Agglomeration and firm-level productivity in African cities

Justin Visagie and Ivan Turok

Working paper 931

May 2026

Agglomeration and firm-level productivity in African cities

Justin Visagie[1] and Ivan Turok[2]

Abstract

Africa's economic prospects are closely tied to the performance of its cities. Yet in many African countries, the expected productivity gains from urbanisation are stifled by chronic infrastructure deficits, governance weaknesses and market inefficiencies. This study examines the relationship between agglomeration and firm-level productivity covering 157 cities across 34 African countries, using comparable cross-country data from the World Bank Enterprise Surveys and the Global Human Settlement Layer. The findings suggest that agglomeration economies are present but the relationship is weak, uneven and highly variable across sectors and countries. Interestingly, agglomeration elasticities are substantially larger for services than for manufacturing, underscoring the importance of service industries in Africa's growth. In addition, population density and simple measures of the built environment show little explanatory power once city size is accounted for. Increasing density alone is unlikely to yield productivity gains unless investments in the built environment improve functional efficiency and connectivity.

JEL Classification: R11, O18, O14

Keywords: agglomeration economies, African cities, firm-level productivity, urbanisation, manufacturing, services

1. Introduction

The prosperity of Africa is closely linked to the performance of its cities. The continent's urban population is projected to triple in size by 2050 and two-thirds of Africa's cities are yet to be built (UN-DESA, 2018). Urbanisation presents an inherent opportunity for aggregate productivity growth and rising living standards because of the external economies of scale and efficiencies associated with agglomeration (World Bank, 2009). Urbanisation has featured strongly in the development path of countries in the North and in East Asia.

Yet conditions on the ground in many African cities differ from other contexts because average incomes are much lower than they were elsewhere at the same level of urbanisation (Jedwab et al., 2015; Leipziger et al., 2015; World Bank, 2015). Consequently, African cities suffer from chronic underinvestment in infrastructure leading to congestion, contagion and serious shortfalls in public

[1] Associate Professor, Southern Centre for Inequality Studies, University of the Witwatersrand; justin.visagie@wits.ac.za; https://orcid.org/0000-0002-2526-231X

[2] Distinguished Research Fellow, Human Sciences Research Council; Professor, Dept of Economics and Finance, University of the Free State; iturok@hsrc.ac.za; https://orcid.org/0000-0001-5520-2492

services (Collier and Venables, 2015, 2017; Bryan et al., 2020). Weak land and property rights, coupled with poor governance, also tend to result in market inefficiencies and failures (Lall et al., 2017).

Therefore there are good reasons to be sceptical about the economic benefits of Africa's urbanisation, particularly given persistent infrastructure deficits, weak governance, and limited industrialisation (Buckley and Kallergis, 2014; Kayizzi-Mugerwa et al., 2014; Collier, 2017; Lall et al., 2021; Turok et al., 2023). At a conceptual level, none of the mechanisms associated with agglomeration economies – such as specialisation and synergies among firms, the spread of knowledge and innovation, business collaboration and competition, labour market pooling, and sharing of infrastructure – are guaranteed by urban population growth. Physical proximity between people or enterprises does not necessarily translate into improved economic connectivity, increased trade and stronger learning, particularly if there are systemic institutional failures and underinvestment in the built environment (UN-Habitat, 2014; World Bank, 2013; Glaeser and Henderson, 2017; Lall et al., 2021). On the contrary, rapid population growth can generate crowding, congestion, crime and pollution. These agglomeration diseconomies offset, and may even surpass, the potential benefits (Grover et al., 2022; Burger et al., 2022).

A related concern is that many African countries are urbanising without industrialising or, in some cases, even experiencing premature deindustrialisation (Rodrik, 2016). Without strong growth in tradable sectors, cities may function more as centres of consumption than production (Gollin et al., 2016; Jedwab et al., 2025). Spurred by access to oil, minerals and other natural resources which can sustain local consumption demand, cities may expand without harnessing the economic advantages of agglomeration. As a result, African cities may not translate their growing size into productivity gains in support of structural transformation, and simply become ever larger consumption hubs.

Empirical evidence on the relationship between agglomeration and firm-level productivity in Africa remains limited (Bryan et al., 2020; Collier and Venables, 2017; OECD, 2022). A recent meta-analysis by Grover et al (2023) on agglomeration economies in developing countries largely overlooks African cities, including evidence from only one study about population density and agricultural productivity in Malawi. The few studies on African cities report mixed findings regarding the existence and magnitude of agglomeration effects, with considerable variation in the scale of these impacts (Henderson et al., 2021; Jones et al., 2020; Jones et al., 2017). As a result, the unique dynamics of African cities remain significantly under-explored (Turok and McGranahan, 2013; Bryan et al., 2020).

The purpose of this paper is to interrogate the relationship between city size and firm-level productivity in Africa based upon a comprehensive database of African firms and cities from World Bank Enterprise Surveys and the Global Human Settlement Layer. We ask a vital question: Do larger cities offer any productivity advantages for African firms? Furthermore, is the elasticity between city size and firm productivity similar in size to that of countries elsewhere? We also ask whether there are noticeable differences between African countries, and whether there are differences between firms in services and manufacturing? Finally, we consider what influence investments in the built environment and urban density have on productivity.

Our work extends the existing literature in several important ways: First, we assemble data covering 157 cities across 34 African countries – the most extensive coverage of the continent to date. The data is based on a standardised survey instrument from the World Bank's Enterprise Surveys (ES), combined with a common approach to defining and measuring cities from the Global Human Settlement Layer (GHSL). Hence comparisons should be robust to sizeable differences between African countries in average incomes, cultural and linguistic norms, and variable administrative boundaries of cities. Second, we examine the influence of agglomeration in both manufacturing and services sectors, whereas the literature is largely restricted to manufacturing. Most African countries have struggled to industrialise, so it's critical to consider the role of services. Third, we examine evidence at the level of the firm which is robust to the choice of output per worker, wages or total factor productivity (TFP) (available for manufacturers) as the relevant dependent variables. Finally, we explore how a variety of important features of the built environment, such as the urban footprint, density and road network, impact on the assessment. This is a crucial component of the analysis because proximity might not translate into connectivity in African cities if they are characterised by haphazard sprawl and chronic underinvestment in essential infrastructure.

2. Methods of measuring agglomeration economies

There is an extensive literature concerned with the size, scale and nature of the relationship between agglomeration and productivity. While the basic idea is fairly intuitive, the concept of agglomeration economies has many permutations and has developed into a large field of enquiry. This includes differences in what is measured (e.g. measuring density or total city size; examining neighbourhood, precinct, city or regional effects; incorporating static or dynamic impacts; and choice of worker, population or firm as the unit of analysis) and how the underlying relationship is expected to work (e.g. encouraging specialisation; sharing of infrastructure; labour pooling; or diffusion of knowledge).

The following dimensions are vital because they inherently influence the empirical specification and results.

The first concerns the choice of productivity metric. This usually involves measuring labour productivity from the perspective of the worker, such as wages, or otherwise production from the perspective of the firm, such as Total Factor Productivity (TFP). TFP is the preferred measure because it directly captures productivity, while wages are bound up with local prices and capture only one dimension of productivity (i.e. labour productivity). However, both have been used extensively in practice (see Combes and Gobillon (2015) as well as Graham and Gibbons (2019) for a compelling discussion).

A second is heterogeneity when focusing on characteristics such as industry or skill intensity. At a theoretical level, localisation economies are defined as benefits which are external to the firm but internal to the industry and city. This implies measuring the concentration of firms within the same sector or value chain. On the other hand, urbanisation economies are benefits external to the firm and industry but internal to the city. This means measuring the overall size of the city, and sometimes the extent of industrial diversity. In addition, the empirical literature suggests that firms in services are more inclined to cluster and derive larger benefits from agglomeration than firms in manufacturing (Melo et al., 2009; Graham and Gibbons, 2019; Grover et al., 2023). This could be because of the role of face-to-face contact for learning, which is vital in knowledge-intensive services.

A third is the scale at which the relationship between size and productivity applies, as well as the approach to defining the economic mass or density. For instance, agglomeration could be measured as a count of the population or employment for a particular area or otherwise converted into a spatial density. The latter is sometimes preferred because it explicitly controls for differences in geographic boundaries across units (Ciccone and Hall, 1996). More sophisticated measures of density have also been developed, such as the De la Roca and Puga (2017) measure of connectivity which applies a discount factor across spatial units, in order to capture the variability in effective density of cities of similar size but different internal distributions. This also relates to the unit of analysis and whether the study aggregates up to the neighbourhood, district, city or regional level.

A fourth dimension is the time horizon of the impact. The conventional approach is to investigate the immediate or static effects of agglomeration. In other words, at any point in time, to what extent do firms or workers in larger cities command a premium? However, the literature has matured to also consider how the benefits of agglomeration may accumulate over time (Glaeser and Mare, 2001; Hanlon and Miscio, 2014; De la Roca and Puga, 2017). For instance, De la Roca and Puga (2017) find that work experience accrued in bigger cities is more valuable, persists even after leaving, and is amplified for workers of higher skill. This dynamic advantage could account for as much as half of the city-size earnings premium.

Another element is the pathway through which agglomeration benefits operate, such as through the mechanisms of 'sharing, matching and learning' (Duranton and Puga, 2004). In practice, most empirical studies simply measure the aggregate influence of size on productivity without determining how this occurs. Therefore, we ignore whether the benefits arise because of sharing, matching or learning, or some combination of all three. Similarly, positive agglomeration effects can turn into negative effects at some city size threshold, but it is common for only net impacts to be considered. Yet, some studies do attempt to control for negative externalities from agglomeration by including local house prices, land rents and local price indices (Moretti, 2013; Grover et al., 2023).

