

Urban Informality in Sub-Saharan Africa

Profiling Workers and Firms in an Urban Context

Wendy Cunningham

David Newhouse

Federica Ricaldi

Feraud Tchuisseu Seuyong

Mariana Viollaz

Ifeanyi Edochie



WORLD BANK GROUP

Social Protection and Jobs Global Practice &
Poverty and Equity Global Practice

February 2024

Abstract

This paper describes the state of informal sector work in urban Sub-Saharan Africa, using household surveys from 26 countries representing 61 percent of the population of Sub-Saharan Africa and firm surveys from three countries. Five main conclusions emerge. First, the urban informal sector is large and persistent in Sub-Saharan Africa. Approximately 56 to 65 percent of urban workers are informal, half of whom are self-employed. Data from five countries suggest little systematic reduction in the prevalence of informality during the 2010s. Second, heterogeneity in the African informal sector cuts along demographic lines. Women are overrepresented in informal self-employment, men in informal wage work, and youth in unpaid employment. Third, while the urban informal workers are, on average, poorer

and in less-skilled occupations than formal sector workers, the majority are not extremely poor and are in mid-skilled occupations. Fourth, informal enterprises are small and are challenged to survive and grow into job-creating firms. Few find much benefit from registration given the costs, both monetary (taxes) and transactional (information about the registration process). Fifth, access to urban public services (utilities) is weakly associated with the probability of working in an informal job, although access to mobile phones is high across all job types. If thriving urban jobs are to contribute to economic and social development in Africa, it will be crucial for policies and programs to take into consideration the heterogeneity in jobs, the profile of workers, and the urban context.

This paper is a product of the Social Protection and Jobs Global Practice and the Poverty and Equity Global Practice. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at <http://www.worldbank.org/prwp>. The authors may be contacted at wcunningham@worldbank.org, dnewhouse@worldbank.org, fricaldi@worldbank.org, and iedochie@worldbank.org.

The Policy Research Working Paper Series disseminates the findings of work in progress to encourage the exchange of ideas about development issues. An objective of the series is to get the findings out quickly, even if the presentations are less than fully polished. The papers carry the names of the authors and should be cited accordingly. The findings, interpretations, and conclusions expressed in this paper are entirely those of the authors. They do not necessarily represent the views of the International Bank for Reconstruction and Development/World Bank and its affiliated organizations, or those of the Executive Directors of the World Bank or the governments they represent.

Urban Informality in Sub-Saharan Africa: Profiling Workers and Firms in an Urban Context

Wendy Cunningham, David Newhouse, Federica Ricaldi, Feraud Tchuisseu Seuyong,
Mariana Viollaz, Ifeanyi Edochie

Originally published in the [Policy Research Working Paper Series](#) on *February 2024*. This version is updated on *September 2024*.
To obtain the originally published version, please email prwp@worldbank.org.

Key words: Urban, Informality, Sub-Saharan Africa

JEL Classification: J1, J4

1. Introduction

Sub-Saharan Africa is urbanizing, with profound implications for its labor markets. Sub-Saharan Africa's (SSA's) urban population, as a percentage of total population was 42 percent in 2021, following a steady upward trend since 1960 when it was only 15 percent (UN Population Division, 2018).¹ The urban population in SSA is expected to surpass the rural by 2036 (Lam et al. 2019). Rural to urban migration is necessarily accompanied by a shift out of agricultural work and into urban services and manufacturing jobs (De Vreyer and Roubaud 2013).² The general consensus is that urban jobs can be a source of inclusive productivity, namely better jobs for the poor or vulnerable characterized by higher earnings and well-being than in rural areas, due to the many benefits offered in urban areas.³ At the same time, if not well managed, migration to urban areas can be a source of constraints – overcrowding, increased crime, weak land tenure – leading to jobs that are no more welfare improving nor more productive than the agricultural jobs that were left behind (Lall et al. 2017).

The formal private sector in Sub-Saharan Africa is unlikely to grow sufficiently fast in the short run to absorb migrants, new labor market entrants, and existing informal sector workers in urban areas. A myriad of entrenched institutional and economic factors hinders the rapid development of a jobs-intensive formal sector.⁴ The East Asian Miracle, driven by the rapid expansion of foreign investment in manufacturing that employed masses of low-skilled people, does not seem feasible in the short term in today's African economies. Instead, SSA will need to undertake a long-term process of institutional and economic reform to bring growth and better jobs to the region.

Instead, the informal sector – broadly defined – is the main source of jobs in urban SSA. The informal economy can be characterized as “small-scale, semi-legal, often low-productivity, frequently family-based, perhaps pre-capitalistic enterprises” (Maloney 2004) and the jobs within.⁵ Drawing from Bandiera et al (2022), we might define them as being less-organized and operating in low value-added sectors. De Vreyer and Roubaud (2013) use data from 2000-2001 from eight principle Sub-Saharan African cities to estimate that 70

¹ The percentage varies widely between 14 percent in Burundi to 90 percent in Gabon.

² Urban population growth in Sub-Saharan Africa can be attributed to high birth rates as well as urban to rural migration (Lam et al 2019).

³ The literature on the economic benefits of urban agglomerations is well established. For example, urban areas provide more access to public and private infrastructure (water, electricity, sewerage) and services (security forces, banks, schools), greater diversity in production and thus opportunities for trade, larger markets of suppliers and clients, and knowledge spillovers. However, as argued in Lall et al (2017), the economic benefits characteristic of urban zones are often absent in African cities.

⁴ See, for example, Calderon (2022), Coulibaly et al. (2022), Abreha et al. (2021).

⁵ ILO (2018) finds that 5.5% of the African labor force is classified as informal employees (not receiving legally mandated social benefits) working in formal sector firms (registered with authorities). The statistic includes data from SSA and Northern Africa and uses national data (rural and urban).

percent of urban jobs are in the informal sector. ILO (2018) estimates that 76.3 percent of Africa's urban jobs are in the informal sector.

Empirical literature supports the low-productivity and small-scale conceptualization of informal enterprises in Africa. Informal sector enterprises are mostly small operations that employ family members. Many do not operate out of a fixed location and oftentimes they are not a full-time business, while co-mingling household and business finances (Loening, Rijkers, and Söderbom, 2008; Fox and Sohnesen, 2012; La Porta and Schleifer, 2014). Compared to larger enterprises, their capital stock is minimal, and their productivity is low (La Porta and Schleifer, 2014), while the owners have limited management skills compared to formal and larger firms' owners. A small minority of African self-employed register with local authorities (Fox and Sohnesen, 2012) or grow and create jobs.

A growing literature characterizes informal sector workers in SSA, highlighting factors consistent with loosely organized forms of work. Workers in the informal sector are younger,⁶ more female,⁷ less educated, and earn lower income, on average than formal sector workers (Danquah et al 2019, ILO 2018, Basu et al. 2018). Informal jobs are generally characterized by low-income levels in comparison to formal sector jobs (see, inter alia, Grimm et al 2012, De Vreyer and Roubaud 2013, Nguimkeu 2014, Fox and Sohnesen 2012).⁸ The high rate of informal employment in Africa goes hand in hand with the high levels of poverty in the region (ILO 2018).

If thriving urban jobs are to contribute to economic and social development in Sub-Saharan Africa, it will be crucial for policy makers to have a better understanding of the size, characteristics and nature of urban labor markets and, especially, urban informality. The existing literature explores urban workers, formal-informal workers, and household enterprises but has yet to pull them together into a single picture of urban labor markets in SSA. The sheer magnitude of the urban informal sector calls for a deeper understanding of the heterogeneity within to enable effective policy making.

This paper provides a detailed description of the magnitude of and heterogeneity within urban SSA labor markets, with an emphasis on informality, to encourage policy makers to design more nuanced interventions to serve the different segments within. It uses data from 26 harmonized household surveys, representing 61 percent of the SSA population,⁹ and three Informal Enterprises Surveys (IES) to sketch a profile of the urban

⁶ ILO (2018) also finds that informal sector workers are not prime-aged adults, namely they are youth and older workers.

⁷ Danquah et al (2019) find that in a four-country, national sample, women are over-represented among informal wage workers but under-represented among the self-employed.

⁸ Similar patterns are found in Perry et al (2008) for Latin America, and in Kanbur et al (2017) and ILO (2018) globally.

⁹ While there are 1.1 billion Sub-Saharan Africans, only 31 percent are working. The urban workforce is 36 percent of the national workforce. Thus, there are roughly 125 million urban workers in the region. We have data for 85 million urban workers in Sub-Saharan Africa.

informal sector. It starts with a description of the distribution of employment across a formality-employment type continuum in the 26 countries and finds that approximately 56 percent of urban jobs are informal, though this statistic rises to 70 percent if we exclude the two countries with the largest urban populations: Nigeria and South Africa. It then profiles urban workers within each formality-employment type category. The paper turns to firms, using data from three countries to unpack informal sector firm characteristics and heterogeneity therein. The next section focuses on the relationship between the urban environment and different types of informal and formal jobs. It starts with a profile of urban infrastructure and services for each employment type, using data from household surveys and the IES. It then uses geospatial data and a Census extract in one country to understand the role played by urban variables that are not measurable in surveys. Finally, it shares policy-related reflections that emerge from the descriptive analysis.

