Spatial Diffusion of Services

Peter Wankuru Chacha, Luke Heath Milsom,

Benard Kipyegon Kirui, Verena Wiedemann*

Preliminary please do not cite or circulate

27$^{\text{th}}$ February, 2025

Abstract

Growing evidence suggests that economic transformation in many low- and middle-income countries is shifting toward services rather than industry. Historically, this growth has been concentrated in urban centers, risking the exclusion of the much larger rural and small-town populations. Using comprehensive survey and administrative worker- and firm-level data, we document a rapid rise in formal consumer service employment outside Kenya's main cities between 2016 and 2022. These new jobs—primarily managerial and professional—reflect the expansion of productive formal firms that are displacing less efficient ones. Employment growth was particularly pronounced in areas with expanding local markets and earlier access to high-speed internet. A simple stylized model suggests that rising service-sector productivity, increasing local demand, and non-homothetic preferences all contribute to these patterns.

*The authors would like to thank Lixi Xu and Thuy (Peter) Le for outstanding research assistance. We gratefully acknowledge financial support from the Structural Transformation and Economic Growth (STEG) programmes, a joint initiative by the Centre for Economic Policy Research (CEPR) and the UK Foreign, Commonwealth & Development Office (FCDO).

1 Introduction

It is becoming increasingly unlikely that industry- and export-led growth will be the sole pillars of poverty alleviation in low- and middle-income economies in the coming decades [Goldberg and Reed, 2023, Diao et al., 2024]. Services, and in particular consumer services, are a potential alternative source of employment growth [Hsieh and Rossi-Hansberg, 2023, Fan et al., 2023, Rodrik and Stiglitz, 2024]. However, a significant concern is that the growth of the service sector is often biased towards urban areas, both in terms of welfare gains and employment opportunities [Chen et al., 2023, Fan et al., 2023, Eckert et al., 2022]. In this paper, we ask how localities, especially those outside cities, generate productive firms and employment in Kenya, a lower middle income economy, in an era of service-led job growth.

We start by asking whether service-led job growth has already been a reality in Kenya in the past decade. Further, we provide evidence on firm and job characteristics in the services economy. Lastly, we study the relevance of potential mechanisms that drive the observed patterns. To do this we collate a comprehensive dataset covering Kenya from 2016 to 2022. We combine detailed administrative data on the universe of formal firms and workers from tax records, household and labour survey data covering the informal economy, time series data on road construction and broadband rollout, and satellite data on urbanization and economic activity.

We first document two key stylised facts. First, in the aggregate we find that formal employment has rapidly shifted from industry to services. Since 2016 this transformation has happened at a rate of over one percentage point a year. Formal employment fell from 33 percent to 27 percent of total formal employment, and employment in services rose from 55 percent to 60 percent. These formal trends are matched when also looking at the informal economy, from 2014 to 2022 the overall proportion of the work force engaged in managerial and professional work increased by 7 percentage points (from 15 to 22 percent), at the expense of non-specialist skill and agricultural workers.

However, these aggregate trends hide substantial geographic variation. Our second stylised fact is that formal employment growth outside the main cities has been almost entirely in consumer facing services. Between 2016 and 2023 the consumer services formal employment share outside the main metropolitan areas has almost doubled. Almost all new formal jobs in these locations have been in consumer services over this time. This growth in consumer services outside main cities is stark, unprecedented, and suggests the consumer service sector could be a route to growth and development for these areas.

We next turn to characterizing these formal jobs and the firms which are creating them.

We find that formal jobs are higher paid, more professional and managerial, and in larger firms as compared to informal jobs. Next we turn to firms and find that firms growing formal employment in consumer services are characterised by sourcing their final demand disproportionately from consumers and growing their market share. That is, we find that expanding formal firms crowd out those who are not employing more formal workers, implying a shift in the firm-size and firm-productivity distributions.

Next shed some light on the underlying economic mechanisms that can explain the stark patterns uncovered. To formalise the main channels and set ideas we first set-up a very simple stylised model. Local jobs in formal consumer services will increase there are greater incentives to be employed (wages rise), there are greater incentives to employ more people (local demand rises), or the wages you gain can buy more (prices fall). These factors in turn will be driven by local consumer services productivity, local demand, and local income crossed with possible non-homoetheticities. These three economic forces: rising productivity, rising local demand, and non-homoethic preferences are the key mechanisms which could be behind the observed patterns. Finally, due to the local nature of demand and supply of consumer services forutious feeback loops could act as a significant multiplying factor.

Armed with this predictions we next show some indicative empirical evidence for each of the main channels, and rule out some potential alternatives. First, using variation in urbanisation from build-up-area satellite data and variation in remoteness from digitised road networks we show that areas with a higher local demand (more urbanised and more remote) see greater increases in consumer service employment. Second, using internet accessibility as a proxy for consumer service productivity and leveraging plausibly exogenous variation due to the roll out of the fiber and a minimum spanning tree type argument, we show that areas where access to the internet is more helpful saw larger increases in consumer service employment. Finally, leveraging data from a detailed consumer expenditure survey we find strong evidence for non-homotheticities in consumption. We also show empirical evidence for fortuitous feedback loops and rule out the alternative stories of migration and changes in the incentives to formalise.

We contribute to two strands of literature studying (i) the determinants of the contemporary service sector transformation in low- and middle-income countries [Ngai and Pissarides, 2007, Rodrik, 2013, Gollin et al., 2016, Nayyar et al., 2021, Ngai et al., 2022, Fan et al., 2023, Schwartzman, 2025, Amodio et al., 2024], (ii) urbanisation and formalisation dynamics [Harris and Todaro, 1970, El Badaoui et al., 2014, Imbert and Ulyssea, 2023], and [Desmet and Rossi-Hansberg, 2014, Eckert and Peters, 2022, Fan et al., 2023] While we are unable to follow the common approach in the first strand of literature to study long-term transformation over the several decades, we will add evidence from a contemporary transformation in a lower middle income economy and empirical documentation of micro-level formal sector dynamics beyond aggregate employment shares. Importantly, we utilize detailed data on both firms and workers to understand better the role firm dynamics play in a specific aspect of structural transformation [Donovan and Schoellman, 2023], the service sector diffusion in our case.

Our setting highlights the relevance of economic growth and productivity shifts for formalisation relative to cases where increased pressure from authorities to formalize jobs can have an adverse impact on the number of formal jobs [de la Parra and Bujanda, 2024].[1] We find a decline in formal manufacturing jobs, accelerated by the Covid-19 pandemic, but not overall decline in sales or the number of manufacturing firms. This is consistent with the trends documented by Diao et al. [2024] in Kenya's neighbouring countries Tanzania and Ethiopia where large firms expand but without increasing employment. The latter jobs get absorbed by smaller firms without much growth in productivity [Diao et al., 2024].

The remainder of this paper proceeds as follows. Section 2 presents the data used in this analysis, section 3 presents stylized facts, section 4 presents some indicative evidence on the main channels, section 5 presents a quantitative model, and finally section 6 concludes.

2 Data

2.1 Administrative data covering the formal economy

We rely on several administrative data sets collected by the Kenya Revenue Authority, which can be linked using anonymized firm identifiers. The two key data sets are monthly transaction-level records on sales and purchases from value-added tax (VAT) records and employer-employee data from pay-as-you-earn (PAYE) returns. Not all firms that file VAT have formal employees on record; many only employ casual workers and family members who do not appear in the PAYE data. Similarly some entities that report PAYE are exempt from filing VAT. The over 76,000 VAT-paying private sector firms accounted for 64% of all formal private sector employment in 2022[2], while their value added accounts for 36% of Kenya's

[1] de la Parra and Bujanda [2024] also show that having a formal job acts as a signal of worker quality to future employers.

[2] Using broad sector classifications reported by firms, we find that VAT-paying firms account for 65% of employment in agriculture, 66% in industries, and 63% in services.

GDP [Wiedemann et al., 2024].

A key advantage of linking VAT data with employer-employee data is that we are able to directly measure the extent to which firms cater to consumers versus other firms.[3] That is, rather than having to rely on self-reported sectors we can use transaction-level information to classify firms as consumer services if they do indeed mainly sell to consumers, and businesses services if they sell mainly to businesses.

2.2 Additional data on labour markets and consumption patterns

We further rely on various survey data (see Table A1 in the appendix) covering consumption and employment patterns in Kenya since 2002 to complement our understanding of the sectors' transformation. Two types of survey data are used. First, surveys with national aggregate employment statistics: the annual Economic Survey conducted by the Kenya National Bureau of Statistics (KNBS). Second, surveys with individual employment information, including the Kenya Integrated Household Budget Survey (KIHBS), the 2019 Kenya Population and Housing Census (KPHC), the Kenya Continuous Household Survey (KCHS), and the Kenya Demographic and Health Survey (KDHS).

2.3 Data on urbanisation, the road network, and fiber broadband

To look at the drivers of changes in employment patterns over space we leverage data from a variety of non-traditional sources.

First, we use data on urbanisation over space and time from processed satellite imagery. In particular, we use 10m by 10m pixels from the ESRI 10m Annual Land Cover data covering 2017 to 2023 [Karra, 2021]. As described in the data documentation, these maps are derived from the European Space Agency Copernicus Sentinel-2 imagery at 10m resolution. Each map is a composite of land use/ land cover predictions for 9 classes throughout the year to generate a representative snapshot of each year. These raw images are processed by the Impact Observatory (https://www.impactobservatory.com/) they use billions of human-labeled pixels to train a deep learning model for land classification. The global map was produced by applying this model to the Sentinel-2 annual scene collections on the Microsoft Planetary Computer. Each of the maps has an assessed average accuracy of over 75%. In this paper, we are interested in the built-area classification available in this dataset.

[3] Fan et al. [2023] highlight that this is their preferred approach. In the absence of such data, they rely on a firm survey that allows the to classify firms into consumer and producer service firms based on the sector of operation and firm size.

This covers: human-made structures; major road and rail networks; large homogeneous impervious surfaces including parking structures, office buildings, and residential housing; examples: houses, dense villages/towns/cities, paved roads, and asphalt.

Second, we use data on the road network from OpenStreetMaps (OSM), a prominent open-source mapping platform that aggregates public road information and volunteer contributions. We obtain yearly snapshots of Kenya's OSM road network data from 2016 to 2022 from Geofabrik, a publicly accessible OSM archival server (https://download.geofabrik.de/africa/kenya.html). This allow us to obtain the development of roads in Kenya over the study period at the national, with the total network length increasing from 100,100 km to 360,600 km from 2016 to 2022. Subsequently, we use the Open Source Routing Machine (https://project-osrm.org/), which is a routing engine based on the contraction hierarchy algorithm, to construct a panel of travel time matrix between firms or sub-county headquarters based on the shortest network path. We also use administrative road data from the Kenyan Road Board's Road Inventory Condition Survey in 2018 and 2023 (https://maps.krb.go.ke/kenya-roads-board12769/maps/110400/ to obtain granular information on road class and condition changes. Unfortunately, this data source has insufficient time variation to construct the travel time matrices. Hence, we only use RICS data as a cross-reference and basis to build our hypothetical optical fiber expansion network later on.

Third, we use data on the fiber broadband network rollout from various sources. To track the backbone structure of the fiber optic network, we rely on the Africa Transmission Maps. We complement these with granular data provided by Kenya's largest telecommunication firm, which has been responsible for much of the expansion of the network accessible to the private sector along the extensive margin (reaching new locations, rather than extending capacity in already connected ones). We obtained data on fiber optic cable operated by the firm itself, fiber leased from other operators including the government, and the geolocation of nodes that act as local service hubs. To develop our instrument for the roll-out of the government network, we further rely on the list of government sites connected as part of the expansion of the strategic government network.[4]

[4] This list was published by the General Auditor: `http://www.parliament.go.ke/sites/default/files/2022-11/Forensic%20Audit%20Report%20on%20National%20Optic%20Fibre%20Backbone%20Infrastructure%20and%20IP%20Based%20HIPATH%204000%20Network%20Projects%20from%20the%20Auditor-General%20for%20August%202022.pdf` *last accessed: 23rd July 2024*

2.4 Geographic classification

We classify administrative units in Kenya into five groups based on their level of urbanization, namely cities, metropolitan areas, large towns, small towns and rural areas.[5] In most of our applications we group categories two and three together into "metropolitan areas and large towns" and categories four and five into "small towns and rural areas". As with any such classification, there will always remain an element of arbitrarity although we have tried to remain as objective as possible. We rely on such classifications only when presenting stylised facts, and in these cases the qualitative take aways are robust to reasonable alternative classifications as discussed in the appendix.

1. City: subcounties with a population density of at least 1500 people per km^2 and a settlement with more than 500,000 inhabitants.[6]

2. Metropolitan area: subcounties neighboring the subcounties classified as cities.

3. Large town: Subcounties with at least 300 people per km^2 and a settlement with more than 150,000 inhabitants.[7]

4. Small town: Subcounties with urban centres of at least 5,000 inhabitants or a population density of more than 300 people per km^2, but which have not been classified as metropolitan areas, large towns, or cities.

5. Rural: Subcounties with a population density of less than 300 people per km^2 and without an urban center of at least 5,000 inhabitants.

3 The changing nature of formal employment in Kenya

In this section, we describe four stylized facts that motivate our work. In all cases, they pertain to our main study period covering seven years from 2016m1 to 2022m12. In constructing these facts we rely heavily on the administrative matched employee-employer data

[5] We largely follow UN Statistical Commission's *Degree of Urbanization* classification. https://unstats.un.org/unsd/statcom/51st-session/documents/BG-Item3j-Recommendation-E.pdf The UN's classification, however, only distinguishes between three groups, we hence expand and adapt it for our purposes.

[6] Cities can span multiple neighboring subcounties like Nairobi and Mombasa, which have 11 and 6 subcounties respectively. We include those subcounties with a population density of 1500 people per km^2 and above.

[7] The town of Eldoret is split across several subcounties, each capturing a relatively small share of the town. We use the four subcounties with the highest population density. In Kisumu, we define the town as the three subcounties Kisumu Central, West and East.

and the VAT records data. For additional indicators such as occupation codes, we leverage survey data.

The first two facts pertain to the macro-level trends we document. First, that formal employment has been shifting rapidly in the aggregate from industry to services. Second, that this pattern is not constant across space: in small towns and rural areas we only see rapid growth in consumer-facing services. Fact three turns the lens on the characteristics of formal jobs, showing that they are better paid, more managerial and professional and in larger firms. Finally, fact four leverages our firm-level data to characterize expanding firms. We find that expanding consumer-services firms outside main urban areas disproportionately sell to final consumers, and are expanding their market share at the expense of smaller firms.

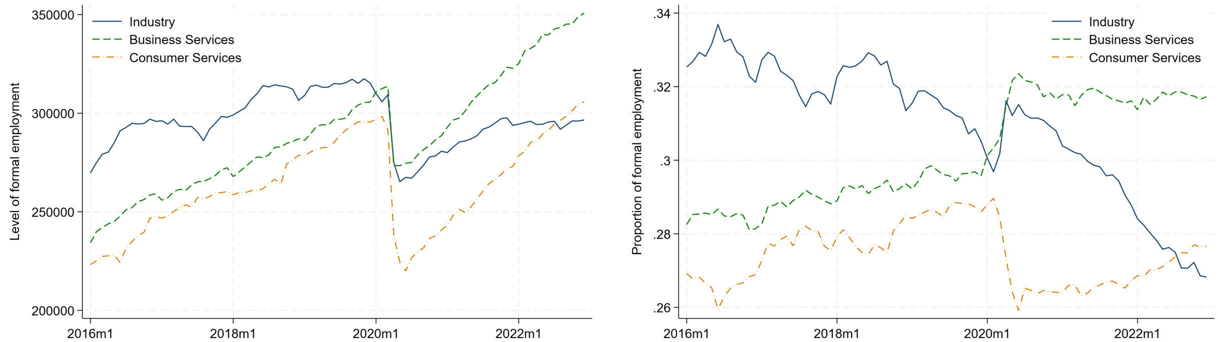
3.1 Formal employment has rapidly shifted from industry to services

Despite the proportion of those in the labor force formally employed remaining constant (see figure A6 in the appendix), the composition of formal employment has changed rapidly over the study period. Figure 1 (b) shows that since 2016 formal employment has shifted from industry to services at the rate of around one percentage point a year. Figure 1 (a) shows that as this is within a context of rapid population increases and rapid formal employment despite the shift across sectors the number employed in industry remains on an upward trajectory.

Figure 1 also splits services into consumer and business services. Notwithstanding the pandemic which hit consumer services harder, a similar upward trajectory on aggregate can be observed in both services sub-categories. The impact of the pandemic and lockdowns in early 2020 are also clearly visible in figure 1.

Figure 1 Change in formal employment by sector

(a) In Levels (b) In Shares



Notes: This figure shows the change in formal employment over time in Kenya using administrative data from PAYE records. The left-hand-side panel shows change in employment in levels, the right-hand-side panel shows changes in employment as a share of total employment. In each case we show employment in industry, business services, and consumer services separately.

Next, we consider whether this shift in the composition of formal sector work is seen in a similar fashion across the country, or if it is instead concentrated in specific geographies.

3.2 Formal employment growth outside the main cities has been almost entirely in consumer-facing services

Figure 2 splits the aggregate trend for each sector (industry, consumer services, business services) into three geographies. The geographic classification is discussed in detail in sub-section 2.4. Broadly we consider three mutually exclusive geographies (i) cities, (ii) metropolitan areas and large towns, and (iii) small towns and rural areas.

Figures 2a and 2b show the trend over time for each geography in formal industry employment in levels and shares respectively. This figure reveals a clear pattern. Within cities industrial employment is increasing in absolute numbers and remaining flat as a share of overall formal employment. However, outside of the main cities, the trend is one of stagnation in levels leading to rapid reductions in industrial employment as a share of total formal employment. Between January 2016 and December 2022 the share of formal employment in industry outside of the main cities fell by around 30 percent, or 10 percentage points — a remarkably fast shift over just seven years.

Figures 2c and 2d analogously show the change in business service employment in levels and shares over time across the three geographies. In levels business service employment is increasing in all three geographies, although to a far greater degree in main cities than outside. This increase is sufficiently fast to be reflected in the shares also, although here we see a particularly fast relative increase in metropolitan areas and large towns vis-a-vis cities or small towns and rural areas. Although we see an increase in levels and shares in smaller towns and rural areas, this is from a very low base, and even at the end of the period total formal employment in this sector in these areas remains below an almost negligible five percent of total formal employment.

