The development of the African system of cities

October 11th 2017

The Development of the African System of Cities

J. Vernon Henderson and Sebastian Kriticos

London School of Economics

Abstract

Sub-Saharan Africa has urbanised at tremendous speed over the last half century, in a process that has dramatically reshaped the economic and spatial profile of the region. Simultaneously, it has challenged much of the conventional empirical wisdom about how and why people move to cities. As we show in this article, the traditional view that countries urbanise alongside structural transformation is challenged in Africa, where urbanisation occurs despite low productivity in agriculture, very limited industrialisation, and a high share of primary sector employment across the urban hierarchy. There appear to be large household income gaps between urban and rural areas inducing migration, and these income premiums apply equally well to farm and non-farm families. Looking across the urban hierarchy, we also discuss how urban primacy can be problematic for economic growth in Africa, how secondary cities are lagging in industrial development, and how growth of employment in tradable services may signal a different path to structural transformation in Africa.

JEL classifications: J00, J31, O13, O14, O18, O55, R00, R11.

Keywords: Urbanisation, industrialisation, primacy, structural transformation, wage premium, Africa.

Acknowledgements: This work was supported by an Africa Research Program on Spatial Development of Cities at LSE and Oxford which is funded by the Multi Donor Trust Fund on Sustainable Urbanisation of the World Bank, and by the UK Department for International Development.

***Corresponding Author Email**: J.V.Henderson@lse.ac.uk

1. Introduction

Traditional economics has long viewed the process of urbanisation through the lens of structural transformation. There, the transition to modern economic growth is intrinsically linked to rural-urban migration through a declining share of employment in rural agriculture and a shift towards manufacturing and service industries in cities driven by urban-rural income differences (Rosenstein-Rodan, 1943; Nurske, 1953; Lewis, 1954; and Rostow, 1960). Urbanisation results from either or both productivity gains in agriculture which release rural labour and *push* employment towards cities, or the rise of industrial sectors – for example driven by enhanced international trade – which stand to gain from agglomeration economies and *pull* resources towards cities.

In recent years however, this convention in the literature has been challenged as research attention has turned towards Sub-Saharan Africa (hereafter Africa). Rather than the traditional *push* vs *pull* forces that drive urbanisation (agricultural revolution vs industrialisation), a growing literature following Collier et al (2009), Jedwab (2013) and Gollin, Jedwab and Vollrath (2016) instead argues that the income effects of natural resource exploitation have driven urbanisation without industrialisation across many African countries. Others suggest that rural deprivation has induced significant migration to African cities for reasons such as civil wars (Fay and Opal, 2000), deficient rural infrastructure (Collier et al, 2009), and climatic variability (Henderson, Storeygard and Deichmann, 2017). It has also been argued that natural population increase – through the combination of higher fertility rates and declining urban mortality – has played a leading role in driving urbanisation and concentration from within African cities (Jedwab et al, 2017; Jedwab and Vollrath, 2017).

These new investigations are motivated in part by facts about Africa's urbanisation and lack of structural transformation. First, contrary to the assumptions of standard structural change models, urbanisation in Africa has proceeded without the development of significant and

competitive manufacturing sectors; in fact, as Table 1 shows, the manufacturing sector has been in decline from an initially low level and has fallen to 8% today – far behind the levels achieved by other developing regions such as East Asia and Latin America. Second, while services contribute to a large portion of African GDP, we will see that there is a relative lack of development of tradable service employment, such as financial and business services.

Despite this, Africa is urbanising at tremendous speed. Since 1960, the level of urbanisation has risen from 15 percent to around 40 percent today and is projected to reach 60 percent by 2050 (UN Habitat 2010). African urbanisation parallels that of Asia as shown in Figure 1, noting that these are the only two remaining world regions which are under 60% urbanised. What is more, urbanisation in Africa is occurring at a much lower income level than in other continents historically. Just recently, Africa passed the 40% urbanisation level with an average GDP per capita of about $1,000, whereas Asia hit 40% in 1994 with a comparative GDP pc of $3,617, and Latin America in 1950 with a GDP pc of $1,860 (Lall, Henderson, and Venables, 2017). This rapid urbanisation at low income levels in Africa is typically accompanied by very high concentrations of people in primate cities, presenting the added issue that low incomes imply a general lack of institutional development and infrastructure investment needed for very large cities to function effectively, so as to harness the benefits of urbanisation.

Finally, while urbanising quickly, Africa has experienced practically negligible agricultural transformation in the last fifty years. In Figure 2, we plot the path of cereal yields across world regions since 1960. While the measure is a crude proxy for productivity, it is also one of the simplest and most internationally comparable. What we see is that relative to the rest of the world, Africa has not only performed consistently worse, but has shown limited evidence of catch-up with other regions since 1960. The continent does exhibit some sustained growth in yields after the 1990s, which is encouraging, but average productivity still remains vastly lower

than other emerging regions – particularly East Asia, which had almost 240 percent higher cereal yields than Africa in 2015.

In sum, Africa is urbanising without structural transformation based on either increased productivity in agriculture or development of an industrial sector. In this review, we will examine the new literature on Africa which attempts to sort out why urbanisation is occurring, as well as the relevant parts of the traditional literature in urban economics. We will develop facts from two databases to help us evaluate this literature and give a specific perspective on African urbanisation and its challenges. We will look at what people are doing in African cities and find that there is still a heavy role of agriculture in cities and little evidence of industrialization anywhere in the urban hierarchy. In reviewing traditional urban economics topics, we also analyse the high growth rate of primate compared to secondary and tertiary cities and the high urban-rural income gaps for families, which far exceed individual wage gaps, and apply equally well to agricultural versus non-agricultural families in cities.

2. Structural transformation

2.1 The literature

While the traditional focus on structural transformation between two sectors, urban and rural, remains characteristic of mainstream economics today, recent work considers urbanisation in the context of a continuum of locations with differing degrees of industrialisation, thus covering national geography at a much finer spatial scale (see Michaels, Rauch and Redding 2012 and the review in Demset and Henderson, 2015). That said, recent work still emphasises structural transformation as accompanying modern economic growth (Caselli and Coleman 2001; Herrendorf and Schoellman 2011).

In the case of closed economy models, current theories generate this transition as driven by technological improvements in agriculture, accompanied by limited income and price elasticities for farm products (Demset and Henderson, 2015). On the empirics, Gollin, Parente and Rogerson (2007) support this idea by presenting evidence that improved agricultural TFP is crucial to the differential timing of take-off across countries. One might argue however, that allowing for international trade as a source of food may diminish the need for improvement in agriculture as a precondition for urbanisation (Matsuyama, 1992; Glaeser, 2014). This seems unlikely though, as even in the modern world, food consumption in most countries is overwhelmingly supplied from domestic farming. For instance, Gollin, Parente and Rogerson (2007) demonstrate that among low-income countries, net food imports accounted for only 5% of total caloric consumption in the year 2000.

In most models the link between urbanisation and the changing structure of economic growth is only indirect: insofar as one equates agriculture with rural and industry with urban, without modelling an urban sector or cities per se. However, Lucas (2004) explicitly models rural-urban migration and the shift from a traditional technology (with no productivity growth) to a modern technology, by suggesting human capital has no productivity advantages in rural areas, but migrants invest in human capital in the urban sector because it is productivity enhancing. Henderson and Wang (2005) analyse a similar transition in a context where endogenous productivity growth through human capital formation is higher in urban than rural areas and the urban sector consists of an endogenous number of cities.

2.1.1 Urbanisation without industrialisation

Given the facts we cited in the introduction, with neither agricultural push nor industrial pull, why do people move to cities in Africa? In a key paper Gollin, Jedwab and Vollrath (2016) model cities growing in response to the increased consumption of non-tradables associated with

increases in resource income spending, in line with consumer city notions (Glaeser, Kolko, and Saiz, 2001). They specify a small, open economy model, where workers are either employed in food production, non-tradable goods, natural resources, or tradable goods. By assumption food production is entirely a rural activity, whereas tradable and nontradable goods are produced in cities. All resource production is sold internationally and has no domestic market. In the model, an exogenous increase in resource export earnings will raise all incomes and therefore increase demand for all types of goods. Crucially, purchases of additional food and tradable goods come from countries on the world market that have comparative advantage in these areas, while increased demand for urban nontradables generates an increase in labour in the urban sector. In sum, the model describes a channel through which resources drive a movement to cities and an increasing labour share in nontradable activities.[1]

One of the key motivating facts in the Gollin et al (2016) paper – which we replicated using data from the World Development Indicators – is to examine two pairwise graphical relationships. One relates the urbanisation level to the share of GDP in manufacturing and services, and the second the urbanisation level to the share of natural resource exports in GDP (see Figure A1 in the Appendix for the results). For Latin America and Asia, there is a clear positive association between the share of manufacturing and services in GDP and the overall level of urbanisation for the countries in these regions. In contrast, there is no such association between industrialisation and urbanisation in Africa. In Africa, urbanisation instead appears to be positively associated with the share of natural resource exports in GDP.

In line with this, within Africa, more urbanised countries are generally the largest exporters of natural resources such as Angola, Gabon, and Nigeria (oil), Botswana, Liberia, and South

[1] In a complementary earlier exercise, Matsuyama (1992) suggested that higher productivity in primary employment may well encourage undue emphasis in that sector, thus locking in a comparative advantage that actually delays industrialisation.

Africa (diamonds and gold), and Zambia (copper) and the Ivory Coast (cocoa). Hence, a major concern for Africa is that resource extraction tends to be heavily mechanised and can crowd out manufacturing by increasing factor prices (Sachs and Warner, 2001); moreover, if manufacturing firms exert positive productivity spillovers that resource firms do not, in principle we can have the so-called Dutch Disease effect (Sachs and Warner, 2001 and Ismail, 2011). That said, recent work by Alcott and Keniston (2014) on US counties argues resource booms can crowd out tradeable manufactures in the short term, but they have little overall impact on long term development.

The role of natural resources seems logical and perhaps applicable to some countries; but upon scrutiny, it seems unlikely to be the main story about Africa urbanisation. A limit to the graphical evidence that we discussed above is that it focuses on bivariate relationships between industrialisation and urbanisation and then between urbanisation and resource exploitation. Once we look at them together, these patterns of correlation are less clear. We estimate a simple regression where:

$$Urbanisation_i = \alpha_i + \beta_1 Industry_i + \beta_2 Resources_i + \delta_i + \delta_i * I + \delta_i * R + \varepsilon_i.$$

The dependent variable is the urbanisation level in country i in 2010. $Industry_i$ is the share of manufacturing and services value added in GDP in country i in 2010, and $Resources_i$ is the average share of natural resource exports in GDP between 1980 and 2010 in country i. δ_i are regional fixed effects, and $\delta_i * I$ and $\delta_i * R$ denote interactions between regional dummies and our covariates of interest. We cover the same 116 countries in Gollin et al (2016), across Latin America and the Caribbean (LAC, N = 26), the Middle-East and North Africa (MENA, N = 17), Sub-Saharan Africa (SSA, N = 46), and Asia (N = 27).

In Table 2, column (1) estimates relationships without any region fixed effects, and suggests that a percentage point increase in either the share of manufacturing and services in GDP or the share of natural resources in GDP is associated with approximately a 1.3-1.4 percentage point increase in urbanisation.[2] In column (2) we include the regional fixed effects and interaction terms, taking Asia as the base region. For Asia, the results show that both industry and resources are strongly, significantly and positively associated with urbanisation. For Africa, once we include interaction terms the net effects on manufacturing and services are small and insignificant, implying that within Africa, countries with more industry are no more urbanised than the region's average country. For natural resources however, there is a small net positive effect on urbanisation (0.66), which is just significant.[3] So any role of natural resource exports in African urbanisation is much lower than in Asia, North Africa and the Middle East. Finally, related to the population growth story as driving urbanisation, if we control for national population growth and regional interactions in Table 2 that results in insignificant coefficients with minimal effect on the other coefficients of interest.

We also looked at an earlier time period with urbanisation in 1985 as correlated with natural resource exports from 1960-1985. There are more missing values in the data and the sample size drops from 93 to 74, nevertheless, interestingly we found that the role of industry and resources switched for Africa. For manufacturing, there was a small net positive and significant effect, but natural resources had no significant effects on urbanisation in Africa, despite the fact that most resource-rich African countries experienced their major resource booms and fastest urban growth before the 1980s. Our conclusion is that neither natural resources nor the

[2] The results are robust to defining industry as the share of manufacturing in GDP rather than manufacturing and services.

[3] Wald tests on the joint significance of B2 and the interaction between industry and the SSA dummy variable is unable to reject the null hypothesis of no effect. For the case of natural resources, it returns a P-value of 0.041.

development of manufacturing are likely to be the main story driving urbanisation in Africa today. So what is?

To answer that we look at what people are doing in African cities versus the rest of the world. The examination will challenge the traditional notion applied to Africa that people move to cities to take up industrial or traded service sector jobs.

2.2 Employment composition across African cities

2.2.1 Africa versus other low income countries

What are the employment activities of individuals in rural and urban areas of Africa? To answer this and some later questions we put together a data set on African cities in 34 countries, discussed below. In this section, we examine the cities in 12 of these countries – for which we have sector-level data on city employment in the most recent census – to show the role of farming and other primary sector activity in cities and how that compares internationally. We note that, since generally over 95% of primary sector employment is in agriculture, we will refer to primary employment as agriculture based. From the most to least populous, the 12 countries are Ethiopia, Tanzania, Uganda, Mozambique, Ghana, Cameroon, Mali, Malawi, Zambia, Sierra Leone, Liberia, and Botswana.[4] As recorded in the relevant censuses we use, these countries have a population of around 219 million individuals: 60 million in urban areas and 159 million in rural areas. Today their populations stand at around 300m. Data are from International Integrated Public Use Microdata Series (IPUMS) (Minnesota Population Center, 2013).

A key feature of these IPUMS data is that they are geo-located at the district level and enumeration gives urban-rural residential status for each household (see the Appendix for

[4] We exclude data on Sudan and South Sudan from our sample due to data issues: the struggle with endemic civil war and their distinct nature which leads the UN to classify them as a North African countries.

details). Table 3 uses this to compare agricultural employment in our 12 African countries to 6 other developing countries around the world. In column (2), Africa's proportion of primary sector activity in urban areas post-2000 ranges from 12.7 – 39.9%, with a mean of 25.7%. Outside of Africa the range is between 4.1 – 14.4 %, with India at 7.4% and a mean of 10.6%. Africa has a clear pattern of a much greater role of farmers in cities. The Africa minimum in the 2000s is above the mean of the other countries and not far from their maximum. Column (3) shows the huge role of primary sector employment in the overall economies of Africa. In the 2000s, agriculture practically forms the entire basis of employment in the rural economy, with 64.5% of working individuals in that sector. This makes Africa an international outlier.

2.2.2 Within the urban hierarchy

We now look within the urban hierarchy to show that outside of the primate city, agriculture plays an even greater role in cities. To do this, we need to define cities and their boundaries. For boundaries of urban areas, we use night lights data;[5] and similar to Storeygard (2016) and Henderson, Storeygard and Deichmann (2017), we define cities as the envelope of the contiguously lit areas in recent years. Population data on cities and urban areas themselves are from Citypopulation.de (which are census based) and the urban area population is the sum of population of all recorded towns and cities in the night lights area. For the cities in these 12 countries, we only keep those urban areas that had at least 2000 inhabitants in the earliest period available for which we have a population number. For the 12 countries above, we have detailed employment breakdown for cities between 1990 and 2010, with typical intervals between each census of around 9-12 years. Further details about methodology, sources, and data availability can be found in the Online Appendix.

[5] The data come from the U.S. Defense Meteorological Satellite Program (DMSP) which has been in operation since the mid-1960s, with a digital archive of global light recordings beginning in 1992. The final dataset excludes all visible light emitted by confounders such as sunlight, moonlight, forest fires, and gas flares.

For the 12 countries, in total, we have 615 cities, separated into four different categories by relative size. We prefer a relative size division because it allows better for heterogeneity in countries and typical city sizes. The primate city is the most populous and always the national capital (or de facto capital) with a mean of 1.8 million in the sample; secondary cities are in the top quartile (excluding the primate) with a mean population of 170,000; tertiary are in the second quartile with a mean of 45,000; and finally, small cities are all urban areas below the median city size with a mean of 19,000.

We ask how the employment composition differs across types of cities to gauge the degree of industrialisation and the role of agriculture as we move through the urban hierarchy. As can be seen in Figure 3, the agricultural sector is lowest in the primate city as one might expect, and then steadily increases as we move down the urban hierarchy.[6] What is astonishing is the extent of agricultural employment amongst urban residents in the second quartile and bottom half of cities by size, at about 40%. Perhaps more troubling is the small role for manufacturing and tradeable services (finance, insurance, real estate and business services) in primate and secondary cities, at around 10-12%, despite the agglomeration opportunities – not to mention their miniscule role in other cities. The chief purpose of cities in the bottom 75 percentiles and even in the upper 25 outside the primate city appears to be housing and servicing ('trade' sector) farmers in the cities and surrounding areas.

Why do so many farmers choose these cities over rural areas to live? We will explore how employment opportunities for household members expand in cities compared to the rural sector, even for farm families. There are also consumption benefits, especially given the

[6] We also looked at the urban-rural split: Figure A2 in the Online Appendix provides the results

general complete lack of service activity in the rural sector and the much better access in urban areas to schools and health care. [7]

2.3 Structural change in cities in Africa

A key question is whether, given the high levels of agriculture in the rural sector and parts of the urban sector, is there change over time? Unfortunately, the only countries in our sample with more than one census period with the required breakdowns on employment are Ghana, Malawi, Mali, Mozambique, Tanzania, and Zambia. For these six, we explore how employment has changed between the 1990s and 2000s census periods.