Finally, empirical assessments usually seek to establish more than the strength of association between agglomeration and productivity. Moving from correlations to casual relationships is not a trivial task because of endogeneity in most standard approaches (Graham and Gibbons, 2019; Combes and Gobillon, 2015). At the level of the locality, it is plausible to assume that places which are more productive are also those which tend to grow in size – the problem of reverse causality. Another common issue is where there are unmeasured variables, as in the case of missing amenities, which simultaneously influence local productivity and the size of the local population. At the level of the worker (or firm), individuals may sort across locations according to factors which are not controlled for (like unobserved skills and ability) – the problem of selection. Selection and reverse causality are both sources of endogeneity which need to be solved for the models to produce unbiased results. This is also an area where the literature has made considerable progress, although the options obviously depend on the type of data and available variables.

3. Evidence of agglomeration economies in developing countries

There is growing interest in the role of agglomeration and productivity in emerging economies (Glaeser and Henderson, 2017; Collier and Venables, 2017). Some studies suggest that the effects of agglomeration on productivity could be even larger in the developing world (Combes and Gobillon, 2015; Glaeser and Xiong, 2017; Bryan et al., 2020). However, a recent meta-analysis by Grover et al (2023) concludes that there is actually no significant difference in the city size premium between developed and developing countries once controlling for a variety of factors in the way in which the studies were constructed.

The evidence about agglomeration and productivity in African countries is much more limited (Turok and McGranahan, 2013; Bryan et al., 2020). There is mixed evidence about the presence and size of agglomeration effects, with large differences in the scale of these impacts (Henderson et al., 2021; Jones et al., 2020; Jones et al., 2017; OECD, 2022). For instance, Grover et al (2023) include evidence from only one African country in their meta-analysis. Consequently, the jury is still out about the relationship between agglomeration and productivity in Africa.

To our knowledge, the most extensive coverage is provided by Jones et al (2020), who examine 43 cities across 22 African countries and another 68 cities in the rest of the world (28 cities in Latin America and 40 cities in Asia) over the period 2006 to 2012. They make use of World Bank Enterprise Survey data combined with satellite imagery, which has the benefit of offering consistent measures of both productivity and city characteristics. A headline finding is that firm-level productivity increases by about 5-6% when doubling city size using the full global sample of developing countries. However, this is not the case when examining African cities, where the results are inconclusive.

Interestingly, Jones et al (2020) struggle to find significant results when measuring population density (rather than total city size) on firm productivity or when exploring more sophisticated measures of internal city density. Instead, they hypothesize that the expected advantage from higher densities could work indirectly by supporting a larger average firm size (because larger firms tend to be more productive). Yet, for African cities, in particular, they find that measures of internal density do not adequately predict the relative scale of employment in the city over time. African cities have fewer and smaller firms relative to the number of people living in the city, and the scale of formal employment is unresponsive to the concentration of population. They conclude that "African cities are not generating the same level of benefits—in terms of increased productivity, wages, and employment generation—as cities in Asia and Latin America" (Jones et al., 2020; p 169).

4. Data and methods

Measuring agglomeration economies across cities in different countries requires consistent data sources and methods of analysis. Geo-coded firm-level data is available from World Bank's Enterprise Surveys (ES) for African countries collected between 2013 and 2020[3]. This ensures that a consistent set of survey instruments are used when measuring firm-level performance. Similarly, we compile population and built environment information for African cities from the Global Human Settlement Layer (GHSL) database (Dijkstra et al., 2021; European Commission, 2021)[4]. The GHSL was developed by the European Commission and applied by six international organisations to create a 'Degree of Urbanisation' classification specifically to overcome the problem of different national definitions of cities and to facilitate international comparisons. It defines cities based on specific population size and density thresholds according to a harmonised methodology using satellite imagery and population census data. Cities are defined as having at least 50,000 inhabitants and with a density of at least 1,500 people per km^2. They are derived by aggregating contiguous cells from a fine-grained population grid where all cells have the same size and shape. Satellite imagery is used to distinguish cells that comprise relatively dense built-up areas from less dense areas and sparsely settled areas. The cities that emerge from this process are continuous built-up urban areas. They typically exclude surrounding towns, villages and semi-dense peri-urban areas but may still include denser suburban areas surrounding the urban centre.

Population density is measured as population per square kilometre using GHSL-defined urban extents. While this captures overall spatial concentration, it does not directly account for effective accessibility between firms and workers (e.g. commuting times or transport networks), and should therefore be interpreted as a proxy for density rather than connectivity.

Figure 1 shows the extent of geographical coverage of the data. The 34 countries included in our sample account for approximately 85% of the population on the continent. There are some limitations to the World Bank's ES. While it usually provides sub-national representation, there are 10 cases where firms were sampled in only one city. In addition, the sample design of the ES is restricted to formal enterprises with a minimum of 5 employees. Hence, the role of agglomeration on productivity in informal enterprises as well as the smallest firms is excluded from the analysis[5]. We also conduct

[3] Geocoded data was made available by the World Bank upon special request and has been assigned a random deviation of within 1km of the original coordinate to ensure anonymity. This condition is sufficient for identifying firms at a city level.
[4] We add an additional 2km buffer onto the GHSL urban boundaries to capture firms which might lie very close but beyond the GHSL city borders (recall the firm data has a 1km deviation from the original coordinate). We further join any urban layers which intersect within a 2km buffer of each other forming larger city-regions. The exception is Egypt where the buffering awkwardly joins Alexandria with Cairo (more than 150km to the South) due to continuous minimum density along the Nile corridor. Therefore, we impose a split to separate out these two mega city regions.
[5] While informal and micro enterprises are important for labour absorption, they tend to be locally traded and are unlikely to have a significant influence on aggregate productive growth or longer term structural transformation.

sensitivity checks by excluding large upper-tail cases such as Nigeria and Egypt in the pooled sample. While this reduces the sample size and affects coefficient magnitudes, the direction of the estimated effects remains consistent. Country-specific heterogeneity, which is often large and significant, is explicitly discussed in Section 5.5.

Table 1 provides descriptive statistics of the variables used in our models divided into city size bands (small, medium, large and very large). The observations (i.e. sampled firms) are fairly evenly spread across each of the city size classes. We exclude firms in rural areas or urban centres of less than 50,000 as out of scope. We also drop cases where a city has less than 10 firms sampled to ensure some minimum representivity. We remove outliers of greater than 3 std deviations – standard practice in the World Bank's ES global indicators – which impacts less than 0.5% of observations. Where necessary, monetary values from the Enterprise Surveys are converted into constant 2009 USD using country-specific GDP deflators and official exchange rates from the World Bank's World Development Indicators, following the standard Enterprise Surveys methodology.

We use three measures of firm-level productivity: output per worker (defined as plants' annual sales divided by workforce), wage per worker (defined as annual labour costs divided by workforce) and TFP (as estimated in the ES's global standardised dataset[6]). While TFP directly captures firm-level productivity, and is the preferred choice in the literature, the ES only collects sufficient information for TFP to be reported for manufacturers.

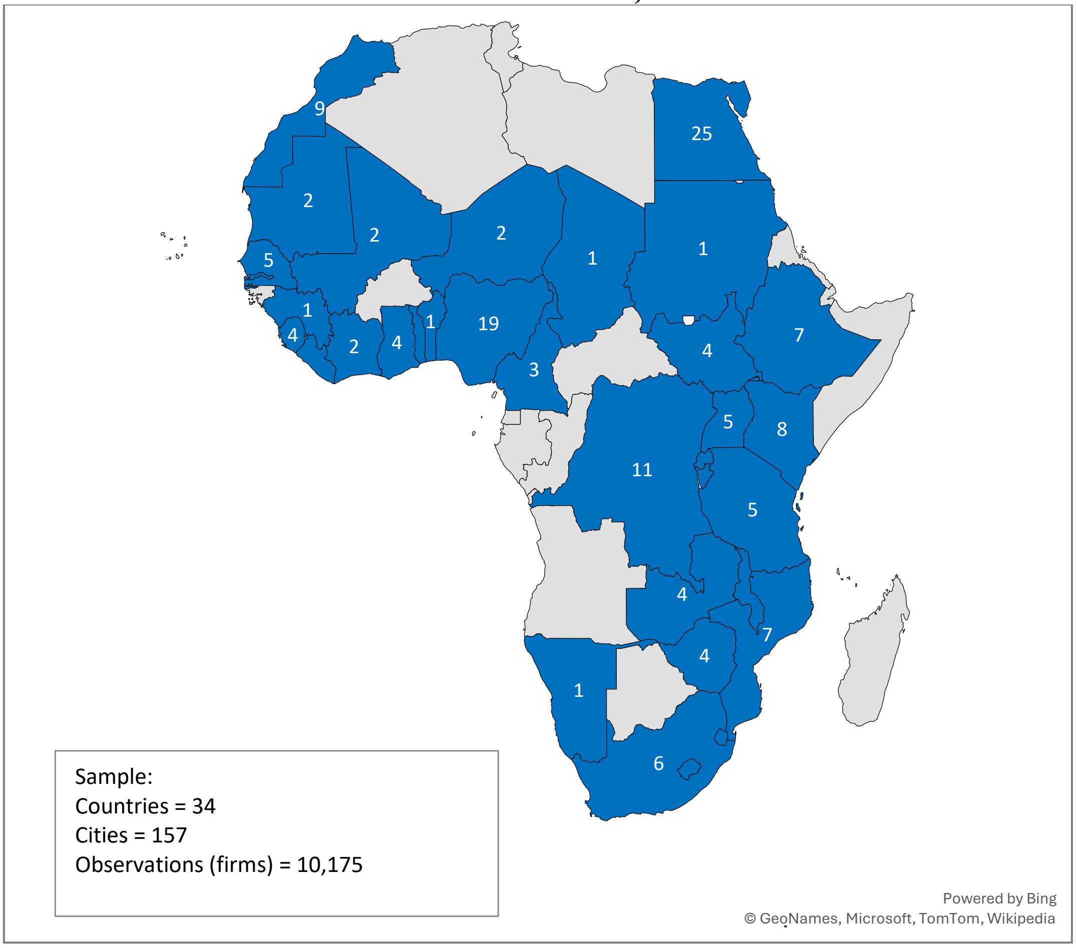
Exploring the detailed sectoral information available within the ES is an important dimension of the analysis. We model the influence of agglomeration on manufacturing and services sectors separately and control for industry sub-sectors as well as looking into industry sub-samples (based on the International Standard Industrial Classification of All Economic Activities (ISIC) revision 3.1). For manufacturers, industry sub-sectors are grouped as: Food/beverage (ISIC 15,16), Wood/paper (ISIC 20-22), Textiles (ISIC 17-19), Industrial materials (such as metals, plastics, minerals and chemicals) (ISIC 23-28), and Higher-value manufacturing (such as vehicles, machinery, electronics and miscellaneous) (ISIC 29-37). For services, industry sub-sectors are grouped as: Construction (ISIC 45), Wholesale/retail (ISIC 50-52), Hotels/restaurants (ISIC 53), Logistics (ISIC 60-63) and IT/telecoms (ISIC 64, 72).