Finally, figures 2e and 2f analogously show the change in consumer service employment in levels and shares over time across the three geographies. Again, in all three geographies, the level trend is upwards, although two differences are of note here. First, the pandemic clearly hit formal consumer services the hardest, and only by the end of 2022 is employment reaching similar levels to those seen just before the pandemic. Second, despite its significantly larger base, the employment change in big cities is not of an order of magnitude different to those outside big cities, in contrast to the case with business services. This feature becomes clear when we consider the change in formal employment shares. Although the proportion employed in consumer services remains fairly constant in big cities, metropolitan areas, and large towns it increases very quickly in smaller towns and rural areas. In this last, and most rural geography, the proportion of formal employment in consumer services has increased by roughly 67 percent or 10 percentage points — an even more remarkably fast shift over just seven years.

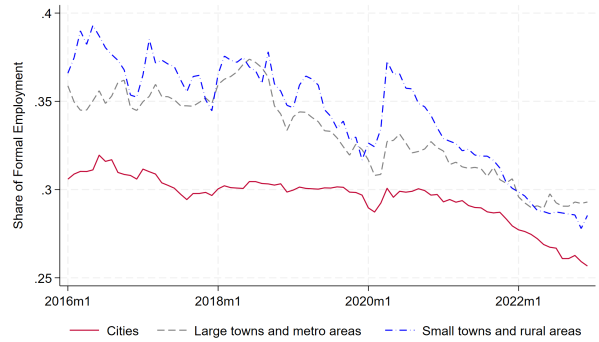
In sum figure 2 paints a clear picture. The aggregate trends uncovered in figure 1 do not hold equally across space. In large and medium-sized cities the trend is from industry to business services, whereas in smaller cities and rural areas, the trend is from industry to consumer services. However, what is clear across the board is that these shifts are rapid, and show little sign of abating.

Finally, we explore whether the trends are sensitive to the way we classify firms into business versus consumer services in Figure A5. Relying on the self-reported sector code of firms to distinguish between the two types of services, results in under-counting employment in business services by 4% in cities, 6% in metro areas and larger towns, and 2% in small towns and rural areas. However, the aggregate shifts in sector shares over time are unaffected by the classification.

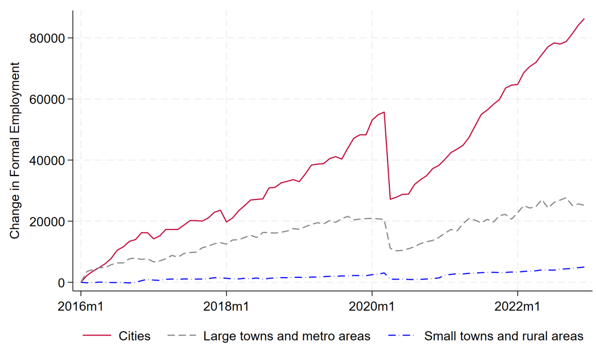
Figure 2 Change in formal employment by sector and geography

(a) Industry in Levels

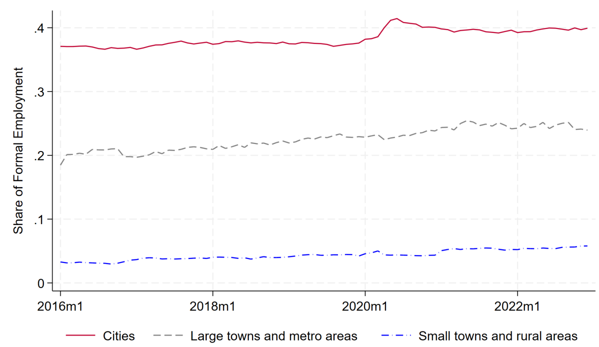
(b) Industry in Shares



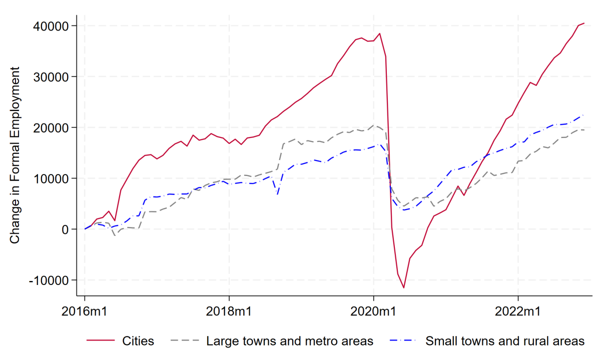
(c) Business Services in Levels



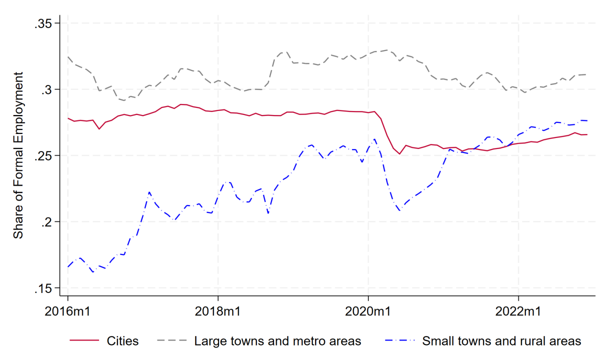
(d) Business Services in Shares



(e) Consumer Services in Levels



(f) Consumer Services in Shares

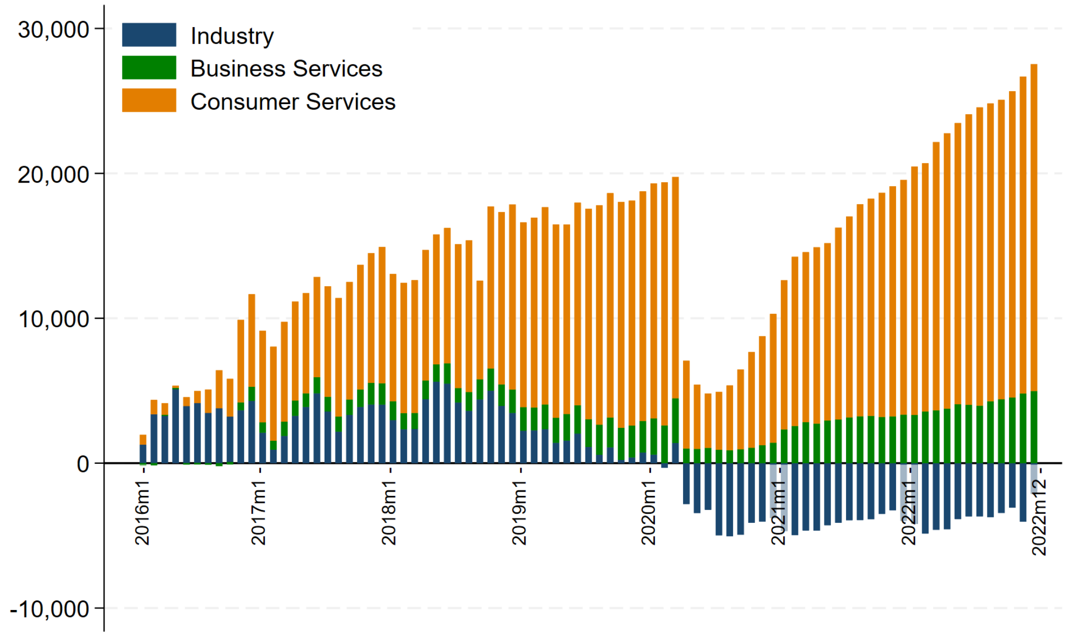


Notes: This figure shows the change in formal employment by sector and geography using administrative data from PAYE records. The left-hand-side figures show changes in levels (normalised to 0 in 2016m1) and the right hand side figures show changes in shares. Panels (a) and (b) show industrial employment, (c) and (d) business services employment, and (e) and (f) consumer services employment. For all figures we display trends for three mutually exclusive geographies: Cities, large towns and metropolitan areas, and small towns and rural areas.

The growth in consumer services has not been even across small towns and rural areas. Figure A9 in the appendix shows that only 30 sub-counties account for over 90% of the growth in consumer services among sub-counties classified as consisting of smaller towns or rural areas. Figure A9 also shows that these high-growth sub-counties are not geographically concentrated in one part of the country, but rather scattered across the whole of Kenya. In subsequent sections, we will investigate in more detail what might be causing these heterogeneous growth rates.

Figure 3 highlights this point in levels by showing formal employment growth in small towns and rural areas broken down into the three main categories. It is very clear from this figure that almost all formal employment growth in these areas has been due to increasing numbers of workers in consumer-facing services.

Figure 3 Formal employment growth in small towns and rural areas



Notes: This figure shows formal employment growth in small towns and rural areas in Kenya. Growth is broken down into that due to employment in industry, employment in business services, and employment in consumer services.

3.3 Formal jobs are higher paid, more professional and managerial, and in larger firms

Above we established striking patterns. Formal employment in Kenya is shifting rapidly from industry to business services in cities, and from industry to consumer services outside main cities. We now turn to better understanding descriptively what it means for formal employment to be expanding, what kind of jobs are these, how well paid, and in what type of firms. Answering these questions is impossible with administrative data covering only formal employment, comparable information on informal jobs is taken from the 2021 wave of the Kenyan Continuous Household Survey (KCHS).

We find that formal jobs are (considerably) higher paid, more managerial and professional, and in bigger firms. Figure 4 summarises these findings.

In sub-figure 4a we plot the distribution of wages in the administrative data covering all formal workers and focusing on formal consumer service workers against the distribution of formal and informal wages found in the KCHS data. Formal jobs are considerably better paid. This holds on aggregate, but also when comparing to only consumer service formal workers or when considering the admin or KCHS data.

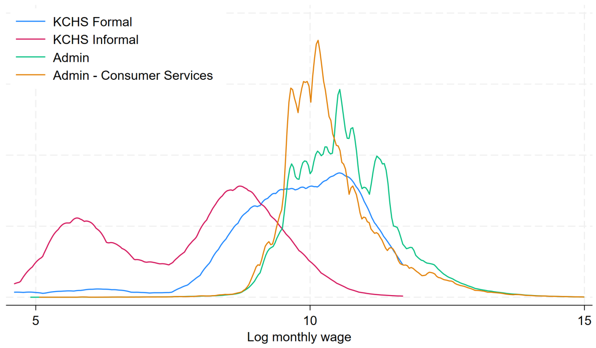
In sub-figure 4b we compare occupations workers in formal vs informal jobs are engaged in using the KCHS data. One can clearly see from this sub-figure that formal workers are far more likely to be employed in a managerial or professional capacity than in informal jobs. This indicates that some of the trends picked up by the DHS survey are driven by the observed shift the formal sector.

In sub-figure 4c we plot the distribution of firm size for workers in formal and informal employment using the KCHS data. This also shows a striking pattern. Informal workers are much more likely to be self-employed or work in firms with 4 or fewer workers than formal workers who overwhelmingly are employed in forms of 5 or more employees. In this figure, the total number of employees at a firm includes both formal and informal workers.

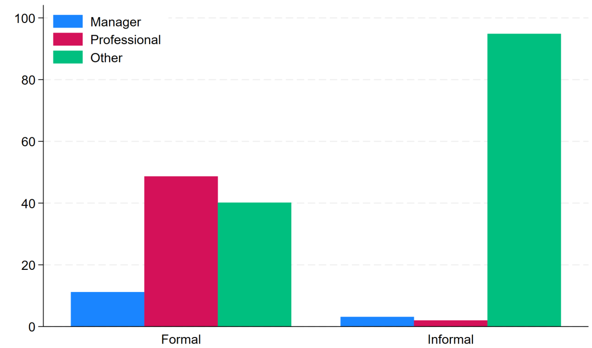
Figure 4 presents a picture of formal jobs being more complex, in larger firms, and better paid. Indeed interviews conducted by the authors corroborate the conclusion that especially outside of the main cities, formal jobs represent the higher echelons of a fairly complex company.

Figure 4 Characterising formal jobs

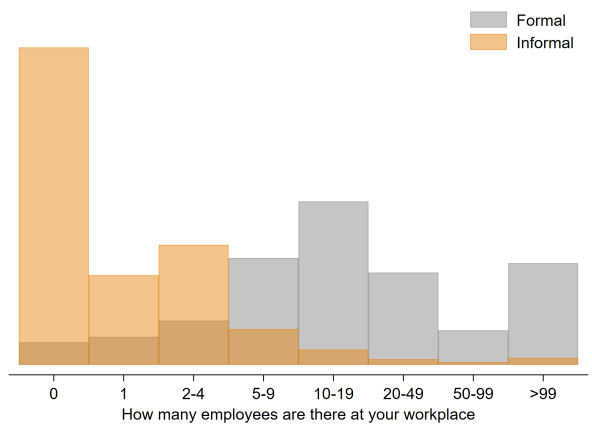
(a) Wage distribution



(b) Occupations



(c) Firm size distribution



Notes: This figure characterises formal jobs via three variables. In panel (a) we show the distribution of log monthly wages for formal and informal workers using data from the KCHS and administrative PAYE records. In panel (b) we show the proportion of workers in managerial vs professional vs other occupations splitting by formality using KCHS data. Panel (c) shows the distribution of firm size for those in formal and informal employment using KCHS data.

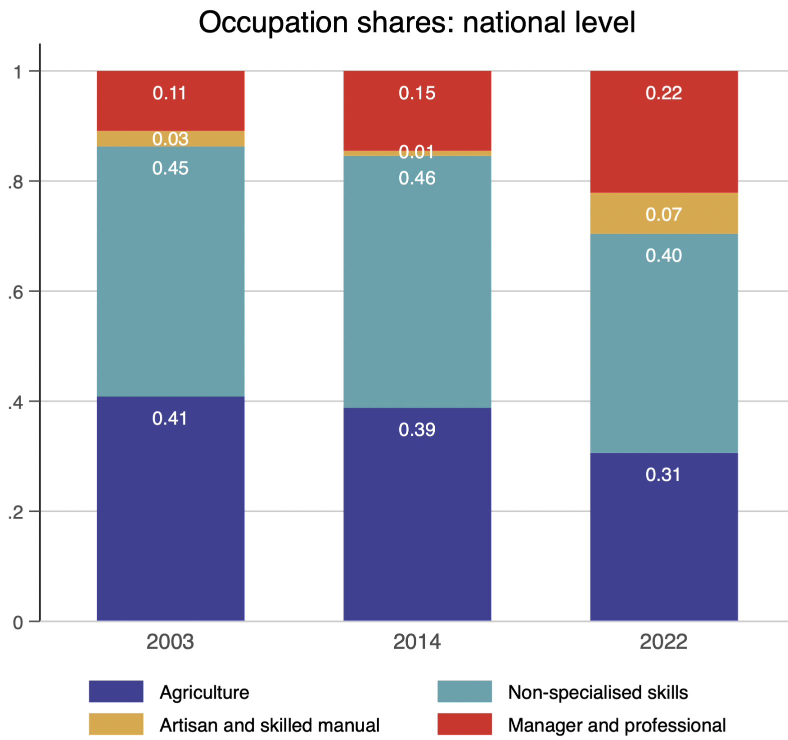
Figure 1 tells a story of rapid sectoral change across Kenya but is confined the sector of operation of formal sector firms. To examine whether this transformation reflects broader trends, we leverage the Demographic and Health Surveys (DHS) to aggregate changes in occupations over time. The DHS are comparable over time and allows us to look at snapshots in 2003, 2014, and 2022.[8]

Figure 5 corroborates these findings using DHS data. It shows that the proportion of workers in managerial and professional roles has increased substantially from 15% to 22% between 2014 and 2022 — at the expense of work in agriculture and non-specialized

[8]Relative to other surveys like the Kenya Integrated Household Budget Survey or the Kenya Continous Household Survey, the DHS data comes with the advantage that several survey rounds were collected with the same questionnaire and classifications at intervals that allow us to observe medium-term transformations. We omit results from the 2008 survey as the same aggregation of responses resulted in implausible sector compositions.

occupations.

Figure 5 Employment in the DHS data

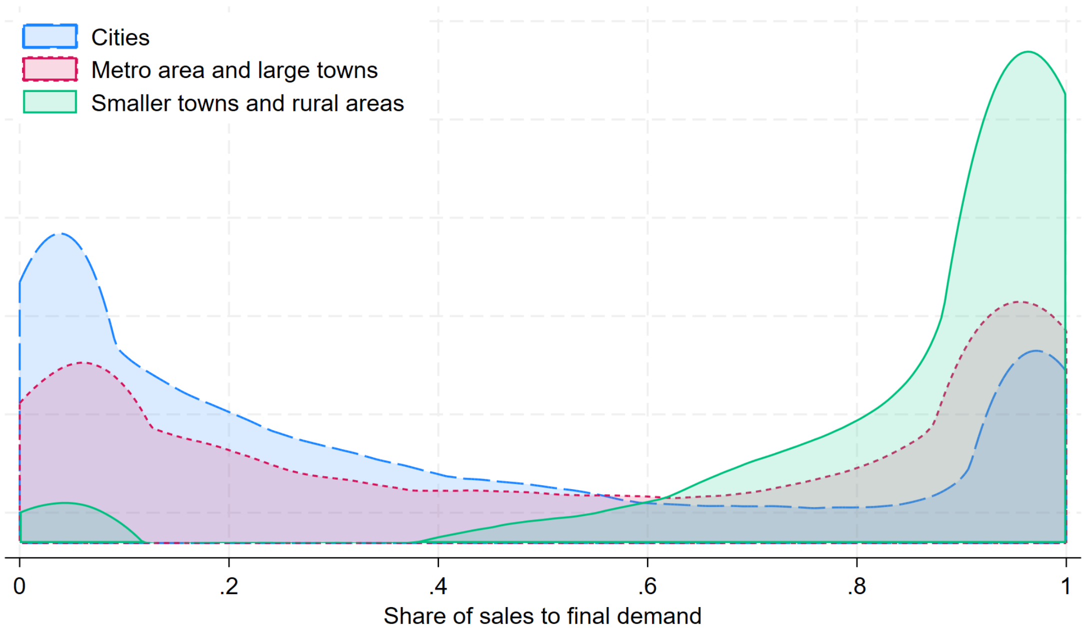


Notes: This figure shows the national-level occupation shares constructed using DHS data from 2003, 2014, and 2022. The occupation classification is as follows: (1) Agriculture: agriculture - self employed; agriculture - employee; (2) Non-specialised skills: sales; household and domestic; services; unskilled manual; (3) Artisan and skilled manual: artisan; skilled manual; (4) Manager and professional: professional/technical/managerial; clerical. Since DHS published datasets for men and women separately, we calculated occupation shares for men and women separately and aggregated them using gender weights generated from the 2009 and 2019 Censuses.