Figure 4 graphs the percentage changes in the total share of employment by industry for urban and rural areas in the six countries. On the whole, the data suggest that the sample experienced some labour reallocations across agriculture, manufacturing, and services, which may be suggestive of structural change taking place. The rural sector is moving from its inordinate high share of farming to have more rural industry and services, although still well behind other parts of the world. In cities farming is declining but industry is stagnant, and the highest gains have been in exportable services, although this is growth from a very low base.

The question is the extent to which such reallocations are growth-enhancing, an issue in the literature. For instance, McMillan, Rodrik and Verduzco-Gallo (2011) study labour productivity growth between 1990 and 2010 for the major world regions. The authors decompose labour productivity growth into within sector growth – from capital accumulation, technological change, or reductions of misallocation – and growth across sectors as employment moves from low-to high-productivity industries. They argue that in high income

[7] However, we note a data issue in defining urban. It is common, particularly amongst farmers in Africa, to take up multiple jobs over the course of a year, particularly before harvesting periods and in off-seasons. Farmers may choose to out-commute from cities or take-up alternative income opportunities in urban areas, despite spending significant amounts of time in rural settings.

countries with supportive institutions, labour reallocations are unlikely to have large effects on economy wide productivity, but in developing nations, there is the prospect that labour reallocations move workers into lower productivity activities such as non-tradable services and other informal sector activity.

Their results show that before 2000, labour in Africa generally moved from high to low-productivity activities, thus reducing overall growth; but in the post-2000 period, structural change tended to be growth enhancing, contributing around 1.4 percentage points to annual labour productivity growth. Amongst individual countries in their data, Nigeria and Zambia exhibited expansions of manufacturing after 2000 and a contraction of agriculture and services – although these are small as compared to Asia. On the other hand, Kenya, Ghana and Senegal had structural change primarily driven by expansion of services.

Similarly, Block (2013) argue that Africa's recent growth reduction has been in line with traditional structural transformation; noticing substantial declines in the share of labour force engaged in agriculture, particularly among rural females over the age of 25 with primary education. The authors estimate greater gains in agricultural TFP since the mid-1970s than we implied in the cereals yield figure earlier, largely reflecting changes in land and labour quality as well as technology adoption. In the industrial sector, Johnson, Ostry and Subramanian (2007) argue that several small African countries have started down the path to the kind of institutional development that many successful East Asian countries had in the 1960s and 1980s. Their view is that Africa's institutional weaknesses could be ameliorated if they take a similar development route to Asia: building a stronger and more dynamic manufacturing export sector which will demand institutional reform.

2.4 Agricultural productivity gap

Although labour appears to be moving out of agriculture in Africa, the large proportion of agricultural employment in the cross-section is still a concern. Large productivity differences between agriculture and non-agriculture are a lost opportunity and have been cited by many as a major source of income differences between developed and developing countries (Gollin, Lagakos and Waugh. 2014; Caselli, 2005). Most notably, Gollin et al (2014) study whether large measured differences in value added per worker in the typical country are merely an artefact of mismeasurement in national accounting practises and other issues, as opposed to real differences in output per worker. Taking into consideration improved measures of inputs, outputs, and value added across sectors – such as accounting for differences in hours worked, human capital, cost of living, and capital intensity – the authors show that the agricultural productivity gap (APG) is substantially lower than national accounts data portray, but even so, output per worker in non-agriculture is roughly twice as high as in agriculture in the typical country, and even higher in developing countries. The implication is that there should be large income gains from workers moving out of agriculture and into other economic activities.

The puzzle in Africa is that, while the continent is urbanising rapidly, urbanisation is not synonymous with a shift from low-productivity agriculture to higher-productivity non-agricultural employment, given the prominence of farmers in cities. Still, an important question is whether urbanisation is income enhancing.

3. Urban Household Income and Wage Premiums in Africa

We now turn to the literature on urban scale economies, which is at the heart of why cities exist and people urbanise. We will compare African wage premiums to those in other countries. Then we will analyse the advantages in Africa of living in cities, the fuzzy line between urban and rural, and the weakness of secondary cities (in the top quartile) in Africa.

3.1 The advantages and disadvantages of cities

Why people live in cities is a fundamental question in urban economics, starting from Marshall (1890) with a literature on denser input–output linkages between buyers and suppliers, better matching of workers and firms in tight labour markets, and localised information and knowledge spillovers (see reviews in Duranton and Puga, 2004; Rosenthal and Strange, 2004; and Combes and Gobillon, 2015). These underlying agglomeration benefits are generally reflected in urban wage premiums (Glaeser and Maré, 2001; Combes, Duranton and Gobillom, 2008; D'Costa and Overman, 2014). However, such premiums may also exist because cities have favourable local endowments of institutions, public capital, or geographical amenities (Combes, Duranton, Gobillon and Roux, 2010), and most critically, because of selection of more productive workers into bigger cities either because they offer greater agglomeration benefits for higher skill workers or because larger cities are more specialised in more skill-intensive industries.

In a paper by Combes et al, (2008) the authors attempt to distinguish between these forces of spatial sorting and agglomeration by exploiting panel data with individual worker fixed effects. Likewise, Young (2014) claims to control for unobservable human capital and concludes that urban premiums simply reflect the efficient allocation and sorting of individuals in response to regional demand for labour skills. In contrast however, De la Rocha and Puga (2017) have argued that the use of worker fixed effects in a static framework understates the role of agglomeration benefits and the interaction between ability and dynamic learning effects of bigger cities. Following individuals over their lifetime, De la Rocha and Puga (2017) study the dynamic advantages of living in bigger cities, as does Wang (2016). Their argument is that OLS estimates of wage gains in a static framework give pretty good reduced form estimates of the overall gains to agglomeration.

Dealing with issues of sorting, selection, and dynamics requires much better data than that available for Africa. Nevertheless, we explore the issue of why agricultural households move to cities by providing estimates of urban-rural household income and individual wage differentials for three African countries – Nigeria, Tanzania, and Uganda. This work builds upon Jones, D'Aoust, and Bernand (2017) who study urban wage premiums in these countries with household survey data from the World Bank's Living Standards Measurement Study (LSMS). We will utilise OLS results, based on the earlier argument that in net, with key biases considered, OLS estimates do a reasonable job of capturing agglomeration benefits (De la Rocha and Puga, 2017), and the fact that with worker fixed effects, identification becomes solely based on "movers" which is particularly problematic given the very short panel nature of our available data.

3.2 Data

The LSMS surveys have detailed and consistent data at the household and individual level on income, education, labour allocation, asset ownership and dwelling characteristics, as well as urban-rural identifiers and enumeration areas, so we can map data to specific urban areas for which we have population sizes. The data sets are the Tanzania Panel Household Survey (2008 and 2010), the Nigeria National Household Survey (2010 and 2012), and the Uganda National Panel Survey (2009, 2010, 2011 and 2012). Note these dates are so close together in each country that they do not give the opportunity to look at dynamics. These sample countries account for around 25 percent of the entire continent's population and feature a variety of income levels, with Uganda below $700 PPP income per capita, Tanzania around $1,000, and Nigeria slightly above $6,000.

3.3 Urban Premiums

3.3.1 Wage premiums and costs for individuals

First we do the traditional exercise of estimating urban wage premiums for individual workers, which is viewed as a productivity factor. We pool survey years and estimate a general specification:

$$w_{it} = \alpha_i + \beta_1 Urban_{it} + \beta_2\, Primate_{it} + x'_{it}\boldsymbol{\beta} + d_t + \varepsilon_{it}.$$

$Urban_{it}$ is an indicator taking the value of one if the individual or household i is living in an urban area at time t outside the primate city; $Primate_{it}$ is for a worker in the primate city. We will also break urban into different relative size categories below the primate. x'_{it} is a vector of time varying control variables representing factors such as household demographics and employment conditions. d_t is year dummy. ε_{it} is the error term including unobservable characteristics of individuals and of the communities they live in.

For analysis at the individual level we restrict our sample to the set of individuals in households that do not own agricultural land, because we have no way to assign farm income to individuals. We look at adults aged 18-65 who are working part or full time with income. Table 4 presents the results with primate city and all other urban distinguished from rural. The controls are footnoted in the table but cover basic characteristics of workers like age and education and whether an individual reports hours worked and if so how many.[8]

Columns (2), (4), and (6) add to the respective columns (1), (3) and (5), occupation fixed effects, from which there is generally but not always a decrease in the urban and primate city premiums. The question is whether occupational controls are appropriate. Occupational opportunities are greater in bigger cities which itself is an agglomeration benefit, so people can more easily shift into higher paying occupations; controlling for occupation then removes this effect. On the other hand, it may be that people have set occupations, but higher level ones are

[8] Restricting the sample to full time workers as in Jones et al (2017) gives similar results.

found disproportionately in bigger cities. And occupation fixed effects may help control for some unobserved individual characteristics and thus sorting effects.

In Table 4, we find significant premiums, ranging from 0.11 to 0.89 for the primate city and 0.08 to 0.234 for urban in general. However, the numbers for Nigeria are particularly small with a range of 7-14% in the four cells. How do these compare internationally? In France and Spain the primate city premium over rural areas has been estimated as 55 percent (De la Rocha and Puga, 2017) and 60 percent (Combes et al, 2008) respectively. When comparing urban and rural areas, the urban premium is estimated to be around 23-32 percent in the USA (Glaeser and Maré) and around 9 percent in the UK (D'Costa and Overman, 2014). For the three countries, African numbers are in line with the range in other countries.

One usual procedure when data are available is to then compare cost-of-living differentials between urban and rural, to see if wage premiums are more than sufficient to compensate for cost-of-living and amenity differentials (see Roback, 1982; Duranton and Puga, 2004; and Combes, Duranton and Gobillon, 2016). Of course, a true spatial equilibrium analysis à la Roback would consider evaluation of all consumer and producer amenities, breaking out the demand and supply side of city residents and looking at compensating differentials whilst worrying about perfect population mobility assumptions which underlie the interpretation of differentials (Demset and Rossi-Hansberg, 2016). Higher wages generated by productivity differentials in urban areas then also offset higher costs-of-living in urban areas and could offset lower non-monetary amenities.

Our data on African cities does not offer enough to do the more full blown analysis that is done on some countries. However, a recent paper by Gollin, Kirchberger and Lakagos (2017) uses data on health, housing quality, crime and pollution for 20 developing African countries, and shows that for almost all measures in all countries, amenities are either constant or increasing

with population density.[9] Hence, their results suggest that African countries are reallocating workers to densely populated areas which, on average, offer both higher wages and amenities.

Here we simplify and focus on cost-of-living differentials as driven by housing costs, with a crude evaluation of real income differentials. Unfortunately only for Uganda, do we have sufficient data on house rents; in the other two samples, almost all rural households report as owner occupiers. To investigate, we ran a hedonic regression of rents for households, controlling for the number of habitable rooms, the building materials, and access to basic services, recognizing that results on Uganda may have limited external validity. Table A1 of the Online Appendix reports the results. Rents in urban areas are 56% higher than in rural, and the primate differential from rural is essentially the same at 49%. In comparison, Ugandan urban-rural wage premiums are only 18-28%, although they are 28-49% for Kampala. Noting that housing is about 30% of the household budget in Uganda, there may be little real wage differentials between the urban and rural sector, although Kampala may offer some net real income gain.

3.3.2 Household income premiums

Section 3.3.1 conducts a standard urban exercise. But for Africa it misses a key point. Because so many farm households live in cities and there are extended families living together, it is important to think about the income advantages of households moving to cities, quite apart from individuals' wage differentials.

We construct measures of income for the household. For the household we add together all income from self-employment, labour income, and capital or land income respectively. In the surveys, income receipts of various forms are reported such as cash and in-kind wage payments, business incomes, remittances, incomes from rent of property and farmland, private

[9] The three countries studied in this section, Uganda, Nigeria, and Tanzania, are all part of the sample of countries studied in Gollin, Kirchberger and Lakagos (2017).

and government pensions, and sales revenue from agricultural produce. These receipts are also reported for a variety of time intervals over which they take place, so to be consistent, we convert all income receipts to monthly intervals. Note again, land income from crop sales or rents is generally only available at the household level, making it difficult to ascribe these income sources to any particular household member for an individual level analysis. Second, farm income is largely reflective of average rather than marginal productivity, where the latter is the main interest in typical wage regressions. In focusing on income premiums at the household level, we study income gaps for all types of households.

Panel A of Table 5 reports results for each of our three countries, with the dependent variable as the logarithm of monthly net income for each household. In columns (1), (3), and (5), controls include a variety of human capital indicators for the household head as well as household size, year fixed effects, the household's urban-rural status, and whether the household lives in the capital city. For each country, in columns (2), (4) and (6) we add in occupational fixed effects for the household head, whether the household owns land, and, if so, the total size of household land holdings.

In columns (1), (3) and (5), the household income premium ranges between 63-85%. Primate city premiums are higher going from 69%-134%. Uganda is at the low end indicating more limited urbanisation benefits. Urban income premiums fall noticeably to 35-46% and primate ones fall to 36-100%, once we add agricultural and occupational controls in columns (2), (4), and (6). However, the key finding is that these premiums in Table 5 are much higher than individual wage premiums in Table 4. For urban wages, with and without occupation fixed effects, premiums are 11-28% (vs now 63-85%) and 8-18% (vs now 35-46%) respectively. For primate city they are 11-89% (vs now 69-134%) and 14-67% (vs now 36-100%) respectively. For Nigeria in particular the household premiums are 35-132% while for wages they are only 7-14%.

We also looked at urban versus rural differentials for 'landless' vs 'landed' households, based on whether they own agricultural land. The results are presented in Table A2 of the Online Appendix. Across all columns, although not reported here, the urban income premiums would be greater for farm land households in half the cases and vice versa for the others. For primate cities again they would be higher in half of the relevant cases. In short there is little difference between the two groups. This is a wrinkle on our thinking about urbanisation and structural transformation. Urbanisation is not simply moving out of agriculture. It is a much more nuanced process, whereby farm households gain from urbanisation the same as non-farm households.

In Table 5 panel B we look at how premiums vary within the urban sector. Theory and evidence suggest that the agglomeration effects should be stronger in larger cities (Glaeser and Mare 2001, Yankow 2006, Gould 2007, and Baum-Snow 2012, with a review in Combes and Gobillon, 2015). As for the employment composition, we define secondary cities as above the 75^{th} percentile in the city population distribution excluding the primate, tertiary as between the 50^{th}-75^{th} percentiles, and small cities as below the median. We base these rankings on the population distribution for the first census available in each country, which is 1991 for Nigeria, and 2002 for Tanzania and Uganda.

The sign of city premiums in almost all cases is positive and significant, confirming that household income in rural areas is lower, on average, than income in any of the four city categories. The key issue is that the income benefits of living in secondary cities over rural areas are noticeably smaller than the gains from living in tertiary cities across all columns. They are even much smaller than in the smallest cities in two of the samples. We also did this hierarchal specification for the wage premium results in Table 4, obtaining similar patterns. Both suggest a lack of advantage of being in the top quartile of the city size distribution – an

issue we will focus on in later sections – and suggest that smaller cities may be stronger facilitators of the transition from rural to urban life than secondary cities.

3.4 The advantages of bigger cities for families

A final question which we now turn to is why there are such large differentials between the family and individual level. There are two possibilities why families might benefit more from agglomeration. One is that labour force participation is higher in bigger cities because of greater job opportunities. We looked at this but there seems no consistent pattern of adults 18-65 working more in cities than in the rural sector, based on whether a person worked in the past 7 days. Everyone in a rural household puts in time working on the farm. The other possibility is the choice of occupations available, which may be particularly relevant to women in rural areas who are often constrained to household work.

As a measure of occupational diversity, we conducted a simple count of the number of unique occupations held by individuals within the same household and regressed that on controls (not shown), including number of people in the household, education and age of the household head and year FE's. Table 6 panel A and B show results for this count for the urban-rural-primate split and then for the split across the urban hierarchy. In Panel A, the columns show significant increases in within household occupational diversity in cities and in the primate even more so. Panel B shows that, within the urban hierarchy, gains for tertiary and smaller cities are generally larger than for secondary cities, consistent with prior results. We also looked at differentials for landed households but found little and mixed evidence of differentials.

3.5 Summary remarks

We have shown that the urban household income and wage premiums in these three African countries are positive and significant under a range of different specifications. We found large household income differentials, which were at least as high for families that own land, again

raising this confounding issue in Africa of famers living disproportionately in cities. In comparing household and individual level outcomes, there are much higher returns to locating in urban areas at the family level rather than the individual. This implies that urban areas in these African countries are able to offer employment opportunities for the family unit that are not available in the rural sector. We showed that families in urban areas have much greater occupational diversity. Finally we noted that income and wage premiums are typically higher in primate and small cities, compared to secondary and tertiary cities. This weakness of secondary cities is a theme in much of the rest of the paper.

4. Looking Within the Urban Hierarchy

So far we have focused on structural transformation and urban income and wage premiums, with comparisons within the urban hierarchy. We now delve into that hierarchy and focus on two issues in the literature, primacy and the role of secondary and tertiary cities. Africa's urbanisation is not only characterised by its tremendous speed, but also its heavy concentration in the primate city in most countries.

