DISCUSSION PAPER SERIES

DP16924

Services, Jobs, and Economic Development in Africa

Leonardo Baccini, Matteo Fiorini, Bernard Hoekman and Marco Sanfilippo

DEVELOPMENT ECONOMICS

INTERNATIONAL TRADE AND REGIONAL ECONOMICS

ISSN 0265-8003

Services, Jobs, and Economic Development in Africa

Leonardo Baccini, Matteo Fiorini, Bernard Hoekman and Marco Sanfilippo

Discussion Paper DP16924
Published 17 January 2022
Submitted 10 January 2022

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

Services, Jobs, and Economic Development in Africa

Abstract

This paper presents data and analyzes the structure of employment in thirteen African economies at the administrative unit level, with a focus on the role of services. We provide two novel pieces of evidence. First, we present a descriptive snapshot of changes in the composition of employment over time and across geographies. This reveals evidence of structural transformation towards services and service-related occupations at sub-national level and provides a fine-grained overview of who works in services and where and how this has changed over time. Second, we provide correlations between services and economic development, using per capita nightlight luminosity as a proxy. We document (a) a strong positive association between high skills services and economic development; (b) substantial heterogeneity across industries within services; and (c) a mediating role of market conditions and technology in the relation between services and economic development. Overall, our work highlights an important role of services activities for employment, skills and economic development in Africa.

JEL Classification: O14, O55, O57, R12

Keywords: economic development, structural transformation, growth, servicification, Africa, employment

Leonardo Baccini - leonardo.baccini@mcgill.ca
McGill University

Matteo Fiorini - matteo.fiorini@eui.eu
European University Institute

Bernard Hoekman - bernard.hoekman@eui.eu
European University Institute and CEPR

Marco Sanfilippo - marco.sanfilippo@unito.it
Department of Economics and Statistics, University of Torino

Acknowledgements
A previous version of this paper was circulated under the title “Services and Economic Development in Africa”. We are grateful to three anonymous reviewers, Richard Newfarmer and participants in the plenary section of the September 2021 STEG workshop and a seminar at the African Development Bank for valuable comments. Lorenzo Sileci provided outstanding research assistance. This work was supported by the International Growth Center (Grant XXX-20086).

Services, Jobs, and Economic Development in Africa*

Leonardo Baccini (McGill University and CIREQ)

Matteo Fiorini (EUI)

Bernard Hoekman (EUI and CEPR)

Marco Sanfilippo (University of Turin and Collegio Carlo Alberto)

December 2021

Abstract

This paper presents data and analyzes the structure of employment in thirteen African economies at the administrative unit level, with a focus on the role of services. We provide two novel pieces of evidence. First, we present a descriptive snapshot of changes in the composition of employment over time and across geographies. This reveals evidence of structural transformation towards services and service-related occupations at sub-national level and provides a fine-grained overview of who works in services and where and how this has changed over time. Second, we provide correlations between services and economic development, using per capita nightlight luminosity as a proxy. We document (a) a strong positive association between high skills services and economic development; (b) substantial heterogeneity across industries within services; and (c) a mediating role of market conditions and technology in the relation between services and economic development. Overall, our work highlights an important role of services activities for employment, skills and economic development in Africa.

Keywords: services; economic development; structural transformation; growth; servicification; Africa; employment

JEL codes: O14; O55; O57; R12

* A previous version of this paper was circulated under the title "Services and Economic Development in Africa". We are grateful to three anonymous reviewers, Richard Newfarmer and participants in the plenary section of the September 2021 STEG workshop and a seminar at the African Development Bank for valuable comments. Lorenzo Sileci provided outstanding research assistance. This work was supported by the International Growth Center (Grant XXX-20086). This paper is complemented by an online appendix and a website providing tools for visualization of the micro-data on jobs in services – see https://globalgovernanceprogramme.eui.eu/services-and-economic-development-in-africa/

Introduction

Economic development is generally associated with a shift of the workforce from the agricultural to non-agricultural sectors, a process commonly referred to as structural transformation. A central feature of successful structural transformation of most economies in the 20th century involved a shift of workers from agriculture to manufacturing (Herrendorf et al., 2014). This pattern of economic growth and development may be less applicable to lower-income countries today than it was in the past. An important stylized fact in this regard is that low-income developing countries are moving into services earlier and at a faster rate than was observed for East Asian economies in the 1970s and 1980s. The associated 'premature deindustrialization' (Dasgupta and Singh, 2007; Rodrik, 2016) is a potential source of concern insofar as it implies that low-income countries today cannot rely on an expanding manufacturing sector to create employment opportunities for relatively unskilled workers, drive economic growth and increase per capita incomes.

Arguments why structural transformation involving a rapid increase in services employment may be detrimental to the growth prospects of low-income countries include perceptions that services offer fewer prospects for sustained high productivity growth than manufacturing and less potential for positive spillover effects (innovation; inter-sectoral linkages; positive agglomeration externalities) or scope for scale economies. Other concerns are that services offer less potential to generate 'good jobs' for unskilled workers at the levels needed than manufacturing did in previous periods. Much of the research literature on the role of services in economic development has focused on the question whether services can realize the type of productivity growth observed in manufacturing. An expanding body of evidence documents that many types of services have experienced levels of productivity growth that are similar to, and sometimes exceed, what has been realized in agriculture and manufacturing. Less attention has been devoted to the question whether the shift to services in low-income countries is associated with the creation of jobs, the skill-intensity of services jobs, and where they are located. Because many services are relatively labor-intensive, involve both lower and higher skilled activities, and often are (and notwithstanding technological change likely will remain) less tradable than goods, there is the potential for job generation in services to be dispersed more widely across space (the territory of a given jurisdiction – locality, region, country) than manufacturing.

Our aim is to contribute to the emerging literature that analyzes structural transformation in low-income economies using micro data as opposed to a focus on broad sectoral shifts and aggregate indicators of output and employment at the country level. We present new data on the composition of jobs in services at the sub-national (administrative unit) level in a sample of thirteen African countries, describe how this has changed over time, and how employment in services correlates with indicators of economic development commonly used in the literature. We use the IPUMS International Database published by the Minnesota Population Center, extracting data for all African countries for which at least two censuses are available and that include information on the industry employing an individual.[1] The resulting dataset spans 56 million individuals covering 1,546 administrative units in the 13 African countries for either two or three census waves. The sample is representative of the Africa region. It includes both low-income countries whose GDP per capita is about US$1,000 (Malawi, Mozambique), middle income countries such as Mauritius; resource rich countries (Botswana, Zambia), more diversified economies (South Africa and Morocco), landlocked countries (Rwanda) and countries with sea borders and a strong tourism industry (Mauritius, Tanzania). Combined, the sample countries account for 31.1% and 44.1% of Africa's total population and GDP, respectively.[2]

We find that employment in trade (wholesale and retail) services is leading the shift towards services and that growth in services jobs is associated with a shift to higher skilled types of services, based on a classification that clusters specific services industries according to educational attainment of individuals and intensity of use of more complex types of occupations. Compared to manufacturing, services workers are on average more educated, engage in higher skilled occupations and are more likely to be female – 40% of employees are women, twice as large as in manufacturing. There has been a shift over time towards occupations related to the production of intangible value added and activities classified as higher skilled. This shift is observed across sectors, including manufacturing, with services more likely to develop in urban and densely populated areas and those with greater access to national and international markets.

Exploratory analysis of the relationship between services and proxies for economic development do not reveal an association between overall employment in services at the aggregate level and economic development. This masks significant heterogeneity within services and across administrative units. We document a strong positive (negative) association between high (low) skills services and economic development, with the positive association particularly strong in areas with low incidence of malaria, more natural resources, and good mobile phone coverage. These associations point to a mediating role of market conditions and technology in the relationship between services employment and economic development.

[1] This includes 12 industries that roughly concord to the services sectors in the International Standard Industrial Classification (ISIC).

[2] Data sourced from the IMF WEO (April 2021 edition). Figures for GDP are based on the PPP values for the year 2019 (the last non-estimated figure), and 2015 for population (the latest non-estimated values).

Health services, public services, and, to a lesser extent, business services are positively associated with proxies for economic development, while transport and private household services have a negative association with development indicators.

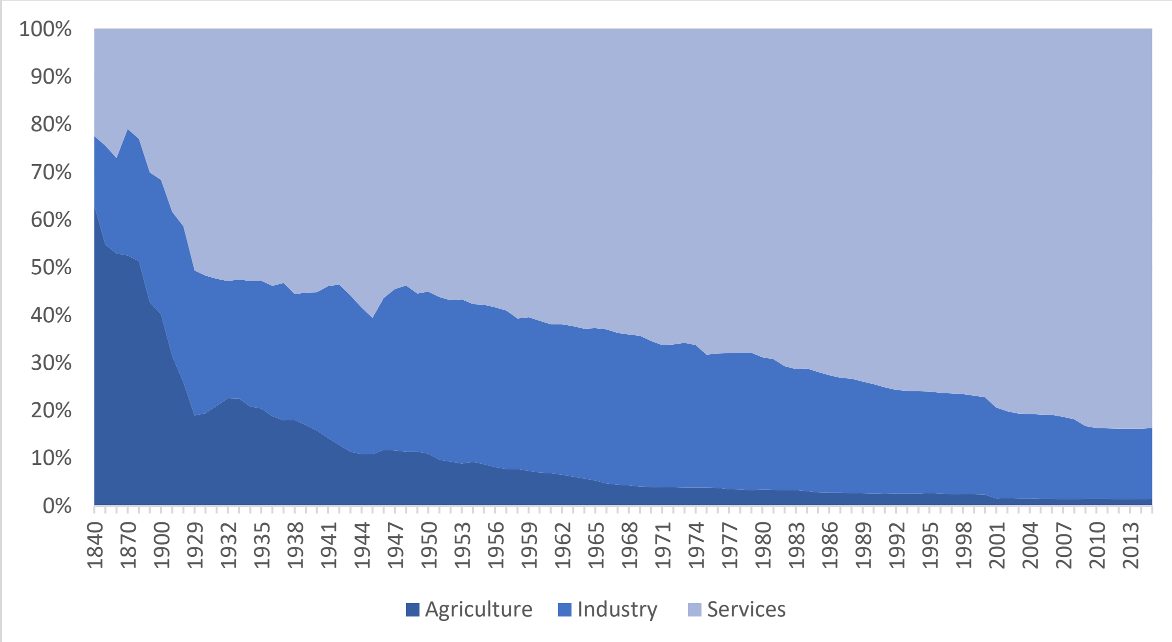
The paper proceeds as follows. Section 1 briefly surveys the related literature on structural transformation and drivers and implications of the shift to services. Section 2 discusses the data. Section 3 focuses on specific services sectors, assessing both graphically and by means of multivariate analyses employment changes across disaggregated services activities. Section 4 reports the results of exploratory analysis of the association between services and indicators of development. Section 5 concludes with suggestions for further research on the drivers and consequences of structural transformation driven by the shift towards services in low-income countries.

1. Services and economic development prospects of low-income countries

The conventional conceptualization of the structural transformation of an economy that underpins economic growth centers around sectoral reallocation of labor and capital to higher productivity activities, involving a steady reduction in agriculture-related employment, with workers absorbed by industry, mostly manufacturing (Kuznets, 1966; Kravis et al. 1983; Syrquin, 1988; Schettkat and Yocarini, 2006; Herrendorf et al. 2014). Figure 1 illustrates the pattern of declining agricultural employment over time as per capita income rises, with offsetting increases in employment shares of industry and services. A key feature of this process is that historically the rate of increase in manufacturing outpaces that of services for some time, after which the overall share of manufacturing peaks and the share of services expands further.

Two major drivers for structural transformation are often distinguished in the economic literature: technological change resulting in differential rates of sectoral total factor productivity growth and a demand effect reflected in differences in income elasticities for goods and services (Duarte and Restuccia, 2010). These drivers result in intra-sectoral reallocation of factors of production as well as inter-sectoral shifts. In practice, there will be significant change in economic activity and employment within sectors – e.g., a decline in personal household services and increases in market-based services activities with associated differential growth performance at the sub-sectoral level. The process of inter- and intra-sectoral reallocation has consequences for both growth and the distribution of income (poverty; inequality) and economic activity across locations within countries, depending on initial endowments, geographic attributes, infrastructure, and economic policy, among other factors.

Figure 1. Sectoral employment shares and economic growth: the US example



Note: Data cover employment in the US and is originally sourced from (Herrendorf et al., 2014)
Source: Authors' elaboration on Our World in Data

In recent decades, many low-income developing countries have experienced structural transformation without industrialization, in that a decline in the share of employment in agriculture is offset by a rise in employment in services, with a relatively constant or even declining share of manufacturing employment (Dasgupta and Singh, 2007). Rodrik (2016) provides cross-country evidence that the share of manufacturing employment (and value added) peaks at lower levels of per capita incomes when compared to the pattern observed in the past for countries that are high-income nations today. African countries are a prominent example, with manufacturing sectors dominated by relatively few large, export-oriented companies that have not created significant employment (most African firms are small and often informal).[3] Using micro data, Diao et al. (2019) also show that the structure of African economies has shifted more rapidly towards services than was the case for the present OECD countries during their industrial development phase and for East Asian economies in the 1970s and 1980s.

There is a vibrant and unsettled debate on whether and how low-income developing countries can rely on other engines of growth and job creation. Central questions in this regard concern the prospects for

[3] For instance, Diao et al. (2021) find that manufacturing employment in Ethiopia and Tanzania is mostly determined by small and informal firms with stagnant productivity growth.

productivity growth in non-manufacturing activities[4] and the ability of services sectors to create jobs for relatively unskilled workers at scale in low-income countries where the type of export-led manufacturing strategy used by successful countries in the past is constrained. Recent work on Africa has emphasized the importance of non-manufacturing activities for both job generation and growth (Newfarmer et al., 2018). Thanks to revolutions in transport and technologies, industries 'without smokestacks', such as horticulture, agro-processing, tourism, e-commerce, digitally enabled business, health and education services, have the potential to drive sustainable demand for jobs, together with rising productivity, leveraging the increasing tradability of many services products (Dihel et al. 2016). Services are very heterogeneous in terms of skill intensity, tradability and scope for productivity growth. Some services offer great scope for the type of productivity dynamics that have characterized manufacturing, others do not (Balchin et al. 2016, Nayyar et al., 2021).

The extent to which the shift to services implies lower productivity growth potential is essentially an empirical question. Young (2014) finds that average productivity growth in services is similar to that in other sectors. Herrendorf et al. (2020) document for the US that this is associated with a large and rising share of services value added in investment expenditure, which has come to exceed the share of goods value-added in investment expenditure. Nayyar et al. (2021) compare productivity levels of services and manufacturing, showing that some of the most innovative services industries are more productive than manufacturing sectors. Buera and Kaboski (2012) and Duarte and Restuccia (2020) have shown that heterogeneity across services in terms of both skills and productivity plays a role in explaining cross-country differences in economic growth.

Arguments that services can play a role similar to manufacturing sector in past examples of successful development are based in part on the view that, due to the changes occurring in the organization of international production, the reduction in transport costs and opportunities offered by new technologies, the nature of services is changing dramatically. Many services activities are (increasingly) tradable, have experienced high productivity growth, and can achieve economies of scale (Gervais and Jensen, 2019; Loungani and Mishra, 2014).[5] Sen (2019) argues that it is important to distinguish between business and nonbusiness services, both because of their different functions and because of the greater prospects for the

[4] To some extent, distinguishing between services and manufacturing is misleading as what matters is productivity growth performance (potential) and employment generation of different economic activities.

[5] Hsieh and Rossi-Hansberg (2021), for example, argue that the availability of a new menu of fixed-cost-intensive technologies in service sectors enables adopters to produce at lower marginal costs in a wide range of local markets. They note that in the US the entry of leading service firms into new local markets has led to substantial unmeasured productivity growth, particularly in small markets. ICT-based technologies and adoption of new management practices permit such services firms to scale up production across many locations, driving productivity growth.

former to affect productivity growth. Sen argues that in low-income countries growth must depend more on the non-business service sector given their role as a major sector of employment outside agriculture and the limited prospects for manufacturing growth and associated business services.

Fan et al. (2021), using data for India for the 1987-2011 period, develop a structural model that distinguishes between consumer and producer services, dividing services sectors in two categories depending on whether their output is absorbed more by consumer demand (e.g., hospitality) or by firms using services as inputs into production (e.g., business services). They investigate whether the rising share of services in the economy is a driver of growth (a source of rising productivity) or is the result of demand factors (the result of rising incomes). Their empirical analysis concludes that in India growth was services-led, driven by productivity improvements in consumer services such as retail and hospitality, which they find drove one-third of the aggregate growth rate. They also find that the distribution of the associated welfare benefits was highly skewed, with high-income households living in urban areas benefitting most. Gollin et al. (2016) highlight the potential role of consumer services in natural resource rich countries, pointing to examples of sub-Saharan African economies where urbanization has occurred without industrialization. They find a strong positive relationship between natural resource exports and urbanization in a sample of 116 developing nations over the period 1960–2010, documenting the rise of so-called 'consumption cities' with economies centered on non-tradable services in contrast to 'production cities' that are more dependent on manufacturing in countries that have industrialized.[6]

The extent to which manufacturing employment has peaked in Africa or is in secular decline is a matter of debate. While the share of manufacturing in employment is low, some researchers argue there is both significant scope for African countries to increase the share of manufacturing in total employment and evidence suggesting that the shift to services is not necessarily associated with manufacturing employment peaking (e.g., Diao et al. 2017; Haraguchi et al. 2017; Mensah 2020; Nguimkeu and Zeufack 2019). Kruse et al. (2021) document that the manufacturing share of employment in Africa increased from an average of 7.2 percent in 2010 to 8.4 percent in 2018, reversing a de-industrialization trend of the decades before 2010 captured in earlier research. They add, however, that trends in manufacturing value added did not reverse, showing that much of this growth is associated with small and less-productive firms.

[6] Eckert et al. (2020) investigate the drivers of urban biased growth in the US by identifying a cluster of "skilled, scalable" services based on each sector's reliance on high skilled-labor and ICT capital. This group includes professional, scientific and technical services; management services; information services and financial services. They show that US cities' comparative advantage in skilled services is a factor behind higher growth in urban areas. Similarly, Fang and Herrendorf (2021) show that removing frictions leading to the underdevelopment of high-skilled services would contribute to sizeable increases in per capita GDP.

While much of the literature on services and development has a focus on productivity, we focus on the employment implications of structural transformation towards services in Africa and the covariation between the shift towards services and development indicators. While a focus on productivity and skill-intensity is central to understand the long-term implications of structural transformation, this might be at odds with developing strategies that maximize job opportunities for workers that move out of the agricultural sector. Understanding which industries and under which conditions can create jobs requires the compilation of granular data – the focus of this paper.

2. The data

We use three types of information: geolocalized employment data across sectors, including up to 12 services sectors; geolocalized indicators of economic development; and relevant features of the economic environment that could shape the relationship between services activities and economic development. This information is merged into one panel dataset, identified at the ADMIN-wave pair, covering 1,546 administrative units in 13 African countries: Benin, Botswana, Egypt, Ghana, Malawi, Mali, Mauritius, Morocco, Mozambique, Rwanda, South Africa, Tanzania, and Zambia. The data span either two or three waves, depending on availability of census waves in the IPUMS International Database. The final panel comprises 3,846 observations[7] and 55,976,623 individuals spanning the period 1982-2013. We describe the construction of the employment data from IPUMS below. The Appendix provides more detail on the construction of the database, available census waves and all variables. A complementary online tool provides access to the data and permits users to generate graphs and maps on sectoral and occupational dynamics in Africa over available census waves in the IPUMS International database.[8]

Employment data come from the IPUMS International Database (Minnesota Population Center, 2019). IPUMS reports information for a repeated cross-section of representative individuals, covering a variable fraction of a country's total population.[9] We aggregate individual-level information to the lowest available administrative designation for each country, in order to obtain a dataset defined at the ADMIN-wave level.[10]

---

[7] When including Mauritius. The number of observations is 3,727 when Mauritius is excluded in analyses using the spatial Gini coefficient and Alesina et al. (2021) variables.

[8] The online tool is available at https://globalgovernanceprogramme.eui.eu/services-and-economic-development-in-africa/.

[9] Coverage fractions and total number of surveyed individuals in each census wave are reported in the Online Appendix.

[10] This operation amounts to aggregating data at the second administrative level for most countries, excluding Botswana, for which the sole administrative division available for aggregation is the first. The second-level administrative unit corresponds to Districts in a majority of cases, with the exception of Benin ("Communes"), Malawi

We only extract data for African countries for which at least two census waves report the "INDGEN" variable,[11] which provides a disaggregated sectoral employment classification, including up to 12 specific services industries.[12]

Given the abundance of temporal inconsistencies in the administrative designations in the IPUMS International Database, reflecting frequent redistricting by national authorities, we rely instead on the temporally and spatially consistent shapefiles provided by Alesina et al. (2021), to which we collapse all indicators included in our analyses.[13] This results in a total of 1,546 unique administrative units across the countries in the sample that are broadly comparable in terms of surface area. We calculate ADMIN-wave level sectoral shares for each of the 17 levels of the INDGEN variable,[14] weighting each individual-level observation by the survey weights provided by IPUMS International. This results in 17 distinct variables measuring the share of each sector in an administrative unit's total recorded sectoral employment.[15] We repeat the same procedure for individual-level occupational characteristics, reported by IPUMS International under the "OCCISCO" variable. This variable includes 11 categories detailing the dwelling covered by each surveyed person within their occupational remit.[16] We thereby obtain 11 distinct occupational shares for each administrative unit, detailing the incidence of each dwelling on total employment. The OCCISCO variable is available in all included census waves, except for Rwanda's first census (1991).