3.4 Growing firms sell mainly to consumers and expand market share.

What are the characteristics of the firms that create these new jobs? First, we find that service firms in rural areas and small towns disproportionately source their demand from final consumers relative to service firms in cities or metro areas 6. This reinforces that services in these areas are not the same as those being provided in large towns, and highlights that demand growth is unlikely to come from downstream productivity growth.

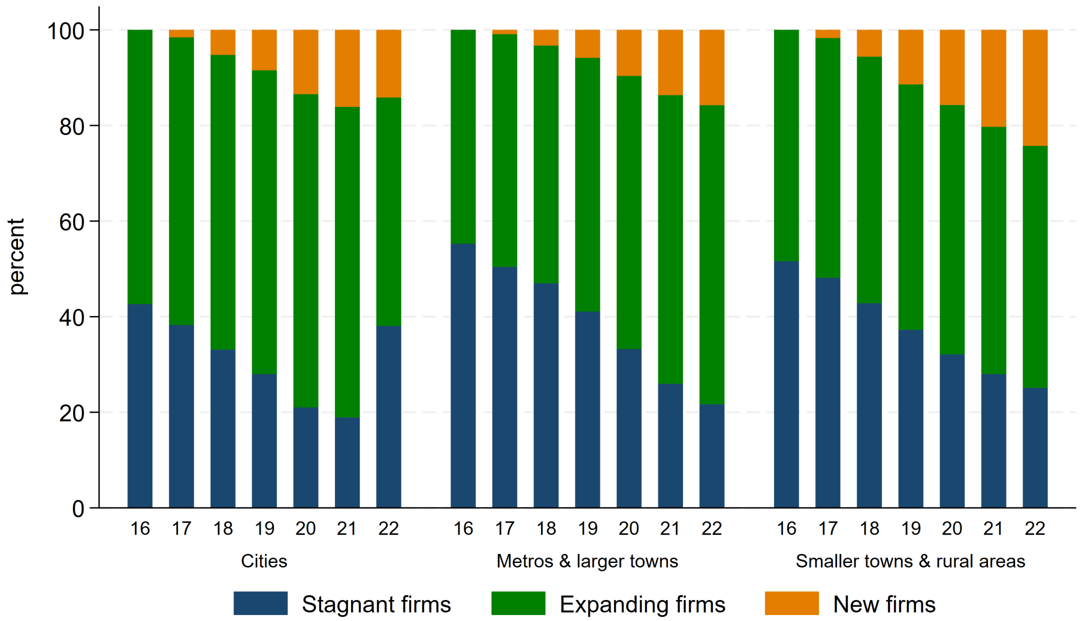
Figure 6 Share of sales to final demand by geography



Notes: This figure plots the share of sales sold to final demand in 2022 by service sector firms registered for VAT.

Second, we show that expanding consumer service firms grow not only in terms of (formal) employment but also market share. Figure 7 shows how expanding and new firms outside cities have significantly increased market share as they grow.

Figure 7 Market share of consumer service firms



Notes: This figure plots the sales share of three types of consumer service firms between 2016 and 2022: firms that remain stagnant or even see a decline in their workforce between 2016 and 2022, those that add new employees, and firms that newly enter.

Table 1 shows in more detail who these expanding firms are. It shows that although most new jobs in consumer services in small towns and rural areas come from new firms — this proportion is less than in more urban areas. That is, a relatively larger proportion comes from existing firms expanding.

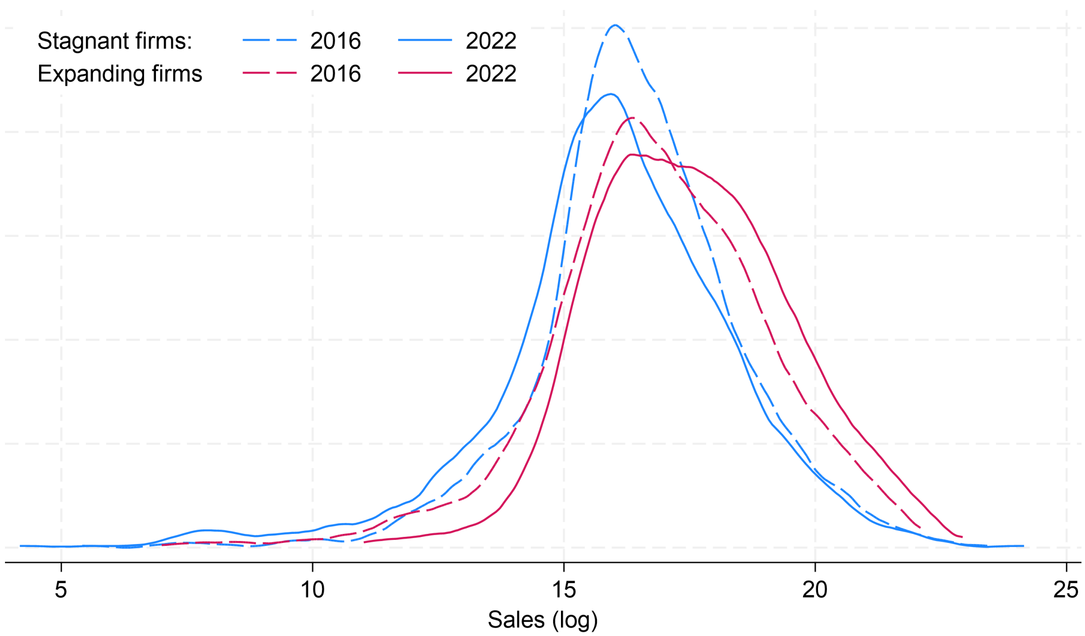
Table 1 What share of net job creation takes place in firms born since 2016?

Sector	City	Metro & large town	Small town & rural
Agriculture	0.46	0.68	0.96
Industry	0.92	0.91	1.15
Business services	0.74	0.65	0.48
Consumer services	0.81	0.75	0.60

Notes: The above table reports the share of net jobs created between 2016 and 2022 that were created in new firms, i.e. those that have not yet been born in 2016. Net jobs created is the difference between the total number of formal jobs in 2022 and 2016. A share greater than 1 implies that net job creation was negative in this sector-region group, but the number of jobs created by new firms was overall positive. We rely on sales shares to classify firms into consumer and business services in this table.

As a result of growing firms expanding market share — the firm-size distribution of expanding firms is becoming increasingly right-shifted. Figure 8 clearly shows this trend. Figure 8 plots firm sales for stagnant and expanding consumer services firms in small towns and rural areas. Firms that expand their workforce between 2016 and 2022 were already larger in 2022. However, the sales distribution of the two types of firms diverges notably in those seven years. Figure A15 in the appendix then shows the aggregate impact these differential growth rates has had on the overall firm-size distribution in small towns and rural areas vs cities and metro areas. This figure shows a clear convergence. The firm-size distribution in small towns are rural areas increasingly looks like that in cities and metro areas/ large towns.

Figure 8 Firm size distribution of stagnant and expanding firms in 2016 versus 2022



Notes: The figure plots the firm size distribution (in log sales) for consumer service firms in small towns and rural areas. We distinguish between expanding and stagnant firms. Expanding firms expand their workforce between 2016 and 2022, while it remains constant or declines for stagnant ones.

4 What is causing the rapid rise of consumer services?

We have documented rapid increases in formal employment in consumer services outside the main urban areas. Understanding the mechanisms behind these changes is crucial to understanding their welfare implications and sustainability. We will first briefly sketch out a model of consumer behavior that formalises and clarifies the main potential channels before showing empirical evidence for each. In future work we will then write down and estimate a quantifiable version of this framework which will allow us to precisely decompose channels and perform counterfactual analysis.

4.1 Stylised model

To set ideas we first write down a stylised and general model which will allow us to focus on the main potential channels. In the appendix section D we detail a micro-foundation of this general model.

We are interested in the number of people employed in the formal consumer services sector in a given location i denoted by L_i. Assuming wages adjust to clear labor supply and demand L_i will be determined by the local wage level w_i. Incentives to work, however, will not only depend on the wage being offered, but also the utility value of consumption that can be purchased with this income. Allowing individuals to consume market and non-market goods implies that the price of consumer services p_i will also be an important determinant of L_i. Additionally, we can allow for non-homothetic preferences whereby L_i will also depend on local income levels I_i. In sum we have $L_i = f(w_i, I_i, p_i)$.

Allowing firms to behave endogenously, local wages will depend on local productivity in the service sector A_i and local demand for service sector goods D_i. Similarly, prices will depend on marginal costs which will be a function of local wages. Therefore we have the following $L_i = f(g(A_i, D_i), I_i, p_i(A_i, g(A_i, D_i)))$. However, local income, and local demand, will also both depend on the size of the local labor force. If more people are working in consumer services incomes and local demand will be higher. This highlights the fortuitous cycles possible in the consumer service sector where local demand is tightly linked to local prosperity. In total we can write a general expression for the number of people working in the formal consumer service sector as the following.

$$L_i = f\left(g(A_i, D_i(L_i)), I_i(L_i), p_i(A_i, g(A_i, D_i(L_i)))\right) \tag{1}$$

This framework highlights a number of key economic forces that could be driving the stylised facts we document.

1. Increases in local market size D_i. As the local population becomes larger or more prosperous local demand for consumer services will rise causing a subsequent rise in local wages and therefore employment. However, local consumer service prices may also increase dampening the effect.

2. Increases in consumer service productivity A_i. As consumer service productivity rises wages increase and consumer service prices fall incentivising a greater shift to employment in consumer services.

3. Increases in local income I_i and non-homotheticities. As incomes rise (due to increases in cash-crop prices or home production productivity) individuals will likely spend a larger proportion of their income on consumer service goods increasing incentives to earn to buy such goods and demand for said goods and therefore wages.

In the remainder of this section we provide some indicative evidence on each of these main channels and some evidence against possible alternative explanations. In the appendix section D we provide a micro-foundation for the stylised model set out above.

4.2 Indicative evidence on the key channels

4.2.1 Increases in local market size

To provide indicative evidence on the local labor market size effect we take urbanisation and remoteness as proxies for the size of local demand.

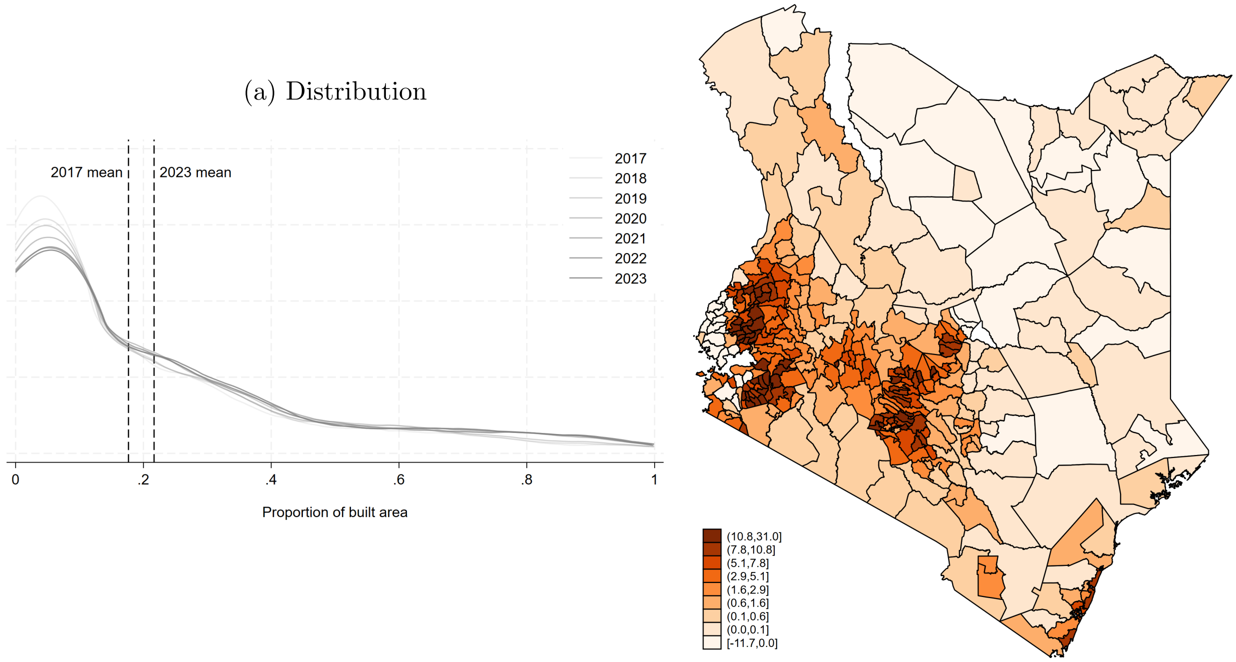
Urbanisation

Urbanisation is closely related to structural transformation and the diffusion and marketisation of services as well as other activities. Rapid future urbanisation is expected in many low and middle-income countries including Kenya, making this a particularly pertinent variable to study in this context.

The urbanisation rate for each sub-county is calculated using satellite data from 2017 to 2023. We use information from the ESRI 10m Annual Land Cover data to calculate the proportion of pixels of built area within each sub-county. This is our main time-varying urbanisation measure. No time-varying and time-consistent alternative measure based on the proportion of population who live in urban areas is available at the sub-county (or county) level.

Figure 9 shows the change in urbanisation rates from 2017 to 2023 across Kenya. Sub-figure 9a shows how the distribution of built-up area changes over time across sub-counties — one can see a clear right-ward shift in the distribution and mean. In sub-figure 9b we show the sub-county level change in urbanisation rates, some sub-counties have increased urbanisation by over 10%, whereas others have flat-lined or even decreased.

Figure 9 Change in the proportion of land built up by sub-county 2017-2023



(a) Distribution

(b) Map

Notes: This figure shows the change in the distribution of built up land in Kenya from 2017 to 2023. In panel (a) we show a kernal density plot of the proportion of built up land across sub-counties in each year 2017 to 2023. In panel (b) we show on a map the change in proportion of built up land from 2017 to 2023 across sub-counties, darker shades indicate larger positive increases in built up land.

Table 2 shows the results from regressing the proportion of a sub-county that is classified as built up against the number of people employed in formal consumer services. That is, it shows the results from estimating regression of the form shown in equation 2, where the fixed effects, weighting, and sample restrictions used are shown clearly in the table.

$$L_{it}^{ConServ} = \beta \cdot \text{PropBuiltUp}_{it} + \text{FixedEffects}_{it} + \varepsilon_{it} \tag{2}$$

In table 2 we restrict attention to the 50% least urban sub-counties in 2016 as we are interested in explaining the phenomenon highlighted in the stylised fact: Rapid growth in formal consumer service employment in small towns and rural areas. However, as shown in the table, results are not sensitive to changing this cut-off. Table 2 shows the regression with various specifications, but the qualitative take-away remains in all columns.

Focusing on our preferred specification in column one of table 2, we see that in relatively rural areas increases in urbanisation are correlated with increases in consumer service employment. In particular, a one percentage point increase in the total area built up is associated with almost six new formal consumer service sector workers on a baseline mean of 340. These results show that urbanisation is positively correlated with formal consumer service employment, but that although although statistically significant, the relationship is relatively weak.

Table 2 Regressing number employed in consumer services against the proportion of built up area

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion built up	5.546**	4.245*	3.990	7.092**	17.89*	9.328***	17.89*	12.37
	(2.596)	(2.474)	(2.425)	(3.443)	(9.336)	(2.980)	(9.256)	(9.634)
Fixed effects	sCnty + Year	sCnty + Year	sCnty + Year	sCnty + Year	Year	sCnty		Cnty + Year
% built up cut off	50%	40%	60%	50%	50%	50%	50%	50%
Weighting				Labor force				
R2	0.982	0.983	0.984	0.979	0.0251	0.982	0.0247	0.179
N	1278	1218	1326	1278	1278	1278	1278	1248

Notes: This table shows the results from regressing the number employed in consumer services against the proportion of built up land at the sub-county by year level. In column one we include sub-county and year fixed effects and only consider sub-counties with less than 50% built up land in 2017. In column two we keep the same fixed effects and reduced the built up cut off to 40% and in column three 60%. Column four keeps the same specification as column one but weights by the 2017 labor force. Column five keeps the baseline specification of column one but only includes year fixed effects. Column six maintains the baseline specification but only includes sub-county fixed effects. Column seven has no fixed effects but maintains the 50% 2017 built area cut off. Column eight includes county and year fixed effects with the 50% cut off. In each regression standard errors are clustered at the sub-county level.

Road building and remoteness

We use information on roads over time from Open Street Maps (OSM) which can be thought of as an open-source version of Google Maps. Open street maps covers roads in Kenya from 2016 including information about where they are, junctions, travel direction, and travel speed. OSM has the advantage that it is an extremely detailed and consistent panel dataset and a complete network with a built-in API that allows us to query travel times between given locations in different periods easily. It has the disadvantage that it is potentially updated with a lag and can include some measurement errors.

From the road network given in each year historical versions of OSM we can construct travel times between each pair of sub-county headquarters. Where sub-county headquarters is the administrative centre of each sub-county. This procedure allows us to calculate sub-county to sub-county travel times which we denote by t_{ij}. Armed with these travel times, we then calculate three measures of connectedness of interest.

- Remoteness. Remoteness is a simple a-theoretic measure of how remote a sub-county is. It is defined as the sum of travel times from the sub-county to all other sub-counties. $R_{it} = \sum_j t_{ijt}$. The remoteness of a location is high if it is far away from many other locations.

- Market potential. A simple a-theoretic measure capturing how connected a sub-county is. $MP_{it} = \sum_j \tau_{ijt}^{-1} L_j$. The market potential of a location is higher if it is better connected to other large locations.