Economists tend to measure urban concentration by an easily available measure, primacy – the share of total urban population in a country's largest city (Rosen and Resnick 1980, Ades and Glaeser 1995, Davis and Henderson, 2003). In Figure 5, we map this measure of primacy across the world using data from the UN for 2015. Africa as a whole has the highest regional level of primacy worldwide. With a few exceptions, all countries below the Sahara have primacy rates above 30 percent and several above 50. However, Figure 5 also shows that some individual countries in Latin America and East Asia have rates of primacy that are similarly high.

The question is whether primacy in Africa is excessive and costly, given that Henderson (2003) shows in a worldwide analysis that excessive primacy detracts from national income growth as resources are squandered in oversized congested cities. In an earlier perspective, Ades and

Glaeser (1995) and Davis and Henderson (2003) suggest that non-democratic countries, such as in Africa, may favour one or two political cities, particularly the national capital, drawing in excessive numbers of people. Correspondingly, Fetzer, Henderson and Nigmatulina (2017) show that, following waves of democratic transition across parts of Africa in the 1990s and early 2000s, secondary and non-capital cities experienced significant improvements in basic services and health and education outcomes.

However excessive primacy in Africa is a more nuanced story. First, African primacy is not unusual conditional on key drivers. A simple regression using WDI data for the 121 countries with urban populations over 1 million and with land size over 500 square kilometres, shows that controlling for country size (population and land area) and GDP pc, Asian countries actually have the highest rates of primacy on average, followed by Latin America and then finally Sub-Saharan Africa (see Table A3 of the Online Appendix for these results). The key is that African countries are small, limiting the ability of many countries to have very large other metropolitan areas. Additionally, primacy is highest in low income countries and tends to fall as income per capita rises, reflecting the Williamson (1965) idea that after initial concentration in one region, countries will start to decentralize production activity as incomes rise.

The second point is that most African primate cities on a global scale are not enormous, except for Lagos at about 21 million for the greater urban area. In small countries like Rwanda the primate city is around 1 million given a national urban population of only about 3 million; while in medium size countries like Tanzania and Kenya, primate city sizes are in the 4-6 million range.

Nevertheless primacy in Africa is probably excessive as it poses serious challenges for the provision of infrastructure and institutions required to accommodate concentrated populations,

implying that congestion and its costs are likely to outweigh the benefits of urban concentration, especially when incomes are relatively low as pointed out earlier. Evidence comes from Castells-Quintana (2017), who uses panel data from 1960-2010 to show that increases in primacy reduce national GDP in countries with poor infrastructure, and this will only stop when at least 50 percent of urban residents have access to basic services such as sanitation and electricity. As the average African country now falls well below that threshold, his results indicate that faltering urban infrastructure contributes to Africa's weak economic growth in the face of rising primacy.

4.1 Growth of cities in the system

Apart from primacy, economists and international agencies worry about balance in the urban system and in particular, the role of secondary cities (Henderson, 1997, Henderson, Lee and Lee 1997, Kolko 1999, Desmet, Ghani, O'Connell and Rossi-Hansberg 2013, and Lall et al 2017). How is the overall urban hierarchy in African countries emerging? As already suggested, small country sizes can mean a limited role for large secondary cities and a sharp departure from city size distributions following Zipf's Law (Gabaix 1999 or Rossi-Hansberg and Wright 2007). Using data from the UN World Urbanisation Prospects, we calculate that the share of Africa's population in non-primate cities with over 1 million population is only 8.5%, while for all low-middle income countries it is 26%. Instead Africa has a high concentration in cities under 300,000, as well as primate cities (with Africa at 28% versus all low-middle income at 17%)[10]. This lack of large secondary city development is a concern we explore below; especially since, these cities are considered to be facilitators of labour mobility, job creation, and the transition from rural to non-rural activities (Christiaensen. and Todo, 2014).

[10] Figure A3 of the Online Appendix provides further detail on how Africa's urban population distribution compares to other world regions.

Moreover in Africa, primate cities are growing very fast while large and medium size cities in the urban hierarchy are faltering. To see this we examine the patterns of growth across different cities in Africa, using the city data set we discussed earlier but now looking at a larger set of African countries. We include data on a range of geographical covariates that are likely to affect city size and growth. In total, we have a sample of 34 African countries and around 1,000 unique cities. Statistics on these cities placed in the urban hierarchy are given in Table 7. For a list of data sources and availability for this section, see the Online Appendix.

We first note that a levels regression of the log of city population in the latest census period against a variety of geographical amenities and industry controls gave us the expected results: cities near major transport junctions such as ports and coastlines are larger, and rainfall has a strong correlation with city population, which further suggests that many of these cities are agriculturally dependent (see Table A4 of the Online Appendix for the results). An indicator of whether a city had an established manufacturing industry in the 1960s and, if so, how many different types of industries they had from Henderson et al (2017) also have strongly positive coefficients, indicating that city size today is correlated with early industrialisation. But our focus is on growth.

In Table 8, in a panel framework, we explore the correlates of population growth over time for our countries.[11] We look at growth across all census periods – from the 1990s, to early 2000s, to late 2000s. Since the time interval between census periods varies across countries, we convert all of our population data to annual growth rates. While geographic factors and early industrialisation are strongly correlated with city sizes, they are less so with growth. Coastal

[11] Typical city growth analysis dating to Glaeser and Scheinkman (1995), Black and Henderson (1999) on the USA, and Henderson and Wang (2007) more globally focuses on local human capital accumulation and growth. For Africa we do not have the detailed data to do such an analysis. However we can examine the role of geographic factors and place in the urban hierarchy within countries.

cities grow faster but not if they have a harbour. Growth is largely parallel, and mostly only differentiated by place in the urban hierarchy.

In columns (1)-(3), primate cities tend to grow faster on average than the rest of the urban hierarchy. Moreover, cities in coastal countries further from the primate city experience faster population growth on average, than those in the shadow of the primate. Once we control for population in the base period in columns (4)-(6) and mean reversion, we see an even larger coefficient on the primate variable. Columns (3) and (6) suggest that tertiary and secondary cities are growing slower than the base, the bottom 50% of cities by size. These patterns also follow those in the raw data, where primate and small cities are growing faster than other cities.

The data suggest that addressing Africa's urbanisation challenges requires a fine balance between tackling the internal and external costs and benefits of primacy. Primate cities are growing the fastest, perhaps because of traditional biases in capital markets and fiscal incentives for firms and government job provision towards capital cities. As such, excessive primacy may arguably be hurting national growth, however, the ability of secondary cities to absorb population may be limited anyway, unless Africa develops more of a manufacturing base.

Over the course of development, tertiary and secondary cities tend to play a central role, as cities which receive industry decentralising from the primate. This decentralisation occurs as production techniques become standardised and firms have less to gain from agglomeration economies of primate cities, allowing them to relocate to somewhat smaller cities where land and labour are cheaper (Duranton and Puga, 2001). In the process, primate (and other very large cities) then focus on traded services, where agglomeration economies are very strong (Arzaghi and Henderson, 2008). This process of decentralisation, or diffusion of manufacturing and concentration of services has been analysed historically in the USA by Desmet and Rossi-

Hansberg (2009) and for India recently by Desmet, Ghani, O'Connell and Rossi-Hansberg (2013), with examples for China and Korea given in Desmet and Henderson (2015). More generally, in most parts of the developed world historically and the developing world today, secondary and tertiary cities are manufacturing oriented (see Kolko 1999 and Henderson 1997 and 1988 on the USA and Brazil, and Henderson, Lee, and Lee 1997 on Korea and Rondinelli 1982 more generally).

In Africa, the problem is that primate and all other cities are yet to have a strong manufacturing base as we will see below, and as suggested by Table 1 and Figure 3. The competitors for population appear to be the smaller more agriculturally based cities. Such cities have a strong role to play in Africa's potential for agricultural and economic development, especially since a recent body of evidence following Ramankutty et al (2002) shows that large reserves of global cultivable croplands lie in Africa. Results from recent land-scape modelling (Ray et al, 2012), randomised field trials (Twomlow et al, 2010), and policy experiments (Sanchez, 2010) have all demonstrated that there is the potential to greatly increase yields in Africa and as such, significant progress could be made through increasing cropping efficiency and closing yield gaps on underperforming land. This could then serve to further accentuate the role of smaller cities where farmers either live or spend incomes, or both. That is a rather different perspective on the role of cities in Africa, at least in the next stages of development.

4.2 The prospect of emerging industry in Africa's non-capital cities

How Africa's larger cities evolve over the next few decades may also largely depend on their ability to jump the hurdle into being internationally competitive in manufacturing or tradeable services (Venables, 2017). Being competitive has many dimensions: greater internal efficiency through better management and infrastructure investments and better transport links, higher

investments in human capital and the like. It is hard to gauge how far these cities are from the hurdle.

However we can ask whether there is any evidence of industrialisation and in particular modern industrialisation in *any* cities in the parts of Africa for which we have comprehensive data. In Figure 6, we rank the entire set of 615 cities that have employment data in the sample of 12 countries used earlier by their shares of employment in manufacturing, finance real estate and business services, and personal trade. Figure 6 shows that 90 percent of cities in our sample have manufacturing employment shares below 13 percent – well below the national *averages* of other world regions. Only a tiny handful have manufacturing employment shares over 20%, a number which would be low for say China and even for India. Around 95 percent of cities have less than 15 percent of the employment share in financial, insurance, real estate, and business services, and most well under 10%. As can be seen in Figure 6, the main employment basis of African cities outside of agriculture is small trade and personal services, which keeps African cities very local in economic scope.

Looking to the future an issue that people raise is whether Africa in structural transformation could skip the 'manufacturing stage' and go directly to traded services, noting now in countries like the USA this is what dominates urban production especially in bigger cities (Kolko, 1999). Note in Figure 4, exportable services are the fastest growing sector in urban areas. To explore this more we broke these services out into distinct categories of financial services and insurance, and real estate and business services. Between the census periods for these six countries, financial and insurance services have increased their employment share in urban areas by around 32%, whilst the real estate and business services share has declined by 19%. In terms of actual growth, financial services registered around 90% growth in employment between the censuses (growing from 87,814 employees to 166,919), whilst real estate and business services grew by only 16% (140,756 to 164,189). These shifting shares were even more dramatic in

rural areas, where the share taken by financial services grew over 100% and that for real estate and business services declined by around 45% in rural areas. For employment growth, rural financial services grew over 400% (6,124 employees to 32,748) whilst real estate grew less than 50% (14,390 to 21,180). See Figures A4 and A5 of the Online Appendix for the results.

5.3.1 Why is African Manufacturing So Far Behind?

There are many factors which could be stalling the development of manufacturing industries across Sub-Saharan Africa. For the heavily resource dependent countries, the mainstream literature has suggested that natural resource exploitation tends to crowd out the manufacturing industry: positive wealth shocks in the resource sector tend to appreciate the exchange rate and dampen the competitiveness of other traded sectors. As noted by Sachs and Warner (2001) resource economies tend to have higher prices, while, according to Harding and Venables (2013), the response to a positive resource shock is to drop non-resource exports in the order of 35-70 percent. Gylfason (2001) argues that since resource extraction tends to be low skill and highly mechanised, it is likely to stimulate a neglect on human capital by crowding out education, innovation, and entrepreneurship. Institutional capacity is also key. For resource economies, this source of wealth has often lead to pressure on authorities to take on large infrastructure projects that are too complex for their institutional capacities (Ismail, 2009). Moreover, these sources of extractable wealth may induce predatory behaviour which leads to widespread corruption of fair institutional behaviours (Kuran, 2004).

For all African countries, a key problem is that current governance and infrastructure tends to be too weak to support a thriving manufacturing sector. Not only is the basic level of key infrastructure – such as highways, ports, power generators – particularly weak, but there is often a lack of ancillary institutions needed to support manufacturing development such as finance, insurance and legal institutions. The risk of factional behaviour on the part of governments may

lower incentives for large-scale investments if business owners fear the possibility of expropriation. As noted earlier, Venables (2017) models this more generally analysing how African cities might break into production of tradeable manufactures if cities were more efficient, with the possibility that some are stuck in a bad equilibrium.

5. Concluding remarks

Agriculture remains a big venture in Africa with capacity not just for improvement in technology and farming practices but also for expansion of farming area and intensity of farming. While there are arguments about consumer cities, urbanisation in Africa, compared to the rest of the world, appears to be less driven by either natural resource incomes or industrialisation. Instead, urban residents in the bottom three quarters of the city size distribution appear to be still heavily engaged in agriculture as their main occupation.

While the literature suggests the productivity gains from structural transformation are particularly large in Africa, the problem is that structural transformation and urbanisation are not well linked: hence the widespread presence of farmers in most cities. We explore household income gaps between urban and rural residents; and find that the gaps for farmers in cities are similar to those in other occupations. Wage gaps for individuals compared to family income gaps appear much smaller, and the data suggest that one of the advantages of living in cities are the opportunities for expanding occupational choice in the household.

Smaller cities near the bottom of the size distribution tend to be growing faster than larger secondary (but not primate) cities, and seem to exist to serve agriculture and to house farmers – with their fortunes perhaps rising with improved productivity in farming in recent years. In comparison larger secondary cities appear to be stagnant and with scant evidence of any real industrialisation in all but a tiny handful of cities. In other parts of the developing world, larger secondary cities tend to be thriving due to industrialisation, but in Africa there is so little

industry nationally that the role and prospective growth of secondary cities is challenged. There may well be a threshold which these cities could pass – with better management and infrastructure investments – that would make them internationally competitive, but given the absolute lack of a manufacturing base in almost all cities, that threshold may be far away.

Primacy in Africa as measured by the share of the primate city in national urban population is extraordinarily high on a world scale, but appears to be explained by small country size and low incomes. The issue is more that primate cities are growing faster than other cities, especially those in the top 50 percentiles of cities by size. The question is why and what are the consequences? High primacy in Africa does not necessarily mean large cities on a world scale, other than Lagos. But high primacy is occurring at low income levels where cities cannot afford appropriate infrastructure and still have deficient institutions and managerial capacity. The issue is that primate cites are extremely costly and recent evidence suggests increasing primacy is hurting national growth in Africa.

References

Ades, A. F. and Glaeser, E. L. (1995), Trade and circuses: Explaining urban giants. *The Quarterly Journal of Economics*, 110(1), 195–227.

Alcott, H. and Keniston, D. (2014), Dutch disease or agglomeration? The local economic effects of natural resource booms in modern America. *NBER* WP# 20508.

Arzaghi, M. and Henderson, J.V. (2008), Networking Off Madison Avenue, *Review of Economic Studies*, 75, 1011-1038.

Baum-Snow, N. and Pavin, R. (2012), Understanding the City Size Wage Gap, *Review of Economic Studies* 79(1): 88-127.

Block, S. A. (2013), The post-independence decline and rise of crop productivity in sub-Saharan Africa: measurement and explanations. *Oxford Economic Papers*, *66*(2), pp.373-396

Castells-Quintana, D. (2017), Malthus living in a slum: Urban concentration, infrastructure and economic growth. *Journal of Urban Economics*, 98, 158-173.

Caselli, F. (2005), Accounting for cross-country income differences. *Handbook of economic growth*, 1, 679–741. ed. by P. Aghion, and S. Durlauf. Elsevier.

Caselli, F. and Coleman II, W. J. (2001), The US structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy*, 109(3), 584-616.

Christiaensen, L. and Todo, Y. (2014), Poverty reduction during the rural–urban transformation–the role of the missing middle. *World Development*, *63*, pp.43-58.

Collier, P., Van Der Ploeg, R., Spence, M., and Venables, A. J. (2009), Managing resource revenues in developing economies. *IMF Staff Papers*, 57(1), 84-118.

Combes, P. P., Duranton, G., and Gobillon, L. (2008), Spatial wage disparities: Sorting matters! *Journal of Urban Economics*, 63(2), 723-742.

Combes, P. P., Duranton, G., Gobillon, L., and Roux, S. (2010), Estimating agglomeration economies with history, geology, and worker effects. in *Agglomeration Economics* (pp. 15-66), E.L. Glaeser (ed), University of Chicago Press.

Combes, P.P. and Gobillon, L. (2015), The Empirics of Agglomeration Effects. In Duranton, G., Henderson, V., and Strange, W. (Eds.), *Handbook of Regional and Urban Economics*, vol. 5A. (Amsterdam: North-Holland).

Combes, P. P., Duranton, G., and Gobillon, L (2016), The Costs of agglomeration: House and land prices in French cities. Working paper

Davis, J. C. and Henderson, J. V. (2003), Evidence on the political economy of the urbanisation process. *Journal of Urban Economics*, 53(1), 98–125.

D'Costa, S. and Overman, H. G. (2014), The urban wage growth premium: Sorting or learning? *Regional Science and Urban Economics*, 48, 168-179.

de la Rocha, J. and Puga. D. (2017), Learning by Working in Big Cities, *Review of Economic Studies,* forthcoming

Desmet, K. and Rossi-Hansberg, E. (2009), Spatial growth and industry age. *Journal of Economic Theory*, 144, 2477–2502.

Desmet, K. and Rossi-Hansberg E. (2013), Urban Accounting and Welfare**,** *American Economic Review*, 103:6, 2296-2327

Desmet, K., Ghani, E., O'Connell, S., and Rossi-Hansberg, E. (2015), The spatial development of India. *Journal of Regional Science*. 55, 10–30.

Desmet, K. and Henderson, J. V. (2015), The geography of development within countries. In *Handbook of Urban and Regional Economics*. G. Duranton, J.V.Henderson. W. Strange (eds), North Holland vol5B.

Duranton, G. and Overman H. (2005), Testing for Localisation Using Micro-Geographic Data, *Review of Economic Studies*, 72 (4): 1077-1106.