("Traditional Authorities"), Mali ("Circles"), Mauritius ("Municipal Wards/Village Council Areas"), Morocco and Rwanda ("Provinces").

[11] The "INDGEN" variable refers to a disaggregated sectoral classification with 17 levels "roughly conforming to the International Standard Industrial Classification (ISIC)."

[12] IPUMS has only one data point for sectoral employment for Burkina Faso, Cameroon, Guinea, Kenya, Lesotho, Senegal, Sierra Leone, South Sudan, Sudan, Uganda, and Zimbabwe, precluding their inclusion. Liberia and Togo are also excluded due to the very long time between the two survey waves reported in IPUMs (1974 and 2008 for Liberia and 1970 and 2010 for Togo).

[13] As Alesina et al. (2021) do not include shapefiles for Mauritius, we rely on the administrative designations provided by IPUMS International for this country.

[14] The categories are: (1) Agriculture, fishing and forestry; (2) Mining; (3) Manufacturing; (4) Utilities; (5) Construction; (6) Wholesale and retail trade; (7) Hospitality (hotels and restaurants); (8) Transportation, storage and communications; (9) Financial services; (10) Public Administration and Defence; (11) Services, not specified; (12) Business services and real estate; (13) Education; (14) Health and social work; (15) Other services; (16) Private household services; (17) Other industry. The "Other services" category includes miscellaneous personal and community services such as sanitation and entertainment, and non-business repairs and rental activities.

[15] "Not in Universe" (NIU) responses that may distort the resulting sectoral shares were dropped from the dataset.

[16] IPUMS codes occupations according to the major categories in the International Standard Classification of Occupations (ISCO) scheme for 1988. For someone with more than one job, the primary occupation is typically the one in which the person had spent the most time or earned the most money. The 11 categories are: (1) Legislators, senior officials and managers; (2) Professionals; (3) Technicians and associate professionals; (4) Clerks; (5) Service workers and shop and market sales; (6) Skilled agricultural and fishery workers; (7) Crafts and related trades workers; (8) Plant and machine operators and assemblers; (9) Elementary occupations; (10) Armed forces; (11) Other occupations, unspecified or not elsewhere classified.

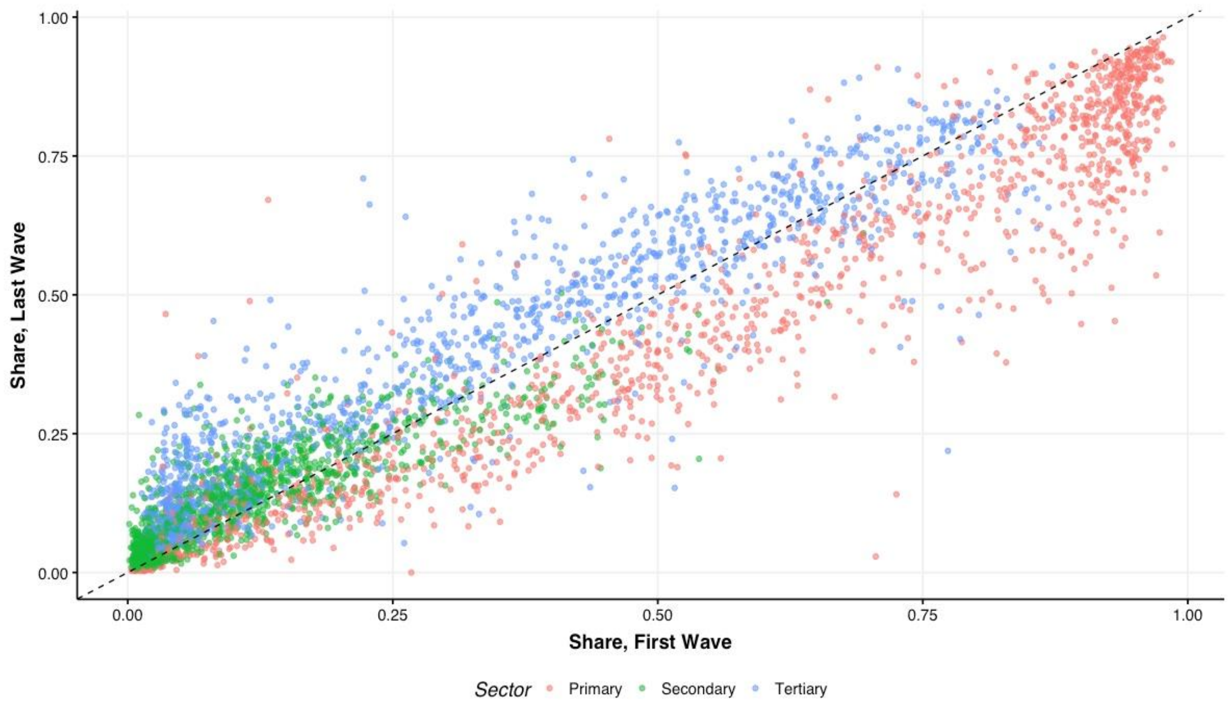
3. Services in African Structural Transformation

The shift towards services noted in the literature discussed in Section 1 is evident in our sample of subnational units in Africa. Figure 2 plots decadal changes in employment shares across the three major sectors – primary, secondary and tertiary.[17] It shows that the drop in primary activities goes mostly hand in hand with an increase in services. Importantly, this pattern of services constituting a major source of employment growth is consistent across our sample, independent of the stage of development (e.g., Egypt vs. Malawi) or the presence of natural resources (Botswana vs. Benin). The data also reveal that the share of the secondary sector generally increases during the period, with many instances in which employment shares in the secondary sector increase from an initial low base, offset by instances where administrative units with initially high shares of secondary sector employment experience a reduction in over time. Most instances where the share of industry increases are in the lower tail of the scatterplot. Conversely, administrative units where industry (the secondary sector) accounts for more than 25 percent of employment in the first wave often see a decline, consistent with Rodrik (2016). Appendix Figure C1 replicates the scatter plot for agriculture, manufacturing and services, showing a virtually identical picture as Figure 2.

The fact that we observe increases in the share of industrial employment in many administrative units suggests secondary sector activities demonstrate dynamism. Whether these activities can grow the share of industrial employment over time is of course an open question. An important research (and policy) question suggested by these descriptive data is what drives the growth in secondary activity at the lower end of the distribution, and whether and how it is distinct from the economic activities that are associated with instances of administrative units that report high shares of industrial employment in the first wave and lower shares in the second wave.[18]

[17] Primary spans agriculture and mining; secondary comprises utilities, manufacturing, and construction; tertiary includes all services sectors.
[18] These data pertain to shares and are not informative about the associated absolute levels of employment.

Figure 2. Structural transformation at the sub-national level

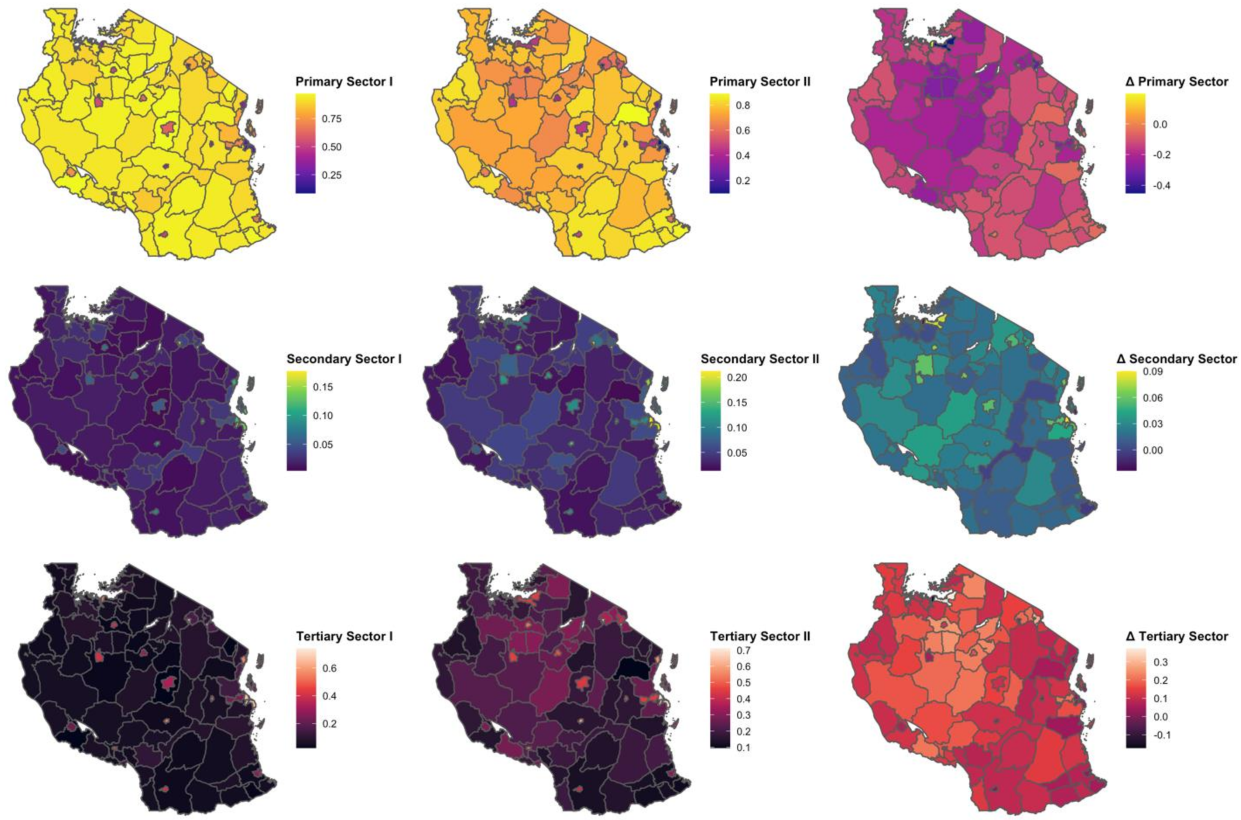


Note: Each dot represents an administrative unit that is observed over two successive waves of the census.
Source: Authors' elaboration on IPUMS.

The Appendix and dedicated webpage[19] provide country specific graphs and maps plotting changes in sectoral employment and occupational dynamics over time. An example is provided in Figure 3, reporting data for Tanzania. It shows how structural change happened across administrative areas within the country over the decade spanning the two most recent population censuses (2002 and 2012). Apart from a few areas, including the capital and other important cities, the rest of the country has been characterized by reductions in the share of agricultural employment that went together with an increase of employment within the services sectors. As is the case for the sample overall (Figure 2), there is also an increase in the share of the secondary sector.

[19] At https://globalgovernanceprogramme.eui.eu/services-and-economic-development-in-africa

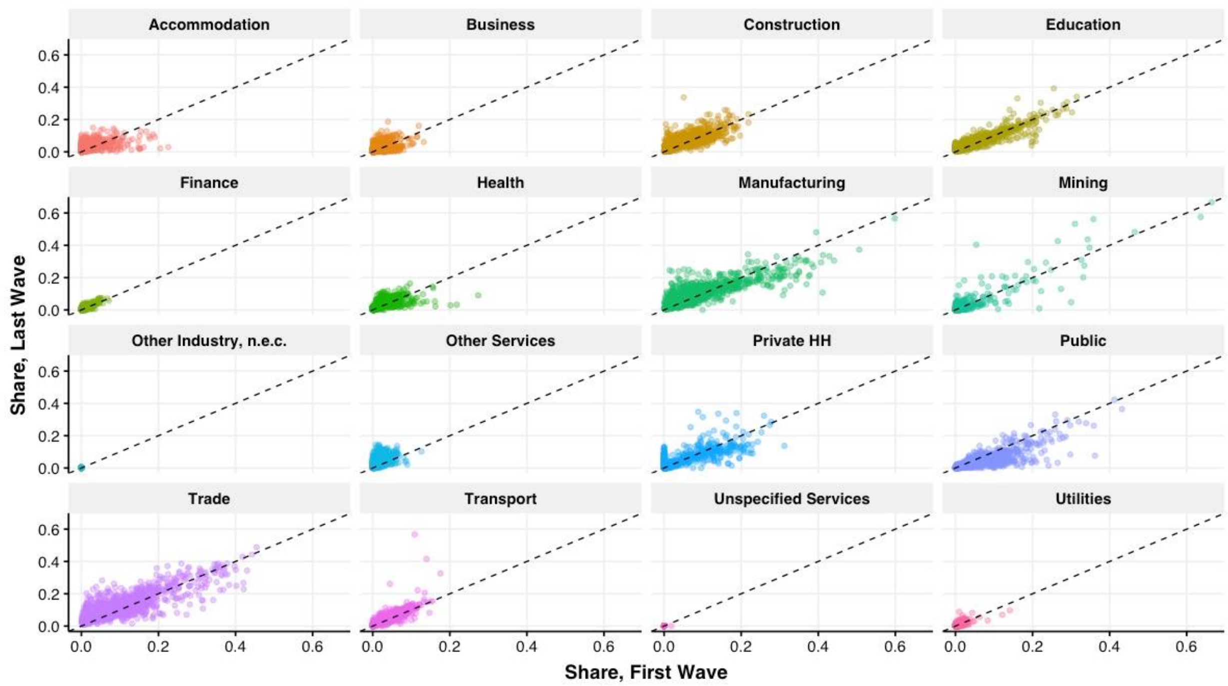
Figure 3. Sectoral dynamics in Tanzania between 2002 (I) and 2012 (II)



Note: the maps report administrative units' share of employment in the three main sectors. The first column reports maps covering the first (I) wave of the census (2002); the second column the next (II) wave; and the third column the difference between the two periods.
Source: Author's elaboration using IPUMS.

Figure 4 reports details for each of the 2-digit industries belonging to the services sector, in addition to other sectors for comparison. Again, all the graphs report information for the entire database and show changes that occurred in the most recent wave compared to the prior one. A few industries within services appear to explain the trends observed in the previous figure. This is most visible in the case of trade. Other services activities (e.g. business and financial services) also increase, but since they start from a very small base, their growth is more difficult to assess through visual inspection. Country-specific data reveal interesting heterogeneity across sectors over time (e.g., a visible decline in public sector employment in Botswana, Tanzania and South Africa – see Appendix).

Figure 4. Structural transformation at the sub-national level: sector specific data



Note: Each dot represents an administrative unit that is observed over two successive waves of the census. Agricultural sector is excluded to allow a better visualization of smaller industries (given that all industries share the same values in the Y-axis).
Source: Authors' elaboration using IPUMS.

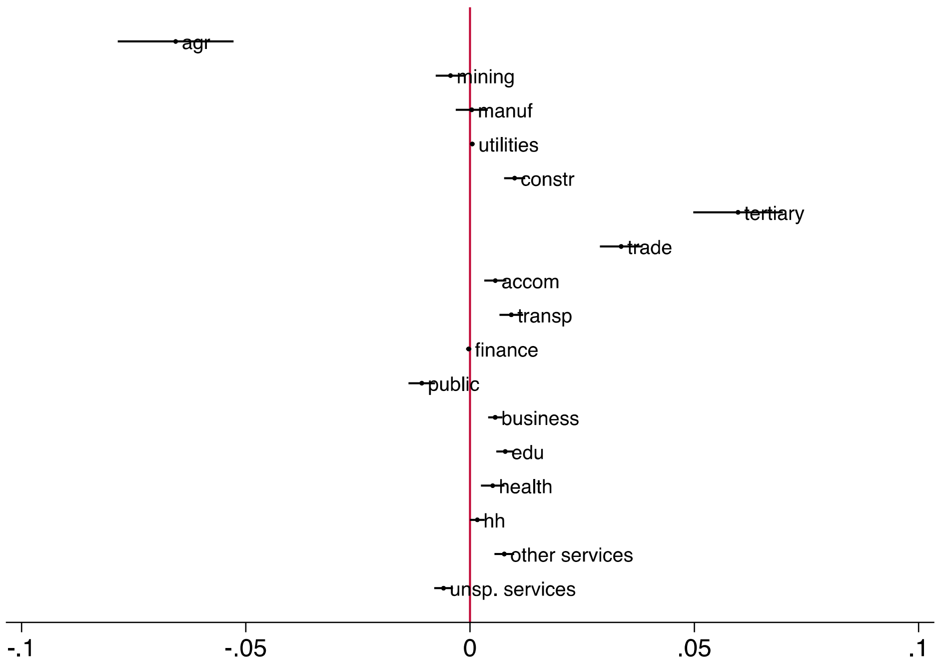
Premature de-industrialization and the rise of services at the sub-national level

To get a better sense of the trends discussed above and the rate of change we replicate a specification proposed by Rodrik (2016) and adopted in other papers (e.g., Kruse et al. 2021), linking sectoral employment to time trends, as follows:

$$Emp\_share_{it} = \beta_0 + \beta_1 pop_{it} + \beta_2 pop_{it}^2 + \beta_3 ntl\_pc_{it} + \beta_4 ntl\_pc_{it}^2 + \theta_i D_i + \gamma_t post\_2000_t + \varepsilon_{it} \quad (1)$$

where the dependent variable is the employment share in agriculture, manufacturing and in all the industries comprising services. All regressions control for both demographic factors and income by means of the inclusion of (log) population and per capita night light and their squared terms, as well as for location *i* (administrative unit) fixed effects. The variable of interest is γ_t, which takes the value of 1 if the survey was run after 2000 and 0 otherwise. The estimated coefficient of this dummy gives us the size of the common shock to all sectors considered in the post-2000 period relative to pre-2000. Results are summarized in Figure 5, which reports the estimated γ_t using coefficients for all the regressions considered. Full results are reported in Appendix Tables A1 and A2. This exercise reveals the drop in agricultural employment has been substantial. On average, across all administrative units, agricultural shares of employment experienced a 6.6 percentage point (p.p.) decline compared to the pre-2000 period. While the changes in manufacturing have been substantially zero, the tertiary sector as a whole has shown an average increase of almost 6 p.p. Within the tertiary sector most industries grew compared to the previous decade. The trade sector records the greatest increase, while the relative size of the public sector shrinks.

Figure 5: Time trends across sectors

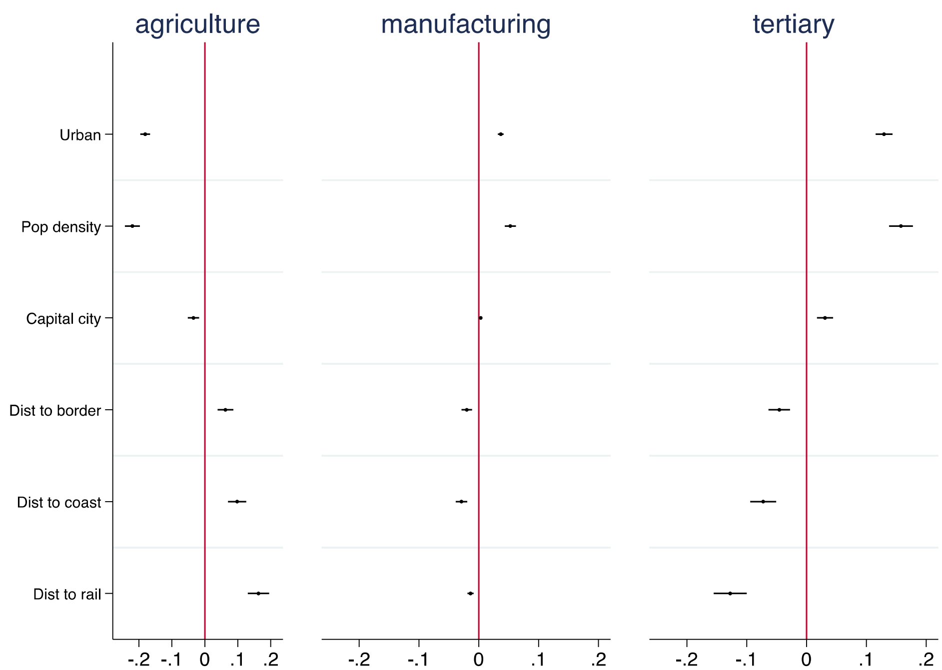


Note: Each point is the estimated γ_t, along with its 90% confidence interval, of different specifications based on equation (1). For detailed results see the Appendix.

Correlates of services employment

To explore whether the distribution of employment across sectors is correlated with specific characteristics of administrative areas, we exploit information on: (1) urbanization, measured as the share of the population living in urban areas; (2) population density; (3) a dummy indicating whether the area hosts the administrative capital of the country; and (4) the distance from the district centroid to the nearest border, coast and colonial railroad. Data on the first three variables are taken from the census data, while the distance measures are from Alesina et al. (2021). The Appendix provides detailed definitions of the variables considered, along with their summary statistics. The variables we examine here mostly relate to historical and geographic factors identified by the existing literature as relevant drivers of African long run economic development (see Michalopoulos and Papaioannou, 2020, for a review). By shaping the direction of economic development, these factors might have determined the distribution of economic activities, and hence been linked with the contemporaneous size of certain types of services. The correlations presented below are useful indicators of potential drivers of services development, but the analysis does not imply anything about causal identification. Results, based on unconditional estimates accounting for country and wave fixed effects are summarized in Figures 6 and 7. All coefficients are standardized to facilitate comparison across the different estimations.

Figure 6. Correlates of sectoral employment



Note: Each point is the estimated coefficient, along with its 90% confidence interval, of a specification in which the outcome of interest (the share of employment in agriculture, manufacturing and services) is regressed against one of the variables reported on the y-axis. A full description of these variables in provided in the Appendix. All regressions include country and wave fixed effects, and standard errors are clustered at the district level.