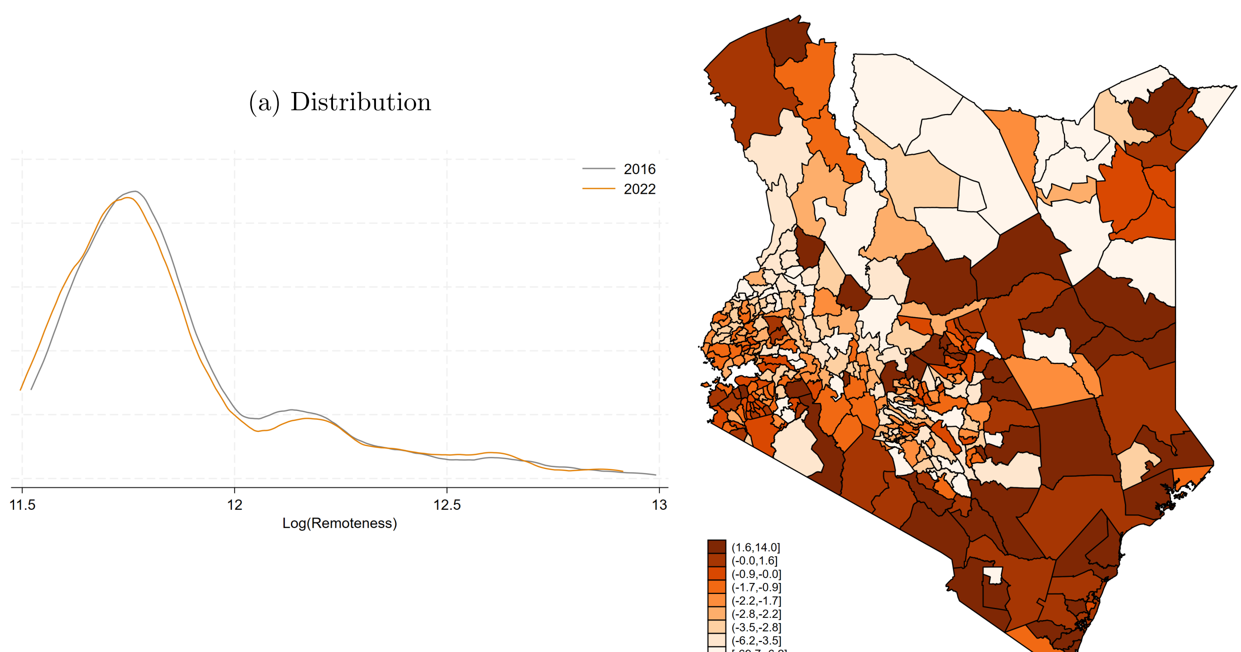
- Market access. A theory-consistent measure of how connected a sub-county is, that takes into account how connected other sub-counties are. $MA_{it} = \sum_j t_{ijt}^{-3.2} \frac{L_j}{MA_j}$. The market access of a location is high if it is better connected to other large locations which are themselves not well connected to alternative markets. To find each location's market access we solve a series of non-linear equations via a fixed-point theorem where existence and uniqueness are guaranteed [Donaldson, 2018].

Focusing on the most intuitive and simplest of the measures considered figure 10 graphically displays the variation in our data. In sub-figure 10a we show the distribution of log remoteness in 2016 and 2022 and in sub-figure 10b we map the change in log remoteness over space. Both sub-figures show considerable variation between 2016 and 2023. Figure A11 visualizes road-by-road changes using an alternative administrative data source from the Kenyan Roads Board.

Figure 10 Change in log remoteness by sub-county 2016-2022

(a) Distribution

(b) Map



Notes: This figure displays the variation in remoteness across sub-counties and time. In panel (a) we show the change over time comparing the distribution of log remoteness across sub-counties in 2016 to 2022. In panel (b) we show the change in log remoteness across sub counties in a map, darker areas indicate larger positive changes in remoteness (sub-counties becoming more remote).

Armed with variation over space and time in connectivity, we next turn to discuss whether such changes are correlated with changes in formal consumer service employment. To do this we run regressions of the form given in equation 3 below. Where $L_{it}^{ConServ}$ denotes local employment in consumer services, $\text{Connectedness}_{it}$ denotes one of the three connectedness measures described above, α_i denotes sub-county fixed effects, τ_t denotes year fixed effects, and ε_{it} is an error term. The object of interest is β_q which can be interpreted as the built-up area quartile q's specific elasticity of connectedness against employment in formal consumer services.

$$\ln\left(L_{it}^{ConServ}\right) = \sum_q \beta_q \cdot \ln\left(\text{Connectedness}_{it}\right) + \alpha_i + \tau_t + \varepsilon_{it} \tag{3}$$

$$\ln\left(L_{it}^{ConServ}\right) = \gamma \ln\left(\text{Connectedness}_{it}\right) + \sum_q \beta_q \mathbb{1}\{\text{Builtup}_{iq}\} \cdot \ln\left(\text{Connectedness}_{it}\right) + \alpha_i + \tau_t + \varepsilon_{it}$$

Table 3 shows the results from running regressions of the form given in 3. Odd columns show the results from running OLS on logged employment, whereas even columns show the results from estimating the equation in levels using PPML. Both tell a similar story that is also consistent across connectedness measures. For areas with little pre-existing built-up area (quartile 1) becoming more connected increases employment in formal consumer services. This effect dissipates and even turns negative for increasingly built-up areas. Taking column two as our preferred specification, a 1% increase in market access is associated with a 0.68% increase in formal consumer service workers for those areas in the lowest pre-period built-up area quartile. Whereas, in the highest quartile a 1% increase in market access is associated with a 0.16% decrease in formal consumer service sector employment.

Table 3 Regressing number employed in consumer services against connectedness measures

	(1) OLS	(2) PPML	(3) OLS	(4) PPML	(5) OLS	(6) PPML
Builtup quartile=1 × Log(MA)	0.738**	0.676***				
	(0.299)	(0.179)				
Builtup quartile=2 × Log(MA)	0.286	0.430*				
	(0.335)	(0.230)				
Builtup quartile=3 × Log(MA)	0.278	-0.420				
	(0.236)	(0.358)				
Builtup quartile=4 × Log(MA)	-0.175	-0.163***				
	(0.444)	(0.0533)				
Builtup quartile=1 × Log(Remoteness)			-2.427**	-3.476***		
			(1.059)	(0.616)		
Builtup quartile=2 × Log(Remoteness)			-1.596	1.531		
			(2.268)	(3.301)		
Builtup quartile=3 × Log(Remoteness)			-5.338***	-1.915		
			(1.766)	(1.785)		
Builtup quartile=4 × Log(Remoteness)			-1.941	1.511		
			(2.801)	(1.366)		
Builtup quartile=1 × Log(MP)					1.707**	1.311**
					(0.665)	(0.562)
Builtup quartile=2 × Log(MP)					0.933	1.837***
					(1.284)	(0.525)
Builtup quartile=3 × Log(MP)					1.163	-1.167
					(0.772)	(1.088)
Builtup quartile=4 × Log(MP)					0.528	-0.549***
					(1.436)	(0.130)
Constant	6.176***	8.920***	37.62***	-3.827	-6.674	13.47***
	(1.304)	(0.344)	(14.26)	(12.77)	(6.539)	(1.603)
Observations	1139	1332	1139	1332	1139	1332

Notes: This table shows the results from regressing the number employed in consumer services against connectedness measures. Each specification includes sub-county and year fixed effects. In odd columns we perform OLS regressions, in even columns we perform PPML regressions. Standard errors are clustered at the sub-county level.

4.2.2 Consumer service productivity

We do not observed local consumer service firm productivity, and so empirically testing this channel is challenging. Instead we leverage variation from one dimension of consumer service productivity: Internet availability. We leverage variation in the roll-out of high-speed internet across Kenya adopting a hypothetical network market access based instrument.

Our analysis relies on multiple sources of spatial and infrastructure data to reconstruct fiber optic networks and simulate hypothetical expansions in Kenya. The primary datasets include:

- OpenStreetMap (OSM) 2023, used to establish a consistent spatial reference for road networks and fiber optic infrastructure.
- ATTM data, providing descriptions of NOFBI optical cables with high accuracy but poor geo-referencing.
- FTTS (Fiber-To-The-Site) nodes, which represent key points of fiber connectivity, and subcounty headquarters, manually geocoded where necessary.
- The NOFBI network in 2016 and 2022, reconstructed using a combination of ATTM and OSM data to improve spatial accuracy.

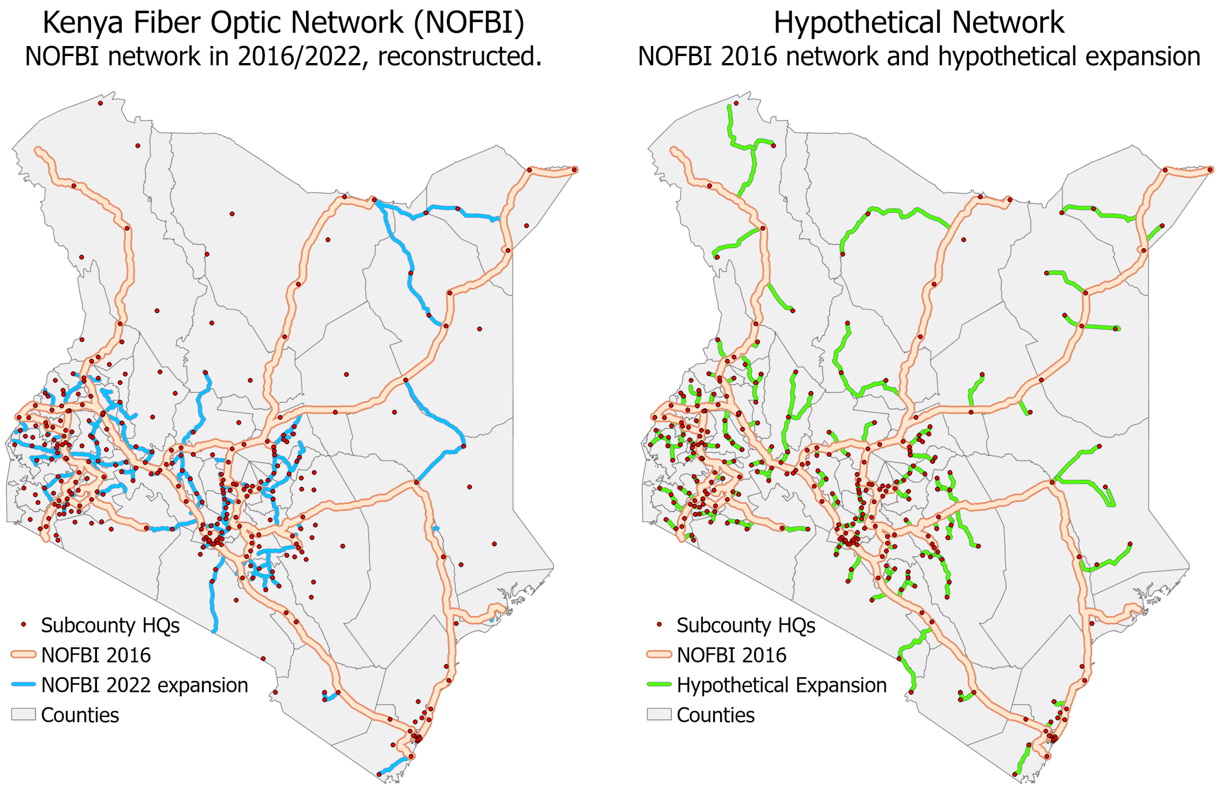
To ensure consistency, we subset the ATTM data to NOFBI and Telkom Kenya-operated lines, integrating pseudo-nodes (intersections) and simplifying nodes within 100m of each other. The NOFBI network is reconstructed by breaking lines into progression increments and ranking connection points using a weighted average progression measure. The final network is re-routed through OSM 2023 to correct spatial inconsistencies.

The instrument is based on a hypothetical fiber expansion network, generated through an iterative algorithm that selects expansion nodes and connects them to the existing network. The methodology follows these steps:

1. Subcounty headquarters and FTTS nodes are processed to ensure accurate representation, removing manually geocoded subcounties within 5km of the baseline network.
2. The selection of unique expansion nodes prioritizes "ODF Cabinet And Unit" or "OFC connection" nodes. For each subcounty, the closest node to the baseline network is selected based on road travel distance.
3. Expansion nodes are ranked using the metric: *Travel distance / Total subcounty population*. The highest-ranked node is iteratively connected to the baseline network, updating the expansion list dynamically.
4. The resulting network consists of 303 unique expansion nodes, with redundant nodes removed (e.g., those within 100m of the baseline network).

The resulting schedule provides a dynamic panel where FTTS nodes and subcounty headquarters are matched to their closest fiber connection each year. Figure 11 shows the actual and hypothetical expansion network as well as the baseline 2016 network.

Figure 11 Actual and hypothetical expansion network



Notes: These figures show the actual and hypothetical fiber expansion networks across sub-counties. The hypothetical expansion network in generated following the procedure described in the text.

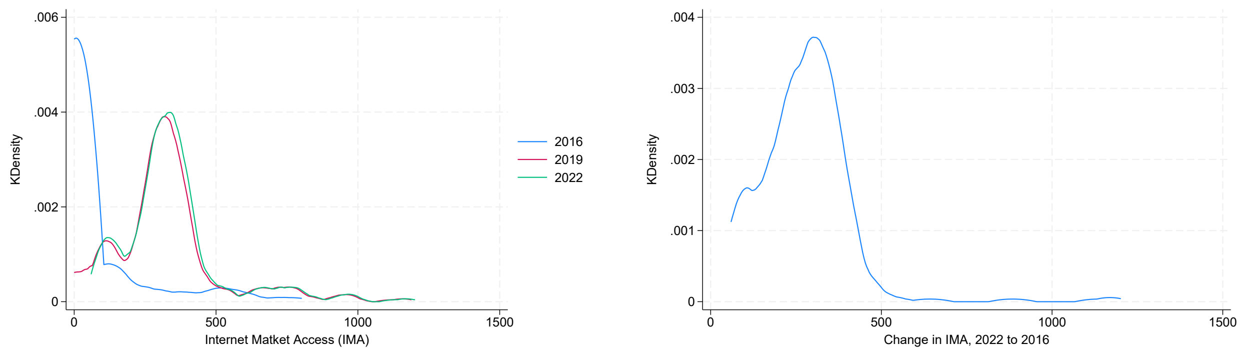
We construct the subcounty-level internet market access measure (IMA) as follows. We calculate the (unrestricted) straight-line distances between all subcounty HQ equipment & FTTS nodes and a version using travel time. We restrict this dataset to one representative point per subcounty HQ keeping the earliest active node if there exists an FTTS node in the county, and then the closest of these nodes to the unique county HQ. The yearly subcounty OD (i, j, respectively) internet market access term is defined as:

$$IMA_{it} = \mathbb{1}\{i \text{ connected}\}_t \sum_j \mathbb{1}\{j \text{ connected}\}_t L_j * (\tau_{ij})^{-\alpha}$$

, where τ_{ij} is the bilateral subcounty distance, using the unique representative subcounty node from the previous step, and α is the market access parameter, currently set as -1.

L_j is the destination subcounty's non-varying population. Figure 12 shows the variation in internet market access across sub-counties over time.

Figure 12 Variation in IMA_{it}



Notes: These figures show the variation in internet market access across sub-counties. The left hand size panel shows the empirical pdf in 2016, 2019, and 2022 whereas the right hand side panel shows the empirical pdf over 2022-2016 differences. Internet market access is calculated as described in the text.

Taking internet access as a proxy for consumer service productivity we can then regress the number employed in consumer services in a given subcounty-year against IMA_{it} instrumenting with internet market access from the hypothetical network and including sub-county and year fixed effects. Table 4 shows the results from these regressions. Column one presents naive PPML regressions, column two describes the strength of the first stage (regressing IMA_{it} on IMA^H_{it} where IMA^H_{it} uses the hypothetical network to calculate internet market access), and column three presents the 2SLS results.

Table 4 Internet market access regression results

	Naive PPML	First-Stage	2SLS PPML
IMA (standardised)	0.234***		
	(0.109)		
N	1295		
Num sub-counties	185		

Notes: This table shows the results from the internet market access regressions. Column one shows the coefficient on standardised internet market access in a PPML regression against formal consumer service employment in a sample of rural and small town sub-counties. Column two shows the coefficient on the instrument for internet market access calculated using the hypothetical expansion network in the first-stage regression. Column three shows the 2SLS PPML regression results. In all cases standard errors are clustered at the sub-county level.

4.2.3 Non-homotheticities and wealth effects

Exogenous increases in local wealth from any channel can cause shifts to consumer service expenditure and therefore local consumer service employment if there are non-homotheticities in consumption. In this manner, observing increasing consumer service employment could reflect a symptom rather than a cause of growth.

In this sub-section, we show evidence of non-homotheticities in the Kenyan economy, whereby as incomes rise, individuals spend a decreasing proportion of their income on food. To this end, we estimate the Engel elasticity, which captures the responsiveness of the expenditure share on food to changes in total household expenditure. A higher absolute elasticity ($|\varepsilon|$) implies that their is scope for demand-side factors to play a role in the expansion of the CS sector: as income increases, the share of food expenditures declines more rapidly, while the share allocated to consumer services grows at a faster rate. Previous estimates of Engel elasticity suggest cross-country variation. For example, Fan et al. [2023] report estimates ranging from 0.33 to 0.39 for India, whereas Boppart [2014] finds values between 0.22 and 0.29 for the United States. These estimates provide a useful point of comparison for our study.