The paper makes three main contributions to the existing descriptive literature on informality in SSA. First, it limits its analysis to informality in the urban space. Second, the analysis considers heterogeneity within the urban labor market by exploring different categories of employment (i.e. self-employed, employer, or wage employee) in the informal and formal sectors. This is in line with recent literature that differentiates by employment categories within the informal sector to explore questions of dualism, differential job quality, transitions between employment types,¹⁰ and heterogeneity within informal enterprises.¹¹ Third, the paper jointly considers employment types and urban variables, such as access to public services and urban density.

The analysis presents five main results. First, the urban informal sector is large and persistent in SSA. More than half of urban SSA workers can be considered informal, reaching 70 percent if we drop the two most populous countries in the sample. Of these, 52 percent are self-employed, while the others are informal paid (33 percent) or unpaid wage employees. Analysis of a sample of five countries find no clear sign that informality is declining in the short-run or that greater levels of short-term economic growth are associated with declines in informality. Second, there is heterogeneity among African informal sector workers. Women are over-represented in informal self-employment, men are informal employees, and youth are disproportionately in unpaid employment. Third, while the urban informal are, on average, poorer and in less-skilled occupations than formal sector workers, the majority are not extreme poor and work in mid-skilled occupations. Fourth, informal enterprises are small (in terms of workers and earnings) and the informal

¹⁰ For the purposes of this paper, employment “categories” are groups of jobs that have the same employer-employee relationship, namely self-employed, employers and employees. Employment “types” adds formality-variables to the definitions, leading to a greater number of groups including formal and informal self-employed, formal or informal employers, and employees who may be formal, informal, or unpaid (where “unpaid” is a subset of “informal employees”).

¹¹ See for example Fox and Sohnesen (2012), De Vreyer and Roubaud (2013), Gindling and Newhouse (2014), Danquah et al (2019), Grimm et al (2012), and Adoho et al (2018) for Africa and Ulyssea (2020) for a global review.

self-employed and employers are similarly challenged to survive and grow into job-creating firms. They believe there is little benefit from registration. Their heterogeneity is found in their business practices. Fifth, after controlling for individual and job characteristics, access to public services is weakly associated with the probability of working in an informal job, though access to mobile phones is high across all job types, suggesting that digital connectivity is a valuable policy focus for informal (and formal) workers.

2. Data and Definitions

We are interested in understanding the nature of jobs, both in terms of the worker's relationship to the firm owner and in terms of job benefits and earnings potential offered by the firm. For the former, we consider three groups: employees who work for others, self-employed who do not have paid employees, and employers. For the latter, we use the firm organization concept put forth by Bandiera et al. (2022) and refer to informal enterprises as those that are less "organized" in terms of business practices, linkages to the organized economy and institutions, spillovers to the household, and value-added, as compared to more organized and higher value-added formal firms.^{12,13}

We use four data sets in the analysis. First, we use the Global Monitoring Database (GMD) to create a profile of urban informal workers in 26 SSA countries. The GMD is the World Bank's repository of harmonized household surveys used for poverty monitoring. We limit the sample to the 26 countries that provide information on employment type, education level, and labor benefits for wage employees, to classify workers as formal or informal.¹⁴ The countries included in the sample represent about 85 million urban workers, equivalent to 68 percent of the total urban workforce of Sub-Saharan Africa.¹⁵ Table A1 in the Appendix lists the survey and year used for each country. Since this paper's focus is urban workers, we restrict the analysis to urban observations and to people 15 years of age or older.

¹² Firm registration status, a common variable used in the literature to proxy firm formality, is inadequate for three reasons. First, most SSA economies have limited reach of, requirements for, enforcement of, and benefits tied to legislation regulating firms and labor markets, making registration status a poor proxy for informal firm characteristics. For example, Kenya does not require self-employed firm owners to contribute to social security and most do not reach the VAT threshold. Second, firm registration requirements differ across countries, resulting in firms of similar nature in different countries being assigned different formality status. Third, registration status is not an effective policy variable to improve the quality of firms and jobs (Levenson and Maloney 1998).

¹³ For measurement purposes, ILO (2018) defines an informal firm based on bookkeeping practices, registration status, place of work, connection with household, and firm size (fewer than 6 workers), which map to many of the Bandiera et al (2022) concepts. These variables are largely not observed in our data so we use proxies to identify informal enterprises and workers.

¹⁴ The definition of employment in the GMD did not change during this period, although the 19th ICLS adopted a new definition of employment in 2013.

¹⁵ Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad, Côte d'Ivoire, Eswatini, Gabon, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Malawi, Mali, Niger, Nigeria, Senegal, Sierra Leone, South Africa, São Tomé and Príncipe, Togo, Uganda, Zambia and Zimbabwe.

Second, we use the World Bank's Informal Enterprise Survey (IES) to profile firms. Firms included in this survey are, by definition, informal, i.e., the businesses are not registered with their governments.¹⁶ The IES is conducted only with manufacturing or services firms in main urban areas. We use the most recent Mozambique, Zambia, and Zimbabwe surveys for our analysis,¹⁷ which follow a sampling strategy that allows for the generation of sample weights.¹⁸

Third, we use geospatial big data and a national Census extract to analyze how the urban environment correlates with employment types. The Mozambique Census was collected in 2017 and provides detailed information on employment type and geo-location at the neighborhood level. A 10 percent extract was shared by the statistical department. Four geospatial big data sets were utilized, each providing data for 1 km by 1 km tessellation (grid) of the country. First, the NASA's Earth Observatory raster dataset¹⁹ provides a measure of annual nighttime light brightness composites.²⁰ Second, the University of Michigan's High Resolution Electricity Access (HREA) data uses the aforementioned nighttime lights data and machine learning techniques to estimate electrification access between 2012 and present. Third, the WorldPop raster provides information on population density, building characteristics, and an urbanization composite at 1 km by 1 km tessellations. Fourth, the Open CellID database provides locations of cell phone towers.

2.1 Defining Employment Types to Classify Workers

Inspired by Danquah et al (2019), we create a classification of six urban employment types using the GMD. We use information on workers' employment category in their main job to classify workers as employees, self-employed or employers. To assign formality status to employees, we use variables that indicate receipt of labor benefits (social security payments, paid annual leave, paid sick leave) and wages, as per ILO (2018).²¹ This relates to

¹⁶ The definition of formal registration varies by country.

¹⁷ In Mozambique, the survey covers three cities (Beira, Maputo, and Nampula). In Zambia, the IES was carried out in three cities (Kitwe, Lusaka, and Ndola), while in Zimbabwe, it covered one city (Harare). Table A1 in the Appendix lists the survey and year used for each country.

¹⁸ The IES collected prior to 2017 did not use a sample frame that could be used to generate sample weights. A key challenge for conducting a representative sample survey of informal sector businesses is the lack of a proper sampling frame of establishments. These businesses, by definition, are not registered with government agencies. However, they tend to cluster in certain geographic areas, such as low-income residential areas. Starting in 2017, the IES implemented Adaptive Cluster Sampling (area-based sampling) in which starting grids (usually squares) are selected. All informal businesses in selected squares are enumerated, using a short-form questionnaire, and a randomly selected subset of the enumerated businesses receive the long-form questionnaire.

¹⁹ A raster is an image file, where each position in the image is an x-y coordinate with a functional mapping to a specific value. They are also called GeoTIFF (Tagged Image File Format) Rasters or tif files.

²⁰ The variable is generated from the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite (~463 meter resolution).

²¹ The employment category and labor benefits variables pertain to the previous week, or the last year if information on the previous week is not available.

applicability of labor legislation on workers. The GMD does not include variables recommended by ILO (2018) to identify enterprise formality, such as firm registration or bookkeeping practices. Instead, based on Fields (2005) idea of upper- and lower-tiers in the informal labor market, and the concept of the informal sector as being less organized and productive, we use level of education and employer status to separate enterprise owners into more formal (organized and higher value-added) and less formal. As per Danquah et al (2019), we consider enterprise owners with at least one paid employee to be formal. Among the self-employed, we use education level to proxy formality status. Self-employed with no more than completed secondary education are classified as lower-tier informal while those with secondary or more are considered upper tier informal. In all cases, we recognize that these are proxies for a broader concept for informality related to scale and productivity of the firm. Though all firms in the IES are informal, as defined by being registered with authorities, we also use the education proxy to create “upper” and “lower” tier categories of informal self-employed.

We justify the use of education as a proxy for informal status in three ways. First, ILO (2018) finds a strong negative correlation between informality (defined as firm registration with authorities) and education.²² Second, education is reported more often than other potential variables to proxy informality. Danquah et al (2019) use occupation to assign formality status, where skilled occupations (ISCO groups 1-4) – that are likely more organized and with higher value added – are classified as upper informal and the rest are lower informal.²³ Due to missing data, if we were to use occupational codes, we would lose 30.5 percent of our sample. Only 8.5 percent of the sample does not report education. In testing education as a proxy for occupation, we find a close match. Specifically, more than 97 percent of those with less than complete secondary education are not in professional occupations (ISCO groups 1-4) as compared to 78 percent of those with at least a secondary school diploma. Third, we are interested in the occupational distribution within employment types, so we allow occupation to be an exogenous variable and use education level as a proxy for degree of formality.