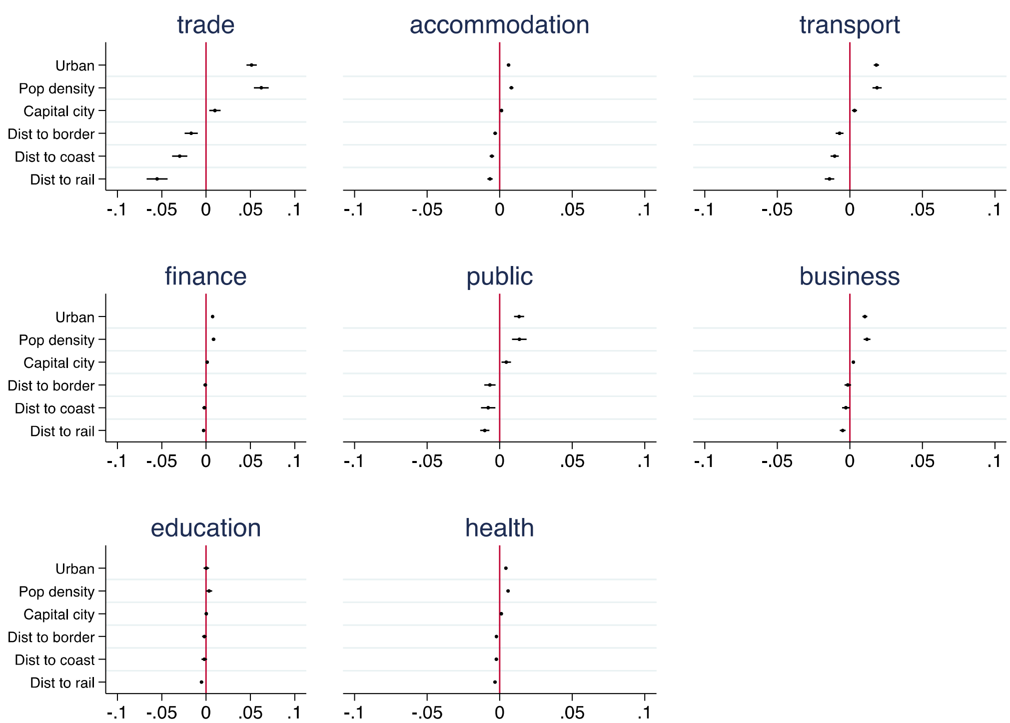
Figure 6 plots the major differences across the three main sectors of the economy. For each variable there is a (symmetric) difference between agriculture and the other sectors, which is more evident for services than for manufacturing. The characteristics of the administrative units appear to shape the distribution of economic activities within each country. Services (and, to a lesser extent, manufacturing) are more concentrated in urban areas, and in more densely populated parts of the country, including in particular areas in and around the capital city, where most administrative activities are concentrated.[20] Similarly, connectivity matters. Services employment is less likely to be high in areas far away from, in order of relevance: a border, a coast (where most of external trade happens) and a (colonial) railway. Surprisingly,

[20] Public sector jobs in a capital city are on average more than two times larger than in other areas of the country.

these variables seem to matter less (though they keep the expected sign) for the location of manufacturing employment.

Figure 7 reports the same correlations for specific services industries. Most of these industries tend to show the same pattern observed for the tertiary sector as a whole, but there is some heterogeneity worth highlighting. Trade and transport activities are most likely to cluster in more densely populated and urban areas, and in ones that are better connected. On the other hand, geography seems to matter less for public services such as health and education.

Figure 7. Correlates of sectoral employment within services

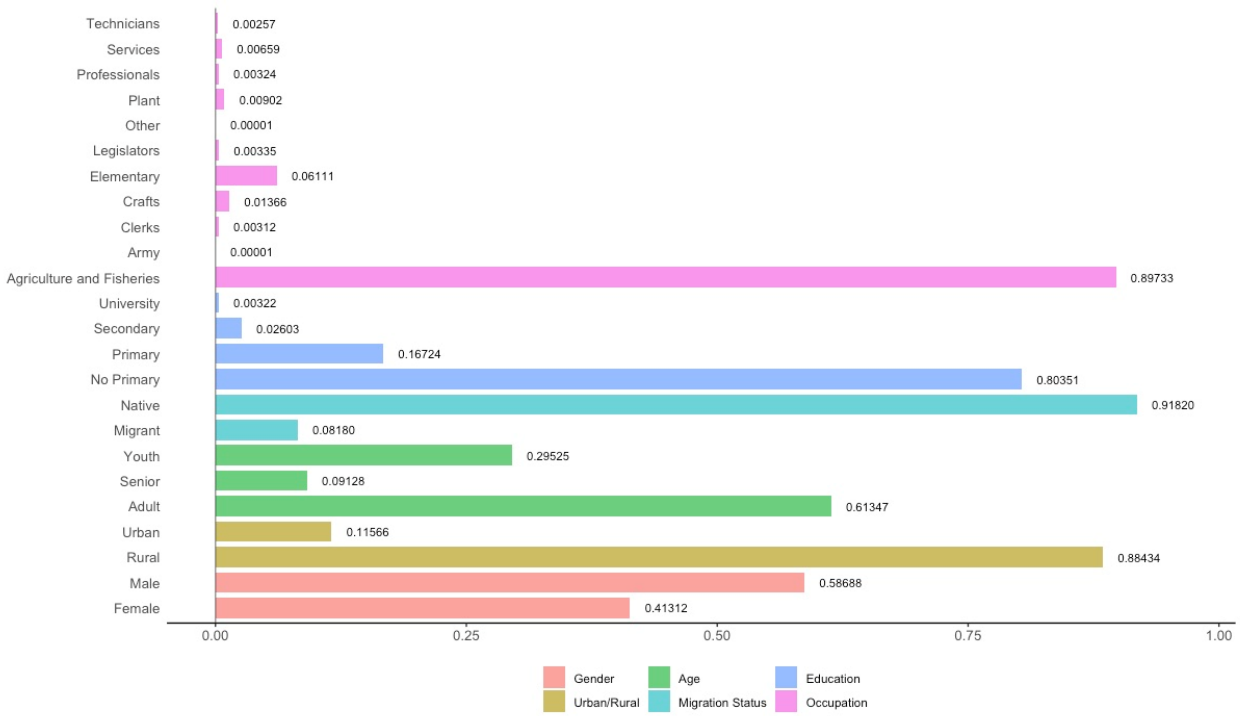


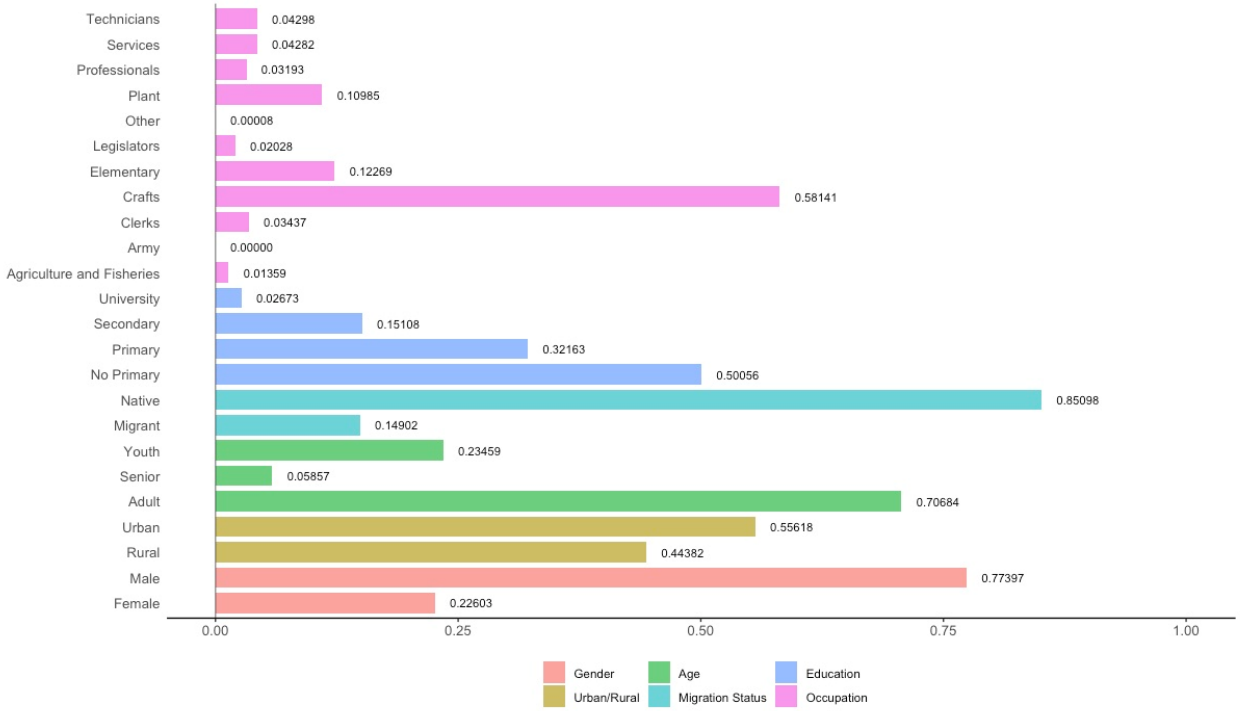
Note: Each point is the estimated coefficient, along with its 90% confidence interval, of a specification in which the outcome of interest (the share of employment in each of the services industries reported) is regressed against one of the variables reported on the y-axis. A full description of these variables in provided in the Appendix. All regressions include country and wave fixed effects, with standard errors are clustered at the district level. Regressions on undefined industries (other and unspecified services) are not included.

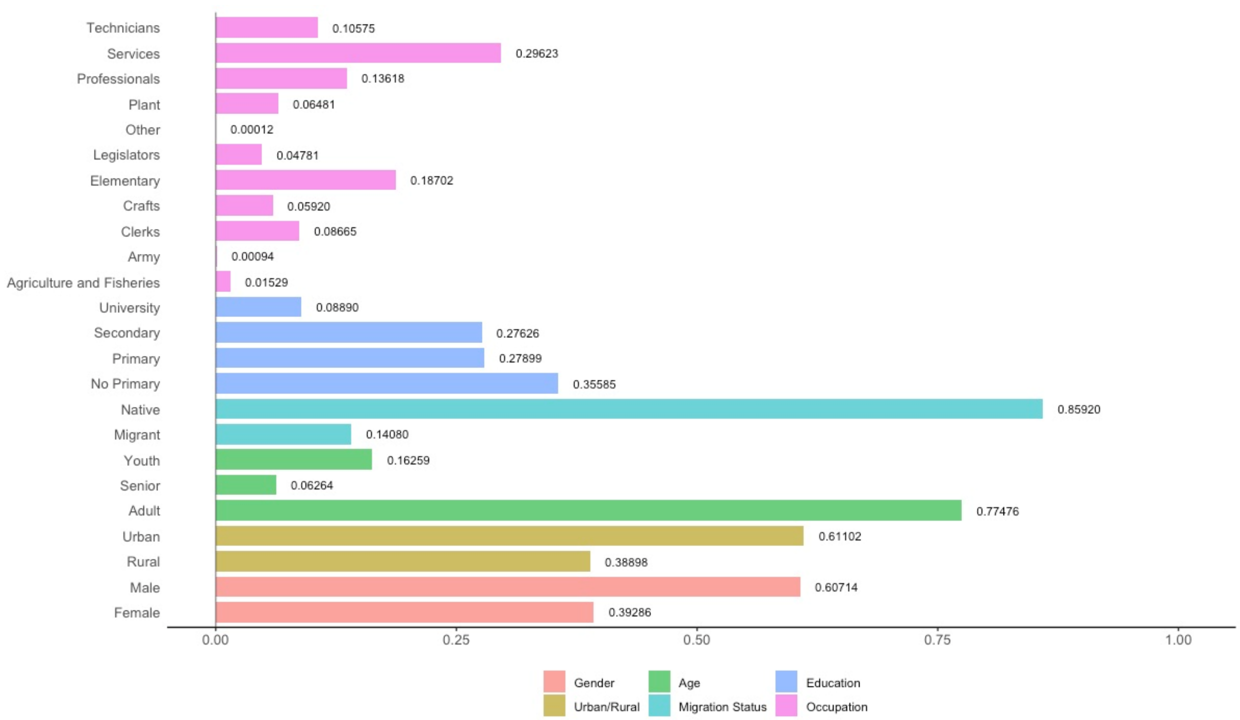
Who works in services?

While services are important sources of employment in Africa, surprisingly little is known about their characteristics as well as their composition. In what follows we focus on the following characteristics: gender; urban/rural residence; age cohorts; migration status; education and occupation. We collapse data over the latest two waves. Figure 8 summarizes the average values across all the countries included in our sample. Some patterns emerge with respect to the type of occupation (some of the more skilled ones are concentrated in the tertiary sector) and education. Regarding the latter, 8.9% of those employed in the services hold a university degree (2.8% in manufacturing) and 27.6% have a secondary school level (15.1% in manufacturing). Services employ a relatively higher share of women, some 40%, almost double that for the secondary sector. On the other hand, industry and services show relatively similar patterns in terms of the share of migrant workers employed (around 14% of the workers are internal migrants). Finally, younger cohorts of workers are less represented in services, compared to both the primary and secondary sectors.

Figure 8. Sectoral composition (%) by groups for primary (top panel); secondary (middle panel) and tertiary activities (bottom panel)







Source: Authors' elaboration on IPUMS

The Appendix and data tool provides all the possible decompositions of the previous graphs by industries and country-industry pairs. As far as the industries within services are concerned, education, public administration, financial and business services employ a relatively large share of more highly educated individuals. Private household services, health and accommodation are female-dominated activities. Again, many services industries employ a low share of youth or migrants.

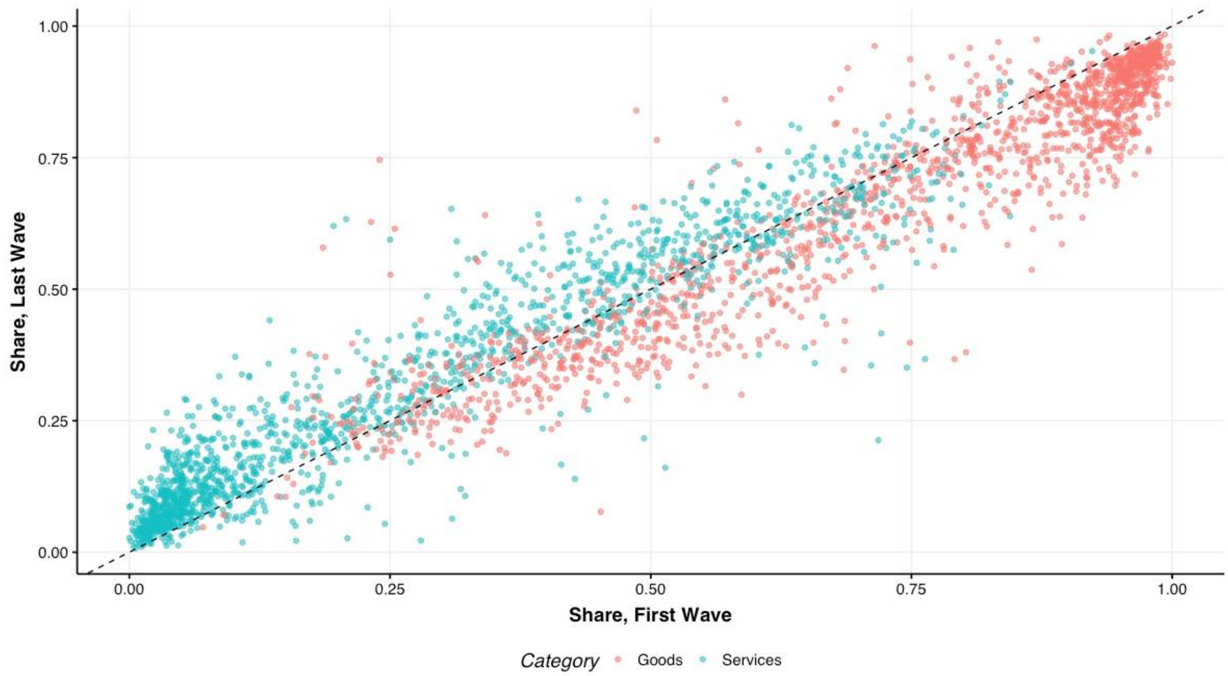
Occupations

There is a rising emphasis on the role of occupations as opposed to industries or sectors in the literature on structural transformation (Lagakos and Shu, 2021). Industries are not necessarily precise categories. For instance, individuals formally employed in the manufacturing sector might perform services-related functions. There is evidence from developed countries of this potential misclassification (Berlingieri, 2014), and data shown in the previous section seem to confirm that this is an issue that can also be relevant in our sample of African countries (for instance, figure 10 shows that some of the workers formally employed in the manufacturing sector are in fact performing service-related occupations as clerks, professionals or technicians, or craft workers in the services).

Duernecker and Herrendorf (2020) have proposed a distinction between categories of occupations that broadly map into the traditional distinction among the main economic sectors, although they are not necessarily attached to a specific sector. The categories are (1) *goods occupations*, which are related to the production of *tangible* value added; and (2) *services occupations*, which are related to the production of *intangible* value added. Employing census data for a large group of developed and developing countries, they show that the typical pattern of structural transformation (i.e., an increase in services as GDP per capita grows) holds also when using such classification. Given that they employ the same data we use in this report, we can replicate their exercise for our sample, and check whether results remain consistent.[21] Figure 9 shows that this is indeed the case. The shift towards intangible activities is evident for the last decades, and this pattern seems true for almost all the individual countries in our sample, with the exception of Benin and Mali. In these two countries, while services are on the rise, goods occupations –especially those related to agriculture—have also kept going up over the last decade.

[21] In the sample, goods occupations include agriculture and industry (elementary, crafts and plant) occupations. Services occupations include the following: armed forces; clerks; elementary services; legislators; professionals; services; technicians.

Figure 9. Structural transformation at the occupational level



Note: Each dot represents an administrative unit that is observed over two successive waves of the census.
Source: Authors' elaboration on IPUMS

The results plotted in Figure 9 suggest that the increases in the shares of secondary sector employment registered by many administrative units reported in Figure 3 involve intangible (services) occupations. To check if this is the case, we run two exercises. First, we replicate Figure 9 with data relative to the manufacturing sector only.[22] Doing this allows us to observe more clearly that service occupations are growing the most in relative term during the most recent decade (see Appendix Figure C2). Second, we run a simple regression in which we link changes in sectoral employment to changes in specific types of occupation. For the manufacturing sector, we find that a 1 percentage point increase in intangible occupations correlates with employment growth (Appendix Table A3), while this is not the case for tangible occupations.[23] This pattern would be consistent with the general trend towards servicification of economic activity in manufacturing.[24] Insofar as this is the case, it may be one factor explaining the patterns of

[22] Considering the whole sample, the share of services (goods) occupations among those employed in manufacturing is about 30% (70%).

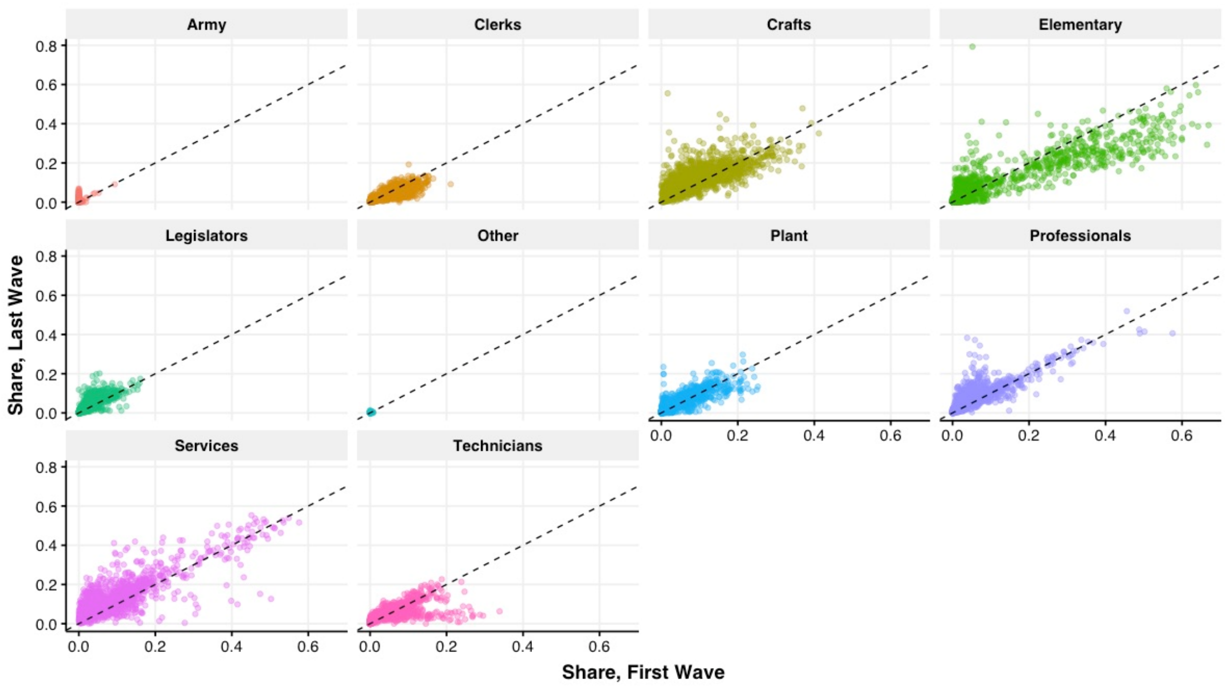
[23] Running the same regression on agriculture and services reveals that changes in employment shares are strongly correlated with the two types of occupations, tangible and intangible.