Following Fan et al. [2023], Engel elasticities are estimated using the following linear regression model, using data on food expenditure from the KIBHS (2015).

$$\ln \vartheta_F^h = \delta_r + \varepsilon \times \ln e_h + x_h' \psi + u_{rh}, \tag{4}$$

Where ϑ_F^h represents the food expenditure share of household h residing in county r, e_h denotes total household expenditure, δ_r captures county fixed effects, and x_h includes household characteristics such as household size and urban or rural residence. Estimates based on data from 2015 are reported in Table 5. The following specifications are considered: (1) A simple regression using the full sample. (2) Excluding the top and bottom 5% of income levels to mitigate the effects of potential misreporting. (3) Using $\ln(\vartheta_F^h - \beta_{\mathcal{F}})$ as the dependent variable, where $\beta_{\mathcal{F}} = 0.05$ proxies the asymptotic food share. This adjustment results in some missing observations due to negative values after subtracting 0.05, which are invalid for logarithmic transformation. The top and bottom 5% of income levels are also trimmed. (4) Estimating separate Engel elasticities for households above and below the median income level. The top and bottom 5% of income levels are trimmed. (5) Allowing the elasticity to vary between counties with high and low urbanization levels, where high-urbanization counties are defined as those in the top quartile of nightlight-based urbanization measures. The top and bottom 5% of income levels are trimmed. (6) Allowing elasticity differences across three broad location types: cities; large towns and metropolitan areas; small towns and rural areas. The top and bottom 5% of income levels are trimmed. (7) Focusing on small towns and rural areas and differentiating elasticity estimates between counties with high and low migration rates to urban centers. The top and bottom 5% of income levels are trimmed. (8) Conducting a pooled regression where the dependent variables are $\ln(\vartheta^h_F - \beta_{\mathcal{F}})$ for food expenditures and $\ln(\beta_{\mathcal{S}} - \vartheta^h_S)$ for service expenditures. Service expenditures exclude medical, education, and transportation-related spending. The top and bottom 5% of income levels are trimmed.

Table 5 presents the regression estimates, highlighting how Engel elasticity varies across different subsamples and estimation approaches. Figure A4 in the appendix plots the empirical Engel curve. Notably, Engel elasticity is more negative for households with higher incomes and for those residing in more urbanized areas, consistent with the hypothesis that demand-side factors play a significant role in the expansion of the CS sector.

Table 5 Estimating Engle Elasticities

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\ln e$	-0.172***	-0.175***	-0.199***					-0.146***
	(0.014)	(0.013)	(0.015)					(0.009)
$\ln e\times$ below median expenditure				-0.105***				
				(0.015)				
$\ln e\times$ above median expenditure				-0.238***				
				(0.021)				
$\ln e\times$ low urbanization					-0.166***			
					(0.015)			
$\ln e\times$ high urbanization					-0.191***			
					(0.023)			
$\ln e\times$ rural/small town						-0.170***		
						(0.014)		
$\ln e\times$ metro/large town						-0.256***		
						(0.011)		
$\ln e\times$ city						-0.111***		
						(0.012)		

Notes: This table shows the results from estimating Engle elasticities using equations of the form given in 4. Each column shows a separate regression each includes county fixed effects. This regression leverages data from the KIBHS in 2015.

This analysis shows strong evidence in favor of non-homotheticities. However, it does not show direct evidence for wealth-effects — that is we have not exploited variation in local wealth. That is, this can be see as necessary but certainly not sufficient evidence for the importance of the wealth-effect channel in explaining the observed patterns.

4.2.4 Fortuitous feedback loops

The stylised model above, and our simple intuition, implies that consumer-services have the specific property that demand and supply are local. This means, as formalised in the model, that they are suseptible to fortuitious feedback loops. This could be a powerful amplification force that could explain the patterns we observe. We can provide some indicative evidence of these forces by regressing formal consumption shares against future growth in formal employment.

Table 6 uses data on formal consumption shares from KIHBS survey 2015/16. We classify consumption into formal and informal by place of purchase. Formal consumption is that done in supermarkets, specialised shops, online, or establishments/ institutions. Informal consumption is that done in open markets, kiosks, general shops, hawkers, or from other households. This data is at the county by urban/ rural level giving us 79 localities which are mapped onto sub-county level classifications in the administrative data. Using this data table 6 shows that areas with higher formal consumption shares in 2015/16 saw larger growth in formal employment and in the number of formal firms between 2016 and 2022. Locations with a greater proportion of spending on formal goods see greater increases in formal employment.

An alternative story explaining fortuitous feedback loops could be one of overcoming fixed costs and therefore of increasing returns to scale. This explaination has recently been posited by Schwartzman [2025]. However we show that this is unlikely to be the case in the Kenyan context. Figure A3 in the appendix shows that unlike Schwartzman [2025] we do not find evidence that consumers switch discontinuously between formal and informal consumption suggesting that the fixed-cost story is less pertinent.

Table 6 Regressing 2016 formal consumption shares against subsequent employment and firm growth

LHS variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Growth: Overall formal emp	251.624*	163.918	211.631	64.305	47.553	35.748	25.323	28.884
	(143.503)	(145.779)	(155.191)	(132.719)	(91.646)	(29.481)	(24.217)	(26.082)
Growth: Overall formal firm	16.718*	10.568	13.090	4.124	2.966	1.926**	1.780***	1.555**
	(9.007)	(9.078)	(9.679)	(8.180)	(5.229)	(0.939)	(0.642)	(0.688)
Growth: Industry formal emp	1.791	6.327	9.589	3.531	2.179	6.498	5.021	7.844
	(12.490)	(12.943)	(13.809)	(13.085)	(12.450)	(13.168)	(12.961)	(13.941)
Growth: Industry formal firm	4.185*	2.652	3.287	1.108	0.837	0.542*	0.517**	0.453**
	(2.231)	(2.247)	(2.396)	(2.043)	(1.386)	(0.287)	(0.198)	(0.212)
Growth: Services formal emp	201.665*	126.455	158.319	50.428	37.398	22.343	20.328**	17.857*
	(110.744)	(111.671)	(119.037)	(101.763)	(70.306)	(15.071)	(8.537)	(9.165)
Growth: Services formal firm	12.066*	7.622	9.440	2.882	2.026	1.319**	1.202***	1.050**
	(6.537)	(6.591)	(7.028)	(5.916)	(3.683)	(0.629)	(0.420)	(0.450)
ln(population)		Y	Y	Y	Y	Y	Y	Y
Rural/Urban			Y			Y		Y
Urban Class				Y				
Urbanisation					Y			
w/o NBO/MSA						Y		
w/o City/Metro/LargeTown							Y	Y

Notes: This table shows the results from regressing formal employment shares in 2016 from the KIHBS survey against subsequent formal employment and firm number growth. The unit of analysis is county by urban/rural in the KIHBS data and matched sub-counties in the administrative data. Standard errors are robust. N=73 in each regression.

4.2.5 Alternative explanations

In the framework and empirical evidence above we have posited, and shown indicative evidence for, a number of key economic forces that could be driving the observed patterns. However, this does not necessarily rule out alternative explainations. In this sub-section we briefly discuss, and show indicative evidence against, some leading alternative explainations for the observed trends.

1. Migration

We use data from the Kenya County Household Survey (KCHS) 2021 to analyse migration patterns. Migration patterns are examined by distinguishing between recent migration and recent migration occurring after 2016. Recent migration is defined as movement from a previous residence to the current residence, restricting the sample to individuals aged 15 to 50 years at the time of migration. Recent migration after 2016 is similarly defined but further constrained to include only those whose movement occurred after 2016, ensuring that the cohort is comparable over time.

The process of in-migration is normalized by the weighted sample size of each county to facilitate comparability across regions. Employment growth is computed as the long difference in employment levels between 2016 and 2021, using the average employment in 2016–2017 minus the average employment in 2020–2021. Control variables include population size and urbanization, the latter measured through nighttime light intensity.

One primary concern in this analysis is the small sample size. In the processed sample, there are 5,593 observations for recent migration and 2,032 observations for migration occurring after 2016. These sample constraints may limit the precision of the estimates and the ability to detect statistically significant effects.

The main results are given in figure A1 in the appendix and indicate a non-significant positive correlation between migration and employment growth. After controlling for population size and urbanization, the coefficient remains close to zero and statistically non-significant, suggesting that the observed correlation is not robust to additional controls. This evidence suggests that migration is not likely to have caused the large shifts in consumer service employment that we find.

2. Ease of reporting: Tax offices

One concern could be that shifts into formal employment do not reflect any underlying economic forces, but rather changes in the cost of formalising (or of not formalising). To investigate this we use geolocated data on tax office locations over time and leverage variation in the roll out of such offices across time and space. Figure A2 in the appendix shows the relationship between the (log) distance to a tax office and the number of formal employments at the sub-county level across the four main employment categories. Figure A2 in the appendix shows no relationship between these variables. Table A5 in the appendix shows the robustness of this null result to various specifications.

3. Tourism

We look into whether the trends are driven by subcounties that are popular destinations for international and domestic tourism. We find close to no overlap between the high growth destinations and tourism.

5 Quantitative model

Work in progress.

6 Conclusion

Leveraging uniquely granular administrative and survey data, we have documented striking patterns of rapid shifts within Kenya's formal sector away from industry and towards services. These shifts however are heterogeneous across space with larger cities seeing a transformation towards business-facing services whereas more rural areas moving rapidly into consumer-facing services. We have shown that these shifts are correlated with urbanization and changes in local connectivity, productivity increases as proxied by internet access, and that preferences are characterised by non-homotheticities. The next step is to understand the causes rather than just the correlates of these changes, quantify welfare implications, and ask what role policy could play.

Appendix

A Additional information on survey data sets

- **Economic surveys** are published annually by the KNBS. They include national-level numbers for formal employment (by industry) and informal employment (by activity or urban/rural classification).

- **The 2015/16 Kenya Integrated Household Budget Survey (KIHBS)** was conducted by the KNBS from September 2015 to August 2016, sampling at the county level and covering 21,773 households. We summarize the formality, sector, and occupation of employment based on information about the main employer, ISIC code, and KNOCS code in the dataset.

- **The 2019 Kenya Population and Housing Census (KPHC)** was conducted in August 2019, sampling at the Enumeration Area (EA) level. We summarize the formality, sector, and occupation of employment based on information about the main employer and the type of main job.

- **The Kenya Continuous Household Survey (KCHS)** provided county-level statistics in 2020 and 2021, covering 19,701 and 16,945 households, respectively. We summarize the formality, sector, and occupation of employment based on information about the main employer, ISIC code, and KNOCS code in the dataset.

- **The Kenya Demographic and Health Survey (DHS)** was carried out in Kenya in 1989, 1993, 1998, 2003, 2008–09, 2014, and 2022. The labour-related statistics are at the county level in 2022 and at the region (province) level in 2003, 2008–09, and 2014. In 2022, 32,156 women aged 15–49 and 14,453 men were interviewed. We summarize the agricultural share and occupation share based on the documented occupation information.

Table A1 Data availability summary, 2015-2022

Data	Formality	Geography	Sector	Occupation	Time
Macro Data					
Economic surveys	Formal	National	✓	-	2015-2022
	Informal	Urban/Rural	✓	-	2015-2022
Micro Data					
Administrative Data	Formal	Subcounty	✓	-	2016-2022
KIHBS	Formal/Informal	County	✓	✓	2015
Census 2019	Formal/Informal	Subcounty	✓	✓	2019
KCHS	Formal/Informal	County	✓	✓	2020, 2021
DHS	-	Region	Agriculture	✓	(2003, 2008, 2014), 2022

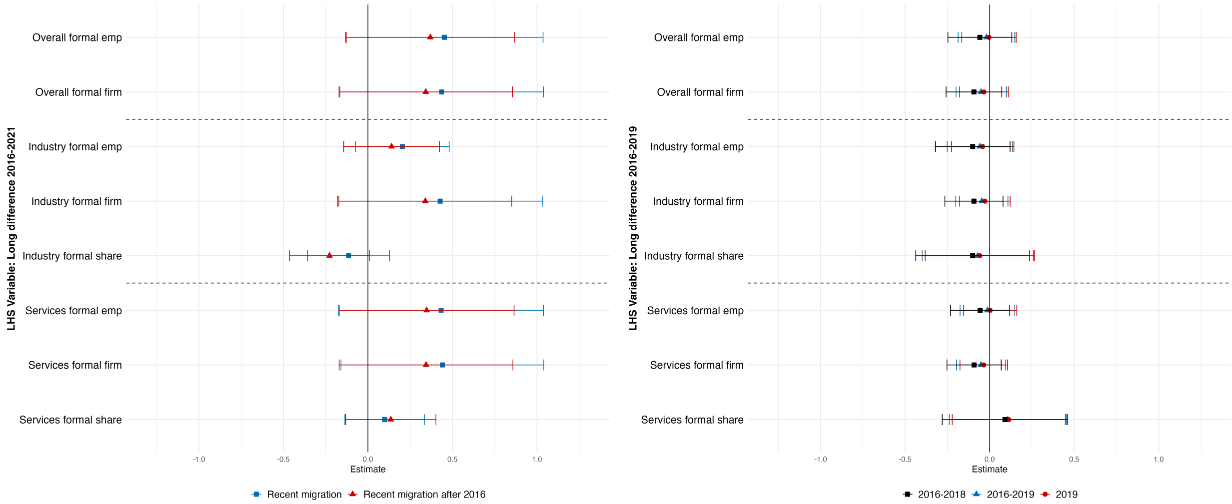
B Additional information on nighttime light data

There are two main data sources of nighttime lights: (i) the Defense Meteorology Satellite Program (DMSP) data, which covers the period from 1992 to 2003, and (ii) the Visible Infrared Imaging Radiometer Suite (VIIRS) data, which covers the period from 2012 to 2022. The VIIRS data are of higher quality and are widely used across various scientific disciplines, although they are not yet as commonly employed in economics [Gibson et al., 2021, Nechaev et al., 2021].

For our analysis, we use the annual VIIRS data for nighttime lights from 2016 to 2022, with a resolution of 464 meters [Elvidge et al., 2021]. We selected the band of average masked Day/Night Band (DNB) radiance values, where the unit of measurement is nanoWatts/sr/cm^2. To measure urbanization based on nighttime lights, we constructed two variables: (i) the average intensity of nighttime lights within a subcounty and (ii) the proportion of pixels with nighttime light intensity above a threshold of 3. This threshold is comparable to that used for DMSP data in other research [Harari, 2020].

C Additional figures and tables

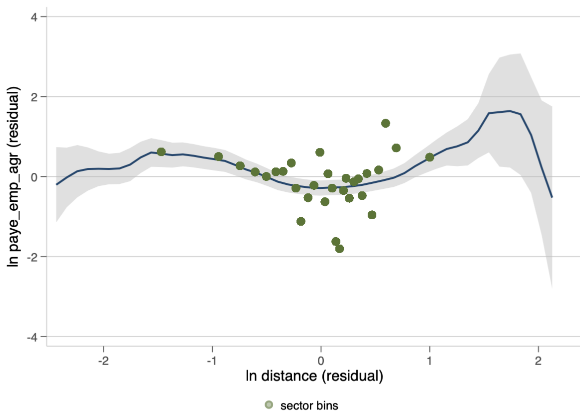
Figure A1 Migration regressions



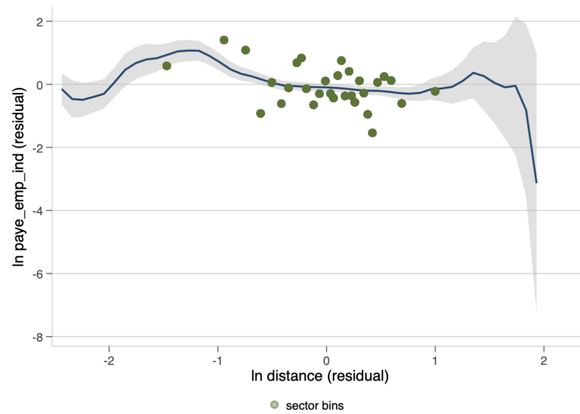
Notes: This figure shows the results from correlative regressions of migration rates against various outcome variables given on the y-axis. Each marker and associated 95% confidence intervals corresponds to a separate regression. Regressions in the left hand side panel include no controls, whereas regressions in the right hand side panel include controls for population size and urbanization. The unit of analysis are counties.