Duranton, G. and Puga D. (2001), Nursery Cities: Urban diversity, process innovation, and the life-cycle of products, *American Economic Review*, 1454-1477

Duranton, G. and Puga. D. (2004), Micro-foundations of Urban Agglomeration Economies. In: Henderson, V, Thisse, F (Eds.), *Handbook of Regional and Urban Economics*, vol. 4. (Amsterdam: North-Holland).

Ellison, G, and Glaeser E. (1997), Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach, *Journal of Political Economy*, 105, 889-927

Fay, M. and Opal, C. (2000), Urbanisation without growth: A not so uncommon phenomenon, WP # 2412. *World Bank Publications*.

Ferré, C., Ferreira, F. H., and Lanjouw, P. (2012), Is There a Metropolitan Bias? The relationship between poverty and city size in a selection of developing countries. *The World Bank Economic Review*, *26*(3), pp.351-382.

Fetzer, T., Henderson, J.V., Nigmatulina, D., and Shanghavi, A. (2017), What happens to Cities when Countries Become Democratic? *Working Paper*, available at https://collaboration.worldbank.org/docs/DOC-23398

Gabaix, X. (1999), Zipf's law for cities: an explanation. The *Quarterly Journal of Economics*, 114, 739–767.

Glaeser, E. L. (2014), A world of cities: the causes and consequences of urbanization in poorer countries. *Journal of the European Economic Association*, 12(5), 1154-1199.

Glaeser, E. L., Scheinkman J., and Shleifer A. (1995),"Economic growth in a cross-section of cities" *Journal of Monetary Economics* 117-134

Glaeser, E. L. (1999), Learning in cities. *Journal of Urban Economics*, 46(2), 254-277.

Glaeser, E. L. and Maré, D. C. (2001), Cities and skills. *Journal of Labor economics*, 19(2), 316–342.

Gollin, D., Kirchberger, M., and Lagakos, D. (2017), 'In Search of a Spatial Equilibrium in the Developing World', Working Paper, available at https://sites.google.com/site/davidlagakos/home/ research/Spatial_Eqm_v5.pdf

Glaeser, E.L., Kolko J., and Saiz A. (2001), Consumer City. *Journal of Economic Geography*, 1 (1): 27-50.0

Gollin, D., Jedwab, R., and Vollrath, D. (2016), Urbanisation with and without Industrialisation. *Journal of Economic Growth*, 21(1), 35-70.

Gollin, D., Lagakos, D., and Waugh, M. E. (2014), The agricultural productivity gap". *The Quarterly Journal of Economics*, 129(2), 939-993.

Gollin, D., Parente, S., and Rogerson, R. (2002), The role of agriculture in development. *American Economic Review*, 160-164.

Gollin, D., Parente, S., and Rogerson, R. (2007), The food problem and the evolution of international income levels. *Journal of Monetary Economics*, 54(4), 1230-1255.

Gould, E. D. (2007), Cities, workers, and wages: A structural analysis of the urban wage premium. *The Review of Economic Studies*, 74(2), 477-506.

Gylfason, T. (2001), Natural resources, education, and economic development. *European economic review*, 45(4), 847-859.

Gyourko, J., Mayer, C., and Sinai, T. (2013), Superstar cities. *American Economic Journal: Economic Policy*, 5(4), 167-199.

Harding, T. and Venables, A. J. (2013), The Implications of Natural Resource Exports for Non-Resource Trade. *CEPR Discussion Papers*: C.E.P.R. 9318.

Henderson, J. V. (2003), The urbanisation process and economic growth: The so-what question. *Journal of Economic Growth*, 8(1), 47–71.

Henderson, J. V., Lee, T., Lee, Y.J. (2001), Scale externalities in a developing country. *Journal of Urban Economics*. 49, 479–504.

Henderson, J. V., Storeygard, A., and Deichmann, U. (2017), Has climate change driven urbanisation in Africa? *Journal of Development Economics* 124, 60–82.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012), Measuring economic growth from outer space. *The American Economic Review*, 102(2), 994-1028.

Henderson, J. V. and Wang, H. G. (2005), Aspects of the rural-urban transformation of countries. *Journal of Economic Geography*, 5(1), 23-42.

Henderson, J. V. and Wang H.G. (2007), Urbanisation and City Growth: the Role of Institutions, *Regional Science and Urban Economics*, 2007, 37, 283-313

Herrendorf, B. and Schoellman, T. (2011), Why is Labor Productivity so Low in Agriculture? Manuscript, Arizona State University.

Ismail, K. K. (2011), The Structural Manifestation of the `Dutch Disease' In Oil-Exporting Countries", mimeo, International Monetary Fund.

Jedwab, R. (2013), Urbanisation without structural transformation: Evidence from consumption cities in Africa. George Washington University, Washington, DC. Processed.

Jedwab, R., Christiaensen, L., and Gindelsky, M. (2017), Demography, urbanization and development: Rural push, urban pull and… urban push? *Journal of Urban Economics*, 98, 6-16.

Jedwab, R. and Vollrath, D. (2017), "The Urban Mortality Transition and Poor Country Urbanization", Mimeo

Jones P., D'Aoust O., and Bernard L. (2017),The Urban Wage Premium in Africa. In: *Johnson-Lans S. (eds),Wage Inequality in Africa. Global Perspectives on Wealth and Distribution*. Palgrave Macmillan, Cham

Johnson, S., Ostry, J. D., and Subramanian, A. (2010), Prospects for sustained growth in Africa: benchmarking the constraints. *IMF Staff Papers*, *57*(1), pp.119-171.

Klein, W. D., Mandt, G. A., and Gagliardo, J. (1992), Defense Meteorological Satellite Program. *6th Conference on Satellite Meteorology and Oceanography* (pp. 450-452).

Kolko, J. (1999), Can I get some service here: information technology, service industries, and the future of cities. *Harvard University, Mimeo.*

Kuran, T. (2004), Why the Middle East is economically underdeveloped: historical mechanisms of institutional stagnation. *The Journal of Economic Perspectives*, 18(3), 71-90.

Lall, S. V., Henderson, J. V., and Venables, A. J. (2017), Africa's Cities: Opening Doors to the World. *World Bank Publications*.

Lewis, W. A. (1954), Economic development with unlimited supplies of labour. *The Manchester school* 22(2), 139–191.

Lucas, Jr, R. E. (2004), Life earnings and rural-urban migration. *Journal of Political Economy* 112(S1), S29–S59.

Marshall, A. (1890),*Principles of Economics*. Macmillan, London.

Matsuyama, K. (1992), Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory* 58(2), 317–334.

McMillan, M., Rodrik, D., and Verduzco-Gallo, Í. (2014), Globalisation, structural change, and productivity growth, with an update on Africa. *World Development*, 63(1), 11-32.

Michaels, G., Rauch, F., and Redding S.J. (2012), Urbanization and Structural Transformation. *Quarterly Journal of Economics*, 127 (2): 535-586.

Nurkse, R. (1953), *Problems of capital formation in developing countries*. New York: Columbia UP.

Ramankutty, N., Foley, J. A., Norman, J., and McSweeney, K. (2002), The global distribution of cultivable lands: current patterns and sensitivity to possible climate change. *Global Ecology and Biogeography*, 11(5), 377-392.

Ray, D. K., Ramankutty, N., Mueller, N. D., West, P. C., and Foley, J. A. (2012), Recent patterns of crop yield growth and stagnation. *Nature Communications*, 3, 1293.

Roback J (1982),Wages, rents and the quality of life. *Journal of Political Economy* 90(6):1257–1278.

Rondinelli, D. A. (1982), Intermediate cities in developing countries: a comparative analysis of their demographic, social and economic characteristics. *Third World Planning Review* 4(4): 357-86.

Rosen, K. T. and Resnick M. (1980), The Size Distribution of Cities: An Examination of the Pareto Law and Primacy, *Journal of Urban Economics* 8: 165-86.

Rosenstein-Rodan, P. N. (1943), Problems of industrialisation of eastern and south-eastern Europe. *The Economic Journal* 53(210/211), 202–211.

Rosenthal, S. and Strange, W. (2004), Evidence on the Nature and Sources of Agglomeration Economies. ch. 49, p. 2119-2171 in Henderson, J. V. and Thisse, J. F. eds. *Handbook of Regional and Urban Economics*, vol. 4, Amsterdam: Elsevier.

Rossi-Hansberg and Wright, E. M. (2007), Urban structure and growth. *Review of Economic Studies*, 74 (2), 597–624.

Rostow, W. W. (1960), *The process of economic growth*. W. W. Norton, New York

Ruggles, S., Sobek, M., Hall, P. K., and Ronnander, C. (1995), Integrated Public Use Microdata Series: IPUMS-95 Version 1.0. Social History Research Laboratory, Department of History, University of Minnesota.

Sachs, J. and Warner, A. (2001), The Curse of Natural Resources. *European Economic Review*, Vol. 45, pages 827-838.

Sanchez P.A. (2010), Tripling crop yields in tropical Africa. *Nature Geoscience* 3: 299 – 300.

Schultz, T. W. et al. (1968), *Economic growth and agriculture*. McGraw-Hill, New York.

Storeygard, A. (2016), Farther on down the road: transport costs, trade and urban growth in Sub-Saharan Africa. *The Review of Economic Studies*, *83*(3), pp.1263-1295.

Tolley, G., Gardner, J., and Graves, P. (1979), *Urban Growth Policy in a Market Economy*. New York: Academic Press.

Twomlow, S., Rohrbach, D., Dimes, J., Rusike, J., Mupangwa, W., Ncube, B., and Mahposa, P. (2010), Micro-dosing as a pathway to Africa's Green Revolution: evidence from broad-scale on-farm trials. *Nutrient Cycling in Agroecosystems*, 88(1), 3-15.

UN Habitat (2010), *The State of African Cities*. Nairobi: UN Habitat.

United Nations. (2015*), World urbanisation prospects: The 2015 revision*. New York: United Nations, Department of Economic and Social Affairs.

Venables, A. (2017), Breaking into Tradables: Urban Form and Urban Function in a Developing City. *Journal of Urban Economics.*

Wang Z (2016),Wage Growth, Ability Sorting, and Location Choice at Labor-Force Entry: New Evidence from U.S. Census Data. *Journal of Urban Economics*, *96*, pp.112-120.

Williamson, J. (1965), Regional Inequality and the Process of National Development, *Economic Development and Cultural Change*, June, 3-45.

World Development Indicators (2017), http://data.worldbank.org/data-catalog/world-development-indicators

World Bank (2015), *Measuring Living Standards within Cities. Households Surveys: Dar es Salaam and Durban.* Washington, DC: World Bank

World Bank (2016), *Cote d'Ivoire Urbanisation Review. Diversified Urbanisation*. Washington, DC: World Bank.

World Bank (2017), Living Standards Measurement Surveys. http://econ.worldbank.org/WBSITE/EXTERNAL/EXTDEC/EXTRESEARCH/EXTLSMS/0,,contentMDK:21610833~pagePK:64168427~piPK:64168435~theSitePK:3358997,00.html

Yankow, J. (2006), Why do Cities Pay More? An Empirical Investigation of Some Competing Theories of the Urban Wage Premium, *Journal of Urban Economics*, 60: 139-161.

Young, A. (2013), Inequality, the urban-rural gap, and migration. *The Quarterly Journal of Economics*, 128(4), 1727-1785.

Table 1: Average value added by industry sector (%)

	Agriculture		Manufacturing		Services	
	1995-2000	2010-2015	1995-2000	2010-2015	1995-2000	2010-2015
Region	(1)	(2)	(3)	(4)	(5)	(6)
Sub-Saharan Africa	23.4	17.5	10.6	8.4	44	55.7
East Asia & Pacific	9.7	5.5	25.8	23.8	52.1	59.2
Latin America & Caribbean	6.3	5.3	17.7	14.7	62.5	64
OECD members	2.3	1.7	18.5	15	70	73.9
World	6.8	3.9	20.2	16.1	61.5	67.8

Notes: Value added is the net output of a sector after adding up all outputs and substracting intermediate inputs. Agriculture corresponds to ISIC divisions 1-5, manufacturing to divisions 15-37, and services to divisions 50-99.
Source: World Development Indicators (2017)

Table 2: Cross-country estimation of the rate of urbanisation in 2010

	(1)	(2)
Manufacturing and Services (% of GDP)	1.315***	1.728***
	(0.173)	(0.346)
Av share of NR exports in GDP 1980-2015	1.416***	1.766***
	(0.225)	(0.274)
LAC		79.02*
		(47.19)
MENA		57.68
		(60.01)
SSA		101.8***
		(29.53)
LAC × Manufacturing and Services (\% of GDP)		-0.774
		(0.581)
MENA × Manufacturing and Services (\% of GDP)		-0.609
		(0.731)
SSA × Manufacturing and Services (\% of GDP)		-1.516***
		(0.403)
LAC × Av share of NR exports in GDP 1980-2015		-1.594
		(1.034)
MENA × Av share of NR exports in GDP 1980-2015		-0.386
		(0.587)
SSA × Av share of NR exports in GDP 1980-2015		-1.107**
		(0.422)
Region fixed effects	No	Yes
Observations	93	93
R-squared	0.379	0.582

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(2) is the rate of urbanisation in 2010. The sample is limited to countries that were considered developing as of 1960; it excludes Europe and the Neo-Europes and countries smaller than 500sqkm. Robust standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: World Development Indicators (2017)

Table 3: Employment shares in agriculture nationally and in urban areas

	Share of Agri in Urban		Share of Agri in Total	No. Districts
	1990s	*2000s*	*Latest Census*	
Africa				
Botswana	5.9	-	26.2	21
Cameroon	-	21.2	61.9	39
Ethiopia	14.4	-	90.2	58
Ghana	21.9	15.2	44.0	102
Liberia	-	13.7	43.9	20
Malawi	18.8	21.1	65.8	76
Mali	39.3	12.7	69.7	28
Mozambique	43.9	36.1	76.7	85
Sierra Leone	-	39.9	78.3	51
Tanzania	37.8	27.4	65.1	112
Uganda	-	16.2	76.0	69
Zambia	10.2	18.9	75.7	101
India	11.0	7.4	56.6	31
Other				
Brazil	6.6	5.9	15.2	25
Cambodia	33.5	14.1	72.3	23
Malaysia	5.6	4.1	16.4	26
Thailand	5.2	14.3	56.6	25
Vietnam	19.4	14.4	54.0	35

Country Years: **Africa:** Botswana 1991, Cameroon 2005, Ethiopia 1994, Ghana 2000, 2010, Liberia 2008, Malawi 1998, 2008, Mali 1998, 2009, Mozambique 1997, 2007, Rwanda 2002, Sierra Leone 2004, South Sudan 2008, Sudan 2008, Tanzania 2002, 2012 Uganda 2002, Zambia 2000, 2010. **India:** 1993, 2004. **Other:** Brazil 2000, 2010, Cambodia 1998, 2008; Indonesia 2000, 2010; Malaysia 1991, 2000; Vietnam 1999, 2009; Thailand 1990, 2000.
Source: IPUMS-International

Table 4: Estimation of the individual urban net income premium

	Tanzania		Uganda		Nigeria	
	(1)	(2)	(3)	(4)	(5)	(6)
Primate City	0.891***	0.673***	0.490***	0.279***	0.105***	0.138***
	(0.0877)	(0.0902)	(0.0493)	(0.0532)	(0.0371)	(0.0375)
Urban	0.234***	0.0801	0.281***	0.175***	0.109***	0.0708**
	(0.0843)	(0.0822)	(0.0377)	(0.0398)	(0.0301)	(0.0295)
R-squared	0.204	0.259	0.238	0.283	0.117	0.167
Observations	2359	2358	5584	5132	8289	8124
R-squared	0.209	0.263	0.239	0.285	0.119	0.170
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes	No	Yes

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(6) is the natural logarithm of total net individual monthly income. Controls are education recorded or not, if recorded level of education, age, age squared, gender, hours worked recorded or not, if recorded number of hours workers. The sample is limited to individuals from households that do not own land. Columns (1)-(2) are for Tanzania, (3)-(4) Uganda, and (5)-(6) Nigeria. Robust Standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: World Bank Living Standards Measurement Surveys

Table 5: Estimation of the household urban net income premium

	Tanzania		Uganda		Nigeria	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: All households: Urban rural and primate split						
Primate city	1.344***	1.000***	0.691***	0.361***	1.317***	0.859***
	(0.0839)	(0.0965)	(0.0610)	(0.0665)	(0.0716)	(0.0715)
Urban	0.677***	0.388***	0.626***	0.463***	0.852***	0.352***
	(0.0569)	(0.0648)	(0.0454)	(0.0469)	(0.0377)	(0.0396)
Panel B: Urban hierarchy split						
Primate City	1.340***	0.978***	0.674***	0.324***	1.362***	0.876***
	(0.0839)	(0.0970)	(0.0610)	(0.0664)	(0.0748)	(0.0756)
Secondary	0.440***	0.149	0.263***	0.0588	0.725***	0.371***
	(0.111)	(0.119)	(0.0763)	(0.0769)	(0.0691)	(0.0692)
Tertiary	0.717***	0.382***	0.307***	0.156**	1.283***	0.682***
	(0.0963)	(0.0966)	(0.0647)	(0.0663)	(0.0615)	(0.0629)
Small	0.755***	0.469***	0.971***	0.800***	0.697***	0.199***
	(0.0718)	(0.0759)	(0.0695)	(0.0667)	(0.0500)	(0.0499)
Observations	6809	6809	10079	9523	19055	17930
R-squared	0.189	0.266	0.367	0.431	0.059	0.179
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes	No	Yes