All firm level variables are derived from the ES. This includes firm size, age, top manager's experience, whether the firm is foreign owned, involved in innovation or exporting, the quality of business local environment (sales lost to power outages) and skills of the workforce.

City variables are derived from the GHSL[7] (Urban centres database) including city population size, surface area and population density. In addition, we make use of a GHSL (Built-up height database) spatial raster (100x100 meters) which reports on average building heights for cities (share of raster's with average building height of four storeys or more). We also do the same for road networks based on a GHSL (Built-up characteristics database) spatial raster (10x10 meters) derived from OSM highways.

[6] Refer to World Bank (2017) for a discussion of the TFP methodology
[7] See European Commission (2024) for technical descriptions of databases, variables and methods published as the GHSL.

Figure 1: African countries and number of cities included the sample (country shaded with number of cities)



Source: World Bank Enterprise Surveys, available African countries 2013 - 2020

Notes: The sample includes country, # of cities and # of firms as follows: Benin= 1 city, 105 firms; Burundi= 3 cities, 135 firms; Cameroon= 3 cities, 231 firms; Chad= 1 city, 125 firms; Côte d'Ivoire= 2 cities, 175 firms; Democratic republic of congo= 8 cities, 378 firms; Egypt= 21 cities, 1868 firms; Eswatini= 1 city, 38 firms; Ethiopia= 7 cities, 540 firms; Gambia= 1 city, 107 firms; Ghana= 4 cities, 439 firms; Guinea= 1 city, 41 firms; Kenya= 7 cities, 542 firms; Lesotho= 1 city, 75 firms; Liberia= 3 cities, 93 firms; Malawi= 3 cities, 209 firms; Mali= 2 cities, 84 firms; Mauritania= 2 cities, 58 firms; Morocco= 8 cities, 434 firms; Mozambique= 6 cities, 320 firms; Namibia= 1 city, 75 firms; Niger= 2 cities, 69 firms; Nigeria= 17 cities, 742 firms; Rwanda= 1 city, 40 firms; Senegal= 4 cities, 296 firms; Sierra Leone= 3 cities, 135 firms; South Africa= 6 cities, 667 firms; South Sudan= 4 cities, 572 firms; Sudan= 1 city, 120 firms; Tanzania= 5 cities, 226 firms; Togo= 1 city, 87 firms; Uganda= 5 cities, 286 firms; Zambia= 4 cities, 368 firms; Zimbabwe= 4 cities, 495 firms.

Our general approach is to build up to a full set of specifications by progressively adding controls. Model (1) presents a simple OLS regression of output per worker on city size. Model (2) adds firm-level controls as well as industry and country fixed effects, while Model (3) estimates the same specification using a two-stage least squares instrumental variable approach (2SLS-IV). Models (4) and (5) replace city size with population density in OLS and IV specifications respectively. Models (6) and (7) extend the analysis by including built environment controls, with Model (7) instrumenting for city size including our most comprehensive set of controls. This approach allows us to assess the robustness of the results across alternative specifications and measures of agglomeration.

Therefore the models are a variant of the following general equation:

$$\ln{(Y)_{ijk}} = \alpha \ln(CSize)_k + \theta Firm_{ijk} + \delta Built_k + \rho_j + \nu_k + \varepsilon_{ijk} \quad \text{....(1)}$$

where the dependent variable, (ln $(Y)_{ijk}$), is the natural logarithm of output per worker (or other productivity measure) for firm (*i*) in industry (*j*), located in country (*k*). The regression estimates the elasticity of output with respect to our main variable of interest, the natural logarithm of city size (ln($CSize$)) (or other measure of density). This coefficient is derived after accounting for a vector of firm-level characteristics (*Firm*) and built-environment factors in each city (*Built*). Furthermore, the model incorporates fixed effects for industry (ρ_j) and country (ν_k) where ε is the error term capturing unexplained variability for firm (*i*)[8].

In the instrumental variable specifications, city size is treated as endogenous and instrumented using lagged population and geographic variables. The first stage is given by:

$$ln(CSize_k) = \pi_1 Z_k + \gamma Firm_{ijk} + \delta Built_k + \rho_j + \nu_k + \varepsilon_{ijk} \quad \text{....(2)}$$

where (Z_k) includes the instrumental variables and firm and built environment controls are added in the relevant specification. The results from the first stage are included in the appendix.

[8] While we would like to include survey-year fixed effects this is not feasible because the Enterprise Surveys are not evenly distributed across countries and years, with several years covering only a single country. In practice, this means year effects are largely absorbed by country fixed effects and add little independent variation.

Table 1: Descriptive statistics

City size rank	Small (<50k)		Medium (50k-2mil)		Large (2mil-5mil)		Very Large (>5mil)		Total	
Productivity	**Mean**	Std dev	**Mean**	Std dev	**Mean**	Std dev	**Mean**	Std dev	**Mean**	Std dev
Output per worker	**36,009**	122,202	**31,756**	104,108	**39,914**	111,882	**35,336**	132,898	**36,067**	117,510
Wage per worker	**4,027**	9,839	**4,108**	11,571	**4,350**	9,576	**3,474**	8,669	**4,031**	9,976
Total Factor Productivity	**2.9**	1.7	**3.0**	1.8	**3.0**	1.8	**2.9**	1.7	**2.9**	1.8
Population										
Population size	**302,189**	133,326	**1,066,296**	413,268	**3,474,534**	821,349	**12,500,000**	6,462,763	**3,947,868**	5,399,975
Population density	**8,010**	6,540	**7,830**	3,074	**8,071**	3,109	**8,638**	2,963	**8,116**	4,242
City surface area	**45.7**	23.1	**153.9**	81.8	**487.0**	199.7	**1,521.2**	658.1	**509.0**	631.2
Built environment										
share of tall buildings	**2.3**	5.7	**2.1**	3.2	**3.9**	4.2	**1.1**	2.6	**2.5**	4.3
share of roads	**1.1**	0.6	**1.0**	0.7	**1.3**	1.1	**0.3**	0.5	**1.0**	0.9
Firm characteristics										
firm size	**48.4**	148.5	**53.3**	110.2	**63.3**	185.1	**57.9**	140.7	**56.1**	151.7
firm age	**14.3**	13.7	**18.7**	15.1	**19.1**	15.5	**22.5**	17.5	**18.5**	15.7
top manger's yrs experience	**13.9**	10.1	**16.5**	10.5	**17.9**	10.7	**20.1**	11.7	**17.0**	10.9
sales lost to power outages	**4.5**	10.5	**5.8**	11.8	**6.3**	11.6	**3.2**	9.8	**5.1**	11.1
share workers without high school	**43.2**	35.7	**41.2**	35.5	**35.2**	35.6	**33.3**	32.7	**38.3**	35.3
dummy: foreign owned	**0.16**	0.37	**0.12**	0.32	**0.13**	0.34	**0.05**	0.21	**0.12**	0.32
dummy: innovator	**0.52**	0.50	**0.43**	0.49	**0.55**	0.50	**0.24**	0.42	**0.45**	0.50
dummy: exporter	**0.09**	0.29	**0.11**	0.32	**0.12**	0.33	**0.08**	0.28	**0.11**	0.31
dummy: form training offered	**0.20**	0.40	**0.26**	0.44	**0.28**	0.45	**0.11**	0.31	**0.22**	0.41
Sample										
Observations (# of firms)	**2,609**		**2,374**		**3,098**		**2,094**		**10,175**	
Count of cities	**74**		**49**		**25**		**9**		**157**	

Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

The 2SLS-IV model is our attempt to deal with endogeneity in light of the challenges in establishing causality (see section 3). We regress a set of instruments on our measure of agglomeration in the first stage, and then use the predicted values as the regressor for agglomeration in the main equation. The effectiveness of our approach entirely depends on the validity of the instruments which need to be both relevant (correlated with agglomeration) and exogenous (not related to productivity or any missing variables which determine productivity).

Following the literature, we use lagged population from 1975 (from the GHSL) as our primary instrument, which is strongly correlated with current city size. We also include geographic variables, namely latitude and average rainfall (measured in the 1990s), which help explain the historical location of settlements. In practice, lagged population size is extremely relevant as an instrument, while geographical variables by comparison have less predictive power. The combination of both sets of instruments is widespread in the literature (see Combes and Gobillon, 2015).

First-stage diagnostics (reported in the main tables) confirm this pattern: lagged population is a strong predictor of current city size, while the geographic variables contribute less explanatory power. Taken together, however, the instrument set passes standard tests for relevance and identification. An exception is for model #5 which uses population density, where lagged population density performs poorly as an instrument.