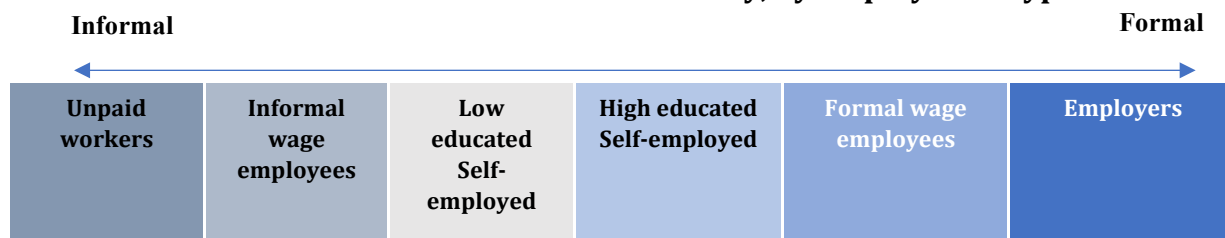
We consider a continuum of urban informality-formality. Figure 1 presents a typology of employment types, from the more informal and less organized jobs (on the left) to the more formal and more organized (formal wage employees) at the right. The characteristics

²² The level of education is closely linked to informality in Africa. Those with no education are mostly informal (94 percent). The rate of informality reduces to 88.5 percent with primary education and further decreases to 68.1 percent for those with secondary education and to 27.0 percent among those with tertiary education (ILO, 2018).

²³ The ISCO (international standard classification of occupations) groups a labor market’s occupations into 10 major groups: managers (1), professionals (2), technicians and associate professionals (3), clerical support workers (4), service and sales workers (5), skilled agriculture/forestry/fishery workers (6), craft related trade workers (7), plant and machine operators and assemblers (8), elementary occupations (9), armed forces (10). [ISCO - International Standard Classification of Occupations \(ilo.org\)](http://ilo.org). We do not include ISCO code 10 in our analysis due to the small sample size.

to locate each employment type on the continuum are drawn from the literature, as discussed below.

Figure 1. Schematic of urban informality, by employment type



Source: Author's elaboration.

Unpaid workers appear at the left of the continuum of employment types. Many are family workers (Filmer and Fox, 2014). They do not have worker protections or a wage, though their labor may be compensated in-kind as a part of the family unit that owns an enterprise or as apprentices.

Self-employed workers are those who work for themselves, typically as household enterprise owners. Most depend on family labor, often unpaid (Filmer and Fox, 2014). Since the GMD does not include information on the registration status of the enterprise, we follow the literature that distinguishes lower and upper tiers of informality (Fields, 2005; Danquah et al., 2019). We define “low-educated self-employed” (LE-SE) workers as own account workers with up to incomplete secondary education; they are informal self-employed in our sample. The “high-educated self-employed” (HE-SE) are the own account workers who have complete secondary education or higher, they are less informal, though not as formal as formal employees.

Wage employees are those who work for someone else in exchange for a pay; we separate them into formal and informal sectors. Using the ILO definition, we consider a worker as formal if he or she reports contributing to the social security system, has health insurance provided through the job, has a written labor contract or is a union member.²⁴ Consequently, a wage employee is considered informal when she or he does not have any of the conditions. We assume that informal employees are to the left of LE-SE on the Figure 1 spectrum since they are at particularly risk to not receiving payment or being the first to lose the job if the firm faces economic hardship and being subject to work conditions decided by the employer and not regulated by the state.²⁵ We assume that formal sector employees are to the right of HE-SE since they are recipients of social benefits and face less employment risk than the self-employed.

Finally, employers are at the far right of the continuum of employment types. The GMD does not include information on the formality status of their businesses, so we use

²⁴ This definition of formal is generous and includes workers who may receive only some benefits but not others. Thus, the size of the informal sector is likely larger than we estimate.

²⁵ While Danquah et al (2019) divide informal wage employees into “upper” and “lower” tiers, we do not introduce this additional division due to sample size limitations.

Danquah et al (2019) and classify them as formal. Since they employ paid labor, we assume that they organize their enterprises more like a firm and less like an extension of their households.²⁶ We do not differentiate between high- and low-educated employers due to a small sample size for many of our variables for analysis.²⁷

Because we use proxies of the formality status of self-employed workers, we propose a lower and an upper bound of urban informal employment. The lower bound includes unpaid workers, informal wage employees and the low-educated self-employed. The upper bound also classifies the high-educated self-employed as informal. While we estimate the size of the informal sector using the upper and lower bound, we limit the rest of our analysis to the lower bound definition.

2.2 Defining Firm Types to Classify Informal Entrepreneurs

The IES is limited to informal firms that are not registered with authorities, so we only consider heterogeneity within the informal sector. Since the sample is only informal firms, we do not need to search for a variable to proxy informality. Instead, to draw a parallel with the worker analysis, we create three sub-groups of informal sector firms using employer status and education of the enterprise owner. We identify the self-employed as those enterprises that employ no paid workers other than the owner. They may employ unpaid workers, who are often family members.²⁸ To proxy the “upper” and “lower” tier of informal self-employed (Fields 2005), we further divide self-employed businesses owners according to their level of education, similar to the procedure used to define employment types using the GMD. Low-educated self-employed are firms where the owners have up to secondary level of education while high-educated self-employed are owners with vocational training or tertiary education.²⁹ Finally, the third group is employers, namely firms that employ at least one paid worker other than him or herself.³⁰ The owner of each firm type is assigned the corresponding employment type, namely low-educated self-employed, high-educated self-employed, and employer.

²⁶ The literature finds that very few self-employed establishments grow into microenterprises (Loening and Imru, 2009; Kinda and Loening, 2010; Grimm et al., 2012; Fajnzylber et al., 2006; Schoar, 2009; La Porta and Schleifer, 2014). Hence this category is mostly comprised of formal firms.

²⁷ We find that 41 percent of employers in our sample have at least completed secondary school and 40 percent are in professional occupations, double the rate of HE-SE who are in professional occupations.

²⁸ The IES of Zambia and Zimbabwe contain information on the number of workers who are family members. The correlation between the number of unpaid workers and the number of family members who work at the firms is 0.66 and 0.58 respectively.

²⁹ We do not use incomplete secondary level as the threshold to define low educational level, as with the workers classification, due to data limitation.

³⁰ Employers in the IES are more educated than those in the GMD. About 85 percent of IES employers have up to a secondary education, as compared to 41 percent of GMD employers.

Figure 2. Schematic of urban informality, by firm



Source: authors' rendering

The IES and GMD samples are not strictly comparable since they differ in several respects. First, the country samples are different. We use data from 26 countries in the worker analysis and from three countries in the firm analysis. Second, all the firms surveyed in the IES are not registered, while in the worker analysis we do not have (nor wish to use) a registration variable, so we use the job category and education level of the respondent to proxy informal firm owners. Third, the IES sample frame is restricted to a few cities that represent the main urban centers while GMD is based on the urban sample from nationally representative household surveys. The IES only has information on manufacturing and service firms, whereas the GMD covers urban agriculture and other sectors as well. Finally, the IES includes firm level variables and aims at analyzing firm level characteristics. In contrast, the GMD is derived from household surveys and therefore focuses on household-related variables, with very limited information on firms.

3. Methodology

We use three main methodologies to describe informal sector workers and enterprises in urban SSA. First, we use simple, unconditional summary statistics to profile the distribution of worker characteristics, firm, and urban variables within employment or enterprise type.³¹ Second, we use Linear Probability Models to estimate conditional correlates to better understand which worker, firm-characteristics, and urban variables are disproportionately represented in each employment or firm type, while controlling for other observed characteristics. Finally, we estimate the conditional correlation between characteristics of the urban space and the spatial distribution of employment types.

3.1 Conditional Correlates of Workers and Informal Firms

For the worker analysis, we estimate a series of linear probability models where the dependent variables (Y_{ic}^g) are indicators of employment type g (g =unpaid worker, low-

³¹ For workers (firms), we use the sample weights, meaning that the sample is representative of all workers (firms) across the sample countries. In other words, each country is not given equal weight. Instead, each worker (firm) is given equal weight, regardless of its country of origin.

educated self-employment, informal wage employee, formal wage employee) for worker i in country c :³²

$$Y_{ic}^g = \alpha^g + \beta_1^g X_{ic} + \beta_2^g J_{ic} + \gamma^g I_c + \varepsilon_{ic}^g \quad (1)$$

The vector X_{ic} includes socio demographic variables: indicators of gender, age cohort, educational level, and marital status. This set of explanatory variables are available for the 26 countries in our sample. J_{ic} is a vector of job characteristics, namely indicators of sector of employment and occupation. The estimated value of the parameters β_1^g indicate how the characteristics included in X_{ic} relate to the probability of being an informal worker of type g with respect to being a worker of any other type, while the estimated value of the parameters β_2^g will provide the same information for job characteristics.