Notes: The dependent variable in columns (1)-(6) is the logarithm of total net household monthly income. Controls are education recorded or not, if recorded level of education, age of hh head, age squared, gender, hh size, hh size squared, and in columns 2, 4 and 6, if household owns land and if so ln of hectares. Robust Standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: World Bank Living Standards Measurement Surveys

Table 6: Estimation of occupational diversity in the household

	Tanzania	Uganda	Nigeria
	Panel A. Urban -rural		
Primate City	0.167***	0.210***	0.207***
	(0.0268)	(0.0297)	(0.0431)
Urban	0.0713***	0.258***	0.140***
	(0.0175)	(0.0210)	(0.0195)
R-squared	0.487	0.184	0.133
	Panel B. Urban hierarchy		
Primate City	0.166***	0.210***	0.221***
	(0.0268)	(0.0297)	(0.0446)
Secondary	0.0141	0.191***	0.100***
	(0.0307)	(0.0373)	(0.0336)
Tertiary	0.0834**	0.343***	0.139***
	(0.0324)	(0.0393)	(0.0341)
Small	0.0901***	0.247***	0.161***
	(0.0209)	(0.0294)	(0.0262)
Observations	7181	10213	9674
R-squared	0.487	0.185	0.133

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(6) is the number of unique occupations held by members within each household. Controls are education recorded or not, if recorded level of education, age of hh head, age squared, gender, hh size, hh size squared. Columns (1)-(2) are for Tanzania, (3)-(4) Uganda, and (5)-(6) Nigeria. Robust Standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: World Bank Living Standards Measurement Surveys

Table 7: Descriptive statistics of population by city type

Population	**Mean**	**Min**	**25th**	**Median**	**75th**	**Max**	**STD**	**N**
Primate	1,981,782	104,837	661,256	1,425,931	3,161,027	7,300,173	1,717,123	34
Secondary	195,682	19,572	63,431	107,500	204,900	2,178,704	263,467	221
Tertiary	56,508	2,443	27,227	39,738	55,958	492,984	59,049	274
Small	26,630	963	13,532	20,016	28,165	230,625	26,007	496

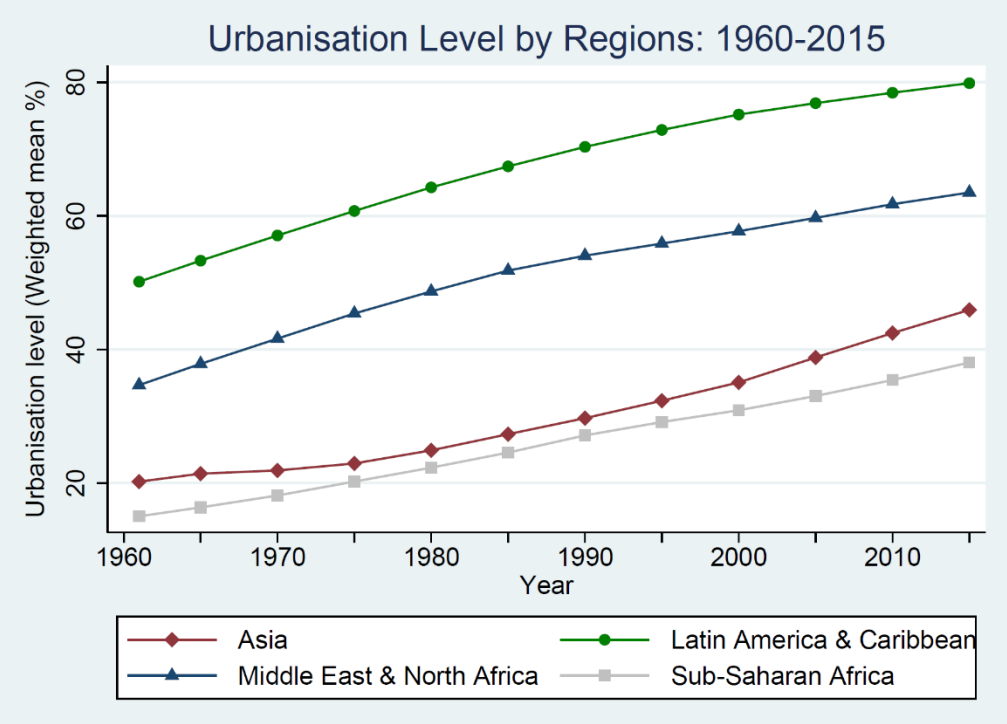
Notes: Table presents summary statistics of population by four city sub-types. Primate cities are defined as the most populous cities, secondary as above the 75th percentile in the city size distribution excluding the primate, tertiary between the median and 75th percentile, and small as all cities below the median.
Source: CityPopulation.de

Table 8: Estimation of inter-census annual population growth

	(1)	(2)	(3)	(4)	(5)	(6)
Coastal	0.00995**	0.0116***	0.00991**	0.0112***	0.0120***	0.0106***
	(0.00390)	(0.00418)	(0.00385)	(0.00385)	(0.00413)	(0.00384)
Harbour dummy	-0.0136***	-0.0121***	-0.0109***	-0.00915**	-0.00860**	-0.00957**
	(0.00406)	(0.00426)	(0.00404)	(0.00405)	(0.00425)	(0.00405)
Di coast	0.000492	0.000416	0.000451	0.000334	0.000404	0.000370
	(0.000463)	(0.000494)	(0.000457)	(0.000457)	(0.000488)	(0.000456)
Primate	0.00621*	0.0159***	0.00138	0.0225***	0.0228***	0.0200***
	(0.00348)	(0.00564)	(0.00350)	(0.00412)	(0.00571)	(0.00642)
Di primate	0.00224***	0.00231***	0.00217***	0.00229***	0.00235***	0.00229***
	(0.000637)	(0.000669)	(0.000627)	(0.000628)	(0.000662)	(0.000626)
Di coast if llocked (100km)	0.000307	0.000652	0.000184	0.000343	0.000589	0.000313
	(0.000630)	(0.000697)	(0.000620)	(0.000620)	(0.000689)	(0.000619)
Di prmt if llocked (100km)	-0.00255***	-0.00338***	-0.00245***	-0.00253***	-0.00342***	-0.00257***
	(0.000860)	(0.00106)	(0.000848)	(0.000848)	(0.00105)	(0.000846)
Ruggedness index	4.63e-08	-2.39e-08	-1.54e-08	5.52e-09	-4.84e-08	-2.34e-08
	(0.000000166)	(0.000000187)	(0.000000164)	(0.000000164)	(0.000000185)	(0.000000163)
Ln(Elevation)	-0.000806	-0.000361	-0.000634	-0.000187	-0.00000452	-0.000321
	(0.000936)	(0.00101)	(0.000924)	(0.000926)	(0.00100)	(0.000925)
Avg Rainfall (mnthly 1960-90)	0.0000395*	0.0000273	0.0000421*	0.0000410*	0.0000330	0.0000435**
	(0.0000223)	(0.0000240)	(0.0000219)	(0.0000219)	(0.0000237)	(0.0000219)
Has Industry (1965)		-0.00392			-0.00135	
		(0.00253)			(0.00254)	
Number of Industries		-0.000638			0.0000370	
		(0.000648)			(0.000651)	
Secondary			-0.00911***			-0.00142
			(0.00170)			(0.00280)
Tertiary			-0.00857***			-0.00508***
			(0.00138)			(0.00170)
Regional cap			0.00115			0.00254
			(0.00163)			(0.00167)
Ln(Initial Population)				-0.00460***	-0.00455***	-0.00437***
				(0.000646)	(0.000809)	(0.00127)
Observations	1712	1390	1712	1712	1390	1712
R-squared	0.204	0.201	0.230	0.228	0.219	0.235
Country-Year FE	Yes	Yes	Yes	Yes	Yes	Yes

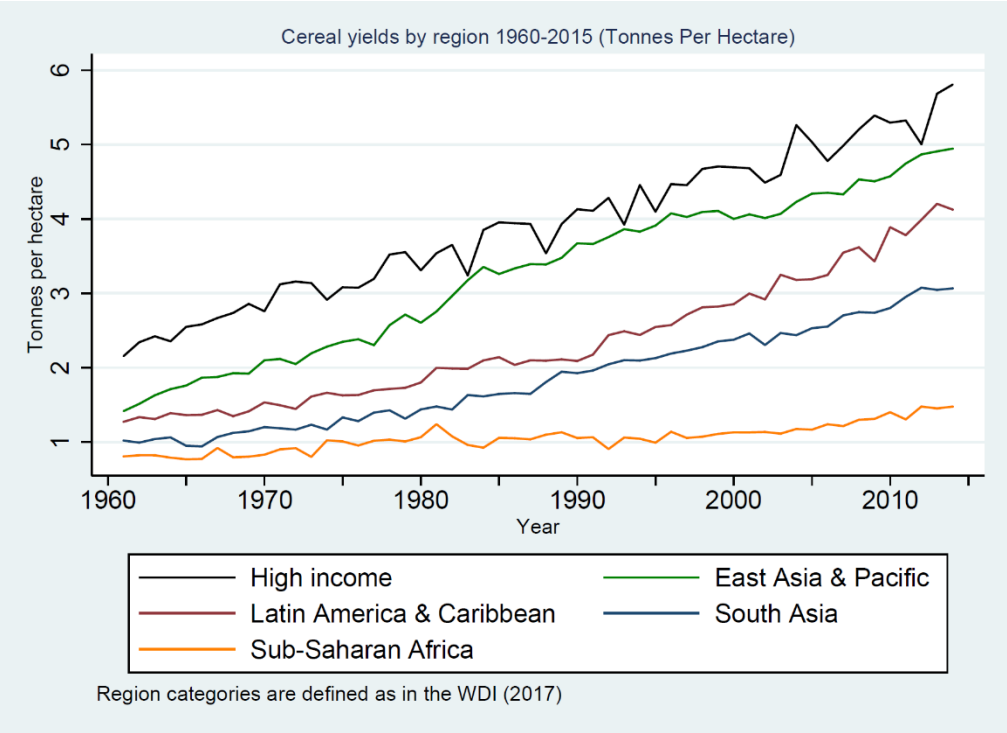
Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(6) is annua city population growth. Columns (4)-(6) control for the natural logarithm of city population in the baseline census period (roughly in the early 1990s). Robust Standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, $p < 0.10$.
Source: Author's own constructed dataset

Figure 1: Urbanisation Level by Regions: 1960-2010



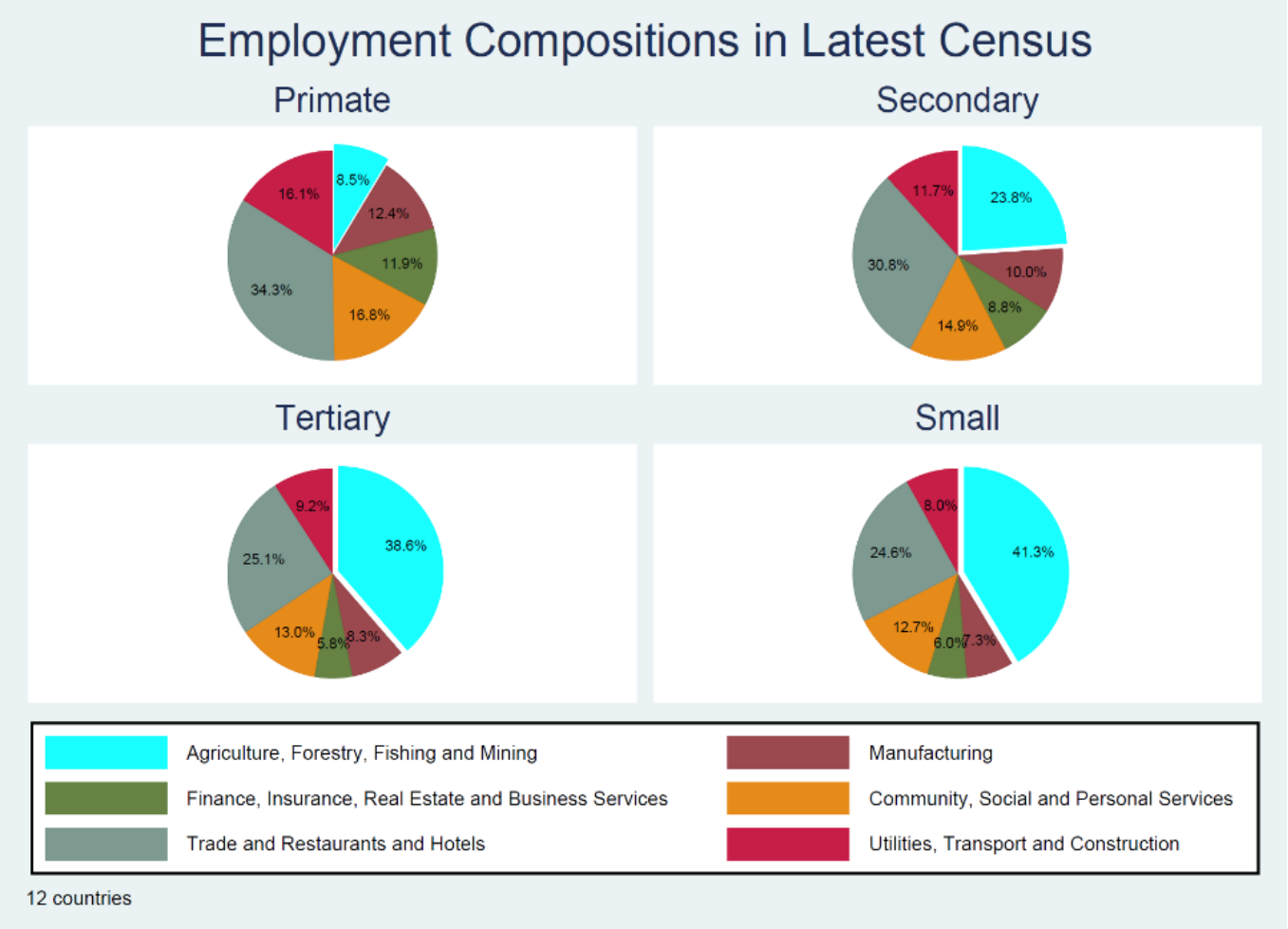
Source 1: UN World Urbanisation Prospects

Figure 2: Growth of Cereal Yields (tonnes per hectare)



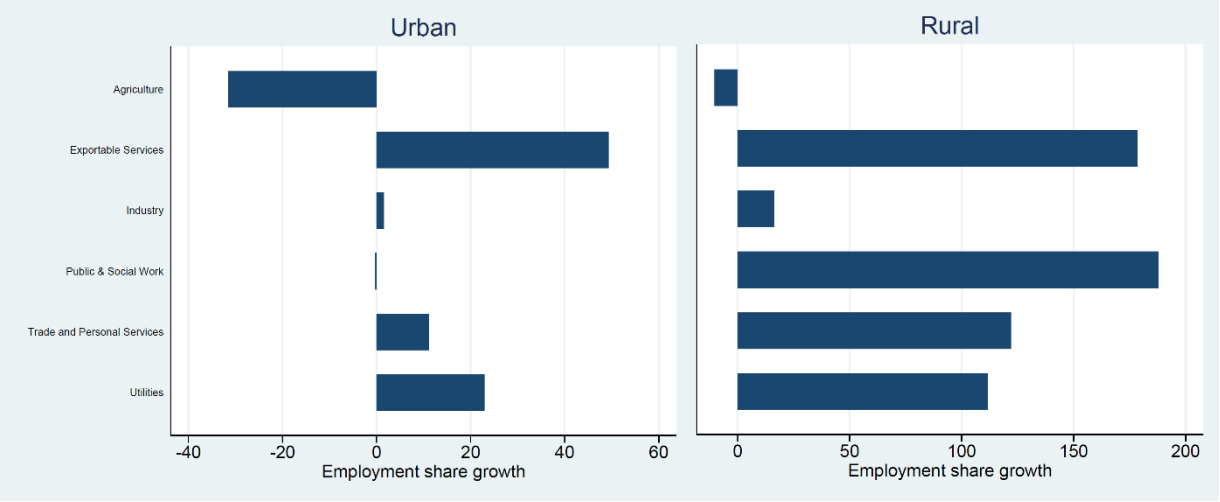
Source 2: World Development Indicators

Figure 3: Employment Compositions across African Cities



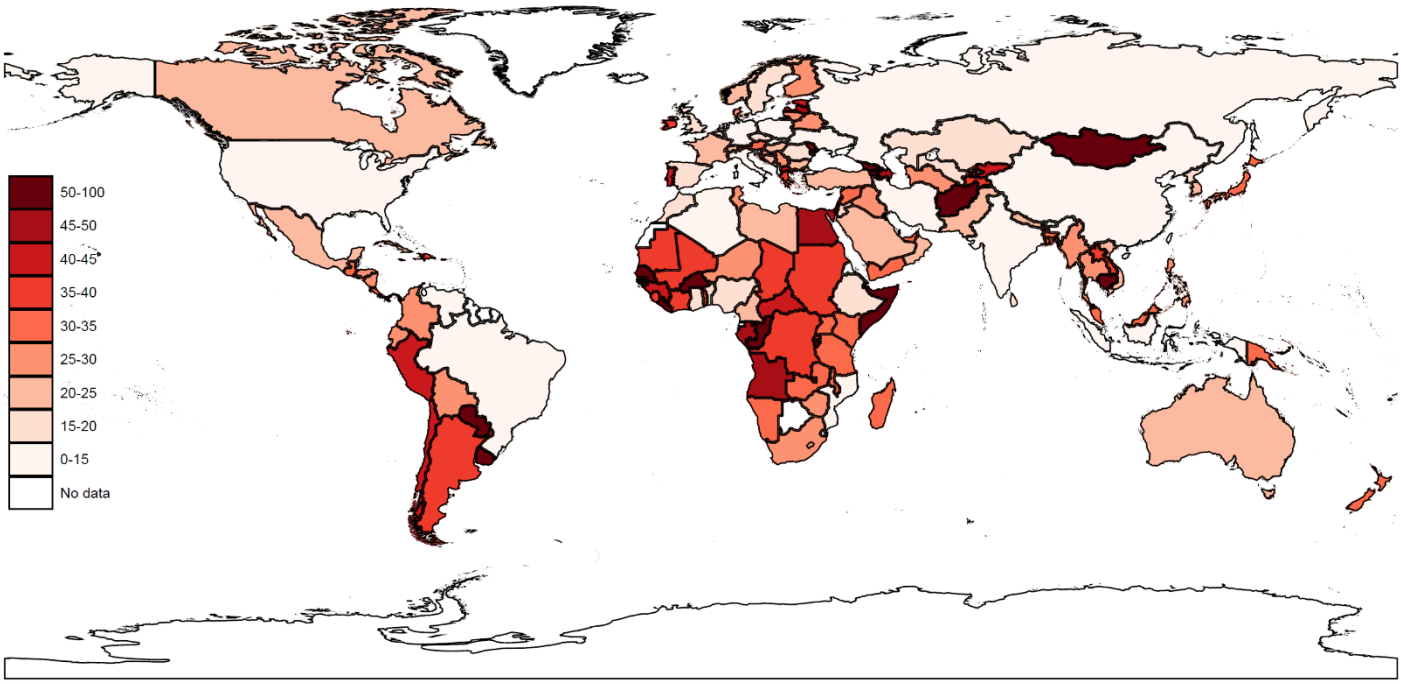
Source 3: Author's own constructed dataset