We also acknowledge potential limitations of the instrument set. Lagged population may capture persistent underlying factors that continue to influence productivity today, while geographic variables may affect firm performance through channels unrelated to agglomeration. These concerns cannot be fully ruled out and should be borne in mind when interpreting the IV results.

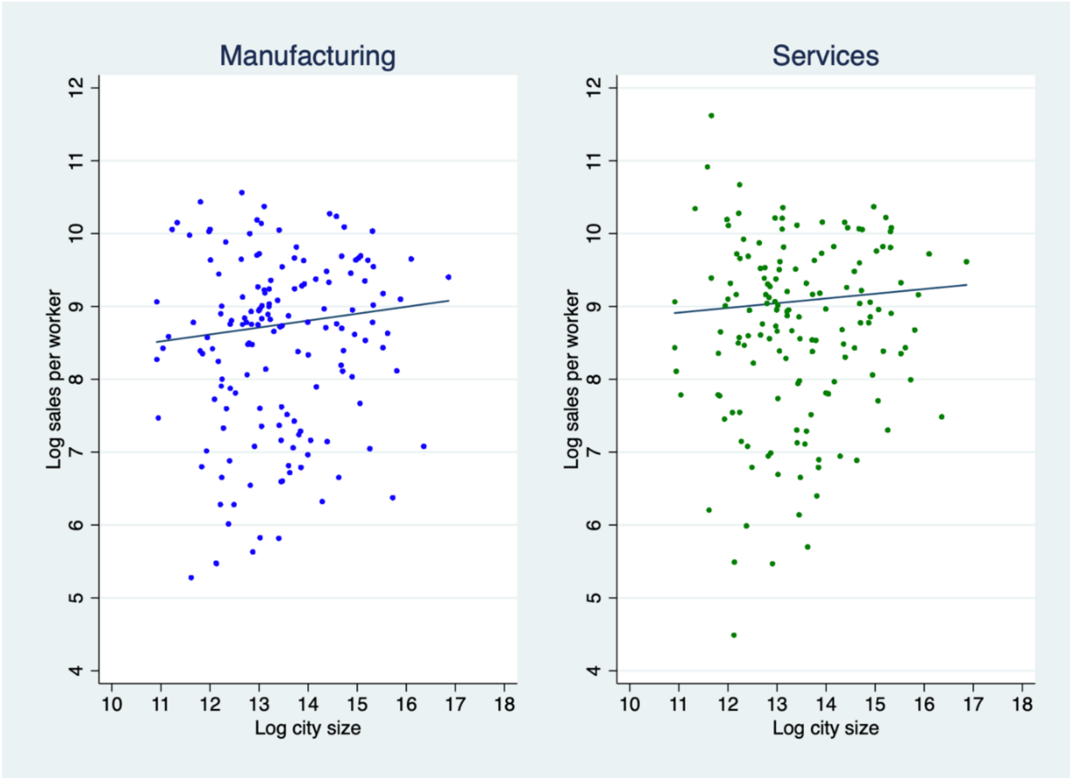
5. Results

The main findings are reported in table 2 (for manufacturers) and table 3 (for services) which summarise the results for models #1-#7 depending on the approach to measuring agglomeration and firm-level productivity. The dependent variable is output per worker but we also explore the impact of choosing wages or total factor productivity as alternative measures of productivity (see tables A1 to A3 in the appendix).

5.1 The raw correlation between city size and productivity in Africa

The simple relationship between output per worker and city size is reported in model #1 and displayed in figure 2 with the associated regression line. The main message is that there is considerable noise in the raw relationship between city size and firm-level productivity. While there is a slightly positive fit to the data for firms in both manufacturing and services, the relationship is loose and statistically insignificant with an R-squared of less than 1 percent. This is equally true of manufacturing and services.

Figure 2: Scatterplot of city size and firm productivity



Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Notes: Observations (dots) represent city-level averages of output per worker. However the fitted line is estimated using the full set of firm-level observations, where city size is repeated for all firms within a city. City averages are shown for clarity to avoid over-plotting.

Table 2: Agglomeration economies for manufacturing firms (dependent variable: output per worker)

Model	#1		#2		#3		#4		#5		#6		#7	
	OLS: pop size (no controls)		OLS: pop size		2SLS: pop size		OLS: pop dens		2SLS: pop dens		OLS: pop size and built env		2SLS: pop size and built env	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
AGGLOMERATION MEASURE														
Instrumented	*No*		*No*		*Yes*		*no*		*Yes*		*No*		*yes*	
log city size	0.076	0.197	**0.057**	0.034	0.043	0.223					**0.065**	0.000	0.055	0.126
log pop dens							0.048	0.672	0.034	0.921				
REGRESSION CONTROLS														
Built env controls	*No*		*No*		*No*		*no*		*No*		*Yes*		*yes*	
tall building share											0.009	0.104	0.008	0.451
road surface share											-0.033	0.394	-0.038	0.582
Firm controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
firm size			**0.000**	0.000	0.000	0.202	0.000	0.210	0.000	0.208	**0.000**	0.020	0.000	0.184
firm age			**0.001**	0.003	0.002	0.447	0.002	0.616	0.003	0.410	**0.002**	0.079	0.002	0.436
top manger's yrs experience			**0.012**	0.003	**0.163**	0.002	**0.013**	0.000	**0.168**	0.002	**0.159**	0.000	**0.160**	0.003
sales lost to power outages			**-0.008**	0.002	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000
share unskilled workers			**-0.007**	0.001	**-0.006**	0.000	**-0.006**	0.000	**-0.006**	0.000	**-0.006**	0.000	**-0.006**	0.000
dummy: foreign owned			**0.557**	0.095	**0.561**	0.000	**0.565**	0.000	**0.568**	0.000	**0.558**	0.000	**0.559**	0.000
dummy: innovator			**0.355**	0.058	**0.357**	0.000	**0.360**	0.000	**0.362**	0.000	**0.356**	0.000	**0.357**	0.000
dummy: exporter			0.500	0.103	**0.500**	0.000	**0.493**	0.000	**0.495**	0.000	**0.500**	0.000	**0.500**	0.000
dummy: form training offered			**0.211**	0.063	**0.213**	0.001	**0.214**	0.001	**0.215**	0.001	**0.211**	0.000	**0.212**	0.001
Constant	**7.777**	0.000	**8.408**	0.000	**8.375**	0.000	**8.738**	0.007	**8.643**	0.007	**8.082**	0.000	**8.230**	0.000
Industry controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
Country controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
Observations	5082		5082		5082		5082		5082		5082		5082	
R-squared	0.0039		0.3354		0.3402		0.3337		0.3385		0.3343		0.3406	
Instrument Diagnostics														
KP rk LM (p-value)					0.006				0.108				0.000	
KP rk Wald F-stat					22.821				5.837				80.082	

Notes: coefficients in bold are statistically significant at a 95% level of confidence; Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Table 3: Agglomeration economies for services firms (dependent variable: output per worker)

Model	#1		#2		#3		#4		#5		#6		#7	
	OLS: pop size (no controls)		OLS: pop size		2SLS: pop size		OLS: pop dens		2SLS: pop dens		OLS: pop size and built env		2SLS: pop size and built env	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
AGGLOMERATION MEASURE														
Instrumented	*No*		*No*		*Yes*		*no*		*Yes*		*No*		*yes*	
log city size	0.062	0.306	0.131	0.131	**0.112**	0.017					**0.126**	0.001	**0.111**	0.016
log pop dens							0.029	0.845	-0.429	0.703				
REGRESSION CONTROLS														
Built env controls	*No*		*No*		*No*		*no*		*No*		*Yes*		*yes*	
tall building share											-0.011	0.420	-0.011	0.423
road surface share											-0.036	0.630	-0.041	0.586
Firm controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
firm size			**0.000**	0.000	0.000	0.726	0.000	0.696	0.000	0.624	0.000	0.713	0.000	0.700
firm age			**0.010**	0.010	**0.010**	0.000	**0.010**	0.000	**0.010**	0.000	**0.010**	0.000	**0.010**	0.000
top manger's yrs experience			**0.014**	0.014	**0.179**	0.000	**0.014**	0.000	**0.179**	0.001	**0.180**	0.000	**0.181**	0.000
sales lost to power outages			**-0.006**	-0.006	**-0.006**	0.085	**-0.006**	0.096	**-0.006**	0.088	**-0.006**	0.087	**-0.006**	0.083
share unskilled workers			**-0.008**	-0.008	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000	**-0.008**	0.000
dummy: foreign owned			0.449	0.449	**0.452**	0.000	**0.460**	0.000	**0.435**	0.000	**0.453**	0.000	**0.454**	0.000
dummy: innovator			0.340	0.340	**0.338**	0.000	**0.340**	0.000	**0.343**	0.000	**0.342**	0.000	**0.342**	0.000
dummy: exporter			0.214	0.214	**0.213**	0.025	**0.199**	0.044	**0.204**	0.035	**0.219**	0.025	**0.217**	0.023
dummy: form training offered			0.149	0.149	**0.148**	0.033	**0.157**	0.027	**0.157**	0.028	**0.144**	0.041	**0.145**	0.037
Constant	**8.251**	0.000	**7.498**	0.000	**7.498**	0.000	**8.995**	0.000	13.046	0.219	**7.357**	0.000	**7.569**	0.000
Industry controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
Country controls	*No*		*Yes*		*Yes*		*yes*		*Yes*		*Yes*		*yes*	
Observations	5093		5093		5093		5093		5093		5093		5093	
R-squared	0.0020		0.2601		0.2668		0.2549		0.2564		0.2675		0.2674	
Instrument Diagnostics														
KP rk LM (p-value)					0.000				0.764				0.000	
KP rk Wald F-stat					20.961				0.428				26.536	

Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates
Notes: coefficients in bold are statistically significant at a 95% level of confidence

5.2 Total city size and firm productivity

We now turn to the relationship between total city size and firm productivity once additional controls are introduced. Model #2 presents the results of city size and output per worker when including firm and country controls, while model #3 further adjusts for potential sources of endogeneity using our instruments. First-stage diagnostics indicate that the instruments for total city size perform well across specifications, suggesting strong relevance.