We proceed by presenting two models for each employment type category. Information on job characteristics is only available for 15 countries. Therefore, we estimate one model that includes all 26 countries but does not include the J_{ic} vector of job characteristics. The second model includes job characteristics as additional regressors but is limited to 15 countries. All models control for country fixed effects (I) and use population weights.³³

For the firm analysis, we estimate similar models consisting of a series of linear probability models where the dependent variables (Y_{fc}^g) are indicators of informal firm type g (g = LE-SE, HE-SE, and employers) for firm f in country c :

$$Y_{fc}^g = \alpha^g + \beta_1^g X_{fc} + \beta_2^g J_{fc} + \gamma^g I_c + \varepsilon_{fc}^g \quad (2)$$

The explanatory variables (X_{fc} and J_{fc}) include the two sets of variables mentioned above: firm owner's demographic characteristics (indicators of gender, age groups and educational levels) and firm characteristics (age, sector of activity, indicator of location and having a bank account). The estimated value of the parameters β_1^g and β_2^g will tell us how the characteristics included in X_{fc} and J_{fc} vectors relate to the probability of being an informal firm of type g with respect to being a firm of any other type.

We proceed by presenting two models for each informal firm type. The first model controls for firm owner's demographic variables only. The second model adds firm

³² Due to the sample size, we cannot include high-educated self-employed or employers in the set Y_{ic}^g .

³³ One of the country fixed effects is dropped from the model since the intercept term is included.

characteristics to the previous model. Both models control for country fixed effects and use firms' weights.

3.2 Correlating Urban Infrastructure and Services with Formality-Employment Type

We begin by estimating the correlation between worker employment types and access to urban infrastructure and services. We adjust Equation (1) so that it takes the following form:

$$Y_{ic}^g = \alpha^g + \theta^g U_{ic} + \beta_1^g X_{ic} + \beta_2^g J_{ic} + \gamma^g I_c + \varepsilon_{ic}^g \quad (3)$$

The dependent variables (Y_{ic}^g) are the same as in equation (1), namely indicators of employment type g (g =unpaid worker, low-educated self-employment, informal wage employee, formal wage employee) for worker i in country c . The new term U_{ic} includes whether the household has access to electricity, to a computer and to improved water, as well as a poverty proxy based on whether a household is in the bottom 40 percent of the national per capita consumption distribution.³⁴ The estimated value of the parameters θ^g will tell us how the urban infrastructure and services variables relate to the probability of being an informal worker of type g with respect to being a worker of any other type. We first estimate Equation (3) without the X_{ic} and J_{ic} terms to assess the correlation between urban variables and type of employment. We then add the socio-demographic and job characteristics (the X_{ic} and J_{ic} terms).

Similarly, to understand how urban infrastructure and services variables correlate with firm type, we re-estimate Equation (2) but add the term U_{fc} . The dependent variables (Y_{fc}^g) are the same as in Equation (2), being indicators of informal firm type g (g = LE-SE, HE-SE, and employers) for firm f in country c . The vector U_{fc} includes indicators of whether a firm needs and has access to water and electricity, and computer access.³⁵

$$Y_{fc}^g = \alpha^g + \theta^g U_{fc} + \beta_1^g X_{fc} + \beta_2^g J_{fc} + \gamma^g I_c + \varepsilon_{fc}^g \quad (4)$$

The urban explanatory variables are available for firms in two countries. We estimate Equation (4) without the X_{fc} and J_{fc} terms and then add the firm owner and firm characteristic as additional regressors. We include a proxy for productivity in the J_{fc} vector to control for

³⁴ The GMD includes data on the aggregate consumption at the household level, derived from household surveys. We use the distribution at the country level and designate all households in the lower 40 percent as "vulnerable."

³⁵ We use a smaller set of infrastructure and services variables in comparison to the descriptive analysis due high correlations between them.

potential endogeneity of access to urban infrastructure and profits. The productivity variable takes a value of 1 if the firm's profits per worker is in the top 10 percent of the sample's distribution of productivity.³⁶

Finally, we use an OLS model to estimate the correlation between employment type and the urban environment in one country. Taking advantage of available geospatial data and the 2017 Census for Mozambique, we explore how the urban environment – crowding, neighborhood economic development, and cell phone coverage – correlates with employment types. To do so, we create a dataset where each observation is an urban *bairro* (neighborhood) in Mozambique, resulting in a sample with 212 *bairros* in five cities. Using data from the 2017 Census, we create variables for the distribution of employment type within a *bairro* and average ownership of a vector of durable goods in each *bairro*. We use publicly available geospatial data to generate a vector of six urban variables: two measure neighborhood economic development (night light luminosity and electrification), three measures of crowding (population density, building density, and urban density), and one measure of cell phone coverage (derived from cell tower density). The methodology to generate each urban variable is provided in Annex 2. The regression specification is as follows:

$$Y_{bc}^g = \alpha^g + \theta^g U_{bc} + \beta^g X_{bc} + \gamma^g I_c + \varepsilon_{fc}^g \quad (5)$$

The dependent variable is the share of total employment that is in employment type g' in *bairro* b in city c . To remain consistent with our definitions used to generate the employment type variable with the GMD and IES data, we use Census data on job category and employment to define g' (g' = low-educated self-employed, high-educated self-employed, unpaid family worker, employer, low-educated employee, and high-educated employee).³⁷ The spatial distribution of each employment type across the five cities is provide in Annex 2.

We redefine U_{bc} and X_{bc} with *bairro*-specific variables derived from publicly available geospatial data and Census data. The vector U_{bc} has six elements: population density, night light density (luminosity), average electrification, cell phone tower count, proportion urban, and building density. The spatial distribution of each urban variable across the five cities is provided in Annex 2. The vector X_{bc} is a vector of durable goods that are used to proxy *bairro* poverty, namely the share of the *bairro*'s population that owns each of the following: radio, landline phone, laptop, iron, electric stove, refrigerator, and car. The term I_c is now the city fixed effect. We use robust standard errors.

4. The Structure of Urban African Labor Markets

³⁶ The 10 percent cutoff follows Grimm et al. (2012).

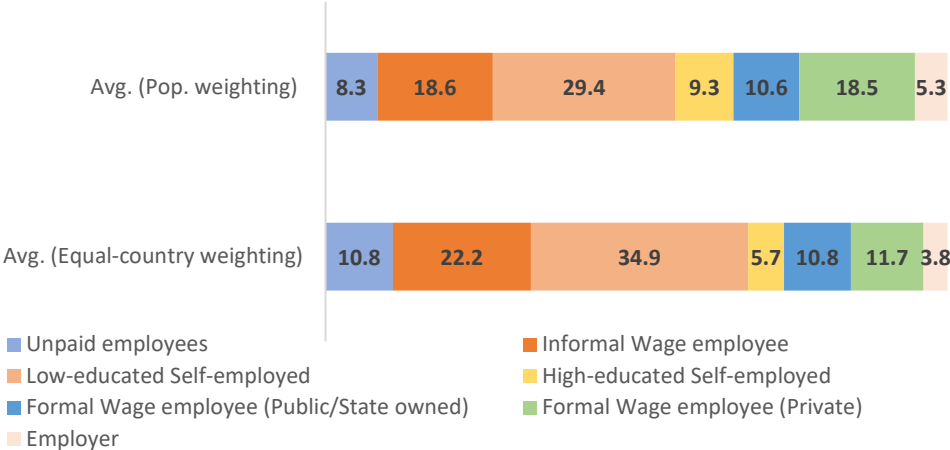
³⁷ The Mozambique 2017 Census does not provide information on social insurance or firm registries. Instead, drawing from Fields (2015), we use education level as a proxy for formal and informal.

We start by examining the prevalence of different types of employment in urban SSA and their dynamics. In particular, we quantify the share of jobs in each type of employment in the region and at the country level. We then briefly explore how the share of jobs correlates with GDP per capita, as well as how the change in the share of jobs correlates with changes in GDP per capita. We conclude that urban informality may decline in the long-run but it is a permanent part of the urban landscape in the short-run.

The informal sector comprises 56.3 percent (lower bound) to 64.6 percent (upper bound) of the urban informal labor market in Africa. Figure 3 presents the average distribution of urban workers by employment type for the 26 African countries in our GMD sample. We present the average using the weights available in the household surveys and an average giving an equal weight to each country. The two methodologies give broadly similar results.

The most prevalent employment type in urban Africa is low-educated self-employed workers, who we define as being in the informal sector, followed by informal wage employees. These two employment types comprise nearly half of urban employment and most of urban informal employment. An estimated 29.4 percent of urban workers in the 26-country sample can be classified as LE-SE. Another 18.6 percent can be classified as informal wage employees. These shares are smaller than the national averages estimated in ILO (2018), reflecting the highly informal nature of rural jobs included in the national estimates.³⁸ The share of urban LE-SE is three times larger than the share of urban high-educated self-employed. Only 18.5 percent of the African urban workforce are formal sector employees in the private sector; the same share as who are informal sector employees.

Figure 3. Distribution of urban employment by employment types, 26 African countries in the SSA region



Source: Own elaboration based on GMD data.

