) Skilled	(3) Unskilled	(4) Agriculture	(5) Manufacturing	(6) Tertiary sector
3G	0.448** (0.178)	0.0795** (0.0364)	0.369** (0.175)	0.0243 (0.233)	0.0714 (0.0774)	0.488*** (0.171)
Observations	150	150	150	150	150	150
R-squared	-0.342	-0.007	-0.257	0.208	0.021	0.226
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Year 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	5.413	5.413	5.413	5.413	5.413	5.413

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 A.4 Robustness, initial conditions**

Variables	(1) Employed	(2) Skilled	(3) Unskilled	(4) Agriculture	(5) Manufacturing	(6) Tertiary sector
3G	0.337*** (0.0710)	0.0620* (0.0315)	0.300*** (0.0862)	-0.336 (0.226)	0.137* (0.0796)	0.653*** (0.194)
Observations	150	150	150	150	150	150
R-squared	-0.381	0.221	-0.431	0.312	-0.426	-0.773
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 A.5 Robustness, accounting for 4G coverage**

Variables	(1) Employed	(2) Skilled	(3) Unskilled	(4) Agriculture	(5) Manufacturing	(6) Tertiary sector
3G	0.278*** (0.0928)	0.0691*** (0.0241)	0.209** (0.0960)	0.0214 (0.287)	0.0882 (0.0659)	0.362* (0.178)
4G	0.156 (0.130)	-0.00910 (0.0240)	0.166 (0.137)	0.183 (0.325)	-0.117* (0.0657)	-0.253 (0.195)
Observations	150	150	150	150	150	150
R-squared	-0.165	0.173	-0.114	0.143	-0.148	-0.086
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	13.22	13.22	13.22	13.22	13.22	13.22

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. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

References

- Alesina, A., S. Hohmann, S. Michalopoulos and E. Papapioannou. 2021. "Intergenerational Mobility in Africa". *Econometrica* 89 (1): 1–35.
- Autor, D.H. 2015. "Why Are There Still So Many Jobs? The History and Future of Workplace Automation". *Journal of Economic Perspectives* 29: 3–30.
- Autor, D.H. and D. Dorn. 2013. "The Growth of Low-skill Service Jobs and the Polarization of the US Labor Market". *American Economic Review* 103: 1553–97.
- Avenyo, E.K., M. Konte and P. Mohnen. 2019. "The Employment Impact of Product Innovations in sub-Saharan Africa: Firm-level Evidence". *Research Policy* 48 (9): 1–14.
- Baccini, L., M. Fiorini, B. Hoekman and M. Sanfilippo. 2021. "Services and Economic Development in Africa", IGC Working Paper No. F-20086-NOC-1.
- Bahia, K., P. Catells, G. Cruz, T. Masaki, X. Pedrós, T. Pfitze, C. Rodríguez-Castelán and H. Winkler. 2020. "The Welfare Effects of Mobile Broadband Internet: Evidence from Nigeria", World Bank Working Paper No. 9230.
- Bahia, K., P. Castells, G. Cruz, T. Masaki, C. Rodríguez-Castelán and V. Sanfelice. 2021. "Mobile Broadband Internet, Poverty and Labor Outcomes in Tanzania", IZA Discussion Paper No. 14720.
- Borusyak, K., X. Jaravel and J. Spiess. 2021 "Revisiting Event Study Design: Robust and Efficient Estimation".
- Buera, F.J., J.P. Kaboski, R. Rogerson and I. Vizcaino. 2021. "Skill-biased Structural Change". *The Review of Economic Studies*.
- CEDEFOP (2014). "Coping with changes in international classifications of sectors and occupations: Application in skills forecasting". Luxembourg: Publications Office of the European Union. Cedefop research paper; No 43
- Gaddis, I., O. Gbemisola, A. Placios-Lopez and J. Pieters. 2020. "Who Is Employed? Evidence from sub-Saharan Africa on Redefining Employment", World Bank Policy Research Working Paper No. 9370.
- Goos, M., A. Manning and A. Salomons. 2014. "Explaining Job Polarization: Routine-biased Technological Change and Offshoring". *American Economic Review* 104 (8): 2509–26.
- Guriev, S., N. Melnikov and E. Zhuravskaya. 2021. "3G Internet and Confidence in Government". *Quarterly Journal of Economics* 136(4): 2533–2613.
- Hjort, J. and J. Poulsen. 2019. "The Arrival of Fast Internet and Employment in Africa". *American Economic Review* 109 (3), 1032–79.
- Hjort, J. and L. Tian. 2021. "The Economic Impact of Internet Connectivity in Developing Countries", PEDL Synthesis Series No. 6.
- ILO. 2012. *International Standard Classification of Occupations: ISCO-08*. Geneva: International Labour Office.
- ILO. 2013. *19th International Conference of Labour Statisticians. 2–11 October 2013*. Geneva: International Labour Office