For manufacturing firms, the main coefficient of interest – the elasticity of city size – is of a similar magnitude across specifications. The estimated coefficients range between 0.020 and 0.058, implying that a doubling of city population is associated with an increase in productivity of around 3–4%. This is comparable to agglomeration effects reported in other contexts (Duranton and Puga, 2004; Melo et al., 2009).

However, the coefficients are not consistently statistically significant across the models. City size is positive and significant in the OLS specification (Model #2), but this significance falls away once instruments are applied in Model #3. This pattern is reversed when using wages instead of sales per worker (Table A1): the coefficient becomes more statistically significant under the IV specification, although it still falls just short of conventional significance levels ($p = 0.104$). Total factor productivity is not statistically significant in any of the models.

Therefore the evidence suggests that manufacturing is only loosely associated with productivity advantages in larger cities.

For firms in services (table 3), by contrast, the impact of city size on productivity is larger and more consistently statistically significant. The coefficients range between 0.131 and 0.112 which means that a doubling of city size increases productivity in services by 8-9%. The results are very similar when using wages instead of sales (Table A2). All of the modelled coefficients are significant at, or close to, the 99% level with the exception of model #2 (table 3) which is just outside the 90% level ($p = 0.131$).

The fact that agglomeration economies tend to be larger in services compared with manufacturing has been noted elsewhere but this is the first time such evidence has been systematically compiled for African countries (Hasan et al., 2017; Grover et al., 2023).

A potential concern is that the larger estimated effects for services may partly reflect differences in prices rather than underlying productivity between urban and rural areas. However, the magnitude of the elasticity (more than double that for manufacturing) and the consistency of the results across alternative measures (wages and sales per worker) suggest that price effects alone are unlikely to fully explain the observed premium for services.

Stepping back from the individual coefficients, the overall explanatory power of the models remains moderate. The R-squared scores vary between 0.25 to 0.40. This means that only about 25-40% of the observed variation in productivity is explained in the models. What makes one firm productive compared to another is fairly elusive. And it is worth noting that compared to city size, factors such as foreign ownership, export status and whether or not a company offers employee training have larger impacts.

5.3 Population density and firm productivity

We have focused so far on total city size as a measure of agglomeration. However, agglomeration effects are often thought to operate through proximity and interaction, suggesting that population density – a common measure in the literature (see section 2) – may be an important alternative to measuring economic concentration within cities.

The effect of switching to population density is reported in model #4 (including firm and country controls) and model #5 (including firm and country controls as well as historic population density as an instrument).

In contrast to population size, population density is not a significant predictor of firm productivity in any of the models, whether examining manufacturers (table 2) or firms providing services (table 3), or when switching between different measures of firm productivity (tables A1-A3 in the appendix).[9] First-stage diagnostics confirm that lagged population density performs poorly as an instrument. In all models, the Kleibergen–Paap rk LM statistic fails to reject under-identification at conventional levels and the corresponding Wald F-statistic indicates a weak instrument (<7).

This is somewhat surprising in light of the consistent significance of population size discussed earlier. While population density and population size are related, they are not necessarily strongly correlated. Figure 3 shows a scatterplot of city population and city density for the 157 cities in the sample. The correlation is statistically significant but fairly weak (with a pairwise correlation coefficient of 0.20).

Bearing this mind, one might still be surprised that higher population density does not correspond with higher productivity. A possibility is that population density does not adequately capture the mechanisms through which agglomeration economies work in terms of sharing, matching and learning. In this sense, African cities are not functionally dense – whether because the infrastructure is inadequate or a large share of the workforce is unskilled and jobless – just crowded.

5.4 Urban form, the built environment and firm productivity

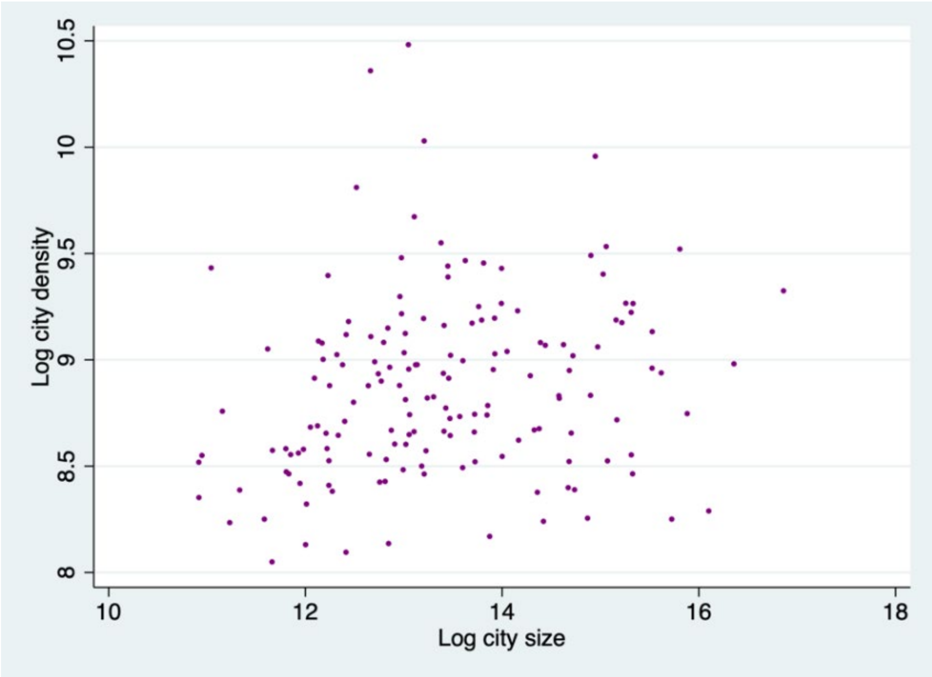
It is arguably the connectivity between firms and workers, rather than sheer population size or crude population density, which lie at the heart of agglomeration economies. Models #6 and #7 attempt to capture features of the urban form and built environment available in the GHSL in order to explore these dynamics.

In particular, we include a measure of the extent of vertical expansion of cities defined as the share of built-up areas with buildings of four storeys or more, derived from GHSL data at a 10m spatial

[9] The single exception is in model #4 in table A3 of the appendix where population density is seen to be both large and significant as a determinant of productivity as measured by total factor productivity. However, the coefficient becomes small and insignificant once applying instruments to control for sources of endogeneity in measuring total factor productivity which is the more robust specification (see model #5 in table A3).

resolution. We also include a measure of road infrastructure, proxied by the share of road surfaces derived from OpenStreetMap (OSM) highways and integrated and published within the GHSL dataset at a similar spatial resolution (European Commission, 2024).

Figure 3: Scatterplot of city size and city density



Source: Global Human Settlement Layer; Authors' own estimates
Notes: n= 157 cities; city density is measured as population per square kilometre

Model #6 includes these built environment variables with industry and country controls while model #7 applies instruments for population size to account for endogeneity. In this case, the built environment controls are weak and mostly insignificant. The slight exception is the share of tall buildings which has a positive impact on productivity measured by sales (table 2) or TFP (table A3) in manufacturing, but the significance falls away in the IV estimates. In addition, including controls for the built environment does not improve the overall model performance compared with just looking at population size or population density. The R-squared values hardly change which is an indication that our new variables have done little to explain more of the variation in firm-level productivity.

The simplest explanation for the poor performance of our built environment variables is that the data quality is too coarse in deriving features of the built environment from satellite imagery. More direct measures of road, rail and network capacity would be interesting to explore in future, although the challenge is assembling this data consistently across African countries.

More fundamentally, connectivity may depend less on observable features of urban form and more on how infrastructure functions in practice. Average building height and road density perhaps fail to capture the ease with which firms and workers interact within cities. There may be different ways to achieve connectivity that go beyond constructing taller buildings and promoting a more compact urban form. Different urban configurations may therefore achieve similar levels of connectivity, implying that outcomes such as congestion, commuting times, and accessibility may be more informative than the built environment measures used here.

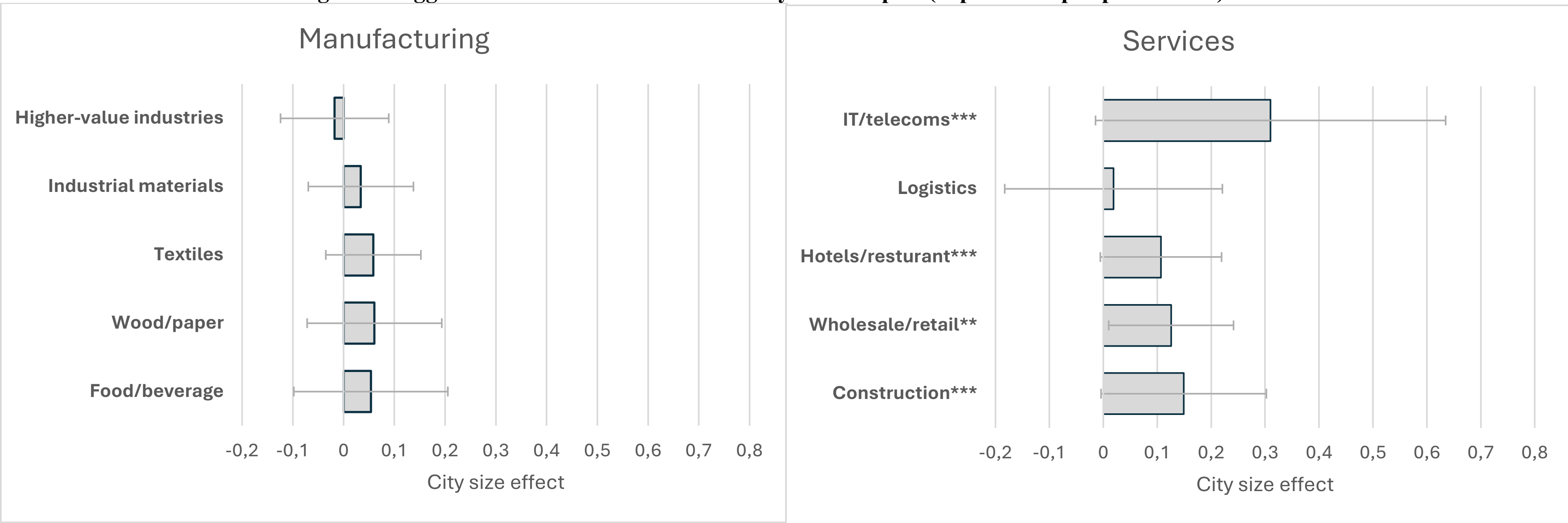
5.5 Industry and country specific variation

The final section returns to the effect of total population size on firm-level productivity in African cities, focusing on the IV specification (model #3), which provide the most consistent and robust results of any of the specifications. Population density and build environment indicators did not add explanatory power once city size is accounted for. Therefore, we focus on total city size as our preferred summary measure of agglomeration.