Figure A2 Residuals of ln(distances) and ln(sectoral employment)

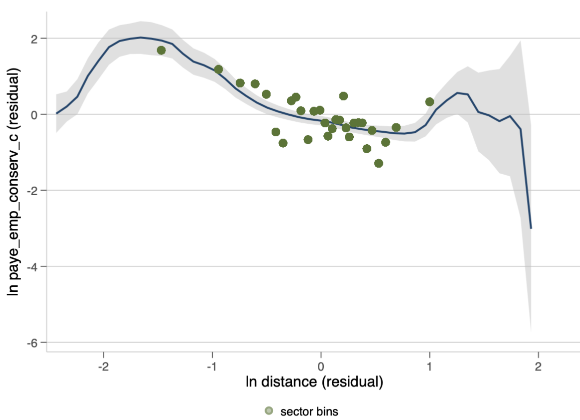
(a) agriculture



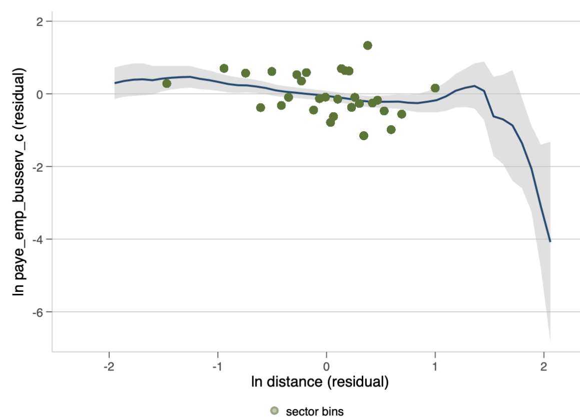
(b) industry



(c) consumer services

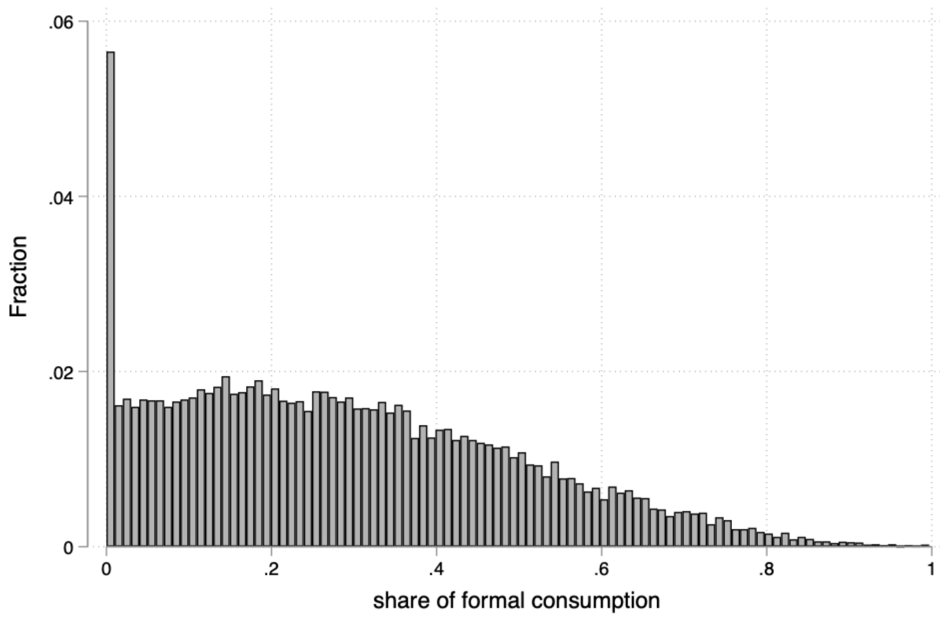


(d) business services



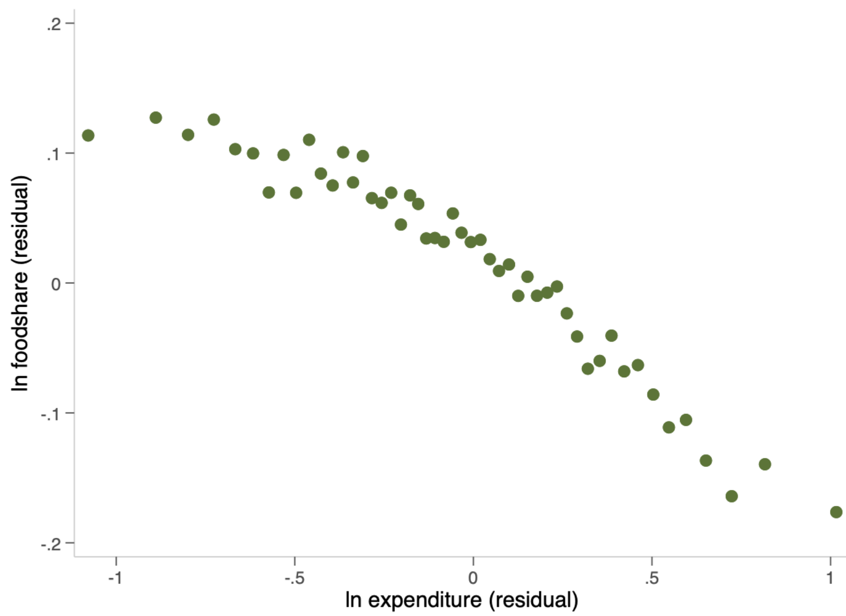
Notes: These figures show the non-parametric relationship between distance to tax offices and formal employment in agriculture (top right), industry (top left), consumer services employment (bottom left), and business services employment (bottom right).

Figure A3 Household level formal consumption shares



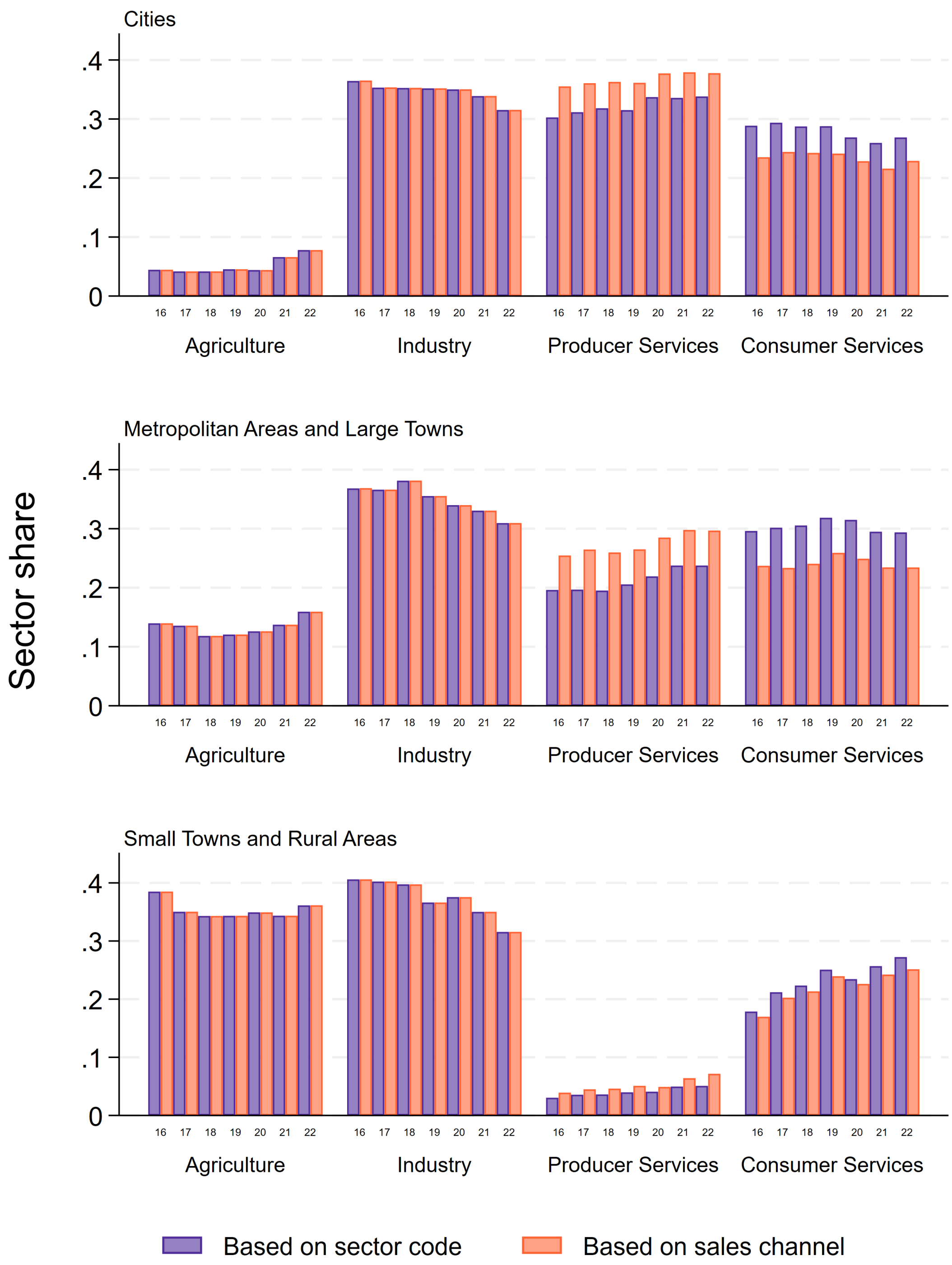
Notes: This figure shows the distribution of formal consumption shares across households. A value of 0 implies that no consumption is from sources labelled as formal.

Figure A4 Engle curve



Notes: This figure plots the empirical engle curve. Log-expenditure and log-food expenditure share are residualised against household characteristics. Data is from the 2015 KIBHS survey.

Figure A5 Sector shares of formal sector employment (by VAT-paying firms)



The above graph plots the share of formal employment in each sector for VAT-paying firms. For service sector firms we plot their classification into consumer and producer services based on their sales channels as well as their sector code.

Figure A6 Overall formal employment as a proportion of the working age population

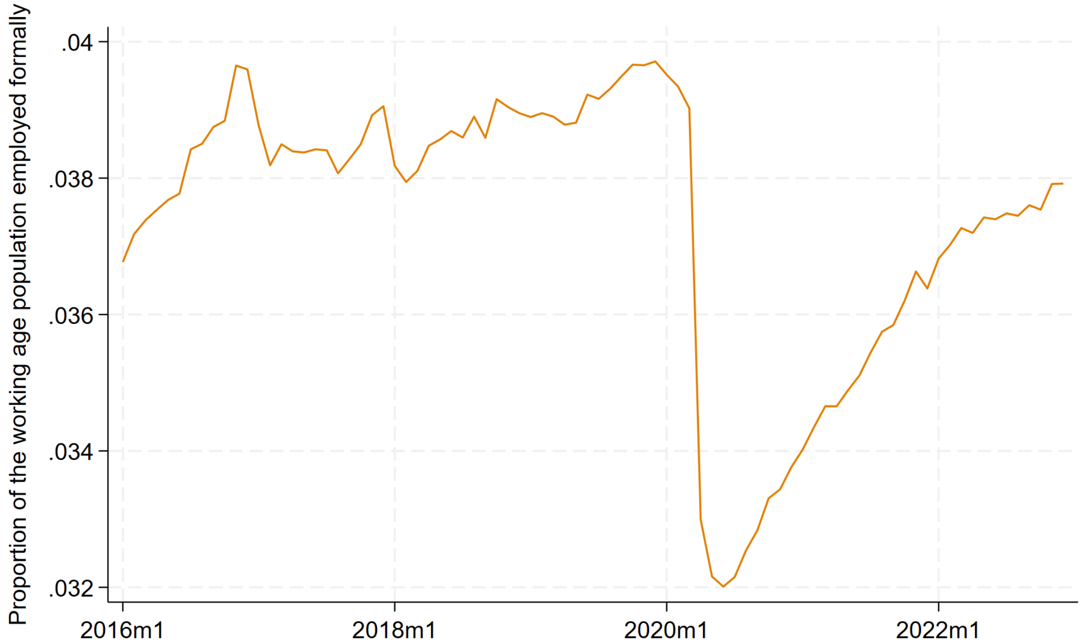
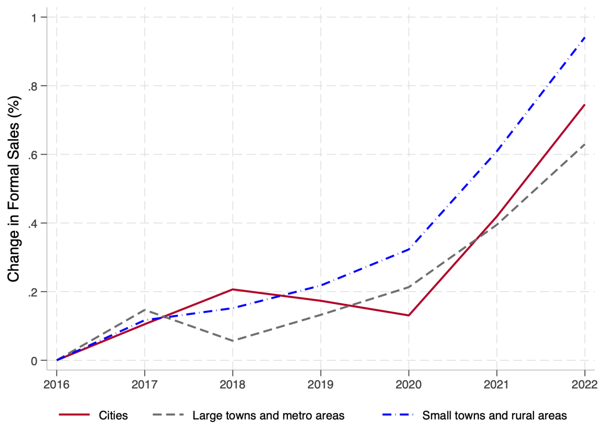
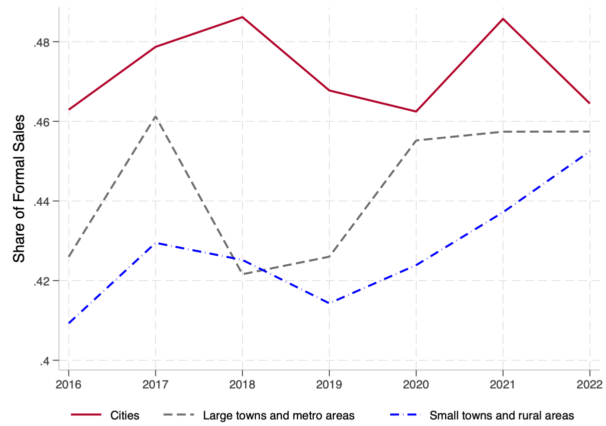


Figure A7 Change in formal sales by sector and geography

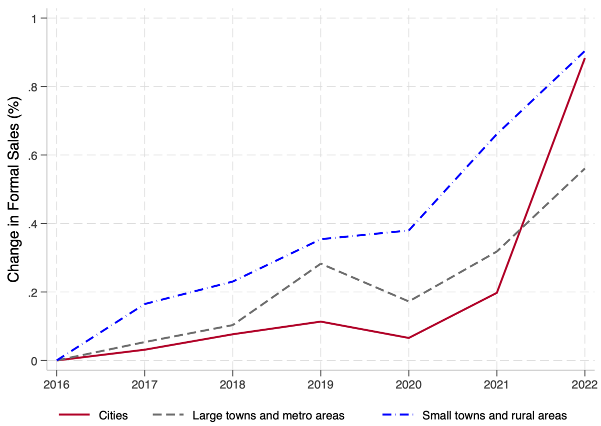
(a) Industry in Percent Change



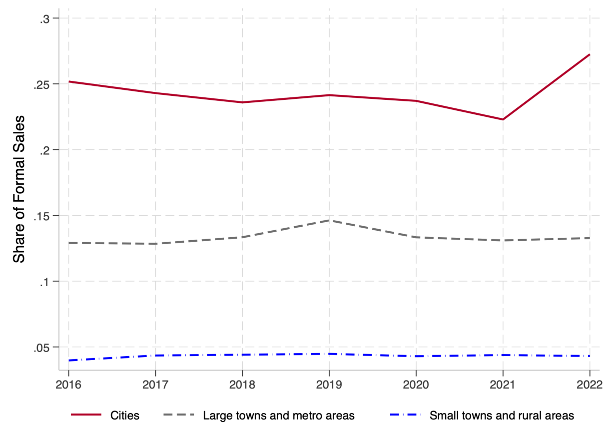
(b) Industry in Shares



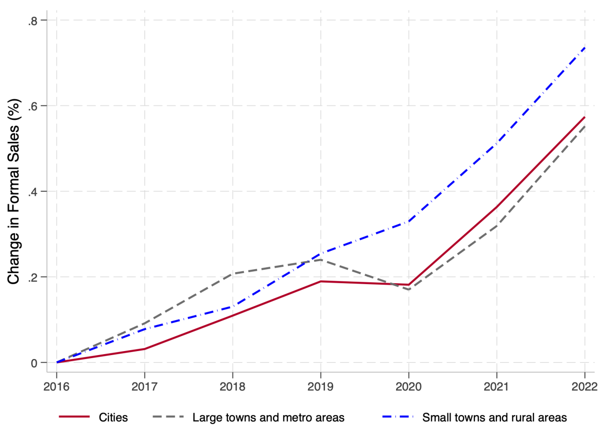
(c) Business Services in Percent Change



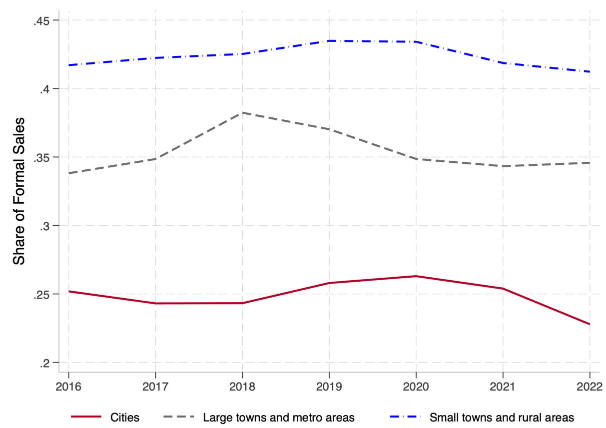
(d) Business Services in Shares



(e) Consumer Services in Percent Change



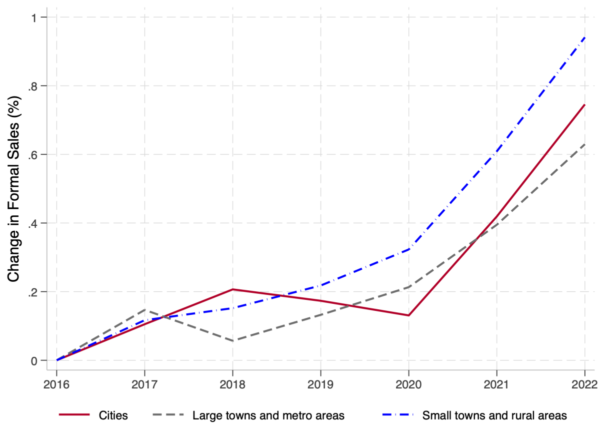
(f) Consumer Services in Shares



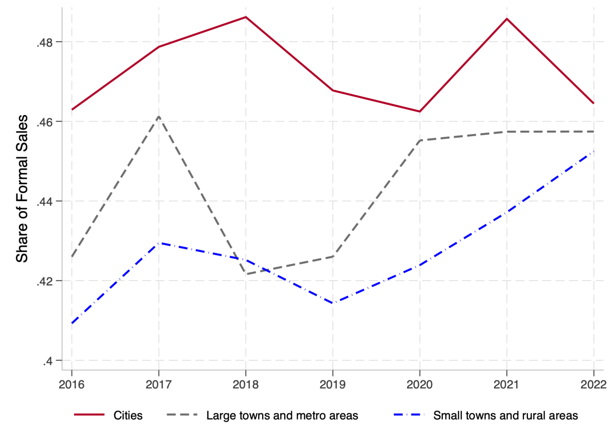
Notes: This figure shows the change in formal sales by sector and geography using administrative data from VAT records. The left-hand-side figures show percent changes (normalised to 0 in 2016) and the right-hand-side figures show changes in shares. Panels (a) and (b) show industrial sales, (c) and (d) business services sales, and (e) and (f) consumer services sales. Firms are classified based on who they sell products/services to. For all figures we display trends for three mutually exclusive geographies: Cities, large towns and metropolitan areas, and small towns and rural areas.