Figure 4: Change in Employment Composition across 6 African Countries



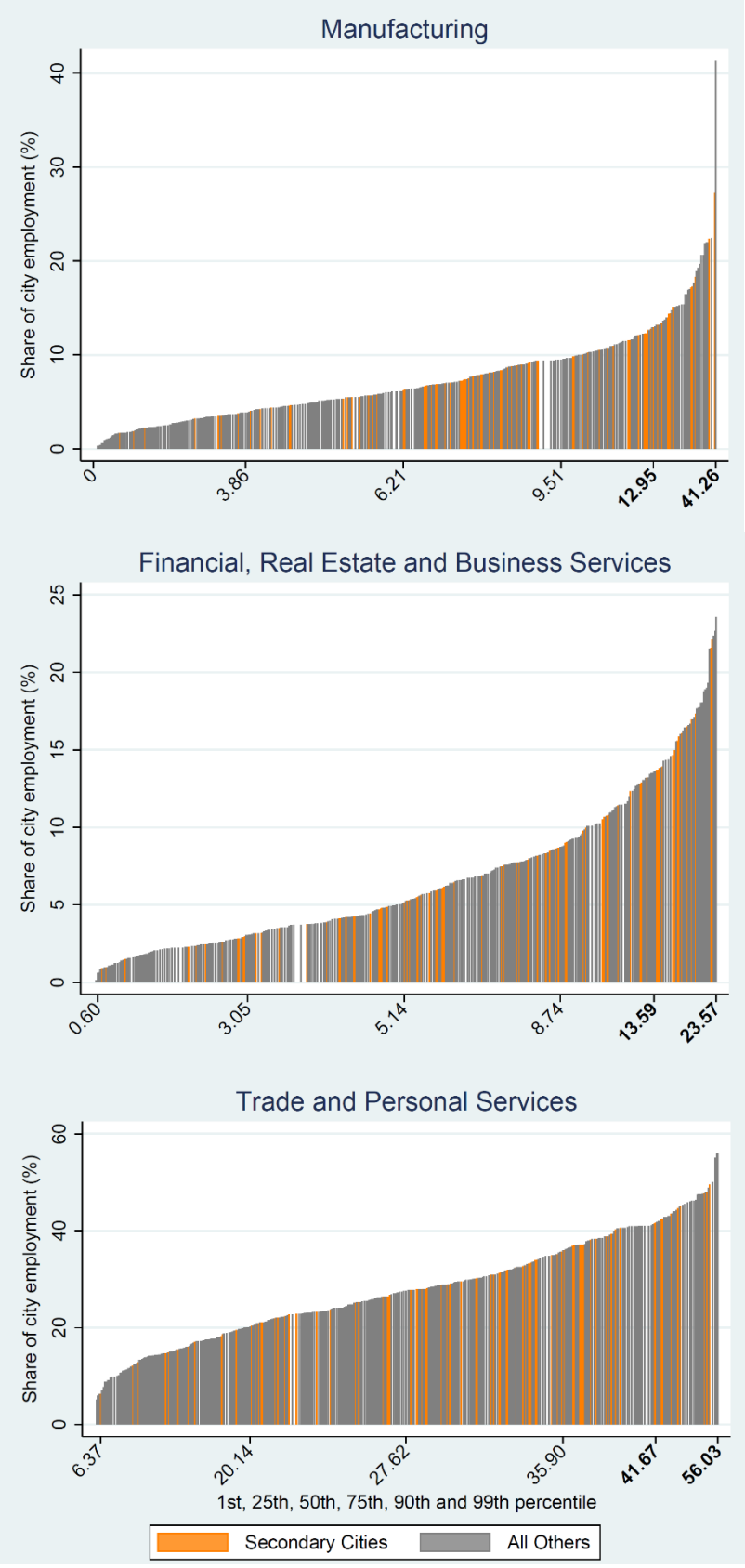
Source 4: IPUMS-International. Country-years: GHA (2000-2010), MWI (1998-2008), MWI (1998-2008), MLI (1998-2009), TZA (2002-2012), ZAM (1990-2000)

Figure 5: Primacy in 2015 across World Countries



Source 5: World Development Indicators

Figure 6: Share of Total City Employment by Industry



Source 6: Author's own constructed dataset

The Development of the African System of Cities: ONLINE APPENDIX

J. Vernon Henderson and Sebastian Kriticos

ABSTRACT

This document provides further results and detail on the data sources used for the paper "The Development of the African System of cities," by Henderson and Kriticos (2018). Appendix A provides results and Appendix B is on data and methodologies.

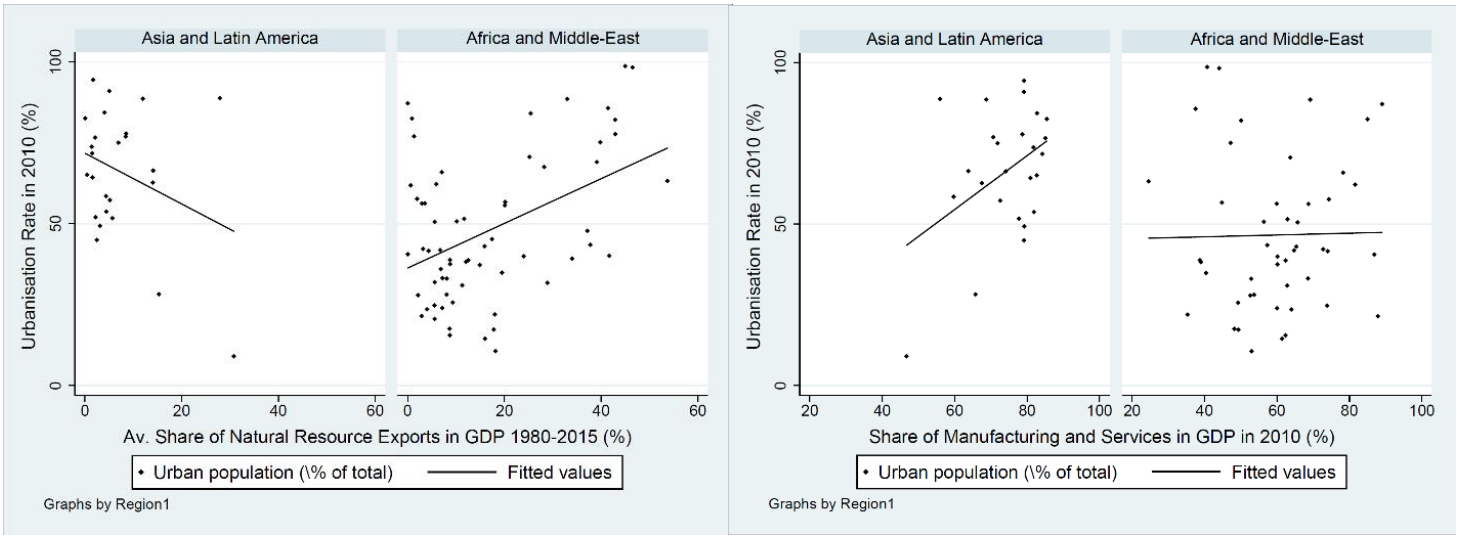
Appendix A: Results

Overview

In this section we provide a set of figures and tables that were excluded from the main text.

A.1 Figures

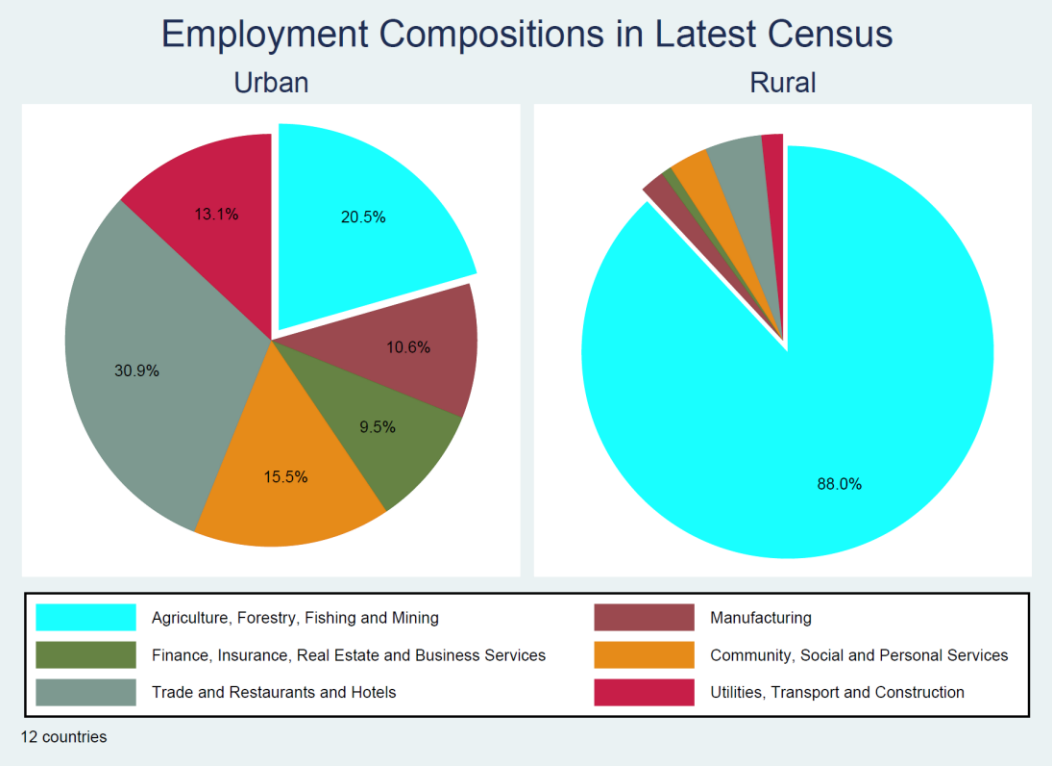
Figure A1: Urbanisation vs Natural Resource share in GDP, and Urbanisation vs Manufacturing and Services in GDP (replication of Gollin et al (2016)).



Source: World Development Indicators

Figure A2: Employment compositions by urban-rural split.

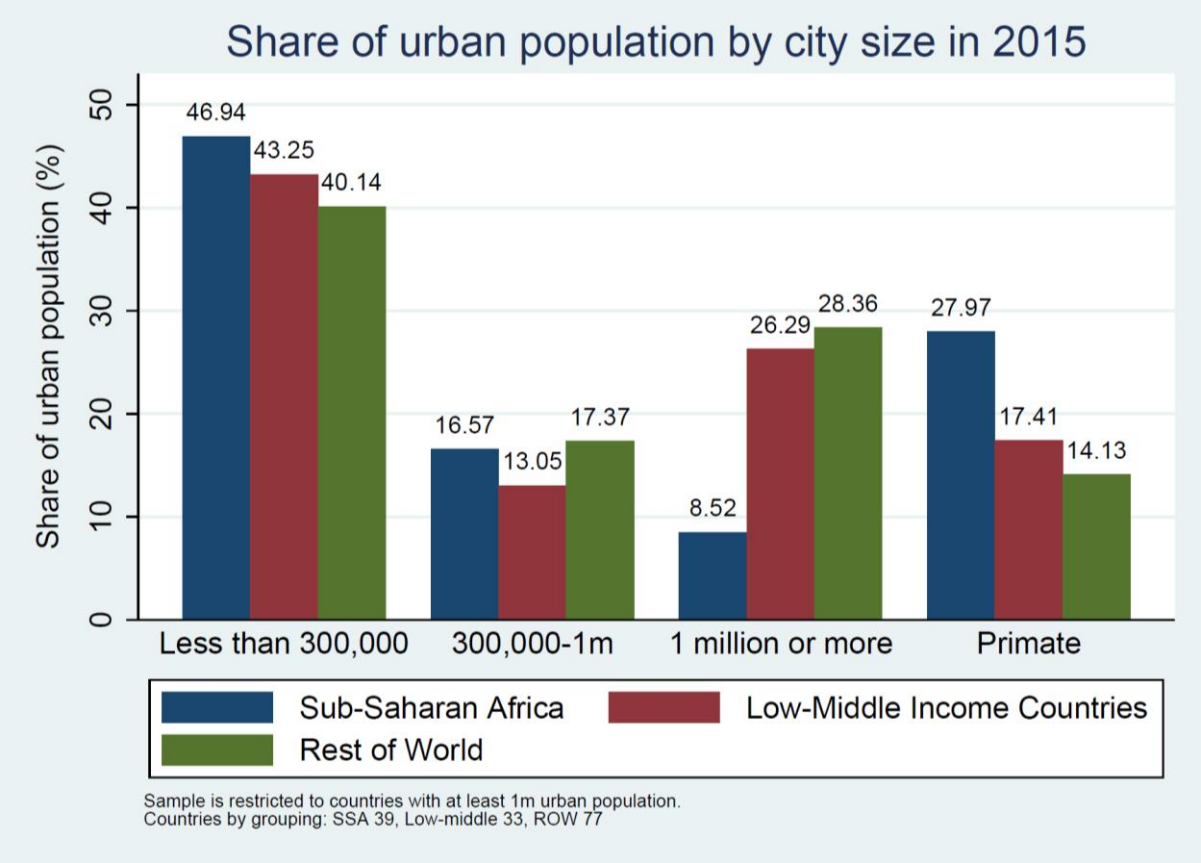
We also looked at how employment compositions varied by urban and rural status rather than looking at the differences across the urban hierarchy as in Figure 3 of the main text. Once again, we see that the primary sector plays a large role in urban employment at the city-level, at 20.5%. Again, the role of tradable service employment and manufacturing industry is very slight at around 9-11%. In the rural economy, agriculture practically forms the entire basis of employment, at nearly 90%, with small trade work again being the second largest sector, followed by community and social work and then manufacturing.



Source: Author's own constructed dataset

Figure A3: Population distribution by city size and region.

As described in section 4.1 of the main text, we looked at the share of population across the city-size distribution by regions using data from the UN World Urbanisation Prospects. As mentioned, Africa appears to have more outliers on each side of the distribution, having both the highest share of population in the largest cities, and the highest shares in cities under 300,000. The area that we place greatest focus on is Africa's starkly lower proportion of citizens in cities with over 1 million population, which in absolute terms, many deem to be secondary cities.



Source: World Urbanisation Prospects

Figures A4 and A5:

As discussed in Section 4.2 of the main text, we looked at how employment composition was changing across census periods for 6 countries in Figure 4 of the main text. These countries are Ghana, Malawi, Mali, Mozambique, Tanzania and Zambia. The results showed that exportable services were the sector recording the fastest growth in relative employment share. In the data, our employment in the exportable services category was the sum of employment from two different categories in the IPUMS data. These are: (1) Financial Services and Insurance Services, (2) Real Estate and Business Services.

Here we break out the exportable services category into its two distinct categories and look at inter-census growth both in the share of each category in the exportable services total, and also the growth in employment that each sector experienced. Figures A4.a and A4.b look at urban areas, while Figures A5.a and A5.b look at rural. Again the country years are as follows: Ghana (2000-2010), Malawi (1998-2008), Mali (1998-2009), Mozambique (1997-2007), Tanzania (2002-2012), Zambia (1990-2000).