Our interest is in whether there are any important differences across sub-samples including sub-sectors in manufacturing and services as well country-specific effects. An alternative approach would be to estimate fully interacted models allowing elasticities to vary by country or firm characteristics, but given limited sample sizes this leads to unstable estimates.

Figure 4 shows the results for different sub-sectors within manufacturing and services.

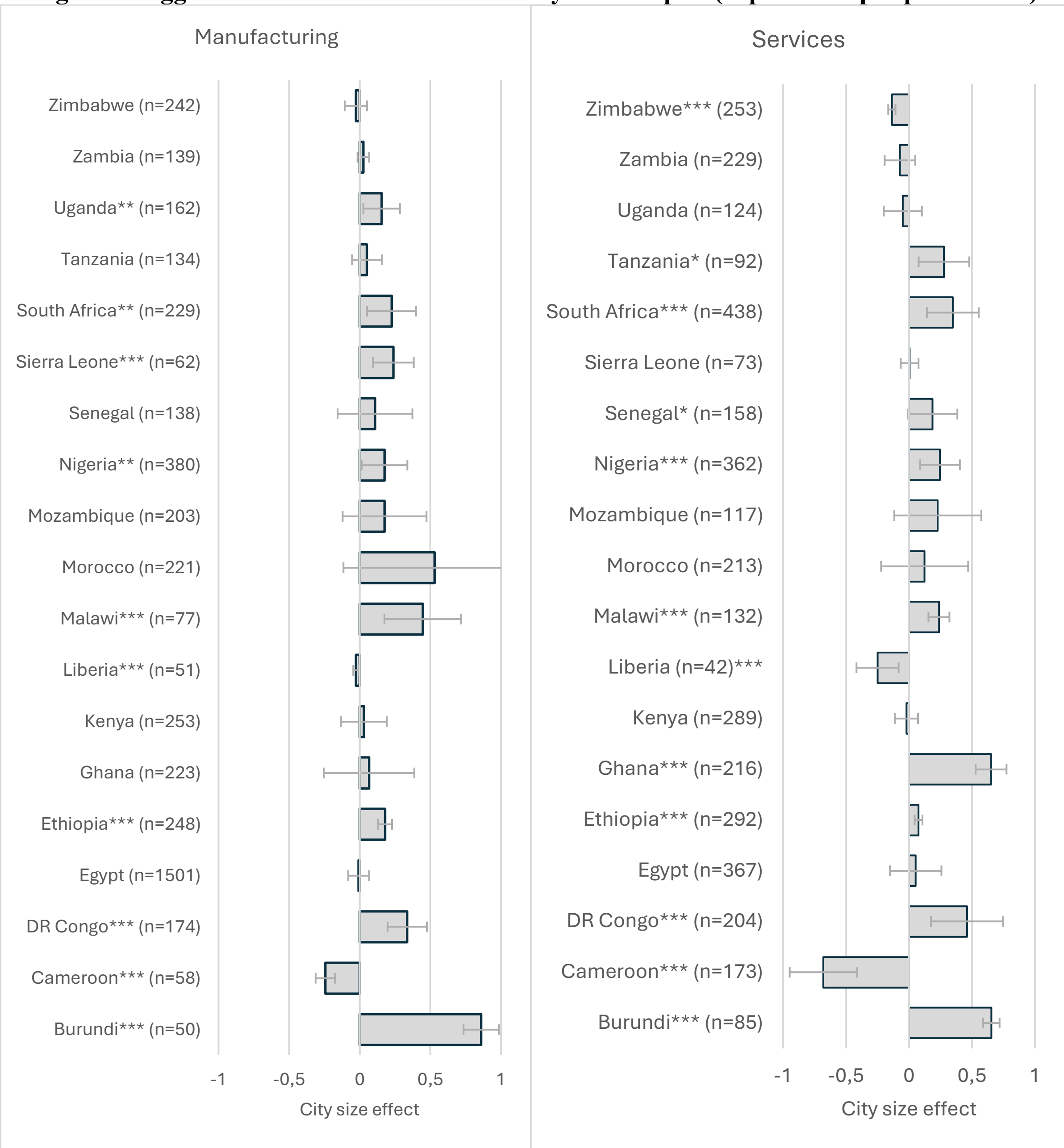
Figure 4: Agglomeration economies for industry sub-samples (dep var: output per worker)



Notes: 2SLS:IV of city population size including firm and country controls; error bars display 95% confidence interval; * significant at 90% level; ** significant at 95% level; ***significant at 99% level

Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Figure 5: Agglomeration economies for country sub-samples (dep var: output per worker)



Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Notes: 2SLS:IV of city population size including firm and industry controls; Countries must have at least three cities; Sudan is excluded due to very large estimates (manu=0.74 & serv=1.35); Error bars display 95% confidence interval; * significant at 90% level; ** significant at 95% level; ***significant at 99% level

In the case of manufacturing firms, the sub-sectors of industrial materials, textiles, wood and paper, and food and beverages each have output elasticities with respect to city size of around 0.05, while higher-value manufacturing shows a slightly negative relationship. The fact that more complex manufacturing activities seem hampered in larger cities is of concern, however none of the reported estimates are statistically significant because of large standard errors. This highlights the fragility of the relationship which makes it difficult to identify sub-sector patterns.

By contrast, the output elasticities with respect to city size for different service sub-sectors are larger and statistically significant. The exception is logistics which has an output elasticity close to zero. Looking more closely at the results, the degree of variability in the estimates is actually similar comparing service sub-sectors with manufacturing sub-sectors, as seen by large error bars in almost all cases. However, because the elasticities are larger for services they tend to be significantly different from zero. Activities such as retail/wholesale and hotel/restaurants are not usually considered tradable so these productivity advantages might sometimes work at a localised scale without much scope to be scaled up – although digital transformation is opening up new opportunities for services to cross-borders.

Figure 5 shows the results for different countries – whether firms in manufacturing or services. In this instance, we are restricted to countries with data points for multiple cities, which is only 19 out of 34 African countries in the global sample. Another challenge is that country-specific samples are inherently small which increases the risk of measurement error. This is compounded by the fact that the relationship between city size and productivity is volatile even in the full sample. Hence we recommend that country-specific estimates are interpreted with caution.

About half of all African countries show statistically significant results in their relationship between city size and economic output per worker. In manufacturing, eleven out of twenty African countries show significant results compared with thirteen in the case of services. The coefficients are also sometimes contradictory with two negative and significant elasticities in manufacturing (Cameroon and Liberia) and three negative and significant elasticities in services (Cameroon, Liberia and Zimbabwe). Only six countries produce positive and significant estimates in both manufacturing and services (South Africa, Nigeria, Malawi, Ethiopia, DR Congo and Burundi), while another five countries produce insignificant results in both cases (Zambia, Mozambique, Liberia, Kenya and Egypt).

While small sample sizes are admittedly responsible for some of the variability between countries, it is also feasible that there are real and fundamental differences in the way in which cities are resourced and governed in different parts of the continent. These differences either enable or undermine the impact of agglomeration on productivity. They would also explain why there is a generally a weak fit across the full sample of countries. The central message is that while agglomeration economies offer potential benefits for African countries, the relationship is tenuous at present, and not all cities and countries benefit.

6. Conclusion

Many scholars and international organisations describe African cities as crowded, disconnected and costly – characteristics typically associated with inefficiency rather than high productivity (Lall et al., 2017; UN-ECA, 2022; UN-Habitat, 2022). Our findings partially support this characterisation, but they also suggest that agglomeration economies are nonetheless present, albeit weak, uneven and concentrated in services rather than manufacturing. Therefore, the question is not whether agglomeration forces exist in African cities, but why they appear to operate so inconsistently.

A key message is that the relationship between city size and productivity in African countries is weak and highly variable with large differences across sectors and countries. This suggests that the mechanisms of agglomeration have the potential to work, but in many instances they don't operate strongly, resulting in a loose relationship. The results are consistent with the view that it is the contextual factors – rather than the absence of the fundamental agglomeration forces – which misfire and undermine the dynamism of urban economies. There is no intrinsic reason why African firms and workers can't benefit from sharing, matching and learning if the conditions are conducive. This is both the challenge and opportunity inherent in Africa's urbanisation.