Figure A8 Change in formal sales by sector and geography (classify firms using sector codes)

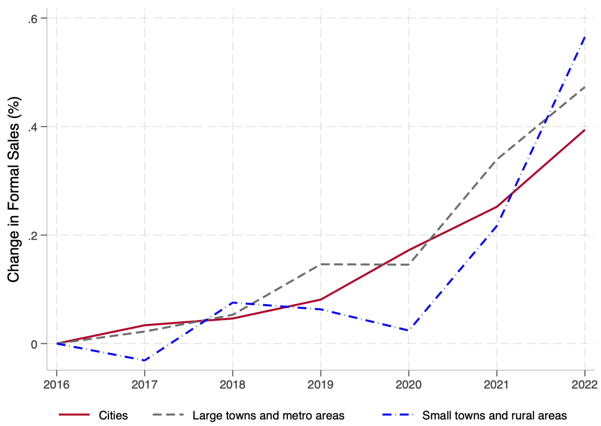
(a) Industry in Percent Change



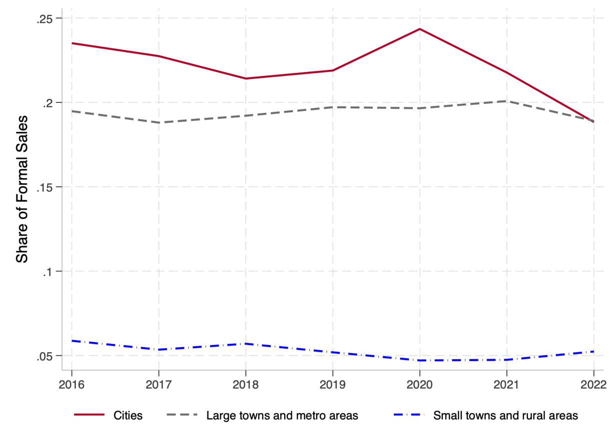
(b) Industry in Shares



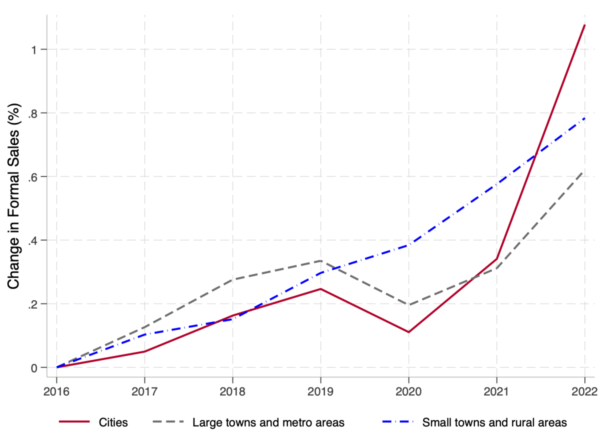
(c) Business Services in Percent Change



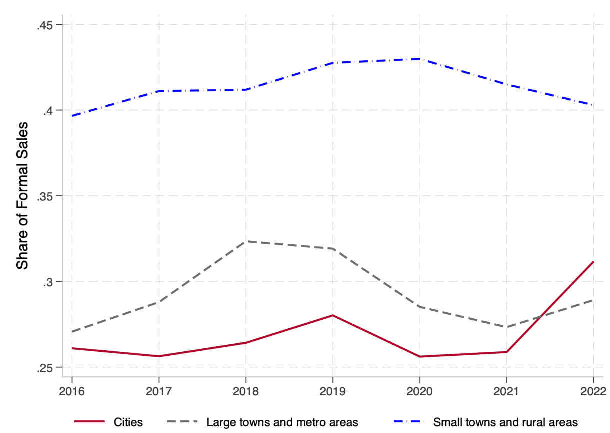
(d) Business Services in Shares



(e) Consumer Services in Percent Change



(f) Consumer Services in Shares



Notes: This figure shows the change in formal sales by sector and geography using administrative data from VAT records. The left-hand-side figures show percent changes (normalised to 0 in 2016) and the right-hand-side figures show changes in shares. Panels (a) and (b) show industrial sales, (c) and (d) business services sales, and (e) and (f) consumer services sales. Firms are classified based on sector codes. For all figures we display trends for three mutually exclusive geographies: Cities, large towns and metropolitan areas, and small towns and rural areas.

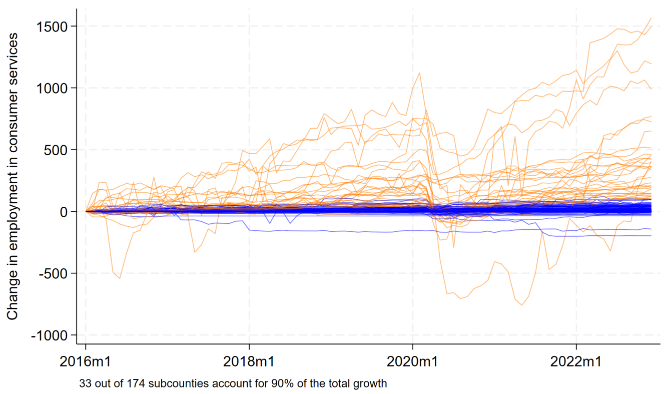
Table A2 Top 5 consumer service sectors in terms of employment shares by geography

Region	**ISIC**	**Description**	**Employment** (%)	**Firms** (%)
Cities	5610	Restaurants and mobile food services	0.15	0.07
	6190	Telecommunication, other	0.05	0.01
	4719	Retail non-specialized, other	0.05	0.01
	5510	Short term accommodation	0.05	0.02
	5590	Accomodation services, other	0.04	0.01
Metro area/larger towns	8610	Hospital activities	0.10	0.01
	5610	Restaurants and mobile food services	0.09	0.09
	5510	Short term accommodation	0.06	0.04
	4773	Retail specialized, other	0.06	0.03
	4719	Retail non-specialized, other	0.05	0.01
Small towns/rural, lower service growth	5510	Short term accommodation	0.21	0.07
	5610	Restaurants and mobile food services	0.10	0.11
	5590	Accomodation services, other	0.10	0.03
	4610	Wholesale on a fee or contract basis	0.09	0.05
	5630	Beverage serving activities	0.06	0.06
Small towns/rural, higher service growth	9601	Washing and dry cleaning	0.19	0.01
	5610	Restaurants and mobile food services	0.09	0.11
	5510	Short term accommodation	0.08	0.05
	4719	Retail non-specialized, other	0.06	0.01
	4711	Retail non-specialized, mainly food/beverages	0.06	0.02

Notes: The table summarised the five most important consumer service sectors for four different geographies: cities, metropolitan areas and large towns, small towns and rural areas with low service sector job growth and those with high service sector job growth. The last two columns report the share the 4-digit sector accounts for within the respective geography - in terms of employment and in terms of number of firms.

Figure A9 In which rural sub-counties are consumer services growing?

(a) Rural sub-county specific trends



(b) High-growth sub-counties with smaller towns or in rural areas

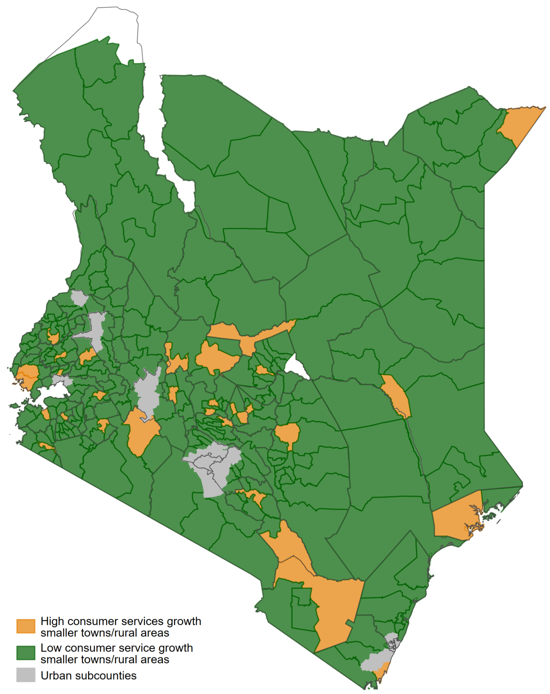
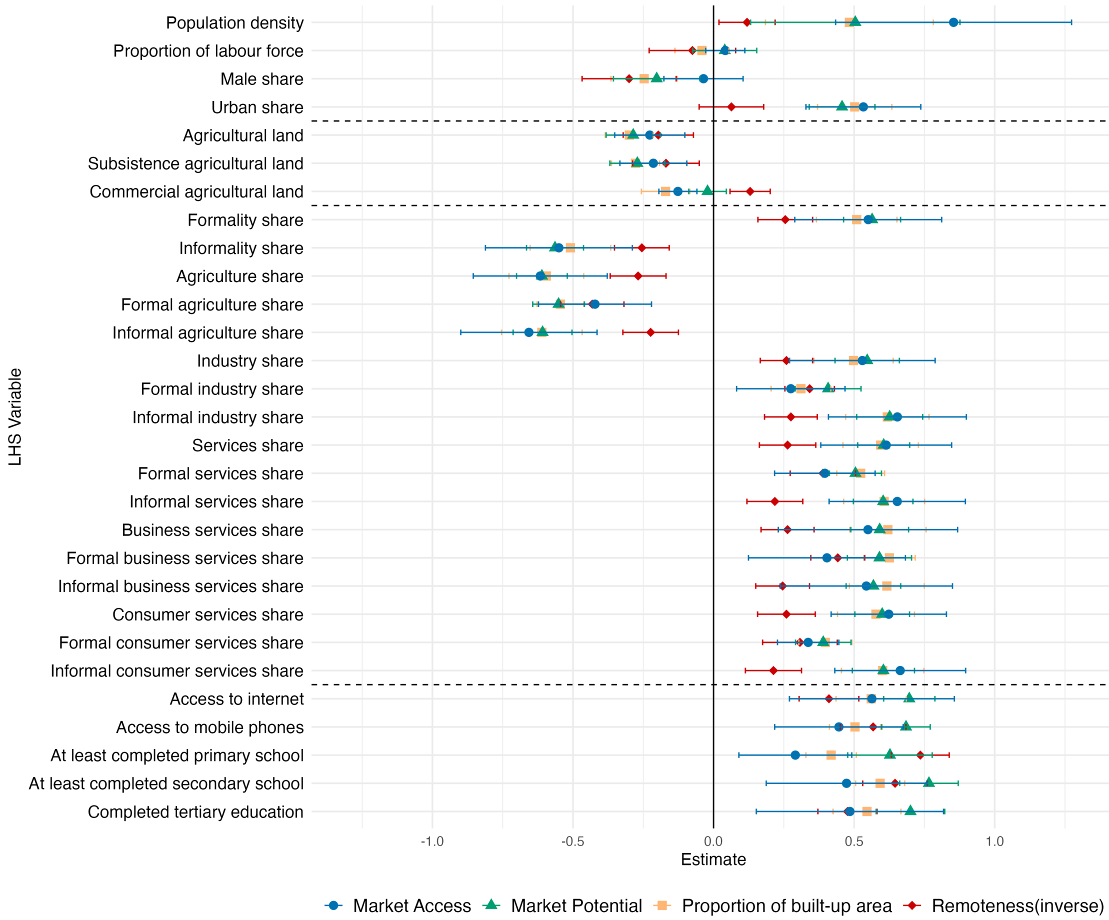


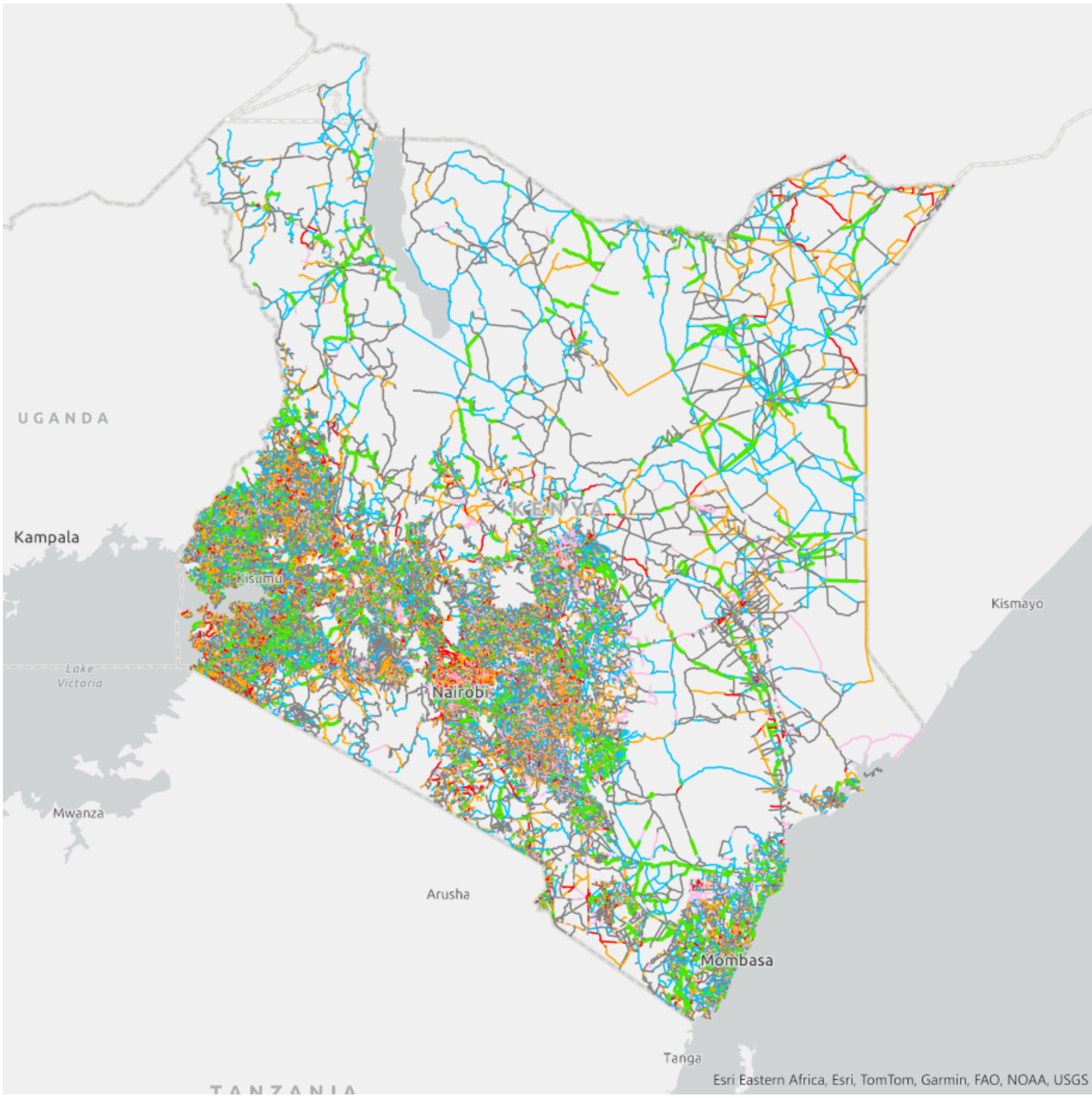
Figure A10 Correlations analysis for urbanisation and connectivity variables



Notes: This figure displays the results of correlation analysis for the proportion of built-up area and connectivity variables (market access, market potential, and the inverse of remoteness). For each main regressor (the proportion of built-up area or connectivity variables), we separately regress a standardized LHS variable on the standardized regressor. The magnitude of the estimated coefficient is reported by a point in the figure, with the 95% confidence interval computed using robust standard errors.

All variables are sub-county level statistics in 2019. The definitions and sources of data are as follows: (1) Urbanisation and connectivity variables: See section 2.5. (2) Sector shares in LHS variables: These were constructed using the 2019 Kenya Population and Housing Census (KPHC). The formality and sector are classified based on information about the main employer and ISIC code. (3) Other LHS variables: All from the 2019 KPHC, compiled by openAFRICA https://open.africa/dataset/2019-kenya-population-and-housing-census.

Figure A11 Road surface quality changes 2018-2023 RICS



Notes: Green implies improvements, red implies worsening.