³⁸ ILO (2018) finds that 28 percent of informal workers in SSA are informal employees, 52 percent are s-e, 18 percent are contributing family workers. Among formal SSA workers, 75 percent are employees, 21 percent are self-employed and 3 percent are employers.

The distribution of workers across jobs differs by sample weighting strategy due to urban population size. Nigeria and South Africa account for 44 percent of the sample and both have amongst the lowest informality rates in the sample (Figure 4), thus skewing the regional average. If we drop Nigeria and South Africa from the sample, the regional informality rate would be 70 – 76 percent, which more closely aligns with most of the countries in Figure 4.³⁹

Except for a few outliers, the country-level distributions follow similar patterns. Informality, dominates in all but three countries,⁴⁰ ranging from 20 percent (South Africa) to 92 percent (Benin).⁴¹ The biggest share of informal sector employment in all but two countries is in LE-SE, ranging from 25-57 percent of all urban jobs.⁴² The share of unpaid workers is usually small, i.e., less than 15 percent of total urban employment, but it also has some cross-country variation ranging from less than 0.5 percent in Eswatini and South Africa to more than 20 percent in Benin, Chad, and Niger. Private sector employment hovers around 10 percent of all jobs in all but five countries in the sample.

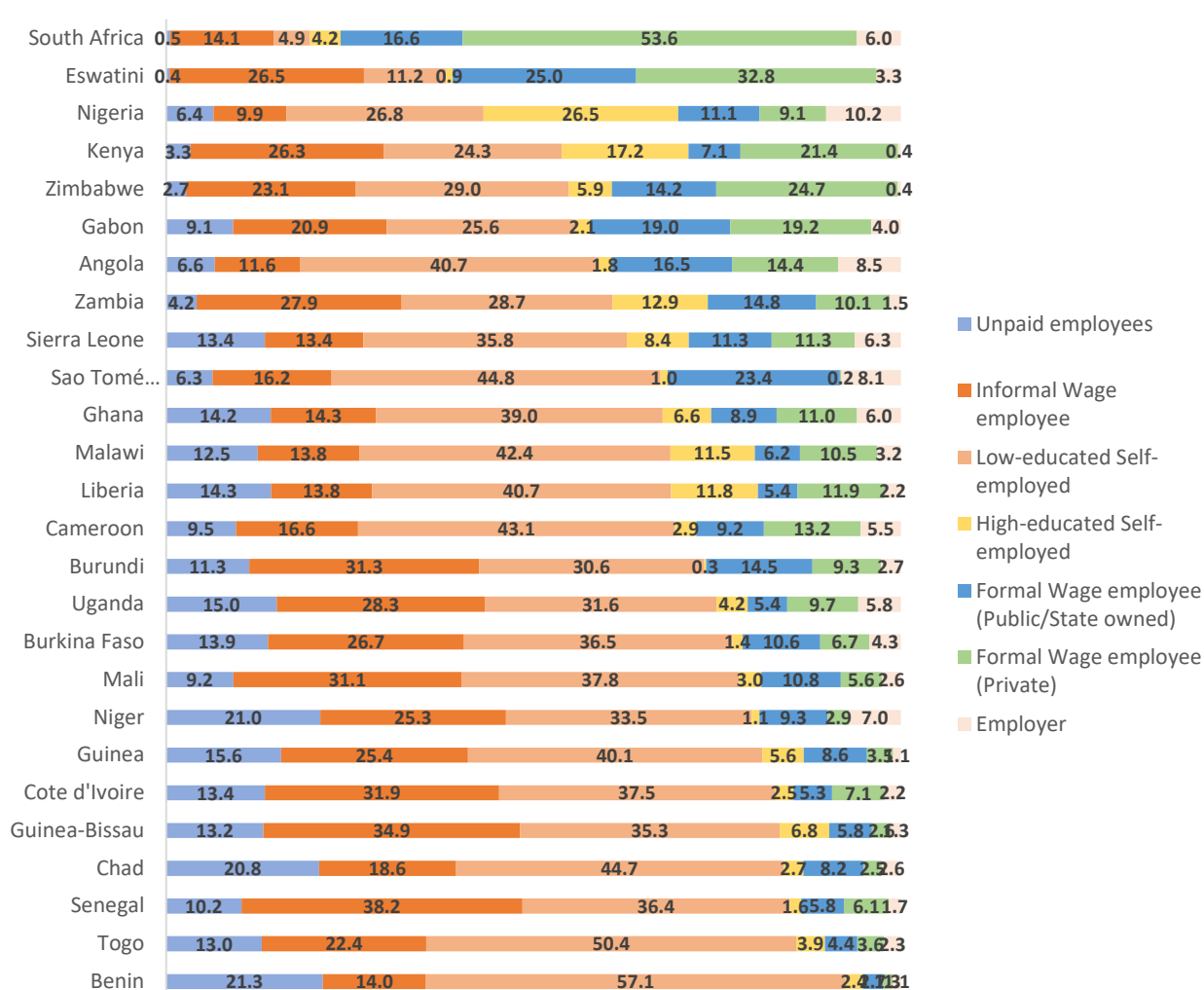
³⁹ Urban informality rates in nine countries that are not in our sample (Botswana, Cabo Verde, Comoros, the Democratic Republic of Congo, The Gambia, Madagascar, Namibia, Rwanda, and Tanzania), reflect the 70-76 percent informality estimates. Of these additional nearly 19 million SSA workers, only 5 percent work in countries with urban informality rates below 70 percent (ILO 2018).

⁴⁰ The exceptions are South Africa Eswatini, and Nigeria, when using the lower-bound definition.

⁴¹ International comparisons suggest that informality in South Africa is much lower than predicted in a country with its level of development (Cunningham, Ngarachu, and Kuddo 2022).

⁴² The only two exceptions are South Africa and Eswatini, where the share of informal wage employees is significantly larger.

Figure 4. Distribution of urban employment by employment types and country, from least (top) to most (bottom) informal



Source: Own elaboration based on GMD data.

Notes: Countries ordered by informality level, from lowest to highest. Rows do not sum to 100% due to rounding errors.

There is a strong negative correlation between informality and GDP per capita.⁴³ Whether measured by the share of the labor force that is informal, the share that is LE-SE or the share that is informal wage employees, countries with a higher share of informal sector employment tend to have lower levels of income per capita (Figure 5). With the exception of Gabon, that has an exceptionally high level of informality given its GDP per capita, the pattern is consistent across the region.⁴⁴

⁴³ This is consistent with a similar exercise in ILO (2018).

⁴⁴ The estimated correlation is statistically significant, with an $R^2 = 0.53$.

Table 1. Urban employment by Employment Types Over time

		Unpaid workers	Informal wage employees	LE-SE	HE-SE	Formal wage employees	Total informality
Cameroon	2007	19.5	15.9	40.4	3.6	20.6	75.8
	2014	10.0	17.6	45.6	3.1	23.7	73.3
	Annualized change	-9.1%	1.5%	1.8%	-2.2%	2.0%	-0.5%
Ghana	2012	18.4	12.7	43.8	1.8	23.4	74.8
	2016	15.1	15.2	41.5	7.0	21.1	71.8
	Annualized change	-4.7%	4.8%	-1.3%	40.4%	-2.6%	-1.0%
Liberia	2014	22.4	17.5	37.9	10.6	11.6	77.8
	2016	14.6	14.1	41.5	12.1	17.8	70.1
	Annualized change	-19.3%	-10.4%	4.7%	6.6%	23.6%	-5.0%
Uganda	2013	7.8	27.9	34.7	6.7	22.9	70.4
	2016	16.0	30.0	33.5	4.5	16.0	79.5
	Annualized change	26.9%	2.5%	-1.2%	-12.7%	-11.1%	4.1%
Zambia	2010	4.5	27.0	27.6	13.6	27.4	59.1
	2015	4.2	28.3	29.2	13.1	25.2	61.7
	Annualized change	-1.5%	1.0%	1.1%	-0.6%	-1.6%	0.9%