Kaplinsky, R. and E. Kraemer-Mbula. 2022. "Innovation and Uneven Development: The Challenge for Low- and Middle-income Economies". *Research Policy* 51 (2): 104394.

Manacorda, M. and A. Tesei. 2020. "Liberation Technology: Mobile Phones and Political Mobilization in Africa". *Econometrica* 88 (2): 533–67.

Mehrabi, Z., M.J. McDowell, V. Ricciardi, C. Levers, J.D. Martinez, N. Mehrabi, H. Wittman, N. Ramankutty and A. Jarvis. 2021. "The Global Divide in Data-driven Farming". *Nature Sustainability* 4: 154–60.

Mensah, J.T. 2021. "Mobile Phones and Local Economic Development: A Global Evidence". <https://dx.doi.org/10.2139/ssrn.3811765>.

MINICOM. 2011. National Industrial Policy of Rwanda. Ministry of Trade and Industry, Government of Rwanda.

NASA Socioeconomic Data and Applications Center (SEDAC). n.d. "Population dynamics".

National Institute of Statistics of Rwanda (NISR). 2018. Labour Force Survey 2017, August 2017 Report.

Newfarmer, R., J. Page and F. Tarp. 2018. *Industries without Smokestacks: Industrialisation in Africa Reconsidered*. Oxford: Oxford University Press.

Nkurunziza, J., A. Broekhuis and P. Hooimeijer. 2012. "Free Education in Rwanda: Just One Step Towards Reducing Gender and Sibling Inequalities". *Education Research International*.

Rodrik, D. 2016. "Premature Deindustrialization". *Journal of Economic Growth* 21: 1–33.

Rwanda Utilities Regulatory Authority (RURA). 2021. Statistics Report for Telecom, Media and Broadcasting Sector as of the First Quarter of the Year 2021.

Suri, T., J. Aker, C. Batista, M. Callen, T. Ghani, W. Jack, L. Klapper, E. Riley, S. Schaner and S. Sukhtankar. 2021. "Mobile Money". *VoxDevLit* 2 (1).

Vivarelli, M. 2021. "Globalization, Structural Change, and Technological Change in Emerging Economies: The Impact on Employment and Skills". In *New Perspectives on Structural Change: Causes and Consequences of Structural Change in the Global Economy*. Oxford: Oxford University Press.

Acknowledgements

This paper has been prepared as part of the “Skills and transitions” research project by the Research Department of the International Labour Organization (ILO). We thank Hannah Liepmann and two anonymous reviewers for their considered comments on preliminary drafts of the paper. We also thank Justice Tei Mensah for kindly making available the data on mobile internet coverage that is used in the paper.

The responsibility for opinions expressed in this article rests solely with its authors, and publication does not constitute an endorsement by the International Labour Office of the opinions expressed in it.

► Advancing social justice, promoting decent work

The International Labour Organization is the United Nations agency for the world of work. We bring together governments, employers and workers to improve the working lives of all people, driving a human-centred approach to the future of work through employment creation, rights at work, social protection and social dialogue.

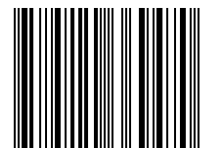
Contact details

Research Department (RESEARCH)

International Labour Organization
Route des Morillons 4
1211 Geneva 22
Switzerland
T +41 22 799 6530
research@ilo.org
www.ilo.org/research



ISBN 9789220369470



9 789220 369470