Figure A12 Change log remoteness 2016-2022 OSM

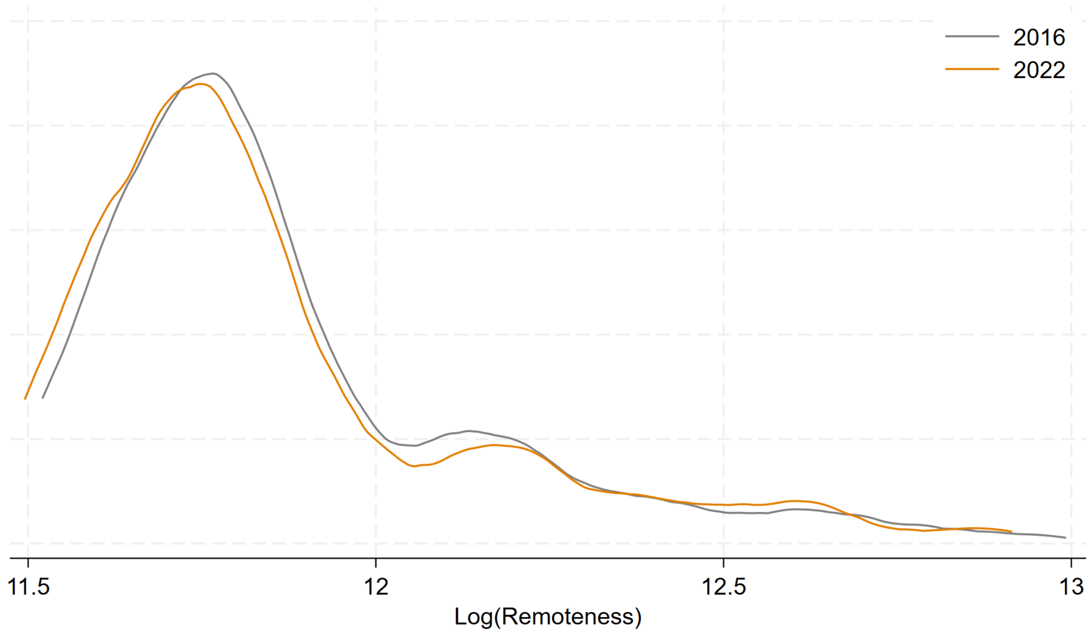


Figure A13 Formal pay over time

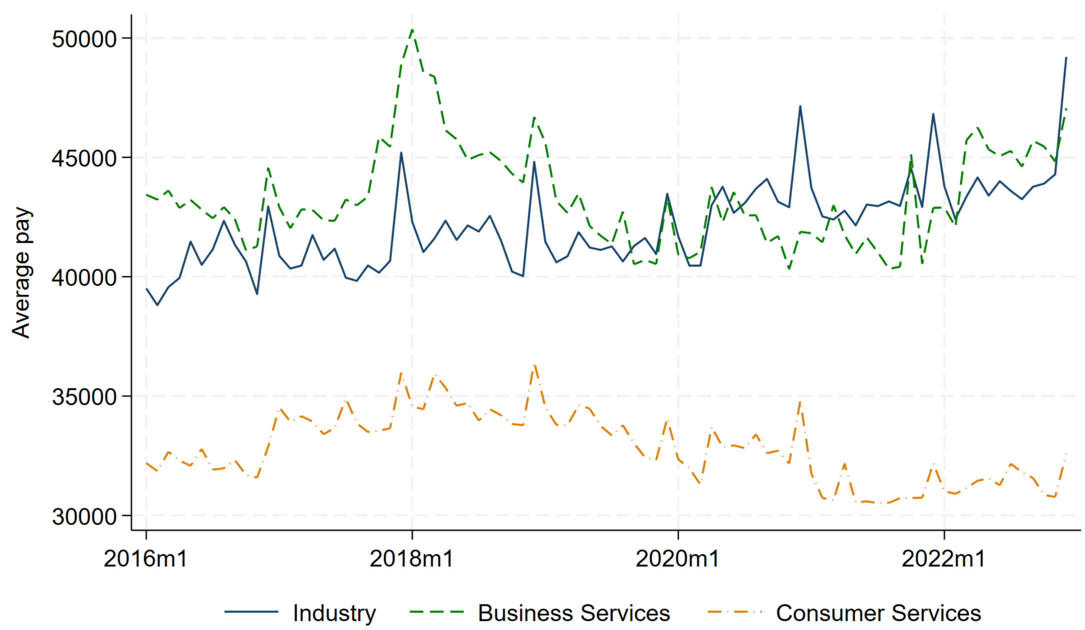


Figure A14 sales ratio over time

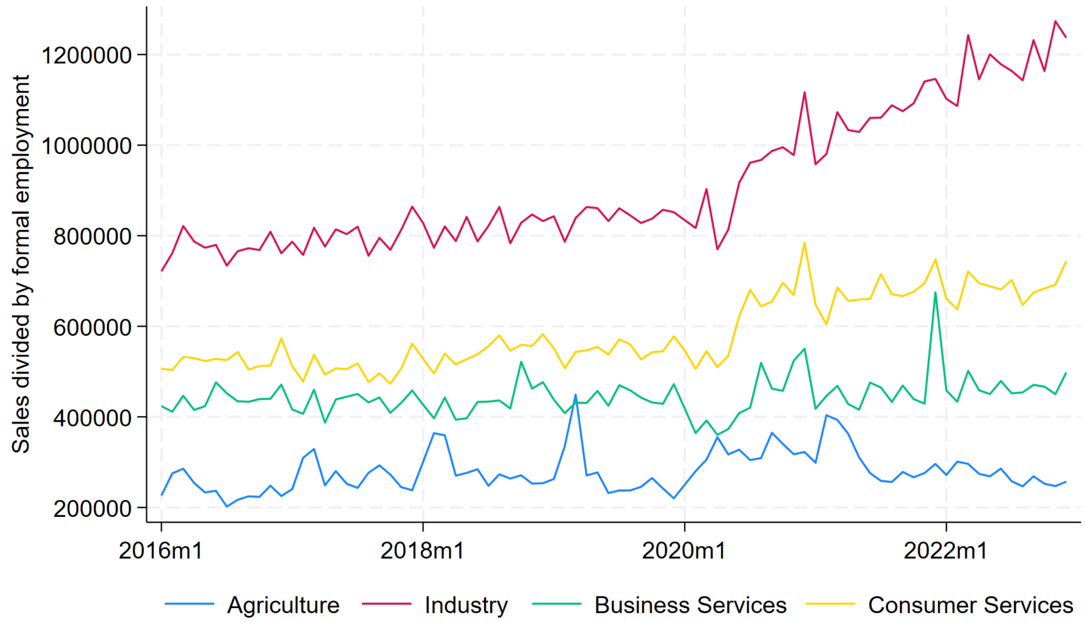


Table A3 Regressing number employed in business services against the proportion of built up area

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion built up	0.599	-0.116	-2.580	0.405	13.95	6.532***	14.00	-2.432
	(2.456)	(2.394)	(3.921)	(4.298)	(10.84)	(2.276)	(10.80)	(6.598)
Fixed effects	sCnty + Year	sCnty + Year	sCnty + Year	sCnty + Year	Year	sCnty		Cnty + Year
% built up cut off	50%	40%	60%	50%	50%	50%	50%	50%
Weighting				Labor force				
R2	0.988	0.988	0.986	0.988	0.0115	0.988	0.0114	0.134
N	1278	1218	1326	1278	1278	1278	1278	1248

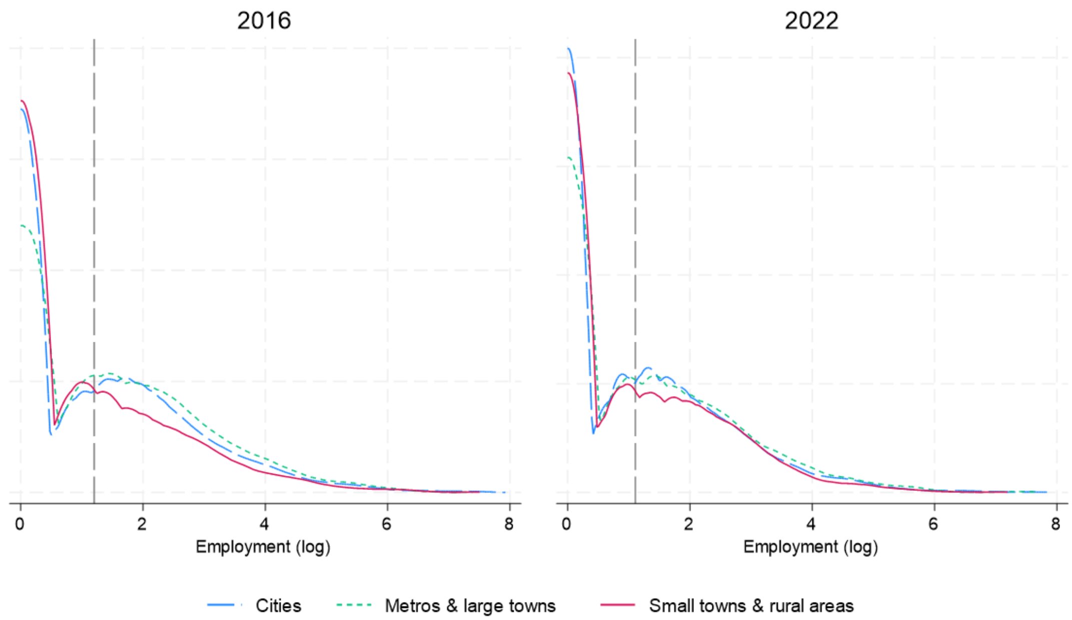
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table A4 Regressing number employed in industry against the proportion of built up area

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Proportion built up	-12.27	-12.04	-11.85	-15.94	19.28***	-11.90	19.07***	16.43
	(8.545)	(9.639)	(7.731)	(10.85)	(6.772)	(8.416)	(6.716)	(11.31)
Fixed effects	sCnty + Year	sCnty + Year	sCnty + Year	sCnty + Year	Year	sCnty		Cnty + Year
% built up cut off	50%	40%	60%	50%	50%	50%	50%	50%
Weighting				Labor force				
R2	0.984	0.982	0.989	0.986	0.0386	0.984	0.0378	0.162
N	1278	1218	1326	1278	1278	1278	1278	1248

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

Figure A15 Firm size distribution - formal employment - of consumer service firms



Notes: The figure compares the firm size distribution (in log formal employment) in 2016 and 2022 for consumer service firms in cities, metropolitan areas, and large towns, small town,s and rural areas. The vertical dashed line represents the mean.

Table A5 Regressing on the growth of employment in consumer services against distance to the nearest tax office

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
distances	-0.0374	16.04	-0.0389	15.37	30.25	-0.0824	-0.0981
	(0.0318)	(10.98)	(0.0317)	(11.06)	(18.85)	(0.113)	(0.112)
Year FE	-	-	✓	✓	✓	-	✓
sCnty FE	-	✓	-	✓	✓	-	-
Cnty FE	-	-	-	-	-	✓	✓
Weight					Pop		
R2	0.0301	0.210	0.0421	0.222	0.222	0.0349	0.0469
N	2004	2004	2004	2004	2004	2004	2004

Standard errors clustered at subcounty level. Urbanisation and population are included as control variables.
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D Micro founded model

In this section, we briefly sketch out a model of consumer behavior. This framework does not endogenise wages or prices across space but elucidates the main mechanisms that could lead to increasing consumer service employment.

Households in location i choose how to spend their time between (i) home production, (ii) cash-crop production, and (iii) wage work. They gain utility from home produced goods and market produced goods, however home-produced goods exhibit a satiation point. Thus, households face the following problem.

$$\max_{L_i^h, L_i^c, L_i^w} \left(\beta\left(\tilde{C}_i^h\right)^\rho + (1-\beta)\left(C_i^m\right)^\rho\right)^{1/\rho} \tag{5}$$

$$s.t. \tag{6}$$

$$\tilde{C}_i^h = \min\{C_i^h, \bar{C}_i^h\} \tag{7}$$

$$C_i^h = Y_i^h = A_i^h L_i^h \tag{8}$$

$$C_i^m = \frac{I_i}{P_i} = \frac{w_i L_i^w + p_i^c A_i^c L_i^c}{P_i} \tag{9}$$

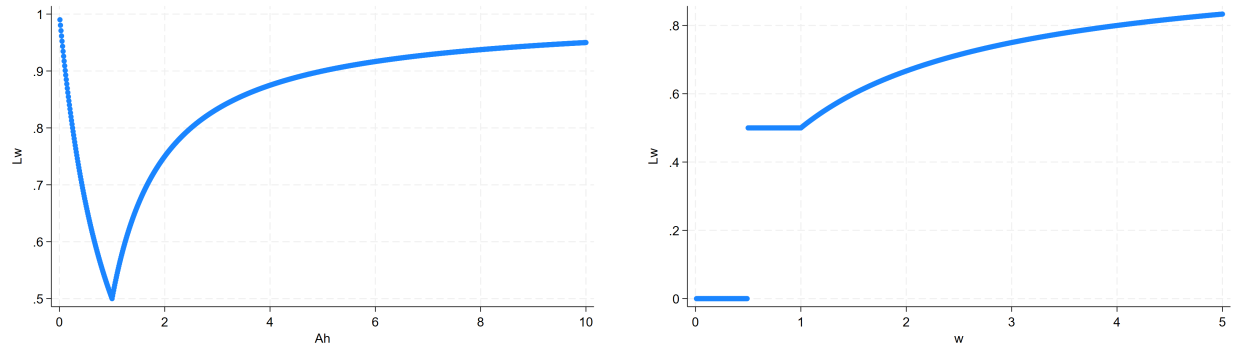
Note that in this set up households will never both produce cash crops and spend time on wage work, indeed if $w_i \geq p_i^c A_i^c$ households will only do wage work, and vice-versa. Noting that $L_i^w + L_i^c + L_i^h = 1$ this simplifies the problem.

We can solve the household problem to find the labor supply to wage work. Two conditions are important. The first is whether or not the household participates in wage work or cash-crop production. Denote this condition by $A_i = (w_i \geq p_i^c A_i^c)$. Second if the satiation point is binding the household will max-out on home production and spend the remainder on their time on wage work or cash-crop production. This occurs if $B_i = \left(\frac{\bar{C}_i^h}{A_i^h} \geq 1 - \left(1 + \left(\frac{\beta}{(1-\beta)}\right)^{1/(1-\rho)}\left(\frac{A_i^h}{w_i/P_i}\right)^{\rho/(1-\rho)}\right)^{-1}\right)$. Then we have the following.

$$L_i^w = \mathbb{1}_{[A_i]}\left(\mathbb{1}_{[B_i]}\left(1 - \frac{\bar{C}_i^h}{A_i^h}\right) + \mathbb{1}_{[B_i^c]}\left(1 + \left(\frac{\beta}{1-\beta}\right)^{1/(1-\rho)}\left(\frac{A_i^h}{w_i/P_i}\right)^{\rho/(1-\rho)}\right)^{-1}\right) \tag{10}$$

We can simulate this for a set of parameter values to see how labor supply adjusts to (real) wages increasing, and home production productivity increasing. Figure A16 shows the results. In the first panel we plot L_i^w against A_i^h. First as A_i^h rises home production is greater rewarded and so household dedicate more time to it at the expense of wage work. However, households quickly reach their satiation point. Beyond this increase home production productivity merely works to free up time that can be spent on wage work. Note that in this figure I assume $w_i \geq p_i^c A_i^c$ always. The right hand panel plots L_i^w against w_i. First, with low wages households prefer to engage in cash crop production. Beyond a certain wage they switch to wage work, but with wages so low they max out on more productive home production. As wages continue to rise they quickly no longer max out on home production and increasing switch away from home production and towards wage work.

Figure A16 Household behaviour simulations



Notes: This figure simulates the model and shows comparative statics for how employment in consumer services changes as home production productivity and wages increase.

This framework highlights the key channels through which employment in consumer services could increase. First, if home production becomes more productive we will see a marketisation of services leading to an increase in local service-sector employment. Second, higher consumer-service wages will incentivse more individuals into employment in consumer services. Consumer service wages will increase if the size of the local market increases or local consumer service sector productivity increases. Third, if incomes rise for exogenous reasons (increasing prices of cash-crops for example), the proportion of income spent on consumer services will increase due to implicit non-homotheticities. This will in turn lead to a more-than-proportional increase in local consumer service employment. Below we show some indicative empirical evidence for each of these channels.

References

Francesco Amodio, Elia Benveniste, Hoang Pham, and Marco Sanfilippo. The local (informal) multiplier of industrial jobs. *mimeo*, 2024.

Timo Boppart. Structural change and the kaldor facts in a growth model with relative price effects and non-gorman preferences. *Econometrica*, 82(6):2167–2196, 2014.

Natalie Chen, Dennis Novy, Carlo Perroni, and Horng Chern Wong. Urban-biased structural change. 2023.

Brenda Samaniego de la Parra and León Fernández Bujanda. Increasing the cost of informal employment: Evidence from mexico. *American Economic Journal: Applied Economics*, 16(1):377–411, 2024.

Klaus Desmet and Esteban Rossi-Hansberg. Spatial development. *American Economic Review*, 104(4):1211–1243, 2014.

Xinshen Diao, Mia Ellis, Margaret McMillan, and Dani Rodrik. Africa's Manufacturing Puzzle: Evidence from Tanzanian and Ethiopian Firms. *The World Bank Economic Review*, page lhae029, 08 2024. ISSN 0258-6770. doi: 10.1093/wber/lhae029. URL https://doi.org/10.1093/wber/lhae029.

Dave Donaldson. Railroads of the raj: Estimating the impact of transportation infrastructure. *American Economic Review*, 108(4-5):899–934, 2018.

Kevin Donovan and Todd Schoellman. The role of labour market frictions in structural transformation. *Oxford Development Studies*, 52(4):362–374, 2023.

Fabian Eckert and Michael Peters. Spatial structural change. Technical report, National Bureau of Economic Research, 2022.

Fabian Eckert, Sharat Ganapati, and Conor Walsh. Urban-biased growth: a macroeconomic analysis. Technical report, National Bureau of Economic Research, 2022.

Eliane El Badaoui, Eric Strobl, and Frank Walsh. The impact of internal migration on local labour markets in thailand. 2014.

Christopher D Elvidge, Mikhail Zhizhin, Tilottama Ghosh, Feng-Chi Hsu, and Jay Taneja. Annual time series of global viirs nighttime lights derived from monthly averages: 2012 to 2019. *Remote Sensing*, 13(5):922, 2021.

Tianyu Fan, Michael Peters, and Fabrizio Zilibotti. Growing like india—the unequal effects of service-led growth. *Econometrica*, 91(4):1457–1494, 2023.

John Gibson, Susan Olivia, Geua Boe-Gibson, and Chao Li. Which night lights data should we use in economics, and where? *Journal of Development Economics*, 149:102602, 2021.

Pinelopi Koujianou Goldberg and Tristan Reed. Presidential address: Demand-side constraints in development. the role of market size, trade, and (in) equality. *Econometrica*, 91(6):1915–1950, 2023.

Douglas Gollin, Remi Jedwab, and Dietrich Vollrath. Urbanization with and without industrialization. *Journal of Economic Growth*, 21:35–70, 2016.

Mariaflavia Harari. Cities in bad shape: Urban geometry in india. *American Economic Review*, 110(8):2377–2421, 2020.

John R Harris and Michael P Todaro. Migration, unemployment and development: a two-sector analysis. *The American economic review*, 60(1):126–142, 1970.

Chang-Tai Hsieh and Esteban Rossi-Hansberg. The industrial revolution in services. *Journal of Political Economy Macroeconomics*, 1(1):3–42, 2023.

Clement Imbert and Gabriel Ulyssea. Rural migrants and urban informality: Evidence from brazil. Technical report, CEPR Discussion Papers, 2023.

et al. Karra, Kontgis. Global land use/land cover with sentinel-2 and deep learning. *IGARSS 2021-2021 IEEE International Geoscience and Remote Sensing Symposium*, 2021.

Gaurav Nayyar, Mary Hallward-Driemeier, and Elwyn Davies. *At your service?: the promise of services-led development*. World Bank Publications, 2021.

Dmitry Nechaev, Mikhail Zhizhin, Alexey Poyda, Tilottama Ghosh, Feng-Chi Hsu, and Christopher Elvidge. Cross-sensor nighttime lights image calibration for dmsp/ols and snpp/viirs with residual u-net. *Remote Sensing*, 13(24):5026, 2021.

L. Rachel Ngai and Christopher A. Pissarides. Structural change in a multisector model of growth. *American Economic Review*, 97(1):429–443, 2007.

Rachel Ngai, Claudia Olivetti, and Barbara Petrongolo. Structural transformation over 150 years of women's and men's work. Technical report, 2022. Unpublished Working Paper.

Dani Rodrik. Unconditional convergence in manufacturing. *Quarterly Journal of Economics*, 128(1):165–204, 2013.

Dani Rodrik and Joseph E. Stiglitz. A new growth strategy for developing nations. 2024. URL https://drodrik.scholar.harvard.edu/publications/new-growth-strategy-developing-nations. January 2024.

Matthew Schwartzman. From street markets to shopping malls: The transformation of the service sector. *mimeo*, 2025.

Verena Wiedemann, Benard K Kirui, Vatsal Khandelwal, and Peter W Chacha. Spatial inequality and informality in kenya's firm network. 2024.