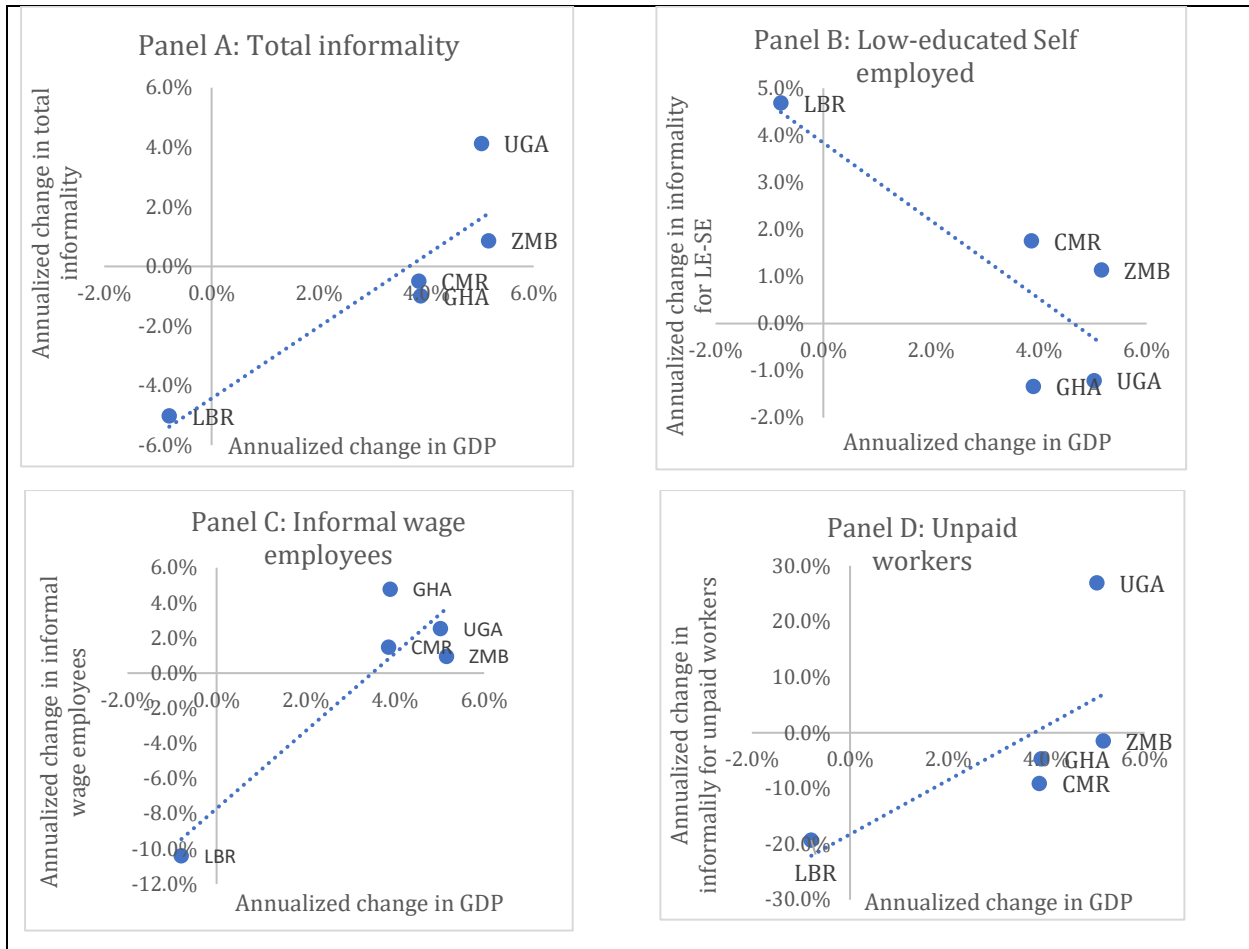
Source: Own elaboration based on GMD data.

The mixed pattern of informality fluctuations across countries is not related to differential rates of economic growth in the short-run. If we correlate the annualized rates of informality increase with the annualized rate of GDP growth over the same period, we find a positive association (Figure 6, Panel A). This is in direct contrast to Kiaga and Leung (2020) and Mehrotra (2020), who found that formality increased in periods of positive economic growth in Africa and Asia, respectively. But our results are consistent with Merotto, Weber and Aterido (2018) who found that informal employment persists even in relatively richer countries, and even when labor is no longer mostly in agriculture or self-employment, and with Fiess et al (2010) and Bosch and Maloney (2008) for Latin America.⁴⁷ Disaggregating by informal job type, we find that informal wage employment is positively correlated with GDP growth (Figure 6, Panel C) and LE-SE is negatively correlated with GDP growth (Figure 6, Panel B), similar to findings by Merotto et al (2018). Due to our small sample size, our results are highly unstable.

poor countries and 55 percent in middle-income countries. The share of informal wage employment remained around 20 percent, while the shares of other types of employment—mostly self-employment, either formal or informal—decreased. (Merotto et al. 2018)

⁴⁷ Fiess et al (2010) argues that economic growth in a closed economy leads to greater consumption of non-tradables, a high share of which are provided by informal enterprises.

Figure 6: Change in GDP and change in informality, annualized



Source: Own elaboration based on GMD data (y-axis) and WDI (x-axis).

In summary, the urban informal sector is large and persistent in Sub-Saharan Africa. The formal private sector employs approximately 18 percent of urban workers while the informal sector is the source of more than half of all jobs. It is dominated by low-educated self-employed workers, as well as informal sector employees. Short-term trends suggest that urban labor markets do not uniformly move toward formality in the short run, though there may be a shift toward more organized employment types as people move away from unpaid labor into other types of informal work. Economic growth does not seem to necessarily lead to a decline in informality in the short-run.

5. Profile of Urban Workers

Given the persistence of urban informal employment, it is useful to characterize urban workers within employment type. We use unconditional statistical comparisons as well as estimate conditional correlations (Tables 2 and 3) as described in the methodology section above. We consider their demographic characteristics including gender, age, education level, and poverty levels. We also consider their characteristics as workers, namely the

distribution by economic sector and occupations. Finally, we compare the characteristics of vulnerable (bottom 40 percent) and non-vulnerable (top 60 percent) workers within each employment type to better understand the profile of poorer workers. Due to sample size, we can only use the GMD to parametrically explore characteristics of unpaid, informal wage, and low-educated self-employed, with a much lighter touch on high-educated self-employed. We complement these findings with an analysis of the IES that allows us to understand more about the characteristics of low- and high-educated self-employed and informal employers.

5.1 Distribution of Demographic Characteristics by Employment Type

Women's informal sector participation is higher than their share of the workforce in urban SSA, and they dominate the non-wage informal sectors. Women are 47 percent of the entire workforce while they represent 51.5 percent of the informal sector (using the lower-bound definition) (Figure 7). Within the informal sector, women are disproportionately represented in two types of jobs: they are 58.1 percent of unpaid workers and 59.9 percent of LE-SE. By contrast, women are less than 40 percent of informal (35.2 percent) and formal (38.2 percent) wage employees. These aggregate trends are reflected across countries.⁴⁸

Married women especially cluster in unpaid and LE-SE. The conditional estimates confirm that women and dependent spouses are more likely to be unpaid or LE-SE than are men and household heads.⁴⁹ As most spouses in the sample are female, it is not surprising that female spouses follow this same pattern (Table 2).

The urban informal sector mainly employs prime-age workers. Those aged 25-55 years are 73.3 percent of the workforce and 66.9 percent of informal sector workers (Figure 7). LE-SE and informal wage employees – the two largest employment types – are mainly prime-aged adults. This age pattern appears in most countries. However, a closer look reveals more nuanced patterns.

Youth crowd into unpaid jobs. Youth (age 15-24) are 16.5 percent of the urban labor force, being particularly over-represented among unpaid workers (61.8 percent) (Figure 7). This result emerges from the conditional analysis, as well, where prime-aged and older workers are 20 percent less likely than youth to hold an unpaid wage job (Table 2). Youth are notably absent from informal self-employment.

Older workers have a propensity for informal self-employment. Older workers make up 10.2 percent of the workforce, but 16 percent of the low-educated self-employed (Figure

⁴⁸ For example, a smaller share of women than men hold informal wage employment in 25 of the 26 countries in the sample.

⁴⁹ The estimates for specifications that includes gender but not marital status, marital status but not gender, and an interactive gender and marital status variables are available from the authors. The lack of statistical significance on the female dummy in the unpaid employee regression is due to the spouse variable. When marital status is not included, the coefficient estimate for females is positive and statistically significant at the 1% level. The coefficient estimates for the female variable are higher in both the unpaid employee and self-employment regressions when marital status is not included.

7). When accounting for their smaller share of the workforce, workers who are older than age 55 are nearly twice as likely as prime-age workers to be low-educated self-employed. In contrast, older workers are 7 percentage points less likely than the rest of the workforce to be informal wage employees. These trends are reflected in the conditional estimates, as well (Table 2). The pattern of increasing informal entrepreneurship with age has been reported in diverse labor markets (Dhillon and Yousef 2009 for the Middle East, Peeters et al. 2009 for Sierra Leone, Cunningham et al 2008 for Latin America).