[24] Servicification denotes the phenomenon of economic production across manufacturing sectors increasingly relying on services, whether as intermediate inputs, as activities performed within the firm or as output sold together with goods (Miroudot and Cadestin, 2017). Econometric analysis of the economic implications of services intensity

growing and declining shares of secondary sector employment observed across administrative units in the 13 countries in our sample. This is a question where further research using more granular information and firm-level data is needed.

Figure 10 plots changes in each occupation covered in our data for the whole sample of countries and over the last two decades. We observe a drop in elementary types of occupations, and a rise of those generally accounted for as high skilled, such as professionals, managers (the "legislators" category) and other services occupations.

Figure 10. Occupation specific change



Source: Authors' elaboration on IPUMS
Note: Each dot represents an administrative unit that is observed over two successive waves of the census. The agro-fishery occupation is excluded to allow a better visualization of smaller occupations (given that all share the same value in the Y-axis).

Clustering services

Economic characteristics, such as tradability, potential productivity growth, economic spillovers, scalability, and skill intensity can vary substantially across services activities. The literature suggests

---

and services-related policy often rely on the servicification channel by investigating the impact of services variables on downstream sectors that use services in their production (e.g., Beverelli et al. 2017).

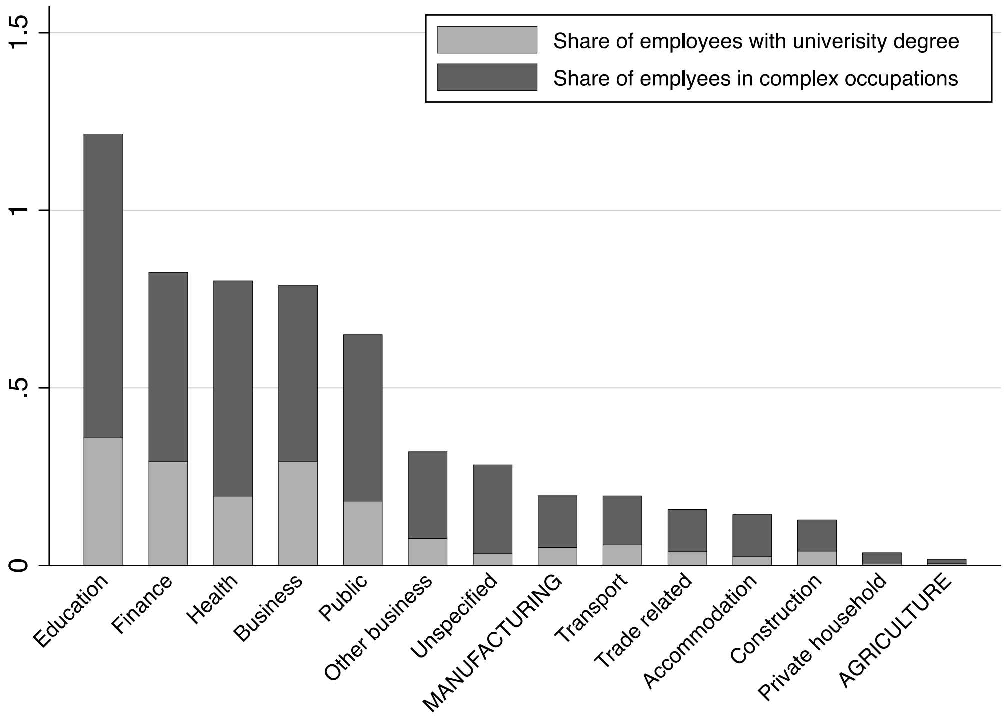
several approaches to cluster individual services sectors in terms specific properties depending on the relevant application. Research on services and global value chains, including the spillover-effects of services trade policy to downstream industries that use services as intermediate inputs, focus on the group of so-called producer services. This includes R&D, financial, business, transport, telecommunications and wholesale trade services, which are used as inputs in virtually all modern production processes (Beverelli et al., 2017). As we are interested in the relationship between services and economic development, we cluster services sectors in terms of employment characteristics which are both measurable in our data and relevant for economic development. These are: (i) intensity in high skilled labor, as captured by the share of employees with a university degree in a sector; and (ii) intensity in complex occupations, measured through the share of legislators/senior officials/managers, professionals and technicians employed in each sector.[25]

Figure 11 plots all services sectors identified in our database ranked in terms of the share of high skilled workers and complex occupations. The chart also reports the same figure for manufacturing and agriculture. We use manufacturing as a divide to identify two groups of services sectors: the cluster of "high skilled, complex occupation" (high skills) sectors, including education, finance, health, business, public services, other business and unspecified services; and the cluster of "low skilled, simple occupations" (low skills) sectors, including transport, trade related, accommodation, and private household services. According to our metric, seven services sectors rely more on skilled-labor and complex occupations than manufacturing. Agriculture has the lowest high skilled labor and complex occupation intensity. While our approach is somewhat discretionary, within the framework of our application it is robust to alternative definitions of the ranking used to identify the two clusters.[26]

[25] The shares are computed at the level of the whole sample. Education categories in the data are identified using IPUMS classification and span less than primary; primary; secondary; and university. Occupational categories classified as complex are legislators, senior officials and managers; professionals; and technicians and associate professionals.

[26] We computed 12 other rankings based on the following metrics: 1. share of employees with university degree; 2. share of employees with university or secondary degree; 3. share of complex occupations; 4. share of complex occupation including clerks; 5. share of employees in resident in urban areas (urban workers); 6. share of employees with university or secondary degree plus share of complex occupations; 7. share of employees with university or secondary degree plus share of complex occupations plus share of urban workers; 8. share of employees with university or secondary degree plus share of complex occupations including clerks; 9. share of employees with university or secondary degree plus share of complex occupations including clerks plus share of urban workers; 10. share of employees with university degree plus share of complex occupations plus share of urban workers; 11. share of employees with university degree plus share of complex occupations including clerks; 12. share of employees with university degree plus share of complex occupations including clerks plus share of urban workers. If we take the average score across all 13 rankings, the 7 high skills sectors correspond to the 7 highest ranked sectors according to the average ranking.

Figure 11: Services sectors, skills and occupations



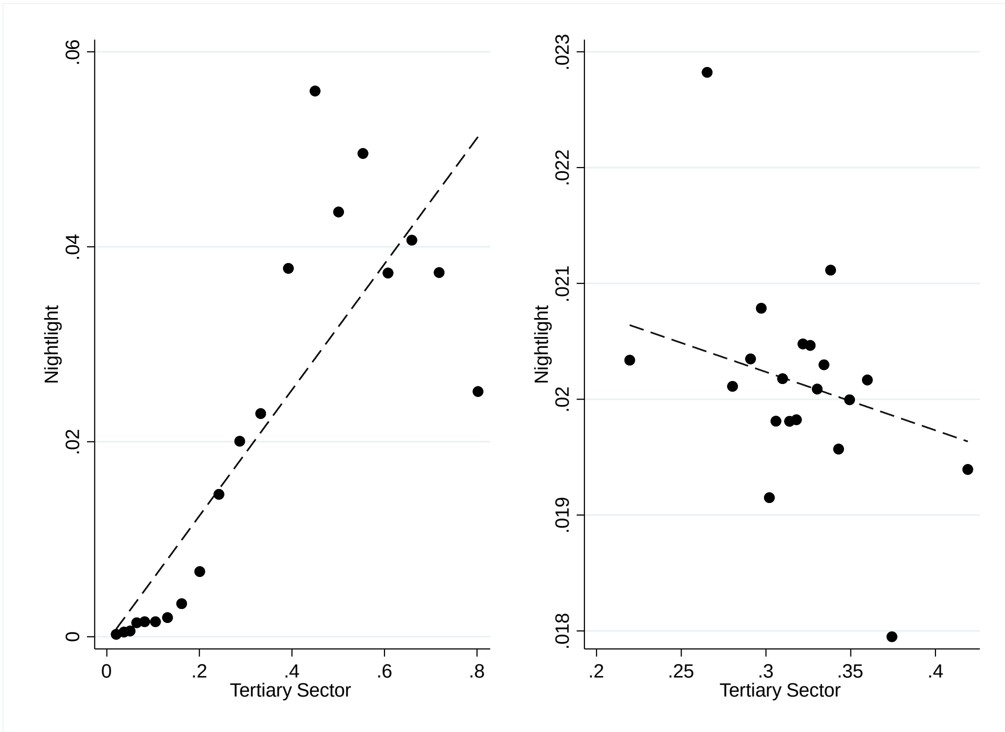
4. Services and development

Figure 12 shows correlations between nightlight *per capita*,[27] our proxy of economic growth, and the share of the tertiary sector, using binned scatterplot. The panel on the left is a barebones correlation, whereas the panel on the right accounts for ADMIN and wave fixed effects. The takeaway messages are clear-cut: 1) there is a strong positive association between growth and services *across* administrative units; and 2) there is a negative (weaker) correlation between growth and services *within* units.[28]

[27] Nightlight per capita are defined as the share of the admin level of nightlight density divided by population. For this exercise, data on night light are those using the DSMP satellite information, which allow to recover information for the earlier periods covered in our sample. For robustness, we have also run all regressions using the recently released data from Li et al. (2020).

[28] In Online Appendix B, we report the same figures for the primary and secondary sector. There is always a negative (positive) correlation between nightlight and primary (secondary) sector with or without fixed effects.

Figure 12 Correlation between nightlight and share of the tertiary sector



Note: binned scatterplot. The left side panel shows the simple correlation; the panel on the right plots correlations accounting for ADMIN and wave fixed effects.

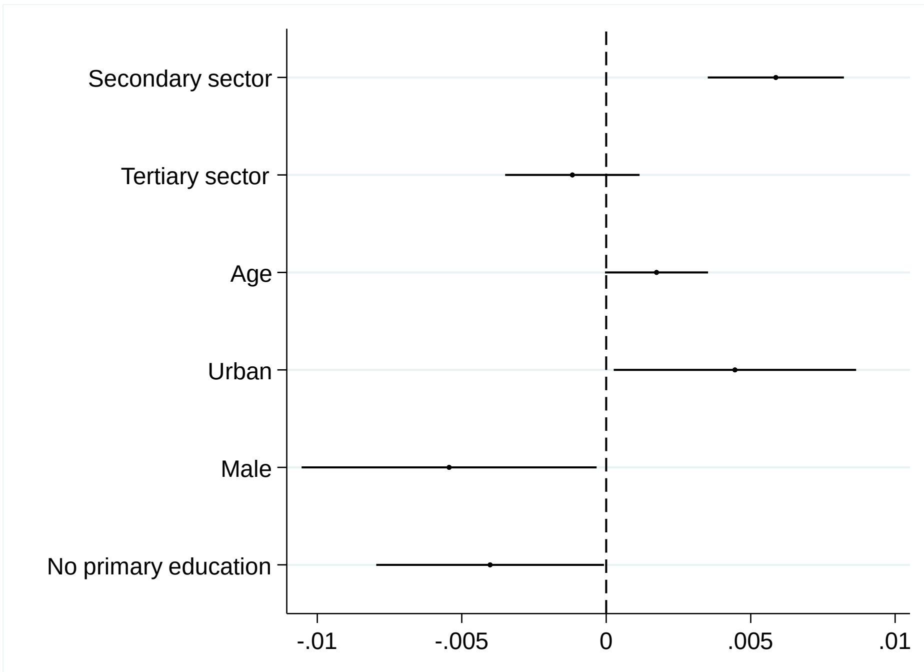
Our reduced form analysis is based on the following baseline model:

$$Nightlight_{it} = \alpha + \delta_i + \tau_t + \beta_1 Secondary\ Sector_{it} + \beta_2 Tertiary\ Sector_{it} + \beta_3 Z_{it} + \varepsilon_{it}, \quad (2)$$

where nightlight *per capita* is the outcome, i.e. sum of nightlights in a ADMIN divided by its population. We observe this outcome in each ADMIN *i* and in each wave *t*. *Secondary Sector* and *Tertiary Sector* are share of worker employed in respectively industry and services in each ADMIN and in each wave. Thus, the share of workers employed in the primary sector is the baseline category in these regressions. Z is a matrix including our controls (share of people living in urban areas, share of male population, average age in an ADMIN, and share of people without primary education), δ and τ are ADMIN and wave fixed effects, and ε are the residuals.

We run OLS regressions weighted by ADMIN’s population. We cluster the standard errors at the level of the ADMIN unit. For ease of comparison among covariates, we standardize all the right hand-side variables. We report here only the results of the baseline model.[29] The estimation sample includes around 3,000 observations, depending on the model. Our unit of analysis is district-census wave. The results of this baseline model are reported in Figure 13. We find that the association between development and the tertiary sector is negative, but not significant at the conventional level. This null correlation may reflect the heterogeneity within the service sector, something we explore in the following section.

Figure 13. Nightlight per capita: Results by sector



Note: OLS regression weighted by population with standard errors clustered by ADMIN units, ADMIN and wave fixed effects. N= 2,909; R^2: 0.965; 90% C.I.

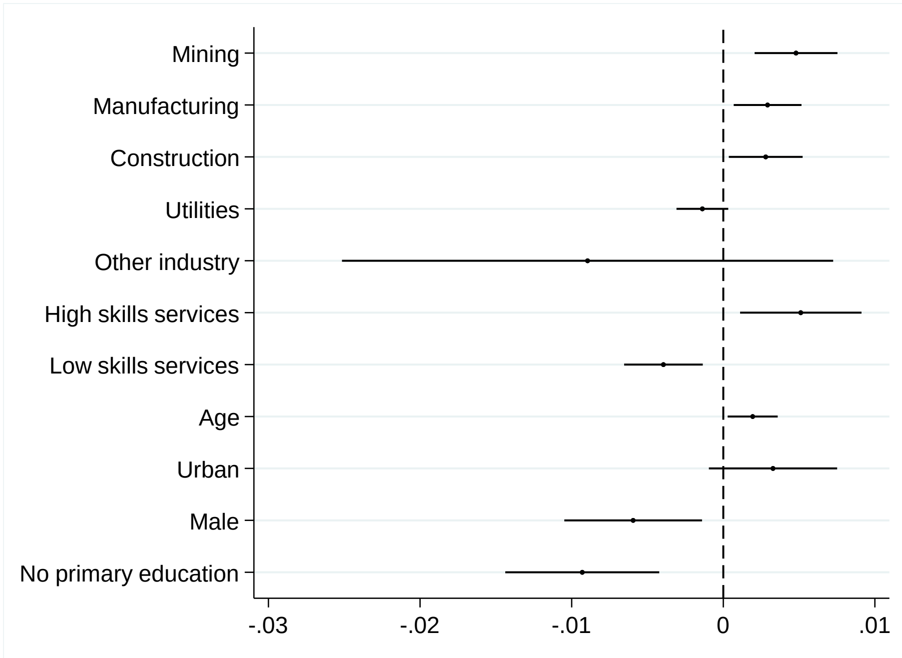
By macro service clusters

In what follows we distinguish between high and low skills services, using the categorization developed above. The regressions show that there is a positive significant relationship between economic growth and the share of workers employed in high skills services (Figure 14). The magnitude of the positive correlation is in line and in fact larger than the magnitude of the positive correlation between nightlight and manufacturing. On the contrary, there is a negative significant association between economic growth and share of workers employed in low skills services. These results help understand the null effect of the baseline analysis.[30]

[29] Online Appendix Table B1 reports all model specifications, with and without fixed effects as well as with and without controls.

[30] We account for the share of population without primary education, which is negatively correlated with development. Thus, our macro service clusters are not a mere proxy of education in this model.

Figure 14. Nightlight per capita: high skill services vs. low skill services

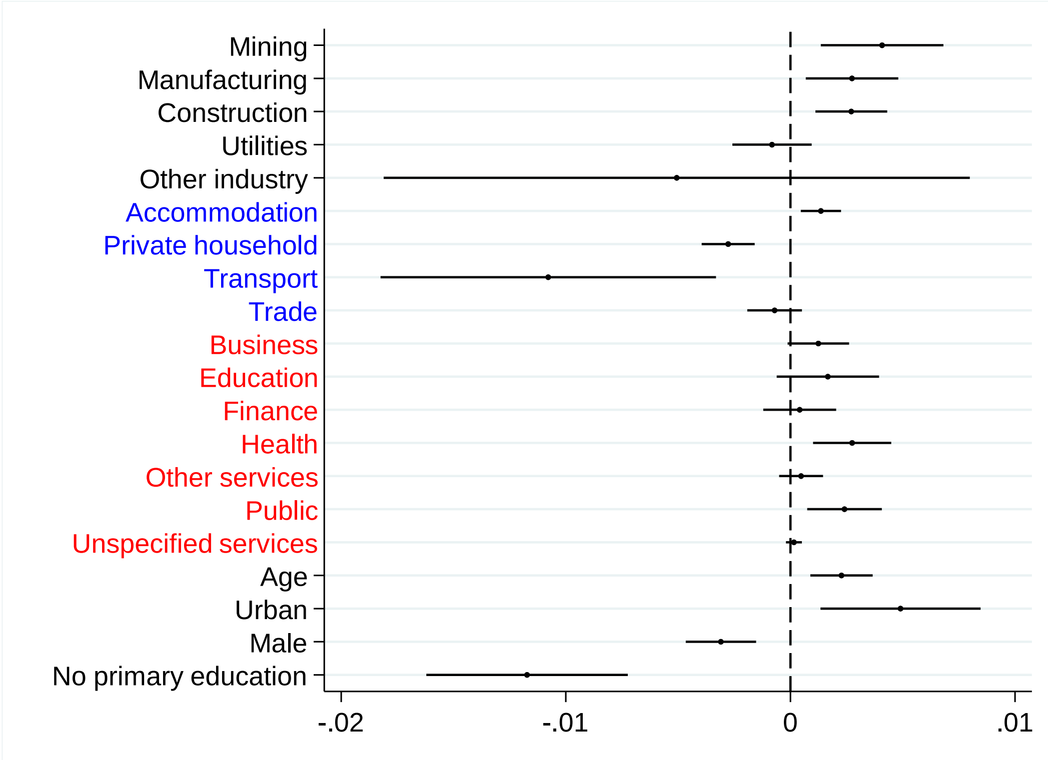


Note: OLS regression weighted by population with standard errors clustered by ADMIN units, ADMIN and wave fixed effects. N=2,909; R^2: 0.967. 90% C.I.

By individual service industries

To determine which service industries are more associated with economic growth, we run our main model (2) including 2-digit industries. Figure 15 reports the results of this analysis. To ease the reading of the results, high skills services are highlighted in red, whereas low skills services are in blue. Strikingly, all the high skill service industries are positively associated with economic growth, though only health and public services are significant. Conversely, private household services, transportation, and trade, which are low skills services, are negatively correlated with growth, though only the first two variables are significant. Accommodation, which are also low skill, are an exception, being correlated positively with growth.

Figure 15 Nightlight per capita: Results by industry

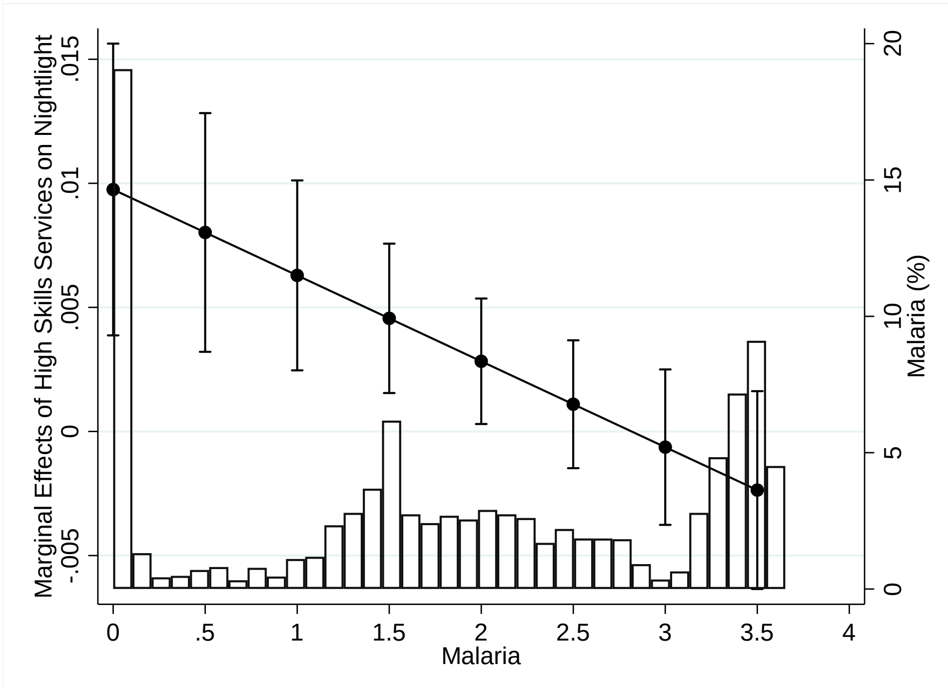


Note: OLS regression weighted by population with standard errors clustered by ADMIN units, ADMIN and wave fixed effects. N=2,909; R^2: 0.972; 90% C.I.