Another way to interpret these findings is that African cities exhibit relative agglomeration benefits. Because the economic environment is challenging in most places, larger cities can still outperform smaller ones, even in the face of weak governance and infrastructure shortfalls. In other words, African cities perform better than other settlements within their nations, even if they fall short of global benchmarks. The implication is that policymakers should not overlook the latent potential of cities to drive economic transformation and growth, despite their current shortcomings.

Yet, the weak relationship between city size and productivity cautions against simplistic views about the transformative power of metropolitan areas. While agglomeration economies exist, increasing city size alone is insufficient to drive large productivity gains. Hence continuing urbanisation will not automatically generate greater prosperity. Instead, firm-level, environmental and institutional factors, such as workforce skills, access to finance and effective regulation, arguably play a more direct role in shaping productivity outcomes. Policymakers should pursue a holistic approach that includes urban investment, regulatory reforms and targeted support for firms.

The stronger agglomeration effects in service industries compared to manufacturing are particularly noteworthy. This suggests that African cities may have some potential to support service-led growth, especially as digital technologies enhance the tradability of services and their role within global value chains expands (Visagie and Turok, 2023). These findings also challenge the over-simplified characterisation of African cities as mere consumption hubs (Jedwab et al, 2025), highlighting the need for a more nuanced understanding of the types of services driving urban productivity. However, it is also plausible that the higher agglomeration premium for services reflects intensified competition in larger local markets, precisely because many services are locally traded. This issue warrants further careful research.

Finally, we find little evidence of an association between firm-level productivity and the characteristics of the built environment in African cities. This may partly reflect limitations in the availability of spatial indicators, which rely on satellite imagery, and may not fully capture the complex realities of urban structure and connectivity. More importantly, it is probably the quality of investment in the built environment which matters more than a particular physical layout. Simply increasing density is unlikely to yield productivity gains unless these investments enhance functional efficiency and improve access to markets (Bryan et al, 2020).

References

1. Bryan, G., Glaeser, E., & Tsivanidis, N. (2020). Cities in the developing world. *Annual Review of Economics, 12*, 273–297.
2. Buckley, R., & Kallergis, A. (2014). Does African urban policy provide a platform for sustained economic growth? In S. Parnell & S. Oldfield (Eds.), *The Routledge handbook on cities of the global south* (pp. 173–190). Routledge.
3. Burger, M., Ianchovichina, E., & Akbar, P. A. (2022). Heterogeneous agglomeration economies in developing countries: The roles of firm characteristics, sector tradability, and urban mobility. *Policy Research Working Papers 9954*. https://doi.org/10.1596/1813-9450-9954
4. Ciccone, A., & Hall, R. E. (1996). Productivity and the density of economic activity. *American Economic Review, 86*(1), 54–70.
5. Collier, P. (2017). African urbanization: An analytic policy guide. *Oxford Review of Economic Policy, 33*(3), 405–437.
6. Collier, P., & Venables, A. J. (2015). Housing and urbanisation in Africa: Unleashing a formal market process. In E. Glaeser & A. Joshi-Ghani (Eds.), *The urban imperative: Towards competitive cities* (pp. 413–436). Oxford University Press.
7. Collier, P., & Venables, A. J. (2017). Urbanization in developing economies: The assessment. *Oxford Review of Economic Policy, 33*(3), 355–372.
8. Combes, P.-P., & Gobillon, L. (2015). The empirics of agglomeration economies. *Handbook of Regional and Urban Economics, 5*, 247–348. https://doi.org/10.1016/b978-0-444-59517-1.00005-2
9. De la Roca, J., & Puga, D. (2017). Learning by working in big cities. *Review of Economic Studies, 84*(1), 106–142.
10. Dijkstra, L., Florczyk, A. J., Freire, S., Kemper, T., Melchiorri, M., Pesaresi, M. & Schiavina, M. (2021). Applying the degree of urbanisation to the globe: A new harmonised definition reveals a different picture of global urbanisation. *Journal of Urban Economics*. https://doi.org/10.1016/j.jue.2020.103312
11. Duranton, G., & Puga, D. (2004). Micro-foundations of urban agglomeration economies. In J. V. Henderson & J.-F. Thisse (Eds.), *Handbook of Regional and Urban Economics* (Vol. 4, pp. 2063–2117). North-Holland.
12. European Commission. (2021). *Applying the degree of urbanisation: A methodological manual to define cities, towns, and rural areas for international comparisons.* Publications Office of the European Union.
13. European Commission. (2024). Data Package 2023, Publications Office of the European Union, Luxembourg, 2024, doi:10.2760/098587, JRC133256
14. Glaeser, E. L., & Henderson, V. (2017). Urban economics for the developing world: An introduction. *Journal of Urban Economics, 98*, 1–5.
15. Glaeser, E. L., & Mare, D. C. (2001). Cities and skills. *Journal of Labor Economics, 19*(2), 316–342.
16. Glaeser, E. L., & Xiong, W. (2017). Urban productivity in the developing world. *Oxford Review of Economic Policy, 33*(3), 373–404.
17. Gollin, D., Jedwab, R., & Vollrath, D. (2016). Urbanization with and without industrialization. *Journal of Economic Growth, 21*(1), 35–70. https://doi.org/10.1007/s10887-015-9121-4
18. Graham, D. J., & Gibbons, S. (2019). Quantifying wider economic impacts of agglomeration for transport appraisal: Existing evidence and future directions. *Economics of Transportation, 19*, 100121. https://doi.org/10.1016/j.ecotra.2019.100121
19. Grover, A., Lall, S., & Maloney, W. F. (2022). *Place, productivity, and prosperity: Revisiting spatially targeted policies for regional development.* World Bank. https://doi.org/10.1596/978-1-4648-1670-3
20. Grover, A., Lall, S., & Timmis, J. (2023). Agglomeration economies in developing countries: A meta-analysis. *Regional Science and Urban Economics, 101*, 103901. https://doi.org/10.1016/j.regsciurbeco.2023.103901
21. Hanlon, W., & Miscio, A. (2014). Agglomeration: A dynamic approach (*NBER Working Papers No. 20728*). National Bureau of Economic Research.
22. Hasan, R., Jiang, Y., & Rafols, R. M. (2017). Urban agglomeration effects in India: Evidence from town-level data. *Asian Development Review, 34*(2), 201–228.
23. Henderson, J. V., Nigmatulina, D., & Kriticos, S. (2021). Measuring urban economic density. *Journal of Urban Economics, 125*, 103188. https://doi.org/10.1016/j.jue.2019.103188

24. Jedwab, R., Christiaensen, L., & Gindelsky, M. (2015). Demography, urbanization, and development: Rural push, urban pull, and … urban push? *Policy Research Working Paper 7333.* Washington, DC: The World Bank.
25. Jedwab, R., Ianchovichina, E., & Haslop, F. (2025). The employment profile of cities around the world: Consumption vs. production cities and economic development. *World Development, 188*, 106883. https://doi.org/10.1016/j.worlddev.2024.106883
26. Jones, P., Collier, P., & Spijkerman, D. (2020). Cities as engines of growth: Evidence from a new global sample of cities. *Journal of Applied Business and Economics, 22*(2), 158–188.
27. Jones, P., D'Aoust, O., & Bernard, L. (2017). The urban wage premium in Africa. In R. B. Freeman & L. F. Katz (Eds.), *African successes: Human capital* (pp. 33–53). https://doi.org/10.1007/978-3-319-51565-6_3
28. Kayizzi-Mugerwa, S., Shimeles, A., & Yameogo, N. (Eds.). (2014). *Urbanisation and socio-economic development in Africa: Challenges and opportunities.* Routledge.
29. Lall, S., Henderson, V., & Venables, A. J. (2017). *Africa's cities: Opening doors to the world.* World Bank. https://doi.org/10.1596/978-1-4648-1044-2
30. Lall, S., Lebrand, M., Park, H., Sturm, D., & Venables, A. J. (2021). *Pancakes to pyramids: City form to promote sustainable growth.* World Bank.
31. Leipziger, D., Freire, M. E., & Lall, S. (2015). Africa's urbanisation: Challenges and opportunities. In C. Monga & J. Y. Lin (Eds.), *The Oxford handbook of Africa and economics, volume 1: Context and concepts* (Chapter 31). Oxford University Press.
32. Melo, P., Graham, D. J., & Noland, R. B. (2009). A meta-analysis of estimates of urban agglomeration economies. *Regional Science and Urban Economics, 39*, 332–342.
33. Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics, 5*(1), 65–103.
34. OECD. (2022). *Africa's urbanisation dynamics: The economic power of Africa's cities.* OECD Publishing.
35. Rodrik, D. (2016). Premature deindustrialization. *Journal of Economic Growth, 21*(1), 1–33.
36. Turok, I., & McGranahan, G. (2013). Urbanisation and economic growth: The arguments and evidence for Africa and Asia. *Environment and Urbanization, 25*(2), 465–482.
37. Turok, I., Scheba, A., & Visagie, J. (2023). Rethinking urbanization and economic development: A synopsis. *Environment and Urbanization, 35*(2), 297–309. https://doi.org/10.1177/09562478231193829
38. UN-DESA. (2018). *Revision of world urbanization prospects.* United Nations Department of Economic and Social Affairs. New York: UN-DESA
39. UN-ECA. (2022). *Cities: Gateways for Africa's regional economic integration.* United Nations Economic Commission for Africa. Addis Ababa: UN-ECA
40. UN-Habitat. (2014). *State of African cities report.* Nairobi: UN-Habitat.
41. UN-Habitat. (2022). *World cities report: Envisaging the future of cities.* Nairobi: UN-Habitat.
42. Visagie, J. and Turok, I. (2023) 'Recognising the role of tradable service exports in the South African economy – An untapped resource?', *South African Journal of International Affairs*, 34(4), pp.661-683.
43. World Bank. (2009). *World Development Report: Reshaping Economic Geography.* Washington, DC: World Bank.
44. World Bank. (2013). *Planning, connecting and financing cities – Now.* Washington, DC: World Bank.
45. World Bank. (2015). *Stocktaking of the housing sector in Sub-Saharan Africa: Challenges and opportunities.* Washington, DC: World Bank.
46. World Bank. (2017). *Firm-level productivity estimates.* World Bank Enterprise Analysis Unit. Washington, DC: World Bank.