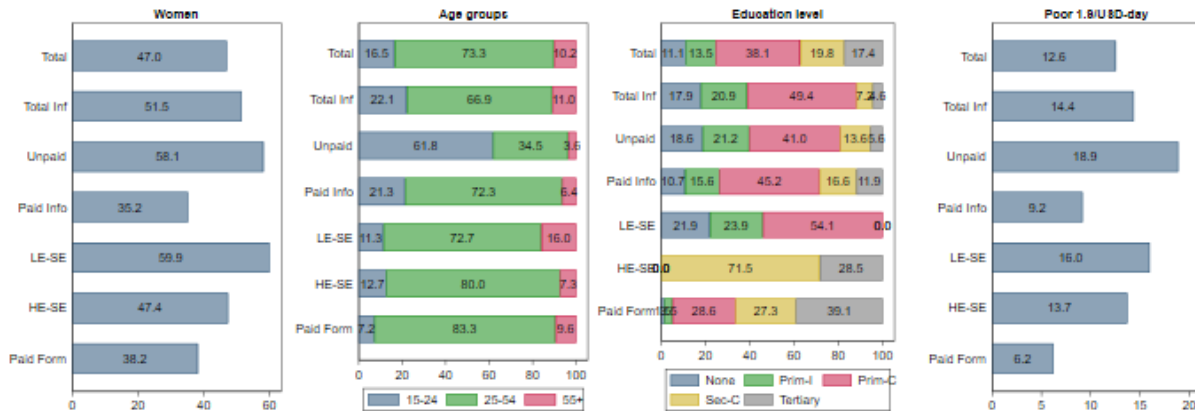
Though most urban informal workers have low educational levels, the relationship is not linear.⁵⁰ A larger share of unpaid and informal wage employees have an incomplete secondary education or less, in comparison to formal wage employees. That said, a measurable share of unpaid and informal employees has completed secondary education (19.2 percent and 28.5 percent, respectively) (Figure 7). The propensity to work in an unpaid job monotonically declines with education level, while the conditional estimates find an inverted U-shape for informal wage employees (Table 2).⁵¹ This may reflect significant heterogeneity in the informal wage employee sector, which may span from very unskilled to rather skilled jobs.

Most urban informal sector workers are not classified as “extreme poor.” Only 14.4 percent of those working in the urban informal sector consume less than USD 1.90 per day per person – the global extreme poverty line when using 2011 PPP exchange rates (Figure 7). This is more than double the share of formal sector workers classified as extreme poor (6.2 percent), though it is still a minority of informal sector workers. Poverty is highest among the unpaid (18.9 percent), followed by LE-SE (16 percent), HE-SE (13.7 percent), and finally, informal employees (9.2 percent).

⁵⁰ Since we define LE-SE and HE-SE by education level, we do not consider them in this analysis.

⁵¹ Those with incomplete secondary school are 6 percentage points more likely (than those with no education) to work as informal wage employees while those with less (incomplete primary) and more (complete secondary) education are only 3 percentage points more likely as those with no education to work as informal employees.

Figure 7. Sociodemographic characteristics by employment types

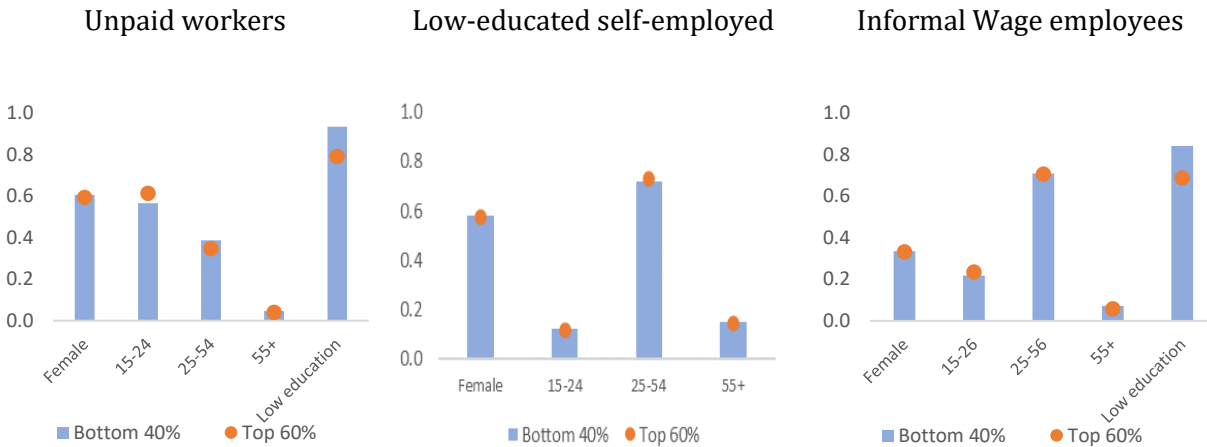


Source: Own elaboration based on GMD. Notes: Cross-country averages using population weights. Countries included: Angola, Benin, Burkina Faso, Burundi, Cameroon, Chad, Cote d’Ivoire, Gabon, Ghana, Guinea, Guinea-Bissau, Kenya, Liberia, Malawi, Mali, Niger, Nigeria, Senegal, Sierra Leone, Sao Tomé and Príncipe, South Africa, Togo, Eswatini, Uganda, Zambia and Zimbabwe.

More and less economically vulnerable workers have similar demographic profiles, regardless of type of informal employment. To better understand the profile of vulnerable informal workers, we separate the sample into workers having low- versus high-welfare levels within each informal type. Workers who are in the bottom 40 percent of the welfare distribution are defined as “low-welfare” and the rest are “high-welfare”.⁵² Within these groups, workers are profiled according to gender, age group, and low- or high- education categories, with the latter defined as secondary complete or higher. In general, worker characteristics are very similar when comparing low- with high-welfare workers within employment type (Figure 8). The only notable difference is that unpaid and informal wage workers in the bottom 40 percent, are more likely than their counterparts in the top 60 percent to not have completed secondary school.

⁵² The welfare distribution is based on data of aggregate consumption at the household level from household surveys.

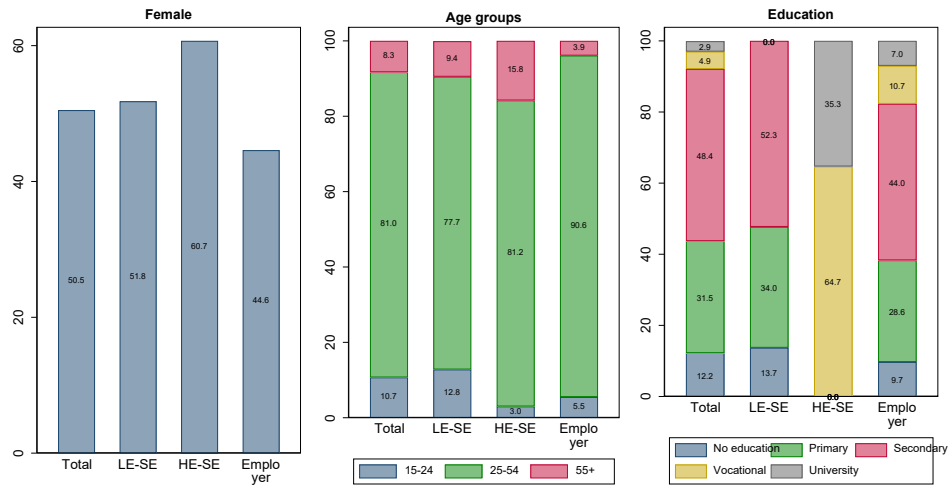
Figure 8. Share of Workers within Employment types and welfare level with the demographic characteristics



Source: Own elaboration based on GMD. Notes: Cross-country averages using population weights. The number of countries used for each variable is the same as the one indicated in previous figures. Up to secondary incomplete education is classified as "low education" and secondary complete or college education is classified as "high education." Each bar is the share of workers in that demographic category and job type who are in the bottom 40% of the welfare distribution. The dot is the share of workers in that demographic category and job type who are in the top 60% of the welfare distribution.

Using the IES to look more closely at informal firm owners, we find that owner characteristics are quite similar across the informal firm types analyzed. Except for women owned firms, who are slightly more likely than men to be HE-SE, both in the unconditional (Figure 9) and conditional (Table 3) estimates, other patterns seem similar. Among those who own informal firms, prime-age adults have lower chances of being LE-SE than young persons and higher chances of being an employer, consistent with the unconditional descriptive statistics (Figure 9) and with the results from the GMD. However, this age distinction is no longer important once controlling for any of the firm characteristics.

Figure 9. Characteristics of Urban Informal Firms Owners, %



Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe.