Investigating heterogeneity across economic environments

The correlation of services and development may be affected by heterogeneity across economic environments. We are interested in three dimensions of this heterogeneity: 1) the mediating effect of (lack of) economic activities at the baseline, which we capture with the incidence of malaria; 2) the mediating effect of natural resources, which we capture with the presence of diamond mines and oil extraction; 3) the mediating effect of technology, proxied by mobile phone coverage. To analyze this heterogeneity, we interact each mediating variable with share of high and low skills services as well as with the other industries, e.g., mining and manufacturing. We then run the same model specification as described in equation (2). To ease the interpretation of the interaction terms, we show the results graphically. We focus on the interaction between high skills services and mediating variables. We start with the incidence of malaria, where low values correlate with thriving economic activities (Acemoglu et al., 2001). Figure 16 reports the marginal effect of high skills services on nightlight *per capita* for different level of incidence of malaria. The positive correlation between high skills services and growth is only significant for low incidence of malaria. On the contrary, for incidence of malaria (roughly) above the mean, the correlation between high skills services and growth is no longer significant, i.e., the line of the marginal effect crosses the 0. In sum, we find convincing evidence that the presence of economic activities at the baseline mediates the positive association between high skills services and development.[31]

Figure 16. High skills services and nightlight per capita: Mediating effect of malaria incidence



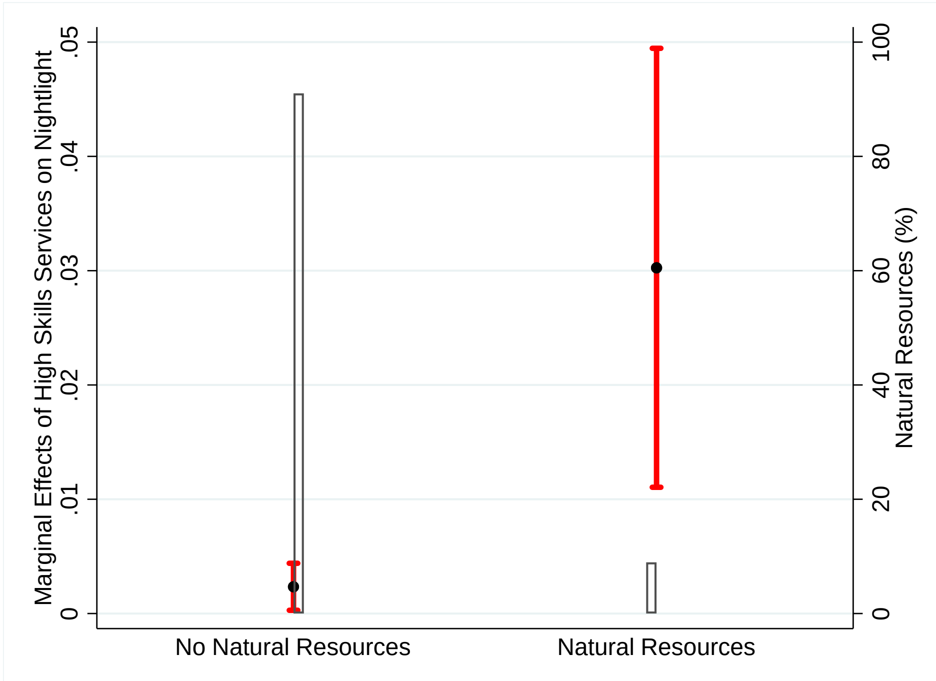
Note: OLS regression weighted by population with standard errors clustered by ADMIN units, ADMIN and wave fixed effects. N= 2,829; R^2: 0.970; 90% C.I.

Turning to the mediating effect of natural resources, we interact high skills services with a dummy equaling one if there is at least one diamond mine or oil well in the ADMIN unit. Figure 17 reports the marginal effect of high skills services on nightlight *per capita* for units with and without mines of diamond. While the positive correlation between high skills services and growth is significant in both regions with and without mines of diamond, the correlation is significantly stronger in regions with mining. This finding suggests that the presence of natural resources and related wealth provides the demand for the expansion of the high skills services, a result in line with Gollin et al. (2016).[32]

[31] In separate regressions, not reported, we show that the incidence of malaria does *not* mediate the effect of manufacturing on development.

[32] Unreported regressions indicate that natural resources do *not* mediate the effect of manufacturing on development.

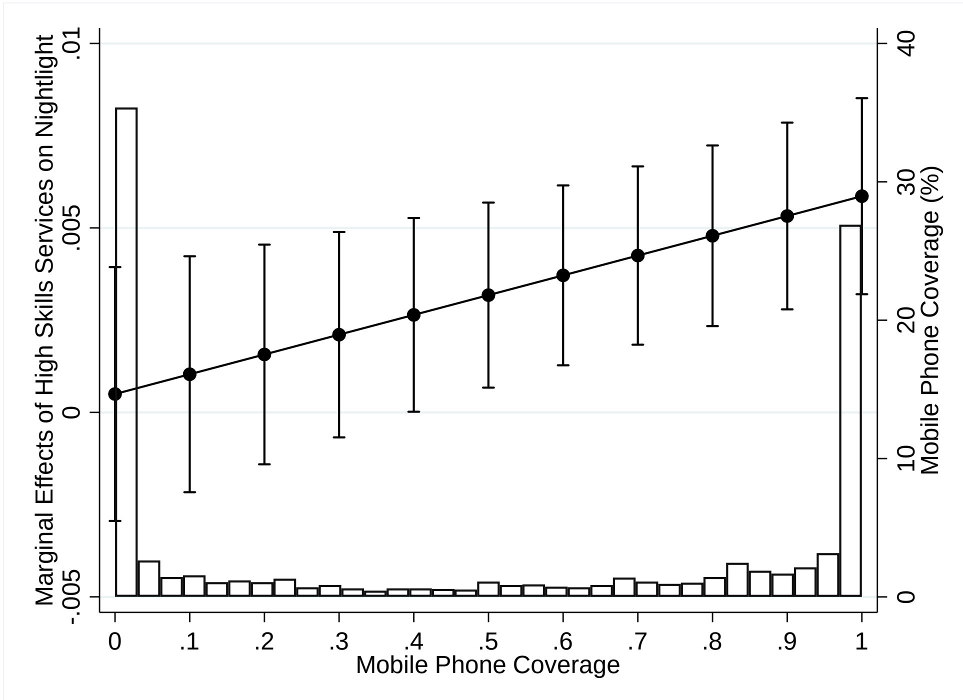
Figure 17. High skills services and nightlight per capita: Mediating effect of natural resources



Note: OLS regression weighted by population with standard errors clustered by ADMIN units, ADMIN and wave fixed effects. N = 2,811; R^2: 0.966. 90% C.I.

Finally, we explore the mediating role of technology by interacting high skills services with coverage of the 3G network using data from Manacorda and Tesei (2020). The rationale is that technology provides the supply for the expansion of services. Figure 18 reports the marginal effect of high skills services on nightlight *per capita* for different levels of internet coverage. The positive correlation between high skills services and growth is only significant for high values (i.e., roughly above the mean) of coverage, whereas is not significant for lower values. This finding points to technology being an important intervening variable to explain the positive association between some services activities and development.

Figure 18 High skills services and growth: Mediating effect of technology



Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMIN and wave fixed effects. Number of observations: 1,958; R^2: 0.975. 90% C.I.

5. Conclusion

This paper contributes to the literature on structural transformation and economic development by providing detailed data on changes in the composition of employment over time at the administrative unit level for thirteen African countries. The analysis reveals novel evidence of structural transformation towards services sectors and service-related occupations at sub-national level, providing a fine-grained picture of changes in sectoral employment shares and occupations. The data depict a clear shift towards services in all the countries in the sample. Across all the administrative units in our sample of countries there is a decline in the total share of employment in agriculture of six percentage points, offset by an equivalent increase in the share of employment in services sectors. As argued by Lagakos and Shu (2021), exploiting the potential of microdata to study issues traditionally analyzed using aggregate or sectoral level information can help to account for the high degree of heterogeneity across services activities as a function of observable characteristics of workers.

Although the overall share of secondary sector employment is stable, there is significant heterogeneity across sub-national geographic units. The data reveal reductions in industrial employment shares in locations where such activity accounted for relatively high shares of total employment in the initial census year, but also numerous increases in areas where industry accounted for relatively low shares of total employment. There is weak evidence that manufacturing employment growth in these locations is associated with servicification, in the sense that employment in occupations associated with intangible outputs appears to grow in many regions where the share of employment in the secondary sector increases between two census waves.

Exploratory analysis of the relationship between services and economic development, using per capita nightlight luminosity as a proxy for growth, reveals no evidence that services in the aggregate are associated with economic development. There is however substantial heterogeneity across different services industries. When we distinguish between high and low skills services sectors by sorting services sectors by intensity of use of workers with a university degree or engaged in occupations that are more complex, classifying sectors that use these categories of workers more intensively than the average observed in manufacturing as high skill, we find that higher-skilled services are strongly associated with development. The positive association between high skill services sectors and development is mediated by geography, institutions, and technology. Greater incidence of malaria, the presence of a mining facility and below average mobile phone coverage in an administrative unit reduces or undoes the significance of the positive association.

Services activities that account for much of the growth of service-sector employment are concentrated in trade (distribution) and transport, low-skilled services that are negatively associated with our nightlights proxy for growth. This pattern underlies the overall null correlation observed between services in the aggregate and development. What explains this pattern is an important question for research. One factor may be that these activities grow more rapidly in early stages of structural transformation, associated with small-scale firms. They may become more productive over time as economies of scale are realized in these sectors and infrastructure improves. Investigating these types of factors requires complementing the census data with firm-level information at the administrative unit level. Doing so is a necessary condition for developing a better understanding of the role services play in structural transformation and the prospects for increasing productivity and employment generation. While lower skill services activities may have less potential to support growth, these activities do generate employment and thus income. Other research has found that the poverty-reducing effects of growth in service activities such as wholesale and retail trade and transportation are similar (and sometimes greater) than that associated with equal growth in agriculture or manufacturing sectors (Dorosh and Thurlow, 2018). From an aggregate growth perspective, however, our results suggest policy should focus on supporting expansion of higher-skill activities.

Our findings suggest several areas on which future research could focus. One is to analyze the evolution within services across the sampled countries to understand better the relationship between high skills services and economic development. A corollary research question pertains to the role of services in explaining manufacturing employment dynamics, including analysis of the extent to which servicification is occurring in manufacturing. More broadly, insofar as high skills services are associated with development, future research to assess the drivers of demand for such services and possible complementarities between services activities would seem apposite. Another area for research suggested by the data concerns the distributional implications of the shift towards services. Existing research has shown that the type of structural transformation towards services matters to understand on development. Using sub-national data for developed and developing countries, Chatterjee and Giannone (2021) find that the development of highly productive types of services is associated with greater regional inequality.[33] Further research on such questions using micro data for African countries is important to understand better the implications of the shift to services and to identify policies that can help address adverse distributional consequences.

References

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

Baccini, L., M. Fiorini, B. Hoekman and M. Sanfilippo (2021). Services, jobs, and economic development in Africa. EUI Working Paper RSC 2021/82.

Balchin, N., Hoekman, B., Martin, H., Mendez-Parra, M., Papadavid, P. and te Velde, D. (2016). *Trade in Services and Economic Transformation*. London: ODI.

Berlingieri, G. (2014). Outsourcing and the shift from manufacturing to services. LSE Centre for Economic Performance.

Beverelli, C., Fiorini, M. and Hoekman, B. (2017). Services trade policy and manufacturing productivity: The role of institutions. *Journal of International Economics*, 104:166–182.

Buera, F. and Kaboski, J. (2012). The Rise of the Service Economy. *American Economic Review*, 102(6),

Chatterjee, S. and E. Giannone (2021). Unequal Global Convergence. mimeo.

Dasgupta, S. and A. Singh (2007). Manufacturing, Services and Premature Deindustrialization in Developing Countries: A Kaldorian Analysis, in G. Mavrotas and A. Shorrocks (eds.) *Advancing Development: Core Themes in Global Economics*, London: Palgrave Macmillan, 435-54.

Diao, X., K. Harttgen, and M. McMillan (2017). The Changing Structure of Africa's Economies. *World Bank Economic Review*, 31(2): 412–33.

[33] This finding can be motivated by the fact that large agglomerations tend to be skill biased also in developing countries – see e.g., Dingel et al. (2021).

Diao X., McMillan M., Rodrik D. (2019) The Recent Growth Boom in Developing Economies: A Structural-Change Perspective. In: Nissanke M., Ocampo J. (eds) *The Palgrave Handbook of Development Economics*. London: Palgrave Macmillan.

Diao, X., Ellis, M., McMillan, M. and Rodrik, D. (2021). *Africa's Manufacturing Puzzle: Evidence from Tanzanian and Ethiopian Firms*. NBER Working Paper 28344.

Dihel, N. and Goswanmi, A. (2016). *From Hair Stylists and Teachers to Accountants and Doctors—The Unexplored Potential of Trade in Services in Africa*. Washington DC: World Bank.

Dingel, J., A. Miscio, D. Davis (2021). Cities, lights, and skills in developing economies, *Journal of Urban Economics*, vol. 125. 103174.

Dorosh, P. and J. Thurlow (2018). "Beyond Agriculture versus Non-Agriculture: Decomposing Sectoral Growth–Poverty Linkages in Five African Countries." *World Development* 109: 440–51.

Duarte, M. and Restuccia, D. (2010). The role of structural transformation in aggregate productivity. *The Quarterly Journal of Economics*, 125: 129–173.

Duarte, M. and Restuccia, D. (2020). Relative Prices and Sectoral Productivity. *Journal of the European Economic Association*, 18(3): 1400–1443.

Duernecker, G. and Herrendorf, B. (2020). Structural Transformation of Occupation Employment. mimeo.

Eckert, F., Ganapati, S. and Walsh, C. (2020). Skilled Scalable Services: The New Urban Bias in Economic Growth. FED Minneapolis Institute Working Paper.

Fan, T., Peters, M. and Zilibotti, F. (2021). Service-Led or Service-Biased Growth? Equilibrium Development Accounting Across Indian Districts. NBER Working Paper 28551.

Fang, L and Herrendorf, B. (2021). High-Skilled Services and Development in China. *Journal of Development Economics*, 151, 102671.

Gervais, A. and Jensen, J. B. (2019). The tradability of services: Geographic concentration and trade costs. *Journal of International Economics*, 118: 331–350.

Gollin, D., R. Jedwab, and D. Vollrath (2016). Urbanization with and without industrialization. *Journal of Economic Growth* 21: 35-70.

Kruse, H., E. Mensah, K. Sen, and G. de Vries (2021). *A manufacturing renaissance? Industrialization trends in the developing world*. WIDER Working Paper 28/2021.

Haraguchi, N., C.F.C. Cheng, and E. Smeets (2017). 'The Importance of Manufacturing in Economic Development: Has This Changed?' *World Development*, 93: 293–315.

Herrendorf, B., Rogerson, R. and Valentinyi, A. (2014). Growth and Structural Transformation. In *Handbook of Economic Growth*, Vol. 2. Amsterdam: Elsevier, 855-941.

Herrendorf, B., R. Rogerson, and A. Valentinyi (2020). Structural Change in Investment and Consumption: A Unified Analysis. *Review of Economic Studies*, 10: 1-37.

Hsieh, C. and E. Rossi-Hansberg (2021). The Industrial Revolution in Services. mimeo.

Kravis, I. B., Heston, A. W. and Summers, R. (1983). The Share of Services in Economic Growth. In *Global Econometrics: Essays in Honor of Lawrence R. Klein*. MIT Press.

Kuznets, S. (1966). *Modern Economic Growth: Rate, Structure, and Spread*. Yale University Press.

Lagakos, D. and Shu, M. (2021). The Role of Micro Data in Understanding Structural Transformation. STEG Pathfinding Paper.

Li, X., Zhou, Y., Zhao, M. and Zhao, X. (2020). A harmonized global night-time light dataset 1992–2018. *Scientific Data*, *7*(1), 168.

Loungani, P. and Mishra, S. (2014). Not Your Father's Service Sector. *Finance & Development*, 51(2): 51-54.

Manacorda, M. and Tesei, A. (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *Econometrica*, *88*(2), 533–567.

McMillan, M., Rodrik, D. and Verduzco-Gallo, Í. (2014). Globalization, structural change, and productivity growth, with an update on Africa. *World Development*, 63, 11–32.
Mensah, E. (2020). Is Sub-Saharan Africa Deindustrializing? GGDC Research Memorandum 186. Groningen Growth and Development Centre.
Michalopoulos, S. and Papaioannou, E. (2020). Historical Legacies and African Development. *Journal of Economic Literature*, *58*(1), 53–128.
Minnesota Population Center. (2019). *Integrated Public Use Microdata Series, International: Version 7.2*. IPUMS International.
Miroudot, S. and C. Cadestin (2017). Services in Global Value Chains: From Inputs to Value-Creating Activities. OECD Trade Policy Papers 197.
Nayyar, G., M. Hallward-Driemeier and E. Davies (2021). *At your service? The promise of services-led development.* Washington DC: World Bank.
Newfarmer, R., Page, J. and Tarp, F. (2018). *Industries without Smokestacks*. Oxford University Press.
Nguimkeu, P. and A. Zuefack (2019). Manufacturing in Structural Change in Africa. Policy Research Working Paper 8992. Washington, DC: World Bank.
Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth*, *21*(1), 1–33.
Schettkat, R. and Yocarini, L. (2006). The shift to services employment: A review of the literature. *Structural Change and Economic Dynamics*, 17(2), 127–147.
Sen, K. (2019). Structural Transformation around the World: Patterns and Drivers, *Asian Development Review*, 36(2): 1–31.
Syrquin, M. (1988). "Patterns of Structural Change." In H. Chenery and T.N. Srinivasan (eds.) *Handbook of Development Economics*, Vol. 1. Amsterdam: Elsevier, 203–73.
Young, A. (2014). Structural Transformation, the Mismeasurement of Productivity Growth and the Cost Disease of Services. *American Economic Review*, 104(11): 3635–3667.

APPENDICES

Appendix A

Table A1: Full results by main sectors based on equation (1)

	(1)	(2)	(3)	(4)	(5)	(6)
	Agriculture	Mining	Manufacturing	Utilities	Construction	Tertiary
ntl_pc	-2.100***	0.198*	0.604***	0.00982	0.416***	0.881***
	(0.354)	(0.106)	(0.129)	(0.0151)	(0.0643)	(0.251)
ntl_pc^2	1.016***	-0.0535	-0.202*	0.0429***	-0.202***	-0.556***
	(0.321)	(0.0488)	(0.119)	(0.0112)	(0.0297)	(0.189)
pop	0.353***	0.0636	-0.0258	-0.0368***	-0.0758***	-0.314***
	(0.129)	(0.0697)	(0.0427)	(0.00591)	(0.0229)	(0.0890)
pop^2	-0.0190***	-0.00209	0.00192	0.00168***	0.00362***	0.0155***
	(0.00539)	(0.00281)	(0.00179)	(0.000256)	(0.000968)	(0.00374)
post_2000	-0.0656***	-0.00436***	0.000364	0.000494**	0.00993***	0.0597***
	(0.00657)	(0.00169)	(0.00182)	(0.000238)	(0.00120)	(0.00508)
Constant	-0.876	-0.442	0.0956	0.203***	0.415***	1.802***
	(0.780)	(0.427)	(0.254)	(0.0342)	(0.135)	(0.538)
Observations	3,135	3,135	3,135	3,135	3,135	3,135
R-squared	0.971	0.920	0.913	0.881	0.899	0.969

Notes: Estimates are based on equation (1). NTL_pc is the value of nighttime lights per capita. pop is the log of population. The squared term of both the former variables is included. Post_2000 is the variable of interest, a dummy taking 1 if the year of the census is successive to 2000. All regressions include district fixed effects. Robust standard errors in parenthesis. District. *** p<0.01, ** p<0.05, * p<0.1

Table A2: Full results by industries within services based on equation (1)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
VARIABLES	trade	accommodation	transport	finance	public	business	education	health	private_hh	other_services	unspec_servic
ntl_pc	-0.0465	0.430***	-0.0739	0.0608***	-0.145	-0.0843**	0.256***	0.148***	-0.219***	0.625***	-0.0792***
	(0.104)	(0.0654)	(0.255)	(0.0149)	(0.110)	(0.0371)	(0.0560)	(0.0435)	(0.0790)	(0.0874)	(0.0242)
ntl_pc^2	-0.0776	-0.162***	-0.215	-0.0228**	0.148*	0.0553***	-0.0416	-0.0990***	0.108**	-0.307***	0.0143
	(0.0615)	(0.0235)	(0.144)	(0.00950)	(0.0860)	(0.0191)	(0.0351)	(0.0273)	(0.0503)	(0.0617)	(0.0127)
pop	-0.206***	-0.0286	-0.0670***	-0.00998*	0.256***	-0.00127	0.0597***	-0.0951***	-0.0949***	-0.111***	0.0222
	(0.0585)	(0.0226)	(0.0259)	(0.00530)	(0.0361)	(0.0186)	(0.0217)	(0.0232)	(0.0240)	(0.0239)	(0.0358)
pop^2	0.00812***	0.00195**	0.00328***	0.000562***	-0.0115***	0.000198	-0.00201**	0.00441***	0.00456***	0.00585***	-0.00167
	(0.00251)	(0.000958)	(0.00104)	(0.000216)	(0.00159)	(0.000790)	(0.000933)	(0.000918)	(0.00107)	(0.00102)	(0.00162)
post_2000	0.0337***	0.00562***	0.00919***	-0.000309	-0.0108***	0.00561***	0.00784***	0.00505***	0.00162*	0.00762***	-0.00591***
	(0.00242)	(0.00126)	(0.00135)	(0.000292)	(0.00150)	(0.000809)	(0.00102)	(0.00134)	(0.000829)	(0.00111)	(0.00106)
Constant	1.371***	0.0721	0.357**	0.0444	-1.362***	-0.00209	-0.386***	0.519***	0.510***	0.499***	-0.0234
	(0.343)	(0.135)	(0.163)	(0.0326)	(0.205)	(0.109)	(0.127)	(0.144)	(0.135)	(0.141)	(0.197)
Observations	3,135	3,135	3,135	3,135	3,135	3,135	3,135	3,135	3,135	3,135	3,135
R-squared	0.904	0.647	0.910	0.921	0.901	0.796	0.941	0.859	0.907	0.716	0.466

Notes: Estimates are based on equation (1). NTL_pc is the value of nighttime lights per capita. pop is the log of population. The squared term of both the former variables is included.
Post_2000 is the variable of interest, a dummy taking 1 if the year of the census is successive to 2000.
All regressions include district fixed effects. Robust standard errors in parenthesis. District. *** p<0.01, ** p<0.05, * p<0.1

Table A3: Correlations between changes in sectoral employment and type of occupation

VARIABLES	(1) Agriculture	(2) Manufacturing	(3) Services
Goods (tangible) occupations	0.751***	0.180	-0.412**
	(0.245)	(0.129)	(0.186)
Services (intangible) occupations	-0.310	0.232*	0.523***
	(0.247)	(0.129)	(0.189)
Constant	-0.0396***	0.00844***	0.0211***
	(0.00204)	(0.00123)	(0.00125)
Observations	1,528	1,520	1,529
R-squared	0.712	0.210	0.867

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Notes: Each column reports a regression linking the first difference in the share of employment in one of the three sectors (agriculture, manufacturing, services) against the first difference in the share of workers classified into goods (tangible) or services (intangible) types of occupations according to the definition by Duernecker and Herrendorf (2020). The latter refers only to those employed within a specific sector. In order to compute the first difference in an homogenous manner, the estimation sample includes the two most recent waves for each country.