Figure A4.a Inter-Census Change in Urban Employment Share for Exportable Services

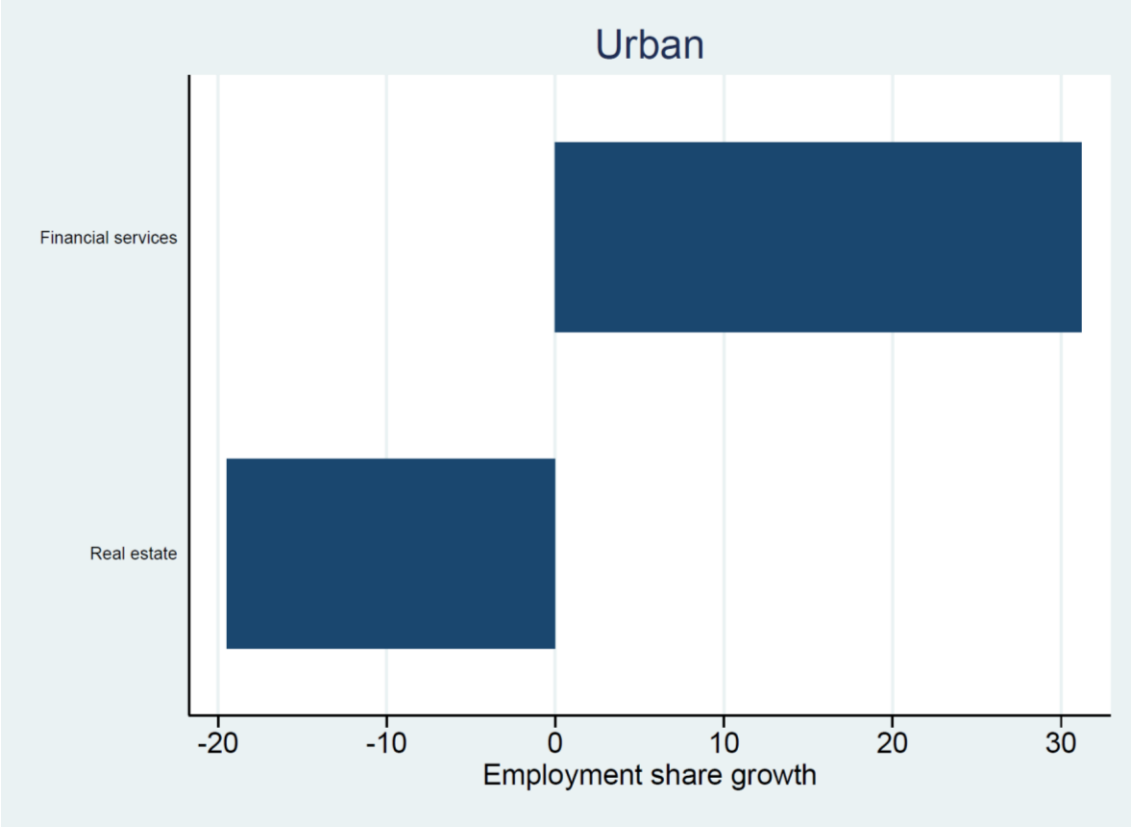


Figure A4.b Inter-Census Change in Urban Employment in Exportable Services

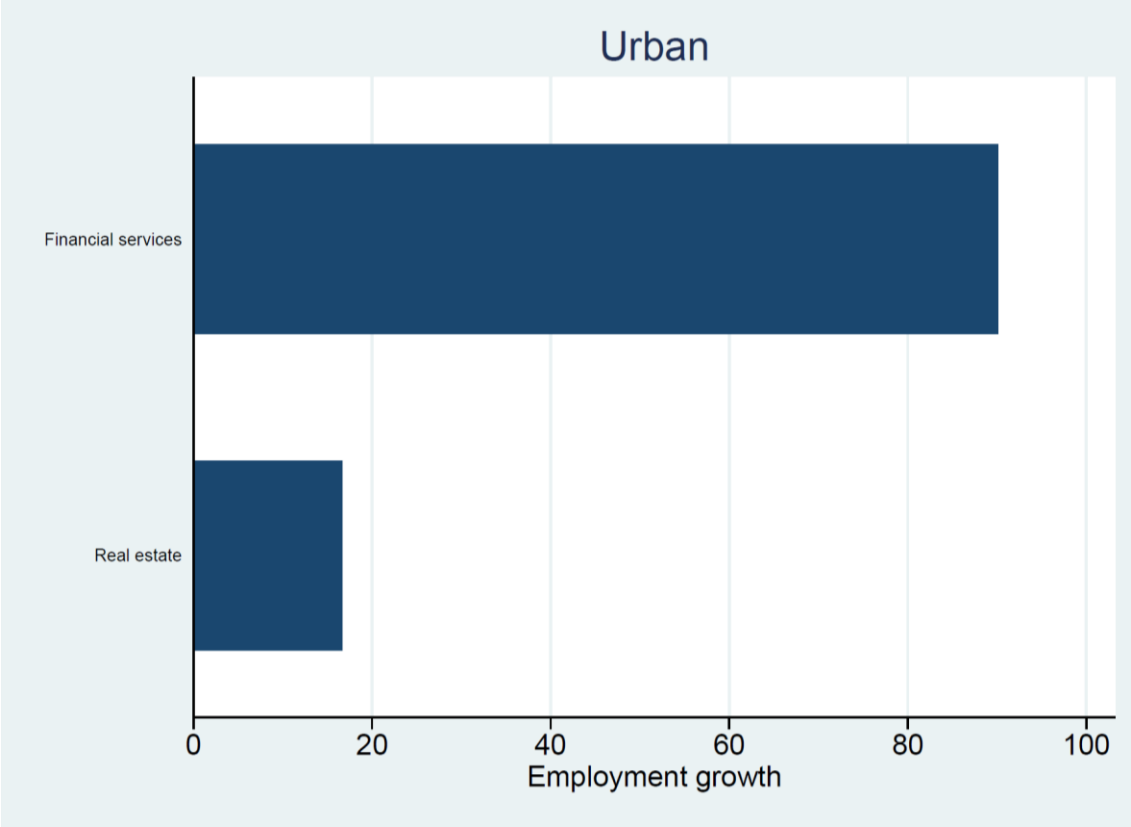


Figure A5.a Inter-Census Change in Rural Employment Share for Exportable Services

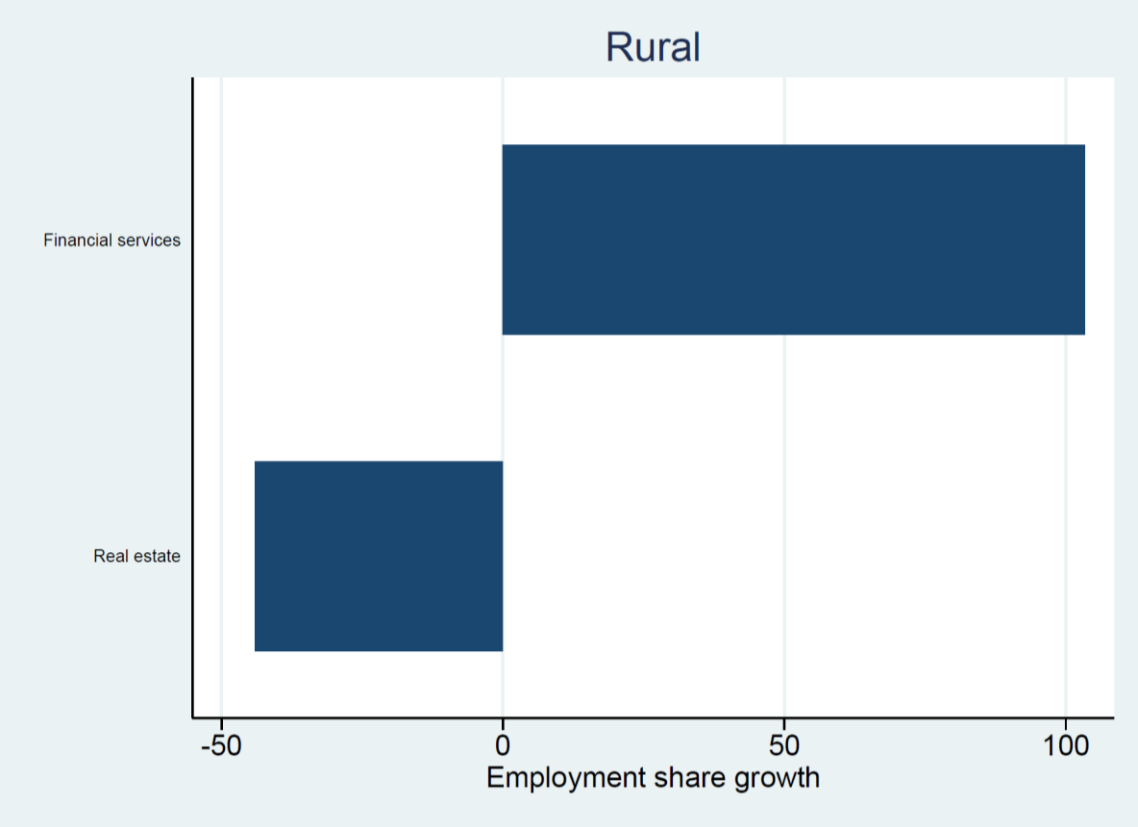
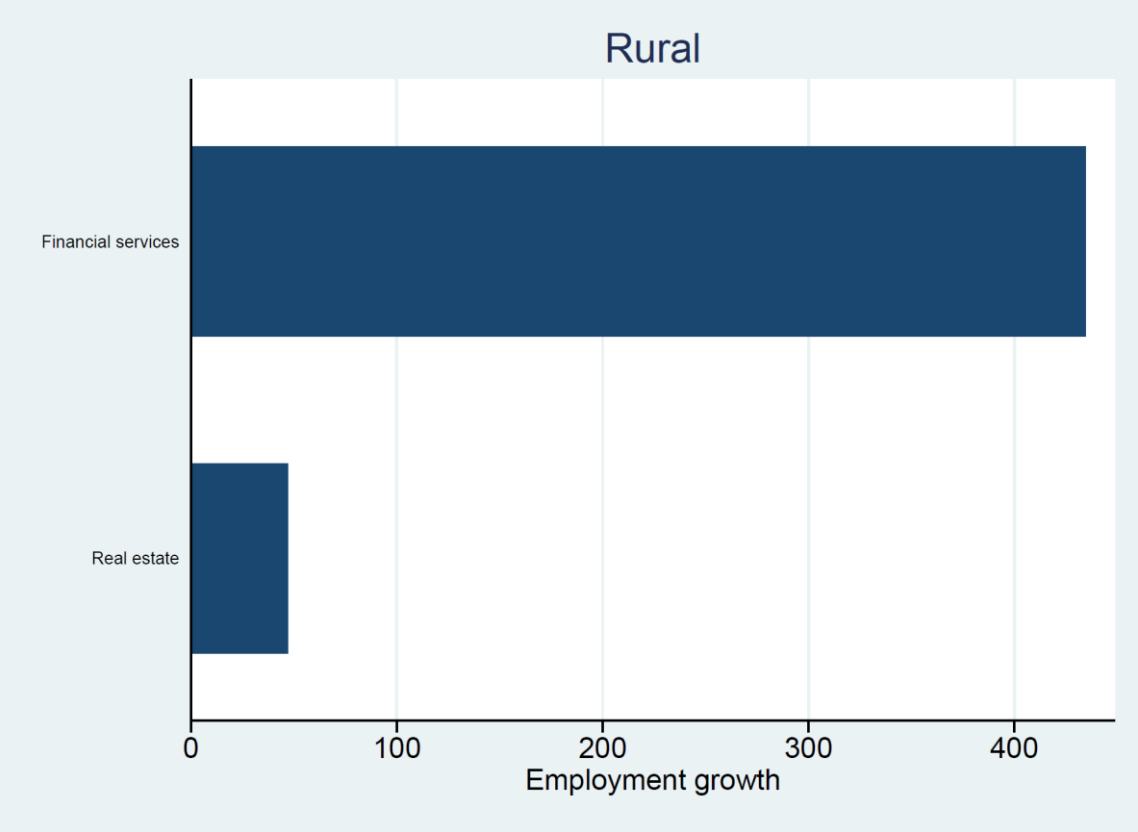


Figure A5.b Inter-Census Change in Rural Employment in Exportable Services



A.2 Tables

Table A1: Estimation of household rents in Uganda

	Ln(Monthly House Rent)	
	(1)	(2)
Primate city	0.491***	0.484***
	(0.0672)	(0.0683)
Urban	0.558***	
	(0.0535)	
Secondary		0.472***
		(0.0736)
Tertiary		0.638***
		(0.0833)
Small		0.562***
		(0.0667)
Constant	8.425***	8.457***
	(0.801)	(0.803)
Observations	3475	3475
R-squared	0.576	0.576
Year FE	Yes	Yes
Region FE	No	No

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(4) is the natural logarithm of total household monthly rent on housing. We control for wall, roof,and floor material as well as toilet type and water source. Columns (1)-(2) are for Tanzania, (3)-(4) Uganda, and (5)-(6) Nigeria. Robust standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: World bank Living Standards Measurement Surveys

Table A2: Estimation of the household urban net income premium for landless vs landed households

	Tanzania		Uganda		Nigeria	
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Landless in agriculture						
Primate City	1.112***	0.931***	0.721***	0.324***	1.017***	0.977***
	(0.115)	(0.116)	(0.0670)	(0.0713)	(0.0720)	(0.0707)
Urban	0.481***	0.256**	0.704***	0.465***	0.568***	0.435***
	(0.106)	(0.105)	(0.0561)	(0.0566)	(0.0432)	(0.0443)
Constant	0.418	0.888**	1.116***	2.489***	2.376***	2.375***
	(0.377)	(0.377)	(0.226)	(0.354)	(0.0989)	(0.100)
Observations	2465	2465	5805	5386	13099	12030
R-squared	0.215	0.311	0.410	0.477	0.058	0.129
Panel B: Landed in agriculture						
Primate City	2.254***	1.831***	0.237	0.595	0	0
	(0.242)	(0.234)	(0.423)	(0.449)	(.)	(.)
Urban	0.595***	0.388***	0.304***	0.361***	0.805***	0.324***
	(0.0830)	(0.0837)	(0.0766)	(0.0738)	(0.0907)	(0.0896)
Constant	2.135***	2.109***	1.690***	3.008***	0.351	0.720**
	(0.261)	(0.258)	(0.369)	(0.442)	(0.275)	(0.284)
Observations	4344	4344	4274	4137	5956	5900
R-squared	0.159	0.236	0.209	0.279	0.039	0.167
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Occupation FE	No	Yes	No	Yes	No	Yes

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(6) is the natural logarithm of total net household monthly income. Controls are education recorded or not, if recorded level of education, age of hh head, age squared, gender, hh size, hh size squared. Panel A limits the sample to households that do not own land, and Panel B limits the sample to households with land. Columns (1)-(2) are for Tanzania, (3)-(4) Uganda, and (5)-(6) Nigeria. Robust Standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.

Source: World Bank Living Standards Measurement Surveys

Table A3: Estimation of the rate of primacy in 2015

	Primacy		
	(1)	(2)	(3)
SSA dummy	5.485*	-0.0881	8.671*
	(2.801)	(3.666)	(4.811)
ln(GDP_pc 2005)		-2.134**	-0.483
		(1.042)	(1.173)
ln(Population 2005)		-5.519***	-6.643***
		(1.357)	(0.985)
ln(Land area sqkm 2005)		-1.519	-1.560
		(1.640)	(1.299)
Asia			15.51***
			(5.364)
LAC			8.569**
			(3.502)
MENA			-0.599
			(3.884)
Constant	30.15***	158.4***	157.6***
	(1.701)	(16.84)	(15.80)
Observations	121	120	120
R-squared	0.024	0.363	0.461

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(3) is the rate of primacy in 2015. The sample is limited to countries with over 1 million urban population and a land size over 500 sqkm. Robust standard errors are presented in parentheses with asterisks indicating *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$.
Source: World Development Indicators (2017)

Table A4: Estimation of city population in latest census

	Ln(City Population)	
	(1)	(2)
Coastal	-0.185	-0.442**
	(0.197)	(0.199)
Harbour dummy	0.680***	0.414**
	(0.208)	(0.205)
Distance to coast	-0.0580***	-0.0276*
	(0.0165)	(0.0164)
Primate	3.598***	1.073***
	(0.192)	(0.312)
Distance to primate	-0.0149	-0.0337*
	(0.0193)	(0.0204)
Distance coast if landlocked (100km)	0.0474	0.0154
	(0.0308)	(0.0361)
Ruggedness index	-0.00000101	-0.00000205
	(0.00000897)	(0.00000946)
Ln(Elevation)	0.0370	-0.00852
	(0.0452)	(0.0447)
Avg. Rainfall (monthly 1960-90)	0.00312***	0.00391***
	(0.000905)	(0.000903)
Has Industry (1965)		0.794***
		(0.124)
Number of Industries		0.144***
		(0.0357)
Constant	10.15***	10.30***
	(0.275)	(0.275)
Observations	1439	1227
R-squared	0.345	0.416
Country-Year FE	Yes	Yes

Notes: Table reports results from OLS regressions, the dependent variable in columns (1)-(2) is the natural logarithm of city population and the dependent variables in columns (3)-(4) is the natural logarithm of city night light intensity. Robust standard errors are presented in parentheses with asterisks indicating *** p < 0.01, ** p < 0.05, * p < 0.10.

Source: Author's own constructed dataset

References for Appendix A

Gollin, D., Jedwab, R., and Vollrath, D. (2016), Urbanisation with and without Industrialisation. *Journal of Economic Growth*, 21(1), 35-70.

Appendix B: Data

Overview

In the following sections we provide further details on the data sources used in the paper. Section B.1 and its sub-parts describe the data from the Integrated Public Use Microdata Series (IPUMS) and the methodology we use to map this district-level data to the city-level in our analyses of urban and rural employment compositions in Section 2 of the main paper. Section B.2 describes the Living Standards Measurement Surveys which are used to study urban premia in Section 3 of the main paper. Finally, section B.3 describes the population dataset we compiled to study city sizes and growth in Section 4 of the main paper.