Appendix

Table A1: Agglomeration economies for manufacturing firms (dependent variable: wages per worker)

Model	#1		#2		#3		#4		#5		#6		#7	
	OLS: pop size (no controls)		OLS: pop size		2SLS: pop size		OLS: pop dens		2SLS: pop dens		OLS: pop size and built env		2SLS: pop size and built env	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
AGGLOMERATION MEASURE														
Instrumented	*no*		*no*		*Yes*		*no*		*Yes*		*no*		*yes*	
log city size	0.058	0.263	0.040	0.126	0.045	0.104					0.038	0.141	0.037	0.163
log pop dens							-0.072	0.489	0.179	0.385				
REGRESSION CONTROLS														
Built env controls	*no*		*no*		*No*		*no*		*No*		*yes*		*yes*	
tall building share											-0.007	0.401	-0.007	0.422
road surface share											0.038	0.403	0.038	0.395
Firm controls	*no*		*yes*		*Yes*		*yes*		*Yes*					
firm size			**0.000**	0.074	**0.000**	0.072	**0.000**	0.075	**0.000**	0.077	**0.000**	0.090	**0.000**	0.085
firm age			0.003	0.266	0.003	0.192	0.003	0.219	0.003	0.193	0.003	0.205	0.003	0.199
top manger's yrs experience			**0.005**	0.025	**0.074**	0.056	**0.006**	0.017	**0.079**	0.039	**0.076**	0.054	**0.077**	0.050
sales lost to power outages			**-0.005**	0.016	**-0.005**	0.014	**-0.005**	0.015	**-0.005**	0.014	**-0.005**	0.014	**-0.005**	0.012
share unskilled workers			**-0.006**	0.000	**-0.0(**	0.000	**-0.005**	0.000	**-0.006**	0.000	**-0.006**	0.000	**-0.006**	0.000
dummy: foreign owned			**0.360**	0.000	**0.360**	0.000	**0.367**	0.000	**0.365**	0.000	**0.361**	0.000	**0.361**	0.000
dummy: innovator			**0.285**	0.000	**0.284**	0.000	**0.291**	0.000	**0.286**	0.000	**0.284**	0.000	**0.284**	0.000
dummy: exporter			**0.252**	0.001	**0.253**	0.001	**0.250**	0.002	**0.244**	0.002	**0.253**	0.001	**0.253**	0.001
dummy: form training offered			**0.128**	0.045	**0.128**	0.042	**0.132**	0.039	**0.129**	0.041	**0.128**	0.045	**0.128**	0.042
Constant	**6.251**	0.000	**6.608**	0.000	**6.434**	0.000	**7.809**	0.000	**5.396**	0.006	**6.522**	0.000	**6.528**	0.000
Industry controls	*no*		*yes*		*Yes*		*yes*		*Yes*					
Country controls	*no*		*yes*		*Yes*		*yes*		*Yes*					
Observations	5082		5082		5082		5082		5082		5082		5082	
R-squared	0.0029		0.3623		0.3679		0.3613		0.3649		0.3682		0.3682	
Instrument Diagnostics														
KP rk LM (p-value)					0.006				0.108				0.000	
KP rk Wald F-stat					22.821				5.837				80.082	

Notes: coefficients in bold are statistically significant at a 95% level of confidence; Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Table A2: Agglomeration economies for services firms (dependent variable: wages per worker)

Model	#1		#2		#3		#4		#5		#6		#7	
	OLS: pop size (no controls)		OLS: pop size		2SLS: pop size		OLS: pop dens		2SLS: pop dens		OLS: pop size and built env		2SLS: pop size and built env	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
AGGLOMERATION MEASURE														
Instrumented	*No*		*no*		*yes*		*no*		*yes*		*No*		*yes*	
log city size	0.070	0.332	**0.136**	0.000	**0.123**	0.009					**0.139**	0.000	**0.123**	0.008
log pop dens	0.070	0.332	**0.136**	0.000	**0.123**	0.009					**0.139**	0.000	**0.123**	0.008
REGRESSION CONTROLS														
Built env controls	*No*		*no*		*no*		*no*		*no*					
tall building share											-0.010	0.375	-0.010	0.381
road surface share											0.044	0.531	0.039	0.586
Firm controls	*No*		*yes*		*yes*		*yes*		*yes*					
firm size			0.000	0.402	0.000	0.421	0.000	0.444	0.000	0.468	0.000	0.429	0.000	0.430
firm age			**0.003**	0.036	**0.004**	0.006	**0.003**	0.038	**0.004**	0.008	**0.004**	0.006	**0.004**	0.005
top manger's yrs experience			**0.014**	0.000	**0.172**	0.000	**0.015**	0.000	**0.177**	0.000	**0.170**	0.000	**0.171**	0.000
sales lost to power outages			**-0.005**	0.023	**-0.005**	0.021	**-0.005**	0.026	**-0.005**	0.022	**-0.005**	0.023	**-0.005**	0.021
share unskilled workers			**-0.007**	0.000	**-0.007**	0.000	**-0.007**	0.000	**-0.007**	0.000	**-0.007**	0.000	**-0.007**	0.000
dummy: foreign owned			**0.304**	0.000	**0.306**	0.000	**0.317**	0.000	**0.307**	0.000	**0.304**	0.000	**0.305**	0.000
dummy: innovator			**0.252**	0.001	**0.250**	0.001	**0.251**	0.001	**0.252**	0.000	**0.251**	0.001	**0.251**	0.001
dummy: exporter			0.023	0.782	0.024	0.770	0.008	0.927	0.012	0.891	0.028	0.745	0.026	0.758
dummy: form training offered			**0.108**	0.045	**0.107**	0.044	**0.117**	0.033	**0.116**	0.037	**0.106**	0.050	**0.107**	0.044
Constant	**6.290**	0	**5.521**	0.000	**5.456**	0.000	**6.702**	0.000	8.316	0.370	**5.241**	0.000	**5.453**	0.000
Industry controls	*No*		*yes*		*yes*		*yes*		*yes*					
Country controls	*No*		*yes*		*yes*		*yes*		*yes*					
Observations	5093		5093		5093		5093		5093		5093		5093	
R-squared	0.0032		0.3012		0.3067		0.2945		0.2988		0.307		0.3070	
Instrument Diagnostics														
KP rk LM (p-value)					0.000				0.764				0.000	
KP rk Wald F-stat					20.961				0.428				26.536	

Notes: coefficients in bold are statistically significant at a 95% level of confidence; Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates

Table A3: Agglomeration economies for manufacturing firms (dependent variable: total factor productivity)

Model	#1		#2		#3		#4		#5		#6		#7	
	OLS: pop size (no controls)		OLS: pop size		2SLS: pop size		OLS: pop dens		2SLS: pop dens		OLS: pop size and built env		2SLS: pop size and built env	
	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value	Coeff	P-value
AGGLOMERATION MEASURE														
Instrumented	*No*		*no*		*yes*		*no*		*Yes*		*no*		*yes*	
log city size	0.020	0.678	0.044	0.212	0.020	0.630					**0.060**	0.055	0.039	0.243
log pop dens							**0.377**	0.019	0.057	0.818				
REGRESSION CONTROLS														
Built env controls	*No*		*no*		*no*		*no*		*No*		*yes*		*yes*	
tall building share											0.013	0.117	0.010	0.211
road surface share											-0.006	0.926	-0.014	0.837
Firm controls	*No*		*yes*		*yes*		*yes*		*Yes*					
firm size			**0.000**	0.015	**0.000**	0.015	**0.000**	0.012	**0.000**	0.017	**0.000**	0.018	**0.000**	0.016
firm age			-0.001	0.648	-0.001	0.738	-0.001	0.604	-0.001	0.750	-0.001	0.743	-0.001	0.769
top manger's yrs experience			0.001	0.781	0.009	0.838	0.001	0.722	0.012	0.806	0.003	0.955	0.006	0.903
sales lost to power outages			-0.004	0.147	-0.003	0.146	-0.004	0.135	-0.004	0.145	-0.003	0.157	-0.003	0.153
share unskilled workers			0.000	0.834	0.000	0.758	0.000	0.999	0.000	0.737	0.000	0.737	0.000	0.713
dummy: foreign owned			0.080	0.372	0.082	0.352	0.084	0.348	0.084	0.337	0.076	0.391	0.078	0.369
dummy: innovator			0.059	0.371	0.063	0.341	0.057	0.390	0.064	0.331	0.057	0.379	0.060	0.347
dummy: exporter			0.158	0.130	0.158	0.124	0.142	0.166	0.155	0.125	0.157	0.135	0.157	0.128
dummy: form training offered			0.091	0.201	0.093	0.187	0.088	0.215	0.093	0.188	0.090	0.212	0.092	0.198
Constant	**2.626**	0	**1.858**	0.000	**2.170**	0.000	-1.012	0.476	1.913	0.401	**1.600**	0.000	**1.899**	0.000
Industry controls	*No*		*yes*		*yes*		*yes*		*Yes*					
Country controls	*No*		*yes*		*yes*		*yes*		*Yes*					
Observations	3994		3994		3994		3994		3994		3994		3994	
R-squared	0		0.3732		0.3802		0.3759		0.3805		0.3812		0.381	
Instrument Diagnostics														
KP rk LM (p-value)					0.025				0.127				0.000	
KP rk Wald F-stat					19.248				6.841				92.521	

Notes: coefficients in bold are statistically significant at a 95% level of confidence; Source: World Bank Enterprise Surveys; Global Human Settlement Layer; Authors' own estimates