Appendix B

Table B1 Main analysis

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS					
	Nightlight *per capita*					
Secondary sector	-0.000	0.006***	0.005***	-0.002	0.006***	0.006***
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Tertiary sector	0.010***	0.003**	-0.000	0.012***	-0.001	-0.001
	(0.002)	(0.001)	(0.002)	(0.002)	(0.001)	(0.001)
Constant	0.018***	0.019***	0.020***	0.017***	0.020***	0.019***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Admin FE	No	Yes	Yes	No	Yes	Yes
Wave FE	No	No	Yes	No	No	Yes
Controls	No	No	No	Yes	Yes	Yes
Observations	3,214	3,213	3,213	2,982	2,909	2,909
R-squared	0.075	0.962	0.963	0.188	0.965	0.965

Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMN and wave fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table B2 Macro service clusters

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS					
	Nightlight *per capita*					
Mining	0.009***	0.004***	0.004***	-0.119***	0.005***	0.005***
	(0.001)	(0.001)	(0.001)	(0.027)	(0.002)	(0.002)
Manufacturing	-0.001	0.004***	0.003***	0.010	0.003*	0.003**
	(0.002)	(0.001)	(0.001)	(0.066)	(0.001)	(0.001)
Construction	0.003***	0.004***	0.003*	0.113*	0.003*	0.003*
	(0.001)	(0.001)	(0.001)	(0.059)	(0.001)	(0.001)
Utilities	0.003	-0.000	-0.001	0.063	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.059)	(0.001)	(0.001)
Other industry	-0.000	0.000	-0.000	0.041	-0.009	-0.009
	(0.001)	(0.000)	(0.000)	(0.215)	(0.010)	(0.010)
High-skill services	0.011***	0.008***	0.007***	0.353***	0.005**	0.005**
	(0.002)	(0.003)	(0.003)	(0.072)	(0.002)	(0.002)
Low-skill services	-0.002	-0.002	-0.003	-0.090	-0.004**	-0.004**
	(0.001)	(0.002)	(0.002)	(0.061)	(0.002)	(0.002)
Constant	0.017***	0.018***	0.019***	0.519***	0.017***	0.015***
	(0.001)	(0.001)	(0.001)	(0.045)	(0.002)	(0.003)
Admin FE	No	Yes	Yes	No	Yes	Yes
Wave FE	No	No	Yes	No	No	Yes
Controls	No	No	No	Yes	Yes	Yes
Observations	3,214	3,213	3,213	2,982	2,909	2,909
R-squared	0.175	0.964	0.964	0.172	0.968	0.968

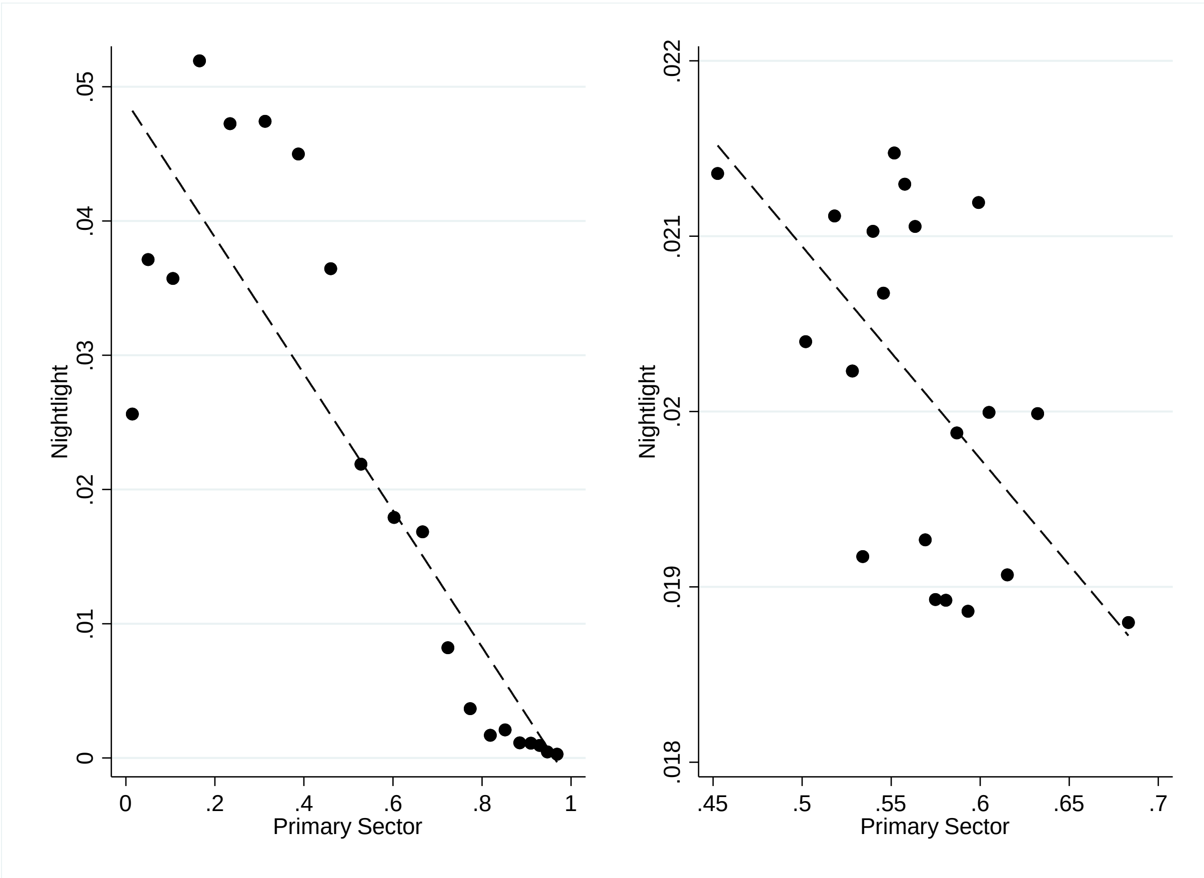
Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMN and wave fixed effects. *** p<0.01, ** p<0.05, * p<0.1

Table B3 Individual service industries

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS					
	Nightlight *per capita*					
Mining	0.008***	0.004**	0.003**	-0.096***	0.004**	0.004**
	(0.001)	(0.001)	(0.001)	(0.021)	(0.002)	(0.002)
Manufacturing	-0.003*	0.003***	0.003***	0.047	0.003**	0.003**
	(0.002)	(0.001)	(0.001)	(0.058)	(0.001)	(0.001)
Construction	0.002	0.003***	0.003***	0.116**	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.058)	(0.001)	(0.001)
Utilities	0.003*	-0.000	-0.001	0.013	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.058)	(0.001)	(0.001)
Other industry	-0.001	-0.001	-0.001	0.053	-0.005	-0.005
	(0.001)	(0.000)	(0.000)	(0.188)	(0.008)	(0.008)
Accommodation	0.001	0.002***	0.002***	0.095**	0.001**	0.001**
	(0.001)	(0.001)	(0.001)	(0.039)	(0.000)	(0.001)
Private household	0.004***	-0.002***	-0.002***	-0.243***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.038)	(0.001)	(0.001)
Transport	0.009***	-0.009*	-0.010**	-0.164***	-0.011**	-0.011**
	(0.003)	(0.005)	(0.005)	(0.060)	(0.005)	(0.005)
Trade	-0.011***	0.002**	-0.000	0.187***	-0.001	-0.001
	(0.002)	(0.001)	(0.001)	(0.066)	(0.001)	(0.001)
Business	0.005***	0.003**	0.002**	-0.026	0.001	0.001
	(0.001)	(0.001)	(0.001)	(0.051)	(0.001)	(0.001)
Education	-0.003*	0.002	0.001	0.251***	0.001	0.002
	(0.002)	(0.001)	(0.001)	(0.066)	(0.001)	(0.001)
Finance	-0.006***	0.001	0.001	-0.010	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.060)	(0.001)	(0.001)
Health	0.004***	0.004***	0.003***	0.102	0.003***	0.003***
	(0.001)	(0.001)	(0.001)	(0.063)	(0.001)	(0.001)
Other services	0.005***	0.002***	0.001	-0.151**	0.000	0.000
	(0.001)	(0.001)	(0.001)	(0.061)	(0.001)	(0.001)
Public	0.005***	0.002**	0.003***	0.191*	0.002**	0.002**
	(0.002)	(0.001)	(0.001)	(0.105)	(0.001)	(0.001)
Unspecified services	0.001***	0.001***	0.000*	-0.023	0.000*	0.000
	(0.000)	(0.000)	(0.000)	(0.019)	(0.000)	(0.000)
Constant	0.018***	0.019***	0.021***	0.513***	0.018***	0.016***
	(0.001)	(0.001)	(0.001)	(0.037)	(0.002)	(0.002)
Admin FE	No	Yes	Yes	No	Yes	Yes
Wave FE	No	No	Yes	No	No	Yes
Controls	No	No	No	Yes	Yes	Yes
Observations	3,214	3,213	3,213	2,982	2,909	2,909
R-squared	0.256	0.968	0.969	0.230	0.972	0.972

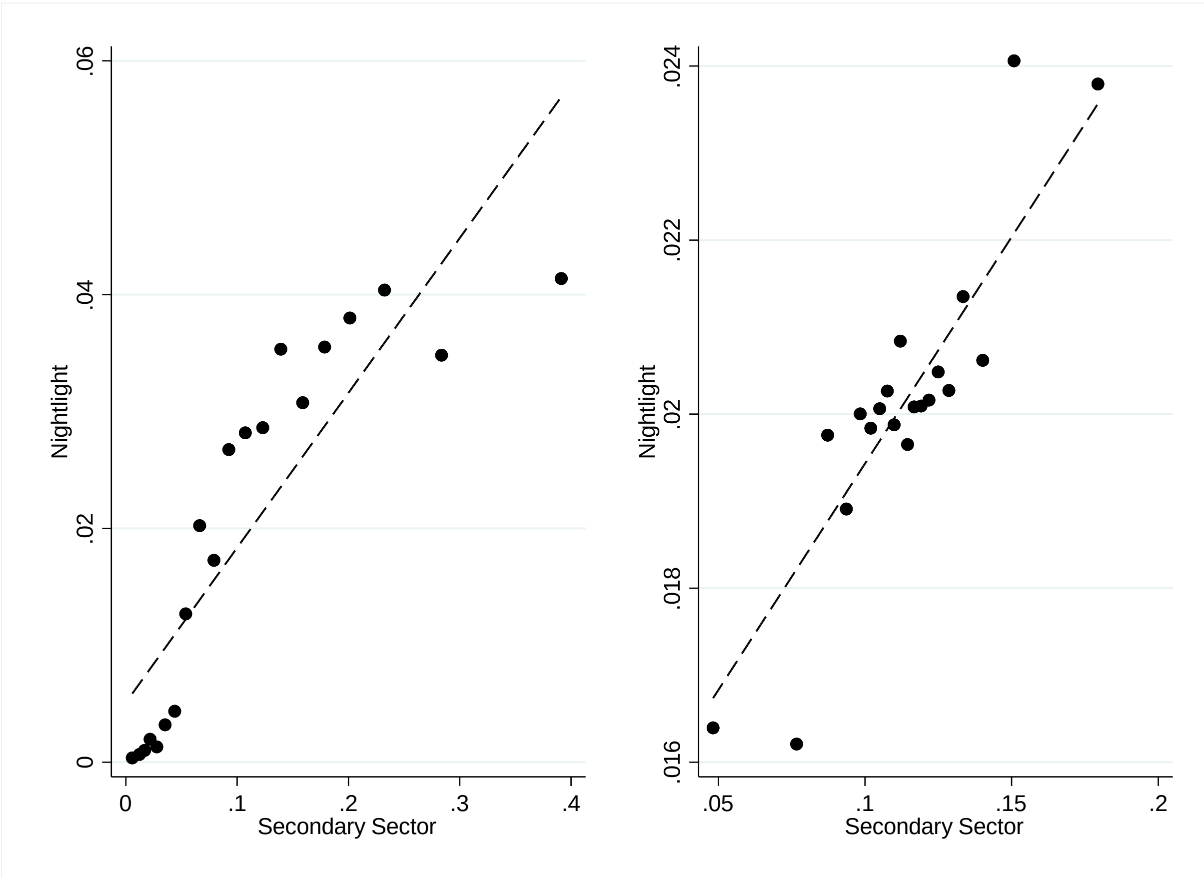
Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMIN and wave fixed effects. *** p<0.01, ** p<0.05, * p<0.1.

Figure B1 Correlation between nightlight and share of the primary sector



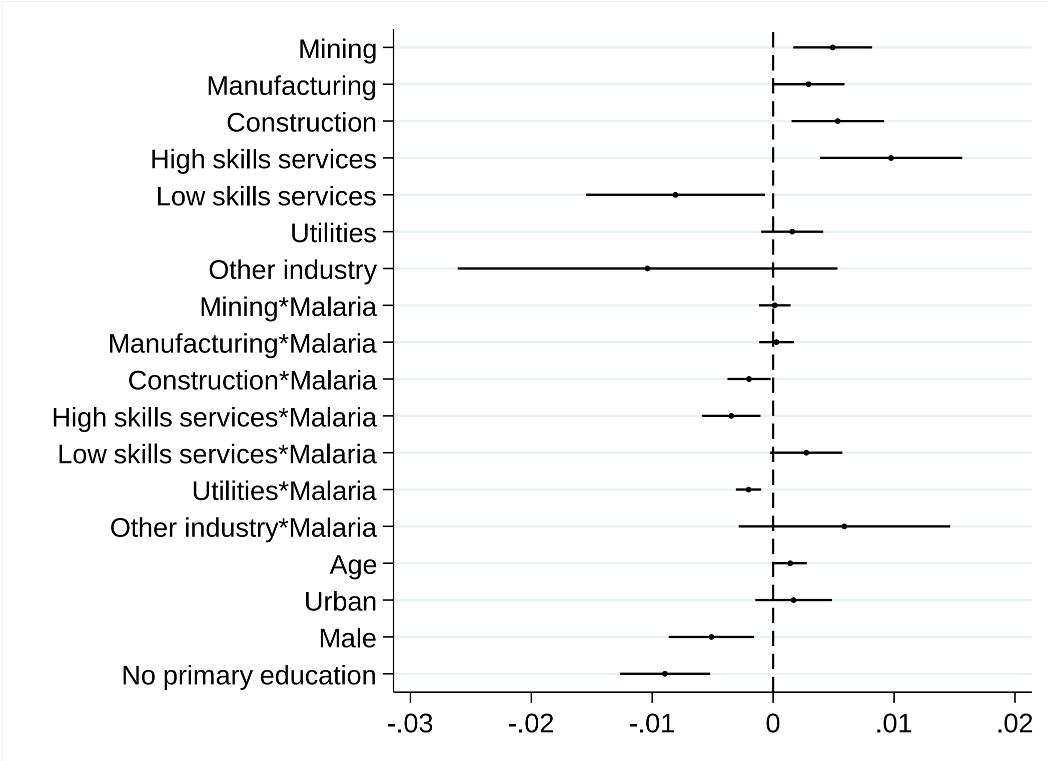
Note: binned scatterplot. The graph on the left shows a simple correlation. The graph of the right shows a correlation accounting for ADMIN and wave fixed effects.

Figure B2 Correlation between nightlight and share of the secondary sector



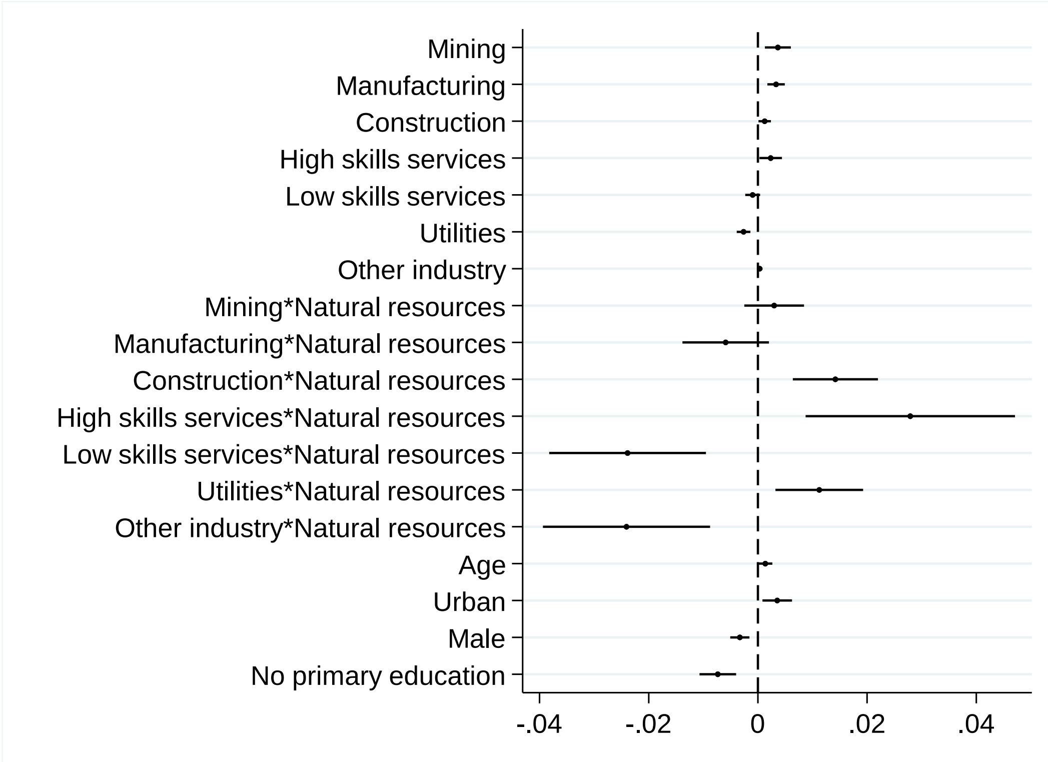
Note: binned scatterplot. The graph on the left shows a simple correlation. The graph of the right shows a correlation accounting for ADMIN and wave fixed effects.

Figure B3 High skills services and growth: The mediating effect of the incidence of malaria



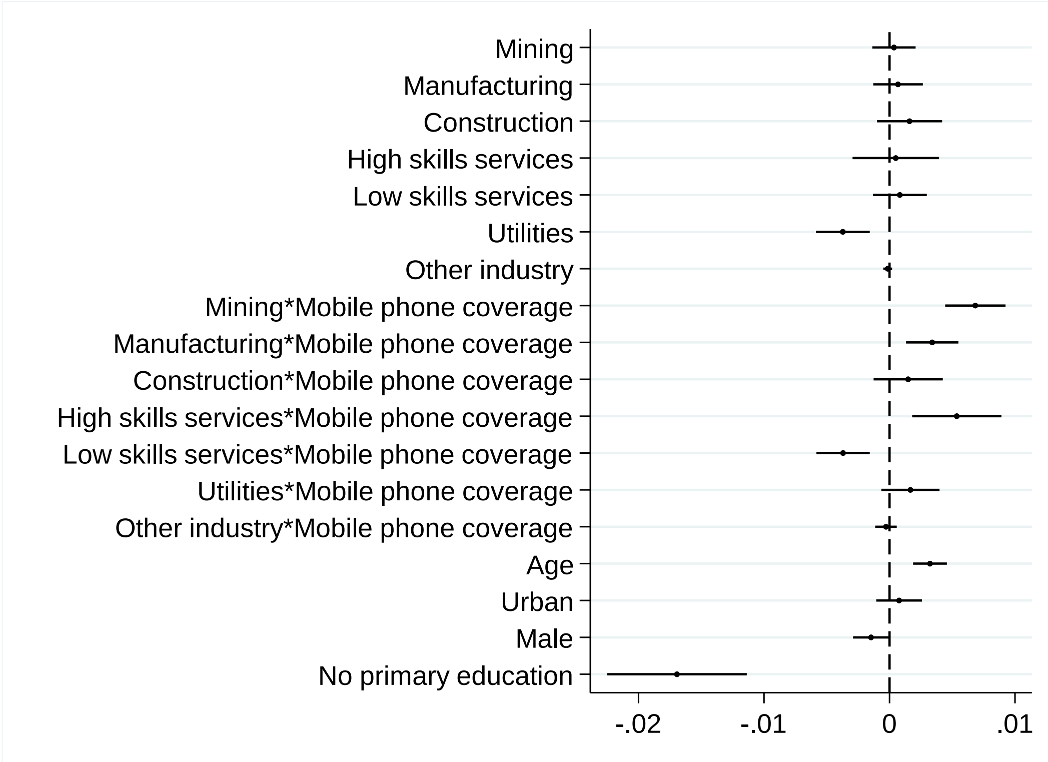
Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMIN and wave fixed effects. The number of observations is 2,829 and the R^2 is 0.970. 90% C.I.