B.1 IPUMS-International Data

The Integrated Public Use Microdata Series (iPums) is a US-based organisation that collects, consolidates, and disseminates micro-data from economic censuses. These data are collected for a number of developed and developing countries in conjunction with the national statistical agencies in the country in question. In Section 2 of our paper, we make use of census data from 12 of these countries; from the most to least populous these are: Ethiopia, Tanzania, Uganda, Mozambique, Ghana, Cameroon, Mali, Malawi, Zambia, Sierra Leone, Liberia, and Botswana. Although we have relevant data for Sudan and South Sudan, we always exclude them from our analyses due to their distinct nature as a principally Sahel country and the troubles with ongoing civil war.

We ended with this subset of 12 African countries because they had available data needed for our analyses of urban employment. These variables of interest, as listed on the IPUMS website, are called INDGEN and URBAN and respectively, they define the industry in which a survey respondent works (such as agriculture or manufacturing – and whether the respondent lives in an urban or rural area as defined by the national census.

In Table A1 below we list all the countries within our constructed dataset and the years in which a census was conducted. The emboldened years are the ones in which both industry (INDGEN) and urban-rural status (URBAN) data was available from IPUMS. For our cross-sectional analyses in the paper, for instance in Figures 3 and 6, we always use the latest census period available. For our analysis of growth across census-periods in Figure 4, we are only able to include the subset of countries which have at least two census periods with the relevant data available (these are: Ghana, Malawi, Mali, Mozambique, Tanzania and Zambia).

Table B1: Countries and Census Years for Dataset in Section 2

	Early Census	Middle Census	Late Census
Botswana	**1991**	2001	2011
Cameroon	1987	**2005**	-
Ethiopia	-	**1994**	2007
Ghana	1984	**2000**	**2010**
Liberia	-	-	**2008**
Malawi	**1987**	**1998**	**2008**
Mali	1987	**1998**	**2009**
Mozambique	-	**1997**	**2007**
Sierra Leone	-	**2004**	-
Tanzania	1988	**2002**	**2012**
Uganda	1991	**2002**	-
Zambia	**1990**	**2000**	2010

Notes: Countries and census years included in the dataset for section 2. Emboldened country years are cases where there is both industry and urban data available to base the analyses in section 3.

B.1.1 Definitions of Urban in the Censuses

As each national statistical office provides slightly varying definitions of urban areas, we provide full details of the urban statistics as given by each census provider in Table A2 below.

Table B2: Urban-rural Status as defined by Country Censuses

Botswana	The United Nations defined urban places as localities with an agglomeration of 5,000 or greater population where at least 75 percent of economic activity is not agricultural.
Cameroon	A city is defined as an entity which has either an administrative function (i.e. headquarters of a district), or has a population of at least 5,000 inhabitants and the following facilities: a complete primary school, a developed health centre, water supply and electricity facilities, and a daily market. A village is a traditional unit headed by a third class chief, and made up of one or more localities or towns (an inhabited site with boundaries identified by a name and recognised by the people).
Ethiopia	Urban areas are defined as localities in which Urban Kebele Administration has been established. Localities with 1,000 or more people whose inhabitants were engaged primarily in non-agricultural activities and capitals of Weredas were also considered urban, irrespective of whether Kebele Administration was established.
Ghana	Urban places are localities with 5,000 or more population.
Guinea	Urban areas are administratively defined. Administrative centres of prefectures are considered urban; the rest of the territory is classified as rural.
Liberia	The samples do not provide a definition of rural and urban. According to the UN Demographic Yearbook for 1974 and 2008, urban is defined as localities of 2,000 or more inhabitants.
Malawi	Malawi defined an urban area as all townships and town planning areas, and all district centres.

Mali	Urban areas are localities with 5,000 inhabitants or more and district centers.
Mozambique	No definition of urban and rural is provided for the Mozambique samples.
Senegal	No definition of urban and rural is provided for the Senegal samples.
Sierra Leone	Towns with a population of 2,000 or more inhabitants are defined as urban.
Tanzania	According to the UN, urban areas in Tanzania are composed of 16 gazetted townships.
Uganda	According to the UN, urban areas in Uganda are composed of cities, municipalities, tons, town boards and all trading centers with more than 1,000 inhabitants.
Zambia	The samples do not provide a definition of rural and urban. According to the UN Demographic Yearbook for 1990 and 2000, urban areas are defined as localities of 5,000 or more inhabitants, the majority of whom all depend on non-agricultural activities.

A key feature of the IPUMS data, is that all information is geo-located – either at the first or second administrative unit in which the household is enumerated – and given this urban-rural identification. This allows us to assign demographic and employment related data collected by IPUMS to city-level statistics using the urban-rural identifier alongside further data on the share of night light intensity emitted by urban areas within districts.

B.1.2 Defining Urban Areas from Night Light and Population Data

We define and construct our city data based on night light data from the US Defence Meteorological Satellite Programme (DMSP). This dataset assigns a digital value to each $1km^2$ grid pixel on the ground based on the average intensity of light emittance for each year in 1992-2012. We define the boundary of a city by a cluster of contiguously lit pixels which form a unique polygon that is surrounded by unlit/dark pixels.

As these light polygons will vary over time for reasons that include both genuine city expansion and temporary events, such as road construction, we need to strike a balance between having a boundary definition that is too large, merging several distinct cities together, and one that is too tight and excludes large parts of the city. To do this, we take the union of lit polygons from the five years 2008-2012, when the polygons are largest, but exclude all pixels that have been lit in only one of those years. We consider the remaining clusters to be consistently lit light unions which are most likely to be cities in the present day.

Final adjustments are made for cases where the light emitted by a given area spreads over water due to the possibility of water reflection. In these cases, we clip boundaries to the point where water begins such as coastlines or river boarders. In addition, light unions that go across country boundaries are separated along the boundaries to become distinct cities on either side of the boundary.

Once our boundaries are defined, we include only the set of light unions that have a recent population estimate. Our population estimates a drawn from Citypopulation.de (census-data), a website documenting population figures and locations for all communes, settlements, and semi-urban settlements across countries and time. When our light unions overlap with populated locations, both datasets are consolidated to form a dataset which defines the size, shape, population, and light intensity for cities in our 12 countries. These light unions are then used as a basis to map district-level data from iPums to city-level statistics.

B.1.3 Mapping iPums Data to the City-Level

From the iPums data, we can use the urban-rural status of recorded respondents to separate the urban and rural economy at the district-level. We then assign district-level urban data to our light unions based on the proportion of the total district night light intensity that is emitted by each distinct light union within that district. Specifically, we convert satellite images from the 1992/3 period to a raster quantifying light intensity within each of our unions as well as the totals in each of our districts. We choose to focus on light emissions in the earliest period available for which we have satellite imagery (1992/3) because this 'fixed stock' of light shares prevents cities from moving within the urban hierarchy over time based on their light emissions, and also limits our sample to the set of cities that were already established in the early period.

These shares are then used as the basis to map census information from iPums such that all urban data is assigned to some city within the district. Our notion is that the share of district lights emitted by each urban agglomeration gives us a crude proxy of the relative weight of importance that the urban area has in terms of population and employment density within its district and is therefore, a useful method to assign total district data to more disaggregated urban areas.

B.2 Living Standards Measurement Surveys

The Living Standards Measurement Surveys have been conducted in a number of developing countries by the World Bank and the national statistical offices of the country in question. In Section 3 of our paper where we study urban premia, we make use of surveys on Tanzania, Uganda and Nigeria. All of these surveys are considered representative of households at the national level, as well as urban/rural and major ecological zones of the countries. In Table A3 below, we provide a list of the surveys we use and the number of households surveyed in each of these rounds.

Table B3: LSMS Surveys

Country	Survey	Year	Sample Size
Tanzania	Panel Household Survey	2008	3,280
		2010	3,924
Uganda	National Panel Survey	2009	3,123
		2010	2,716
		2011	2,716
		2012	3,119
Nigeria	National Household Survey	2010	5,000
		2012	5,000

In each sample, a two-stage probability sampling methodology is used. In the first stage, "Primary Sample Units" (PSUs) are selected based on the probability proportional to size of all of the enumeration areas in geographic zones in the country. In the second stage, households are

then selected randomly from each PSU, after which, each individual within a household is surveyed. All of the LSMS surveys are publicly available for download from the World Bank website, so for further information on any individual survey and its methodology, we refer the reader to the information documents provided by the World Bank.

Although the contents of each survey vary, they all have quite consistent data at the household and individual level on aspects such as income, educational attainment, demographics, labour allocation, asset ownership and dwelling characteristics, as well as urban-rural household identifiers. Agricultural households report on various aspects of farming such as crop choice, inputs on the farm, labour usage and the types of land allocation such as harvesting, grazing or fallow. Amongst non-agricultural households, additional modules are provided on whether they are self-employed with their own business and if so the revenues and various factor costs of that business. In some cases, aggregation of revenues and costs at the household level is already computed in the survey and these aggregations are used where possible. For example, labour income at the individual level is already aggregated in the surveys to include all wage, in-kind and bonus income from all jobs. Elsewhere, input costs of agricultural and non-agricultural businesses are aggregated to the household level.

We calculate income from the survey data and aggregate either to the individual or household level (depending on our analysis) using all available sources of money flowing in. Letting i index an individual or household, this can be summarised as follows:

$$Y_i = \sum_i y_i^{SE} + \sum_i y_i^L + \sum_i y_i^K$$

where y_i^{SE}, y_i^L, and y_i^K represent self-employed income, labour income, and capital income respectively. Households reported receipts of incomes through various forms and over various time intervals. The variables used for income receipts and the time intervals over which they were received are reported as follows:

Table B4: Income sources and time intervals in LSMS surveys

Income Source	Time Interval
Last payment in cash	Hour, Day, Week, Fortnight, Month, Quarter or Year
Last payment in kind (value in LCU)	Hour, Day, Week, Fortnight, Month, Quarter or Year
Net income from business	Week or Month
Remittances in cash	Year
Remittances in kind	Year
Rent of property	Year
Private or govt pensions	Year
Domestic remittances	Year
Rent of farmland	Year or cropping season
Sales of crops	Year or cropping season
Sales of crop residue	Year or cropping season
Sale of livestock products	Year or cropping season

All revenues are converted to a monthly interval. In cases where incomes are reported over the year, quarter, fortnight, or week, the variables are scaled to a monthly value simply by multiplying by the ratio of a month to the time interval in question (for instance a quarter is multiplied by 1/3 to be monthly). In cases where the last income payment is reported based on a day of work, the figure is multiplied by the average days the respondent reports to work each week, and then multiplied by 52/12 to be a roughly monthly figure. In cases where the last income payment is reported based on an hour of work, the figure is multiplied by the average hours the respondent reports to work each day, then the average days they report to work each week, and finally by 52/12 to get a roughly monthly figure.

A similar method is used to convert expenses to a monthly aggregate figure. The reported expenses in the household surveys are as follows:

Table B5: Expenses and time intervals in LSMS surveys

Expense	Time Interval
Wages	Month
Raw materials	Month
Other expenses	Month
Farm inputs	Year
Additional agricultural expenses	Year

We conduct analyses both at the household and individual level, subtracting expenses from revenues to get our net income figure. For the individual analysis we choose to restrict our sample to the set of individuals in households that do not own any agricultural land because farm income was recorded at the household level in the surveys and thus, we have no way to assign farm income to individuals. We look at adults aged 18-65 who are working part or full time with income in our analysis at the individual level.

At the household level, we construct measures of income including income from self-employment, labour, capital and land income. All income is measured before taxes. We choose to include in the calculation of monthly household income, transfer payments such as remittances, gifts and pensions and we subtract transfer payments flowing out of the households. These sources of incomes are likely to be important for the budget constraints of households, particularly in rural communities so they are important in our study of urban-rural income disparities, although we note that our results are robust to excluding remittances.

B.3 Population and Geographical Data

In Section 4 we study population size and growth across African cities. We draw population data from two sources: Citypopulation.de (census-data for 35 countries) and Africapolis (for Nigeria). Citypopulation.de is a website that collects population statistics for countries, administrative areas, cities and agglomerations. In all cases for our countries used, estimates are based on census results and official estimates. However, we rely on the Africapolis database for Nigerian population information, because Nigerian census results for 2006 are disputed and not fully available. In this database, the city population numbers are derived by comparing urban

areas in Nigeria to areas in the neighbouring countries with similar urban morphology, whose population and density is known. Then the density from these cities is applied to the cities in Nigeria and the population is derived from this density. The extents of Nigerian cities are created from satellite images. More on the Africapolis methodology can be found here (http://www.e-geopolis.eu/article237.html).

Table 4 provides the countries and census years for which we have available population data.

Table B6: Countries and Census Years with Population Data

Country	Early Census	Middle Census	Late Census
Angola	1990	-	-
Benin	1992	2002	-
Botswana	1991	2001	2011
Burkina Faso	1985	1996	2006
Burundi	1990	2008	-
Cameroon	1987	2005	-
Central African Republic	1988	2003	-
Chad	1993	2009	-
Cote d'Ivoire	1988	1998	-
Dem. Rep. Congo	1984	2004	-
Djibouti	-	-	2009
Equatorial Guinea	-	1994	-
Eritrea	1984	1997	-
Ethiopia	-	1994	2007
Gabon	1993	2003	-
Gambia	1993	2003	-
Ghana	1984	2000	2010
Guinea	-	1996	-
Kenya	1989	1999	-
Lesotho	1996	2006	2011
Liberia	-	-	2008
Madagascar	-	1993	-
Malawi	1987	1998	2008
Mali	1987	1998	2009
Mauritania	1988	2000	2013
Mozambique	-	1997	2007
Namibia	1991	2001	2011
Niger	1988	2001	2012
Nigeria†	1991	2000	2010
Republic of Congo	1984	2007	-
Rwanda	-	1991	2012
Senegal	1988	2002	-
Sierra Leone	-	2004	-
Swaziland	1986	1997	-
Tanzania	1988	2002	2012
Togo	-	-	2010

Uganda	1991	2002	-
Zambia	1990	2000	2010
Zimbabwe	1992	2002	2012

† *Data is not census-based*

B.3.1 Harmonising Data with Light Unions

All countries from the Citypopulation.de dataset are appended to one dataset and geocoded with Google-Maps geocoder. Missing coordinates are replaced with OpenStreetMaps geocoder or by hand search online using a variety of sources such as Wikipedia (or /http://tools.wmflabs.org/geohack/). Some evident errors (from similar names that yielded duplicate coordinates, or by identifying large cities that were not falling into a light union) were also corrected by hand. Nigerian cities were available with coordinates in the Africapolis database.

The cities from our population database get associated with the light unions if they intersect them within a 5km buffer, which we use to accommodate for coordinate imprecision and making sure the smaller unit satellite settlements are also included, as they are likely to be part of the labour market of the light union. If there are multiple unions that are within a 5km radius of the populated location, then the city is attached to the nearest union. The light unions and the coordinates that do not get attached to any union within 5km are dropped (15.5% of city points, half of which are small Nigerian villages from the Africapolis database, the other half are small cities from the remaining 38 countries.

The population for each city point is added for all points associated with one light union. Thus, our final unit of observation is a lit urban area aggregating all city points inside it within a 5km bugger. In total, we have 2068 urban areas with some population information across 41 cities.

B.3.2 Other Data.

For the regressions in Table 8 of the main text, we also combine our population statistics with a range of geographical covariates as well as information on early industrialisation. In Table A7 below, we outline the sources and definitions for the variables used in our regressions.

Table B7: Variables, Definitions and Sources for Covariates in Table 8 of Main Text

Variable	Definition	Source
Population	Number of individuals within a city.	Citypopulation.de
Annual Population Growth	Calculated as the log difference of population between two census periods, divided by the number of years between both censuses.	Citypopulation.de
Coastal	Indicator variable for whether the city is based on the coast.	Calculated by author in ArcGis
Distance to Coast	Euclidean distance from the urban centre (defined as the most intensely lit area) to the coast.	Calculated by author in ArcGis
Distance to Primate City	Euclidean distance from the urban centre (defined as the most intensely lit area) to the primate city.	Calculated by author in ArcGis

Ruggedness Index	Measure of average city ruggedness based on Nunn and Puga (2012)	Nunn and Puga (2012)
Elevation	Average elevation, meters above sea level	ESRI
Has Industry	Indicator for whether in 1960, an urban area had any of 26 different manufacturing industries as defined by the *Oxford Regional Economic Atlas for Africa* (Ady, 1965)	Ady (1965)
Number of Industries	Number of distinct manufacturing industries an urban area has in 1960 as defined by the *Oxford Regional Economic Atlas for Africa* (Ady, 1965)	Ady (1965)
Regional Capital	We choose the set of regional capitals that was in place in 1992 as defined by the website Statoids, and we can think of these as a group of cities in which various political economy mechanisms were operating.	Statoids

References for Appendix B

Ady, Peter. (1965). Oxford Regional Economic Atlas Africa. Oxford: Clarendon Press.

Klein, W. D., Mandt, G. A., and Gagliardo, J. (1992), Defense Meteorological Satellite Program. *6th Conference on Satellite Meteorology and Oceanography* (pp. 450-452).

Nunn, N., & Puga, D. (2012). Ruggedness: The blessing of bad geography in Africa. *Review of Economics and Statistics*, *94*(1), 20-36.

Steven Ruggles, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. *Integrated Public Use Microdata Series: Version 7.0* [dataset]. Minneapolis: University of Minnesota, 2017. https://doi.org/10.18128/D010.V7.0.