Figure B4 High skills services and growth: The mediating effect of natural resources



Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMIN and wave fixed effects. The number of observations is 2,811 and the R^2 is 0.966. 90% C.I.

Figure B5 High skills services and growth: The mediating effect of technology

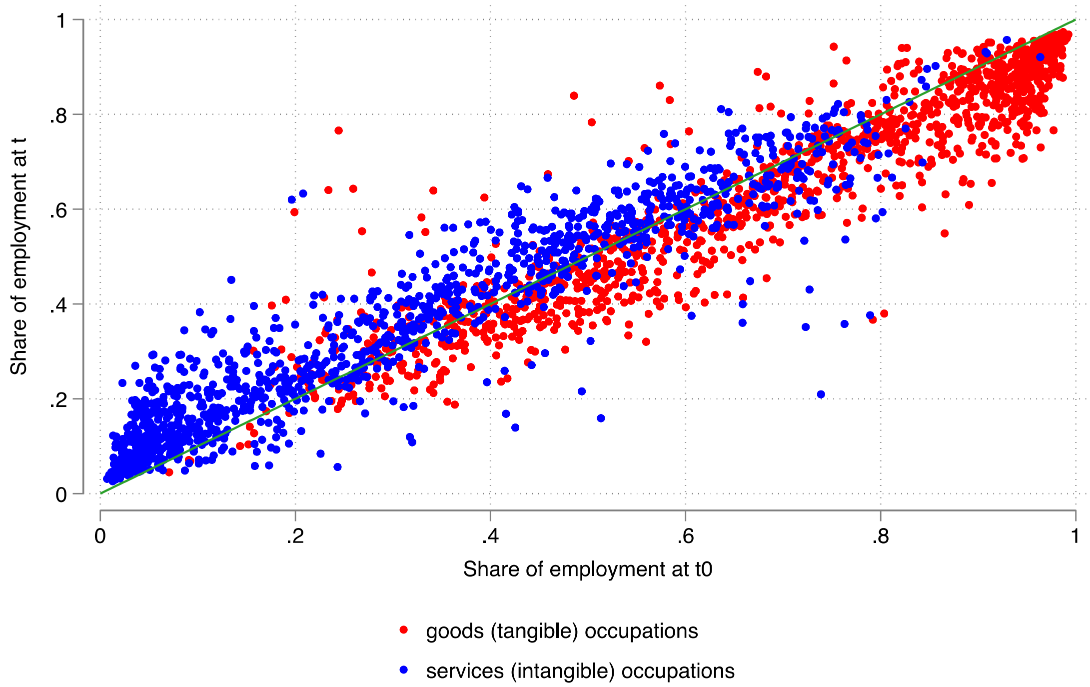


Note: OLS regression weighted by population with standard errors clustered by ADMIN units. The outcome variable is nightlight *per capita*. The model includes ADMIN and wave fixed effects. The number of observations is 1,958 and the R^2 is 0.975. 90% C.I.

Appendix C

Figure C1. Structural transformation at the sub-national level: Agriculture, manufacturing and services

Figure C2. Changes in type of occupations within the manufacturing sector



Notes: Goods and Services occupations are constructed on the basis of the classification of Duernecker and Herrendorf (2020) and relative to those employed in the manufacturing sector only.
Source: Author's elaboration on IPUMS.

Appendix D: Data

1. Sample selection and data availability

To obtain temporally and spatially explicit information on employment in Sub-Saharan African (SSA) countries, we collected data on all SSA countries available on the IPUMS International database (Minnesota Population Center, 2020). The IPUMS International database includes employment information for a representative sample of SSA countries' individuals, covering different fractions of each country's total population. Table 1 lists the countries included in our sample, together with information on the year of which data collection, the coverage fraction and the number of surveyed individuals. The country coverage of the sample is determined by the availability of data on the "INDGEN" variable in at least two census waves per country to be able to track the change in the sectoral composition of employment at the administrative unit level. This constraint led to the exclusion of Burkina Faso, Cameroon, Guinea, Kenya, Lesotho, Senegal, Sierra Leone, South Sudan, Sudan, Uganda, and Zimbabwe, as data for the "INDGEN" variable in IPUMS International was reported in only one wave. In addition, Liberia and Togo were excluded due to the extended period between the only two waves included in the IPUMS International database (1974-2008 for Liberia and 1970-2010 for Togo). This left 13 countries: Benin, Botswana, Egypt, Ghana, Malawi, Mali, Mauritius, Morocco, Mozambique, Rwanda, South Africa, Tanzania, and Zambia.

We aggregate individual-level information on sectoral employment to the lowest available administrative unit level[35] (namely, the second administrative level, except for Botswana, for which only the first administrative level is available). The administrative designations in the IPUMS International Database have numerous spatial and temporal inconsistencies, including frequent redistricting by national authorities.

[35] This operation amounts to aggregating data at the second administrative level for most countries, excluding Botswana, for which the sole administrative division available for aggregation is the first. The second-level administrative unit corresponds to Districts in most cases. Exceptions are Benin ("Communes"), Malawi ("Traditional Authorities"), Mali ("Circles"), Mauritius ("Municipal Wards/Village Council Areas"), and Morocco and Rwanda ("Provinces"). Administrative units are roughly comparable in terms of surface, apart from a few rural districts in Egypt, Mali and Morocco, corresponding to sparsely habited and/or desertic areas.

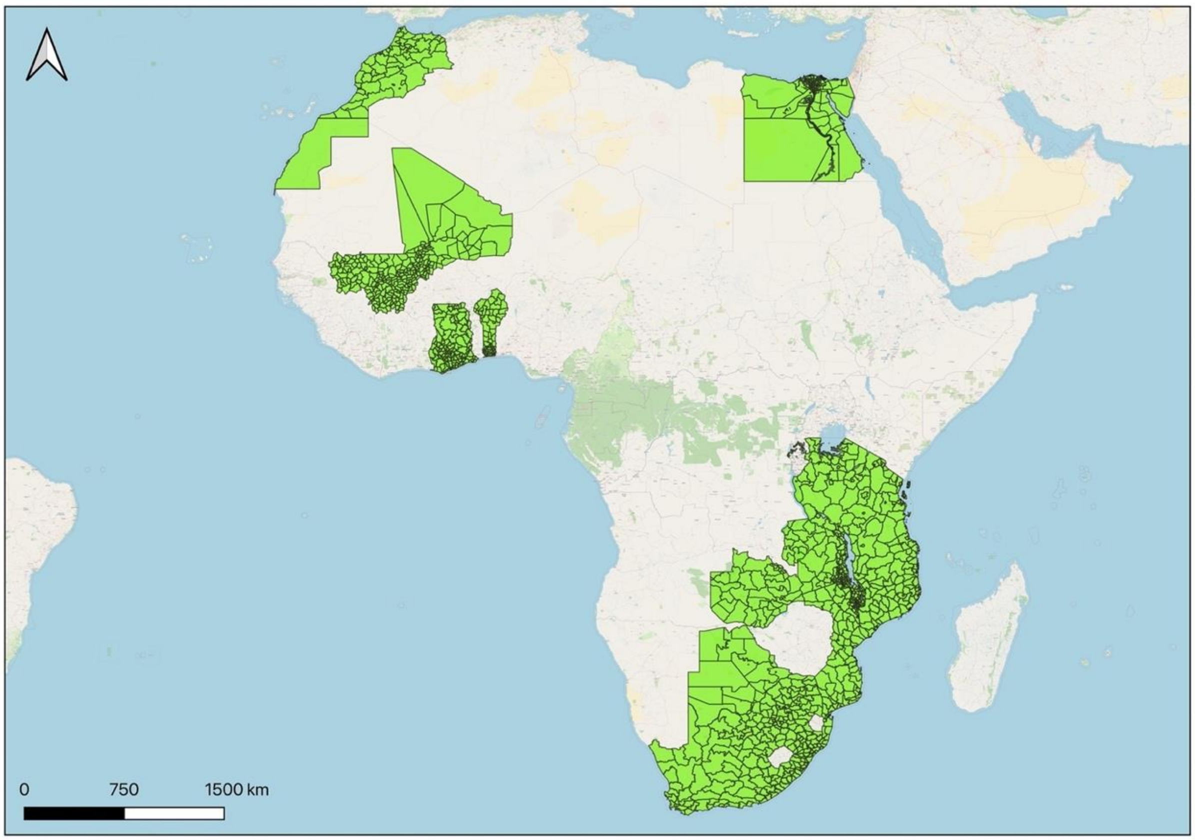
Appendix Table D1: Sample countries, waves, coverage fractions and number of surveyed individuals

Country	Year	Fraction	N_{obs}
Benin	1992	10.0	498419
Benin	2002	10.0	685467
Benin	2013	10.0	1009693
Botswana	1991	10.0	132623
Botswana	2001	10.0	168676
Botswana	2011	10.0	201752
Egypt	1986	14.1	6787077
Egypt	1996	10.0	5902243
Egypt	2006	10.0	7282434
Ghana	2000	10.0	1012062
Ghana	2010	10.0	1405681
Malawi	1998	10.0	519551
Malawi	2008	10.0	680230
Mali	1987	10.0	785384
Mali	1998	10.0	991330
Mali	2009	10.0	1451856
Mauritius	1990	10.0	106710
Mauritius	2000	10.0	119695
Mauritius	2011	10.0	123189
Morocco	1982	5.0	1012873
Morocco	1994	5.0	1294026
Morocco	2004	5.0	1482720
Mozambique	1997	10.0	1551517
Mozambique	2007	10.0	2047048
Rwanda	1991	10.0	742918
Rwanda	2002	10.0	843392
Rwanda	2012	10.0	1038369
South Africa	2001	10.0	3724723
South Africa	2007	2.0	1047657
Tanzania	2002	10.0	3732735
Tanzania	2012	10.0	4498022
Zambia	1990	10.0	787461
Zambia	2000	10.0	996117
Zambia	2010	10.0	1321973

Source: IPUMS.

In the case of Rwanda, the second-level geographic definer is only available in the last two waves (2002 and 2012), while the first level definition is only available for the first wave (1991), thereby making it impossible to harmonize consecutive surveys at the same administrative level. Given these inconsistencies, we instead obtain temporally and spatially consistent second-level administrative shapefiles from Alesina et al. (2021), to which we aggregate all variables employed in the report.[36] The Alesina et al. (2021) database does not include spatial designations for Mauritius; we therefore retrieve the IPUMS International shapefiles for Mauritius and aggregate all available data[37] to these geographic delimiters as a second-best solution. Appendix Figure D1 offers a global representation of the 1,546 unique administrative units contained in Alesina et al. (2021).

Appendix Figure D1: Administrative units



Source: Authors' elaboration on IPUMS and Alesina et al. (2021)

2. *Employment and Occupation Data*

In order to obtain information on sectoral shares defined at the second administrative level for all countries in our final sample, we first need to unpack the INDGEN variable into the 17 levels which constitute its categories: (1) Agriculture, fishing and forestry; (2) Mining and extraction; (3) Manufacturing; (4) Utilities (electricity, gas, water and waste management); (5) Construction; (6) Wholesale and retail trade; (7) Hospitality (hotels and restaurants); (8) Transportation, storage and communications; (9) Financial services

[36] Notably, in the case of Rwanda, the shapefile harmonisation procedure operated by Alesina et al. (2021) results in all IPUMS observations being clustered in a subset of 30 out of 145 Rwandan districts.

[37] Alesina et al. (2021) and Manacorda and Tesei (2020) data sources are unavailable for Mauritius.

and Insurance; (10) Public Administration and Defense; (11) Services, not specified; (12) Business services and real estate; (13) Education; (14) Health and social work; (15) Other services; (16) Private household services; (17) Other industry, not elsewhere classified. We first obtain 17 different binary variables coded as 1 if the individual's survey response belongs to the target category, and 0 if it does not[38]. From these 17 intermediate variables, we calculate each sector's share on overall employment at the second administrative level by collapsing the individual binary indicators to the ADMIN units, weighting each observation by the survey weights provided by IPUMS International.

We also produce coarser definitions of sectoral employment as an alternative to the disaggregated version; here, we lose in terms of granularity what we gain in terms of ability to describe general macro-sectoral trends. We proceed as follows: (1) Agriculture, fishing and forestry, and (2) Mining and extraction are grouped in the Primary sector; (3) Manufacturing, (5) Construction, and (17) Other industry, n.e.c. are aggregated in the Secondary sector; every other sector is considered Tertiary sector. We create individual categorical variables for each macro-sector and we repeat the procedure outlined for disaggregated sectors in order to obtain macro-sectoral shares at the second administrative level.

We proceed similarly for what concerns data on individuals' class of occupation (occupational role), contained in IPUMS International's "OCCISCO" variable. The OCCISCO variable contains detailed information on 11 categories of mansions, coded according to the major categories in the International Standard Classification of Occupations (ISCO) scheme for 1988. For someone with more than one job, the primary occupation is typically the one in which the person had spent the most time or earned the most money. The 11 categories are: (1) Legislators, senior officials and managers; (2) Professionals; (3) Technicians and associate professionals; (4) Clerks; (5) Service workers and shop and market sales; (6) Skilled agricultural and fishery workers; (7) Crafts and related trades workers; (8) Plant and machine operators and assemblers; (9) Elementary occupations; (10) Armed forces; (11) Other occupations, unspecified or not elsewhere classified. Again, we obtain 11 distinct binary variables repeating the procedure used for the sector of employment, and we then proceed to collapse these binary indicators to the desired ADMIN level using IPUMS survey weights in the procedure.

3. *Economic and geographic covariates*

By leveraging the rich information contained in the IPUMS International dataset, we are able to obtain an ample set of covariates which describe patterns of economic development in each administrative unit that

[38] IPUMS records a non-insignificant percentage of "Not in Universe" (NIU) and missing responses. In order not to distort the resulting sectoral shares, we decide to drop NIU and missing observations from the dataset. Our shares are thus calculated using the total number of individuals having answered the INDGEN survey field at the denominator.

we observe. We cover several categories designed by IPUMS, namely: (1) Group quarters (urban/rural status, residence status, and geo-identifiers for administrative units of levels 1 and 2); (2) Family (number of own children and number of children under 5 years old); (3) Demographics (age, gender, marital status); (4) Fertility (children ever born, children surviving, number of births in the past year, children surviving from births in the past year, mortality status of mother); (5) Ethnicity (religion, ethnicity, language); (6) Education (school attendance, literacy, educational attainment, years of schooling); (7) Occupation (besides the abovementioned variables, we include employment status and labour force par-ticipation), (8) Migration (migration status).

We replicate the procedure described in the previous section in order to calculate ADMIN-level shares for each variable of interest. In particular, where available, we compute:

- Shares of urban and rural residents;
- Average number of own children and of children under five years old;
- Average age;
- Shares of female and male inhabitants;
- Shares of single, married, divorced, and widowed individuals;
- Shares of individuals with zero, one, two, three or more children;
- Shares of individuals with zero, one, two, three or more surviving children;
- Shares of individuals with zero, one, two or three or more newborns in the last year;
- Shares of individuals with zero, one, two or three or more surviving newborns in the last year;
- Mother mortality shares (percentage of individuals' mothers alive or deceased);
- Shares for each major religion in the dataset (shares of individuals with atheist, Buddhist, Hindu, Jewish, Muslim, Christian or other religious beliefs);
- Indices of linguistic and ethnic fractionalization as in Alesina et al. (2003)
    - Calculated as $\Phi_j = 1 - \sum_{i=1}^{N} s_{ij}^2$ where s_{ij} is the share of ethnic or linguistic group $i = \{1, \dots, N\}$ in ADMIN unit j.
    - We calculate the share s_{ij} by dividing the instances of each unique ethnic or linguistic group recorded in each ADMIN unit in the raw dataset by the ADMIN unit's total population.
- Shares of individuals attending, not attending, having attended and having never attended school;
- Literacy and illiteracy rates;
- Shares of individuals with no formal education, or with primary, secondary or university education;
- Average number of school years completed;

- Shares of employed, unemployed and inactive individuals;
- Shares of individuals participating and not participating to the labour force;
- Shares of native and immigrant individuals:
    - the migration variable is not perfectly conformable across different country censuses. As can be inferred from Tables 2-13, an individual may be categorized as a migrant if their previous (according to different temporal criteria) residence is different to their current one. We define "immigrants" those individuals whose previous residence is in the same major but different minor administrative unit, or in a different major administrative unit, or abroad, without further distinction on the time horizon on which "previous residence" is established. Indeed, the definition may vary from the more general "previous residence, unspecified", to "previous residence, 10 years prior [to the survey]", "previous residence, 5 years prior [to the survey]", and "previous residence, 1 year prior [to the survey]". Given the large temporal gaps between each wave of the survey, it is reasonable to expect that these discrepancies in categorizing prior residence do not introduce bias in the classification of migrant and native individuals.

IPUMS International provides information on the uniformity of the educational attainment variables. In particular, for school attendance, the data is largely comparable across waves, with slight differences in the age of the respondents and hence in the composition of the universe. Small differences can also be found with respect to the inclusion of correspondence courses, adult literacy classes, and non-traditional studies. With respect to literacy rates, "samples provide differing criteria with respect to the level of ability that should constitute literacy", hence the standards can vary across samples-waves, based on the abovementioned criteria. For educational attainment, in order to ensure comparability across samples and waves, IPUMS International applies the UN standard 6-3-3 classification (primary, lower secondary, higher secondary schooling). Most countries' school system diverge from this standard, but where applicable, data are recoded to adapt to the 6-3-3 system. For countries reporting only degree rather than grade, the answer is classified to the corresponding category. Finally, concerning years of schooling, this variable is top-coded differently across samples, with a maximum of 18+ years. In our sample, Botswana 1991 is top-coded at 17+, Botswana 2001, all Mauritius and all South Africa at 13+, while the remaining country-wave pairs are all top-coded at the standard 18+. The data is hence not perfectly comparable across country-wave pairs unless the diverging observations are dropped from the panel, since the top coding discrepancies will alter downwards the average number of years of schooling for ADMIN units located in countries with lower top values for this variable. To synthesize the information above, Tables 2-13 present the full list of IPUMS covariates together with their availability on a per-country, per-wave basis.

Appendix Table D2: Data availability, Benin

Variable	**Label**	Benin 1992	Benin 2002	Benin 2013
RESIDENT	Residence status	x	x	x
URBAN	Urban/Rural status	x	x	x
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	o	x	x
BIRTHSURV	Children surviving births in the past year	o	x	x
MORTMOT	Mortality status of mother	x	x	x
RELIGION	Religion	x	x	x
ETHNICBJ	Ethnicity, Benin	x	x	x
LANGBJ	Primary language spoken, Benin	o	o	x
SCHOOL	School attendance	o	o	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	x	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATEP	Migration status, previous residence	x	x	x

Appendix Table D3: Data availability, Botswana

Variable	Label	Botswana 1991	Botswana 2001	Botswana 2011
RESIDENT	Residence status	x	x	x
URBAN	Urban/Rural status	x	o	o
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	o	o	o
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	o	x	x
BIRTHSURV	Children surviving births in the past year	o	x	x
MORTMOT	Mortality status of mother	o	x	x
RELIGION	Religion	o	x	x
LANGBW	Primary language spoken, Botswana	o	o	x
SCHOOL	School attendance	o	o	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	x	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATE1	Migration status, 1 year	x	x	x

Appendix Table D4: Data availability, Egypt

Variable	Label	Egypt 1986	Egypt 1996	Egypt 2006
RESIDENT	Residence status	o	o	o
URBAN	Urban/Rural status	x	x	x
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	o	o	o
CHSURV	Surviving children	o	o	o
BIRTHSLYR	Number of births in the past year	o	o	o
BIRTHSURV	Children surviving births in the past year	o	o	o
MORTMOT	Mortality status of mother	x	x	x
RELIGION	Religion	x	x	x
SCHOOL	School attendance	o	o	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	o	o	o
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATEP	Migration status, previous residence	o	x	x

Appendix Table D5: Data availability, Ghana

Variable	**Label**	Ghana 2000	Ghana 2010
RESIDENT	Residence status	o	o
URBAN	Urban/Rural status	x	x
GEOLEV1	1st subnational admin. level	x	x
GEOLEV2	2nd subnational admin. level	x	x
NCHILD	Number of children	x	x
NCHLT5	Number of children under 5	x	x
AGE	Age	x	x
SEX	Gender	x	x
MARST	Marital Status	x	x
CHBORN	Children ever born	x	x
CHSURV	Surviving children	x	x
BIRTHSLYR	Number of births in the past year	x	x
BIRTHSURV	Children surviving births in the past year	o	o
MORTMOT	Mortality status of mother	x	x
RELIGION	Religion	x	x
ETHNICGH	Ethnicity, Ghana	x	x
SCHOOL	School attendance	x	x
LIT	Literacy	x	x
EDATTAIN	Educational attainment, international recode	x	x
YRSCHOOL	Years of schooling	x	x
EMPSTAT	Employment Status	x	x
LABFORCE	Labour force participation	x	x
OCCISCO	Occupation, ISCO general	x	x
INDGEN	Industry, general recode	x	x
MIGRATE5	Migration status, 5 years	x	o

Appendix Table D6: Data availability, Malawi

Variable	Label	Malawi 1998	Malawi 2008
RESIDENT	Residence status	o	x
URBAN	Urban/Rural status	x	x
GEOLEV1	1st subnational admin. level	x	x
GEOLEV2	2nd subnational admin. level	x	x
NCHILD	Number of children	x	x
NCHLT5	Number of children under 5	x	x
AGE	Age	x	x
SEX	Gender	x	x
MARST	Marital Status	x	x
CHBORN	Children ever born	x	x
CHSURV	Surviving children	x	x
BIRTHSLYR	Number of births in the past year	x	x
BIRTHSURV	Children surviving births in the past year	x	x
MORTMOT	Mortality status of mother	x	x
RELIGION	Religion	x	x
ETHNICMW	Ethnicity, Malawi	o	x
SCHOOL	School attendance	x	x
LIT	Literacy	x	x
EDATTAIN	Educational attainment, international recode	x	x
YRSCHOOL	Years of schooling	x	x
EMPSTAT	Employment Status	x	x
LABFORCE	Labour force participation	x	x
OCCISCO	Occupation, ISCO general	x	x
INDGEN	Industry, general recode	x	x
MIGRATEP	Migration status, previous residence	o	x

Appendix Table D7: Data availability, Mali

Variable	**Label**	Mali 1987	Mali 1998	Mali 2009
RESIDENT	Residence status	x	x	o
URBAN	Urban/Rural status	o	x	x
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	o	o	x
BIRTHSURV	Children surviving births in the past year	o	o	o
MORTMOT	Mortality status of mother	x	x	x
RELIGION	Religion	o	o	x
LANGML	Primary language spoken, Mali	x	x	x
SCHOOL	School attendance	o	x	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	o	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATEP	Migration status, previous residence	o	x	x

Appendix Table D8: Data availability, Mauritius

Variable	Label	Mauritius 1990	Mauritius 2000	Mauritius 2011
RESIDENT	Residence status	x	x	x
URBAN	Urban/Rural status	x	x	x
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	o	o	o
BIRTHSURV	Children surviving births in the past year	o	o	o
MORTMOT	Mortality status of mother	o	o	o
RELIGION	Religion	x	x	x
LANGMU	Primary language spoken, Mauritius	x	x	x
SCHOOL	School attendance	o	o	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	x	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATE5	Migration status, 5 years	x	x	x

Appendix Table D9: Data availability, Morocco

Variable	Label	Morocco 1982	Morocco 1994	Morocco 2004
RESIDENT	Residence status	o	o	o
URBAN	Urban/Rural status	o	o	o
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	x	x	x
BIRTHSURV	Children surviving births in the past year	x	x	x
MORTMOT	Mortality status of mother	o	o	o
RELIGION	Religion	o	o	o
LANGMA	Primary language spoken, Morocco	o	x	x
SCHOOL	School attendance	o	o	o
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	x	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATE5	Migration status, 5 years	o	o	x

Appendix Table D10: Data availability, Rwanda

Variable	**Label**	Rwanda 1991	Rwanda 2002	Rwanda 2012
RESIDENT	Residence status	x	x	x
URBAN	Urban/Rural status	o	x	x
GEOLEV1	1st subnational admin. level	x	o	o
GEOLEV2	2nd subnational admin. level	o	o	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	x	x	x
BIRTHSURV	Children surviving births in the past year	x	x	x
MORTMOT	Mortality status of mother	x	x	x
RELIGION	Religion	x	x	x
LANGRW	Primary language spoken, Rwanda	x	x	o
SCHOOL	School attendance	o	x	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	o	x	x
YRSCHOOL	Years of schooling	o	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	o	x	x
INDGEN	Industry, general recode	o	x	x
MIGRATEP	Migration status, previous residence	x	x	o

Appendix Table D11: Data availability, South Africa

Variable	**Label**	South Africa 2001	South Africa 2007
RESIDENT	Residence status	o	x
URBAN	Urban/Rural status	x	x
GEOLEV1	1st subnational admin. level	x	x
GEOLEV2	2nd subnational admin. level	x	x
NCHILD	Number of children	x	x
NCHLT5	Number of children under 5	x	x
AGE	Age	x	x
SEX	Gender	x	x
MARST	Marital Status	x	x
CHBORN	Children ever born	x	x
CHSURV	Surviving children	x	x
BIRTHSLYR	Number of births in the past year	x	x
BIRTHSURV	Children surviving births in the past year	x	x
MORTMOT	Mortality status of mother	x	x
LANGZA1	Language spoken at home, SAF	x	o
RELIGION	Religion	x	o
SCHOOL	School attendance	x	o
LIT	Literacy	x	x
EDATTAIN	Educational attainment, international recode	x	x
EMPSTAT	Employment Status	x	x
LABFORCE	Labour force participation	x	x
OCCISCO	Occupation, ISCO general	x	x
INDGEN	Industry, general recode	x	x
MIGRATE5	Migration status, 5 years	x	o
MIGRATEP	Migration status, previous residence	o	x

Appendix Table D12: Data availability, Tanzania

Variable	**Label**	Tanzania 2002	Tanzania 2012
URBAN	Urban/Rural status	x	x
GEOLEV1	1st subnational admin. level	x	x
GEOLEV2	2nd subnational admin. level	x	x
NCHILD	Number of children	x	x
NCHLT5	Number of children under 5	x	x
AGE	Age	x	x
SEX	Gender	x	x
MARST	Marital Status	x	x
CHBORN	Children ever born	x	x
CHSURV	Surviving children	x	x
BIRTHSLYR	Number of births in the past year	x	x
BIRTHSURV	Children surviving births in the past year	o	x
MORTMOT	Mortality status of mother	x	x
SCHOOL	School attendance	x	x
LIT	Literacy	x	x
EDATTAIN	Educational attainment, international recode	x	x
EMPSTAT	Employment Status	x	x
LABFORCE	Labour force participation	x	x
OCCISCO	Occupation, ISCO general	x	x
INDGEN	Industry, general recode	x	x
MIGRATE1	Migration status, 1 year	x	o

Appendix Table D13: Data availability, Zambia

Variable	Label	Zambia 1990	Zambia 2000	Zambia 2010
RESIDENT	Residence status	x	x	x
URBAN	Urban/Rural status	x	x	o
GEOLEV1	1st subnational admin. level	x	x	x
GEOLEV2	2nd subnational admin. level	x	x	x
NCHILD	Number of children	x	x	x
NCHLT5	Number of children under 5	x	x	x
AGE	Age	x	x	x
SEX	Gender	x	x	x
MARST	Marital Status	x	x	x
CHBORN	Children ever born	x	x	x
CHSURV	Surviving children	x	x	x
BIRTHSLYR	Number of births in the past year	x	x	x
BIRTHSURV	Children surviving births in the past year	x	x	x
MORTMOT	Mortality status of mother	o	o	x
RELIGION	Religion	o	x	x
LANGZM1	Primary language spoken, Zambia	x	x	x
SCHOOL	School attendance	o	x	x
LIT	Literacy	x	x	x
EDATTAIN	Educational attainment, international recode	x	x	x
YRSCHOOL	Years of schooling	x	x	x
EMPSTAT	Employment Status	x	x	x
LABFORCE	Labour force participation	x	x	x
OCCISCO	Occupation, ISCO general	x	x	x
INDGEN	Industry, general recode	x	x	x
MIGRATE1	Migration status, 1 yesr	x	x	x

We also aggregate two additional sets of economic and geographic covariates, derived from Manacorda and Tesei, (2020) and Alesina et al., (2021). In particular, the former dataset is constituted of 50 x 50 Km grid cells spanning the whole African continent. We retrieve grid cell-level information on:

- Mobile phone coverage – 2G + 3G and 3G only;
- Internet penetration;
- Number of protests (from the GDELT Database);
- Number of protests (from the ACLED Database);
- GDP growth;
- Number of lightning strikes;
- Rainfall and temperature;
- Distance to the closest border;
- Percentage of a grid cell covered by an oil field;
- Number of cities;

- Percentage of a grid cell covered by forests and mountains;
- Presence of mines and diamond mines;
- Infant mortality rates;
- A dummy "Capital City" indicator, or a distance to the capital city;
- Length of the electric and primary road networks.

Given the grid cell format, we first need to preprocess the data, which is stored as a STATA dataset, in order to resample its information to our desired administrative level. We thus rasterize the grid cells, thereby transforming each 50x50Km square in a stack of raster pixels containing the values of the variables of interest; we then calculate zonal statistics at the ADMIN level for each layer of the raster stack, therefore obtaining distinct variables for each layer, characterized as the mean or the sum (depending on the underlying variable: e.g. average internet penetration, or total lightning strikes) of pixel values contained in an administrative unit. With 50x50Km pixels and ADMIN units varying in size, there are pixels bound to span over the borders of ADMIN units; through the `exactextractr` package in R version 4.0.3, we are able to aggregate pixels to an ADMIN unit using only the exact fraction of a pixel covered by an ADMIN unit polygon shapefile. All variables described above are only available for a subset of years spanned by our master dataset, 1998-2012.

The Alesina et al. (2021) database is instead available at the exact administrative unit level at which we have assembled our master dataset[39], and is therefore immediately conformable to it. We retrieve information on:

- Mean terrain ruggedness;
- Mean stability of malaria transmission;
- Mean agricultural suitability;
- Dummy variables for diamond mines and oil fields contained in the ADMIN unit;
- Distance[40] to capital;
- Distance to border;

[39] Which indeed uses the Alesina et al. (2021) georeferenced ADMIN units, as described above.
[40] All distance metrics are calculated from the ADMIN centroid.

- Distance to nearest colonial railroad;
- Distance to nearest catholic mission;
- IPUMS urban shares, birth cohorts 1960 and earlier;
- IPUMS shares of agricultural, manufacturing and services workers for birth cohorts 1960 and earlier.

4. *Geolocalized Indicators of Economic Development*

We leverage on a burgeoning literature in the field of Economic Geography, dating back to Henderson et al., (2012), in constructing disaggregated indicators of economic performance based on remotely sensed data, primarily night-time lights (NTLs). NTLs have been shown to be a useful proxy for economic growth and size, especially for developing countries[41]. One frequent caveat against the use of NTLs in longitudinal economic analyses has been retraced in time-series inconsistencies in NTLs trends due to satellite intercalibration issues (Gibson et al., 2021)[42]; however, a recently published dataset by Li et al. (2020) has attempted to mitigate these issues and is therefore viable for timeseries and panel analyses. In particular, the authors employ two generations of satellite instruments, the pioneering DMSP-OLS (which collected data between 1992 and 2013) and its successor, VIIRS (still in orbit), in order to construct a longer data series spanning 1992-2018. Their dataset improves on pre-existing NTL data products on two fronts: first, they intercalibrate the internal 1992-2013 DMSP-OLS time series at the global scale; second, they successfully attempt to reduce discrepancies between the DMSP and VIIRS data series by converting the VIIRS radiance information to DMSP-like NTL data, thereby providing the scientific community with a dataset exhibiting temporal consistency and extending the span of historical DMSP NTL data to 2018, with the trade-off of losing some granularity.[43]

We obtain the full (1992-2018) dataset from Li et al. (2020) and compute zonal statistics at the second administrative level for each of the ADMIN-waves pairs included in our sample, obtaining two main

[41] In developed countries, the top-level saturation of the NTLs data due to the widespread presence of night-time illumination masks cross-sectional and temporal heterogeneity, while this effect is less pronounced for developing countries presenting large swaths of unlit territory.

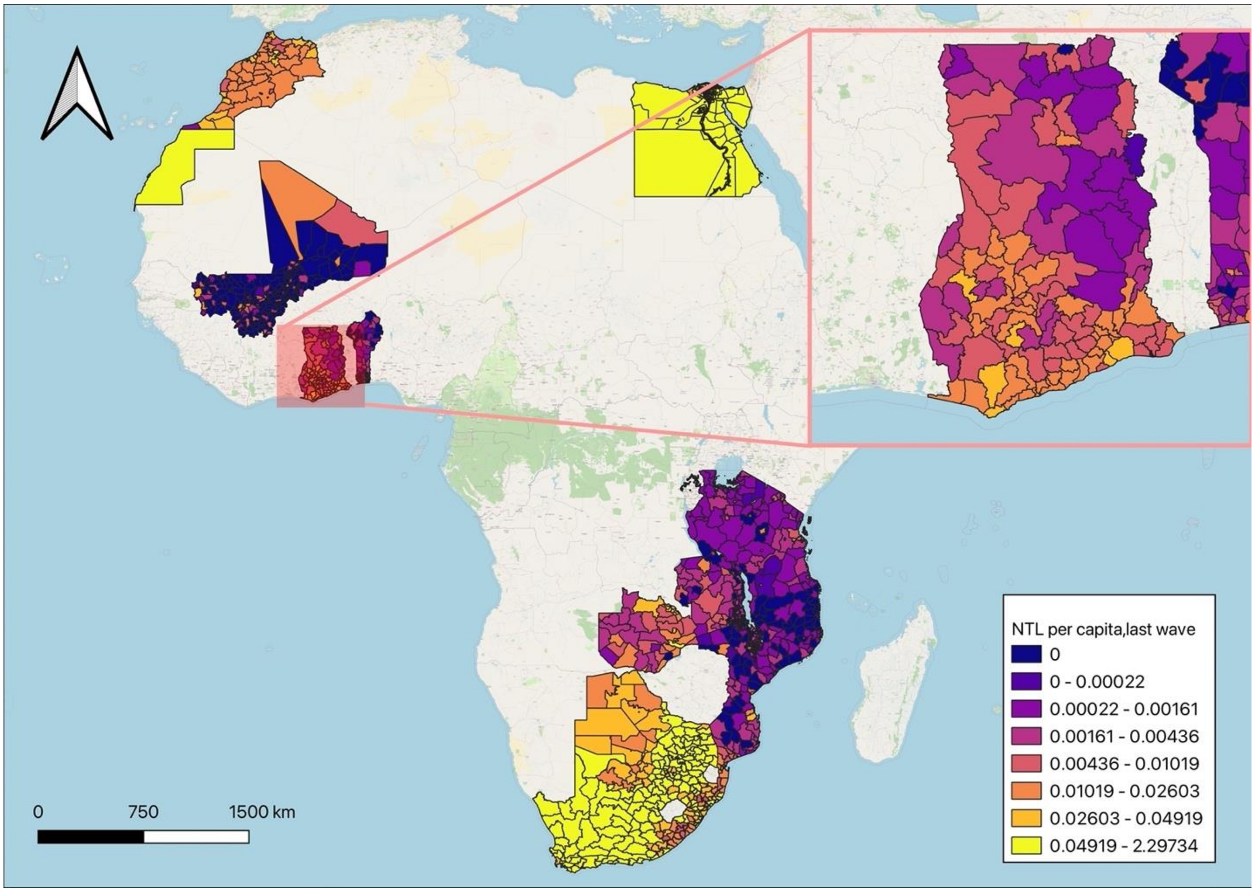
[42] Indeed, DMSP-OLS, the NOAA program which has produced the frequently used 1992-2013 NTLs time series, has launched six individual satellites whose orbits have progressively lowered over time and whose sensors are not perfectly identical, resulting in inconsistencies when employing their measurements across years.

[43] Due to the conversion between VIIRS and DMSP NTL, the target resolution is that of DMSP-OLS, coarser than VIIRS. Some caveats remain, mostly due to uncertainties in NTL pixels with values lower than 10.

variables: mean NTLs and the Sum of Lights (SoL), which we then divide by each ADMIN's total population to obtain an indicator of nightlights per capita. Appendix Figure D2 provides some descriptive evidence on the sub-national variation of this variable, showing enough variation both between- and within-countries.[44]

[44] The relatively higher values of Western Sahara is due to the fact that it reports—due to its higher area—relatively high values of NTL (which is summed across pixels) and a small population.

Appendix Figure D2. Nightlights per capita, last wave

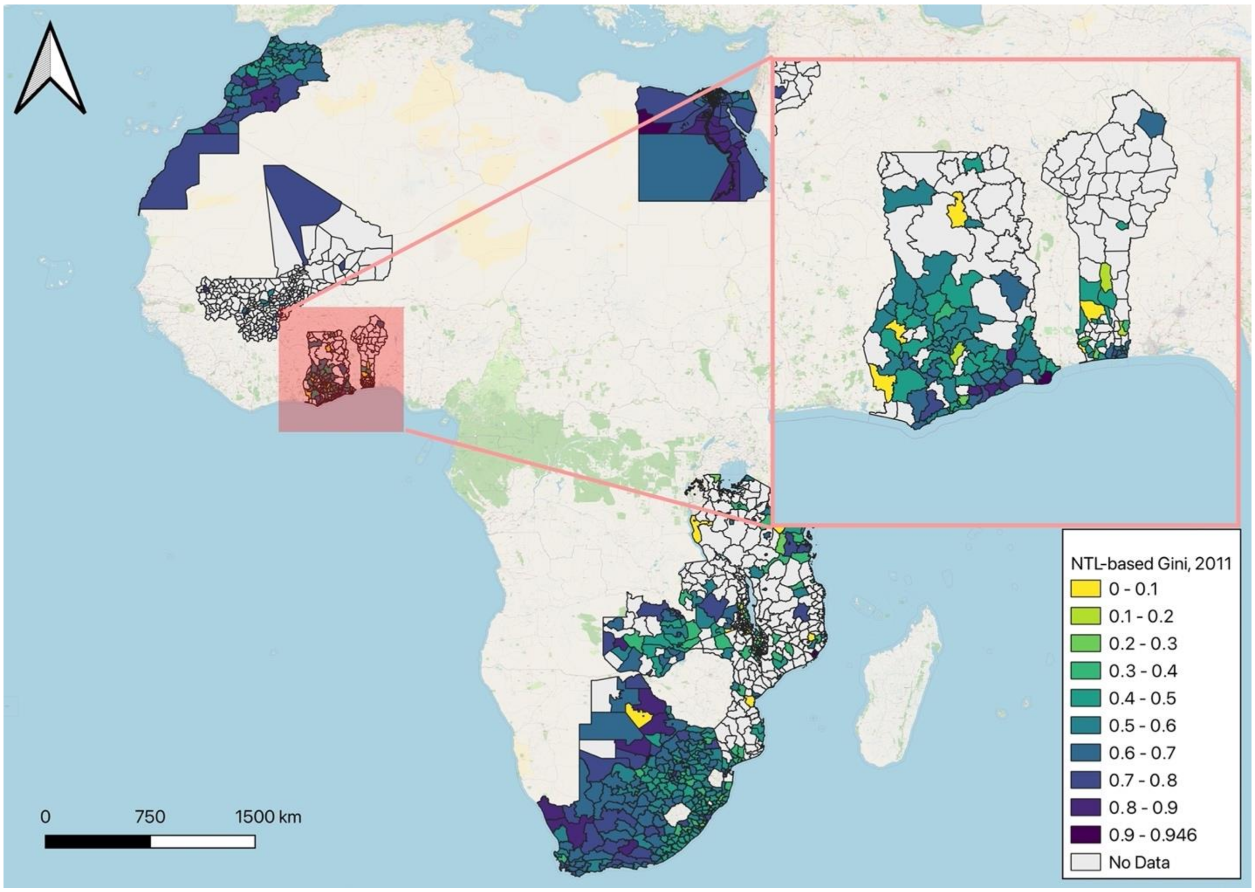


Source: Author's elaboration on Li et al. (2020)

5. Remotely-sensed Spatial Inequality Measurements

Our fourth and final set of variables is represented by measurements of spatial inequality, which represent one of the latest developments in the use of remotely sensed socioeconomic data. In particular, we exploit the Li et al. (2020) harmonized NTLs dataset described in section D.2. in conjunction with the recently released LandScan dataset (Oak Ridge National Laboratory, 2019), a 1Km^2 pixel-level raster population database covering 2000-2018, and the Gridded Population of the World v3 (Center for International Earth Science Information Network-CIESIN-Columbia University) dataset, a 4.6Km^2 pixel-level raster population database for the years 1990 and 1995. We calculate 1Km^2 pixel level NTLs sums for 1992-2018 from the Li et al. (2020) dataset, and population counts for 1990, 1995 and 2000-2018 by leveraging the combination of LandScan and GPW v3; we then calculate population counts for 1992-1994 and 1996-1999 by linearly interpolating population values from 1990 to 1995 and from 1995 to 2000. We thus obtain a pixel-level population and nightlights panel dataset comprising over 10 million observations over the countries in our sample. In order to obtain remotely-sensed measurements of inequality, we employ the procedure described in Elvidge et al. (2012) and further refined by Mirza et al. (2021) as follows: we first calculate average lights per person (LPP)[45], by dividing pixel-level NTLs by pixel-level population counts; we then calculate inequality in the distribution of LPP for each administrative units by calculating Gini coefficients for the range of pixels contained in each administrative unit in our sample. Appendix Figure D3 provides evidence on the distribution of the inequality measure at the administrative level and for the year corresponding to the last census wave available.

Appendix Figure D3. Spatial Gini indicator based on Mirza et al., (2021)



Source: Author's elaboration on Li et al. (2020), LandScan and GPW v.3.

[45] As in Mirza et al., (2021) we first censor both pixel-level NTLs and population in order to exclude zeroes, an operation which prevents placing excessive weight on uninhabited regions or places without detectable NTLs.

References

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

Li, X., Zhou, Y., Zhao, M. and Zhao, X. (2020). A harmonized global nighttime light dataset 1992–2018. *Scientific Data*, *7*(1), 168.

Manacorda, M. and Tesei, A. (2020). Liberation Technology: Mobile Phones and Political Mobilization in Africa. *Econometrica*, *88*(2), 533–567.

Minnesota Population Center. (2019). *Integrated Public Use Microdata Series, International: Version 7.2 [dataset]*. IPUMS International.

Mirza, M. U., Xu, C., van Bavel, B., van Nes, E. and Scheffer, M. (2021). Global inequality remotely sensed. *Proceedings of the National Academy of Sciences*, 118(18), e1919913118.