5.2 Distribution of Job-Related Characteristics by Employment Type

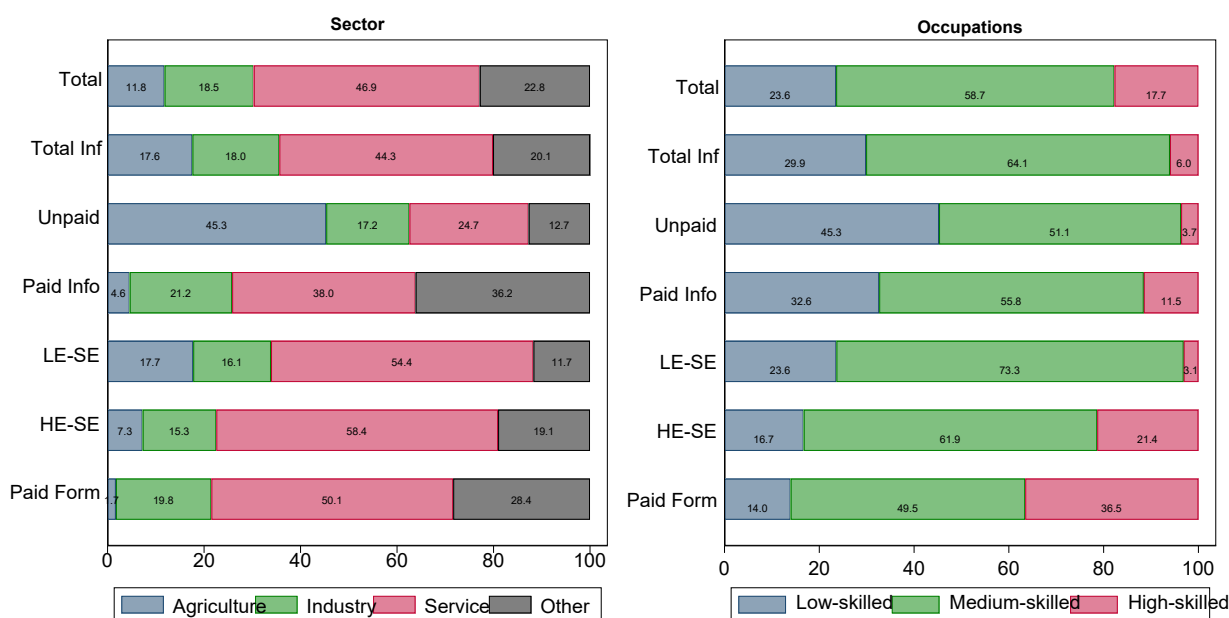
Urban informal and formal workers are employed in similar sectors, with some variation across informal employment types. The sector of employment can be roughly categorized into four groups: agriculture, industry, services, and other (not classified).⁵³ Figure 9 shows that 44.3 percent of urban informal jobs are in the services sector, only six percentage points less than the share of urban formal wage jobs in the service sector. Looking within the informal sector, LE-SE and HE-SE are more active in the services sector (54.4 and 58.4 percent, respectively) than are informal wage employees (38 percent), perhaps reflecting crowding into low value-added self-employment activities. That said, the conditional analysis shows that, while industry and service sectors are more likely than the agricultural sector to employ informal wage workers, unpaid workers are mainly employed in agriculture, even in this urban sample (45.3 percent).⁵⁴ This is not surprising since agriculture is often a subsistence activity to directly provide food for household consumption. For all three types of informal employment, workers in the lower welfare group have higher shares of employment in agriculture and lower shares in other activities, especially in services, than workers in the higher welfare group (Figure 11).

⁵³ “Other sector” includes Professional, Scientific and technical activities; Administrative and support service activities; Education; Health; Art, entertainment and recreation; Activities of households as employers; and Activities of extraterritorial organizations and bodies.

⁵⁴ When looking at the employment distribution by sector at the country level, we note a few deviations from the regional pattern. For instance, services is the most important sector for unpaid workers in Cameroon, Eswatini and Zimbabwe while industry dominates in Guinea, Senegal and Togo. Agriculture is the main sector of employment for LE-SE in São Tomé and Príncipe, being a small island state with only 70,000 people living in its urban capital, and with an economy still predominantly driven by agriculture (and tourism).

Most informal workers, and a fair share of formal wage employees, work in medium-skilled occupations (Figure 10). We define medium-skilled occupations as those that require job-specific skills, excluding highly skilled professionals.⁵⁵ About 64 percent of informal sector workers are in medium-skilled occupations, as compared to 49.5 percent of formal wage workers. High-skilled occupations generally increase with level of formality, where one-third of formal wage employees are employed in high-skilled occupations. Notably, both low- and high-skilled occupations are found across all employment types, both in the aggregate and at the country level. The conditional correlates find that medium-skilled occupations are particularly aligned with LE-SE, while low-skilled occupations are more aligned with being an informal employee (Table 2). As expected, more vulnerable workers have higher shares of employment in low-skilled occupations and lower shares in medium- or high-skilled occupations than workers with a high level of welfare (Figure 11).

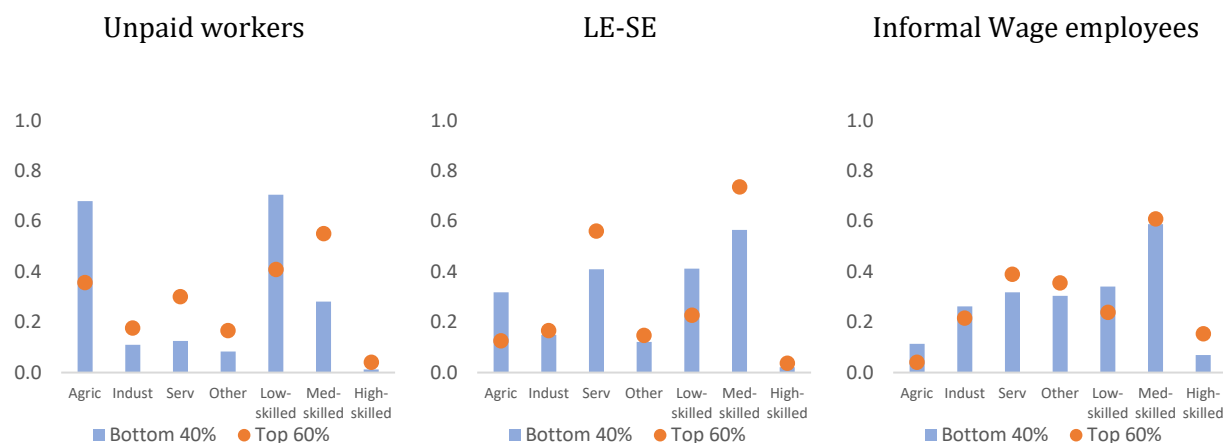
Figure 10. Jobs characteristics by employment types



Source: Own elaboration based on GMD. Notes: Cross-country averages using population weights. Countries included in the sectoral averages: Angola, Burundi, Benin, Burkina Faso, Chad, Cote d’Ivoire, Cameroon, Gabon, Ghana, Guinea, Guinea Bissau, Mali, Niger, Senegal, Sierra Leone, Sao Tomé and Príncipe, South Africa, Eswatini, Togo, Uganda, Zambia and Zimbabwe. Countries included in the occupation averages: Angola, Benin, Burkina Faso, Chad, Cote d’Ivoire, Cameroon, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Niger, Senegal, Sao Tomé and Príncipe, South Africa, Togo, Zambia and Zimbabwe.

⁵⁵ Low-skilled occupations include agricultural, forestry and fishery workers and elementary workers; medium-skilled occupations include clerical and support workers, service and sales workers, craft and related trades workers, and plant and machine operators and assemblers; high-skilled occupations comprise managers, professional and technicians and associate professionals.

Figure 11. Share of Workers within Employment types and welfare level with the job characteristics



Source: Own elaboration based on GMD. Notes: Cross-country averages using population weights. The number of countries used for each variable is the same as the one indicated in previous figure. Up to secondary incomplete education is classified as "low education" and secondary complete or college education is classified as "high education".

Table 2. Determinants of Urban Informal Types of Workers

VARIABLES	Unpaid workers		Low-educated Self-employed		Informal Wage employees	
	(1)	(2)	(1)	(2)	(1)	(2)
25-54 years old	-0.192 [0.008]***	-0.222 [0.006]***	0.110 [0.007]***	0.140 [0.006]***	-0.013 [0.007]*	-0.004 [0.007]
55 years or older	-0.192 [0.008]***	-0.247 [0.006]***	0.189 [0.010]***	0.212 [0.008]***	-0.090 [0.009]***	-0.065 [0.007]***
Education - Primary	-0.014 [0.004]***	-0.003 [0.005]	-0.026 [0.007]***	-0.008 [0.007]	0.031 [0.006]***	0.015 [0.006]**
Education - Primary complete, Secondary incomplete	-0.042 [0.004]***	-0.008 [0.004]*	-0.118 [0.006]***	-0.129 [0.006]***	0.061 [0.005]***	0.041 [0.006]***
Education - Secondary complete	-0.054 [0.006]***	-0.018 [0.005]***	-0.613 [0.009]***	-0.463 [0.006]***	0.027 [0.009]***	-0.002 [0.008]
Education - Tertiary	-0.064 [0.004]***	-0.018 [0.005]***	-0.600 [0.006]***	-0.485 [0.007]***	-0.033 [0.007]***	-0.028 [0.008]***
Female	0.004 [0.004]	0.023 [0.003]***	0.105 [0.005]***	0.131 [0.004]***	-0.057 [0.005]***	-0.090 [0.004]***
Spouse	0.065 [0.004]***	0.059 [0.004]***	0.017 [0.007]**	0.013 [0.006]**	-0.076 [0.006]***	-0.050 [0.005]***
Other relative or not	0.181 [0.006]***	0.193 [0.005]***	-0.146 [0.007]***	-0.180 [0.006]***	0.052 [0.007]***	0.049 [0.006]***
Industry		-0.227 [0.006]***		-0.149 [0.009]***		0.333 [0.009]***

Services	-0.258 [0.006]***		-0.077 [0.008]***		0.311 [0.008]***	
Other	-0.260 [0.006]***		-0.231 [0.008]***		0.446 [0.008]***	
Occupation-Medium skill	0.023 [0.003]***		0.148 [0.006]***		-0.220 [0.007]***	
Occupation-High skill	0.021 [0.004]***		0.040 [0.007]***		-0.262 [0.008]***	
Constant	0.235 [0.009]***	0.455 [0.008]***	0.304 [0.010]***	0.239 [0.010]***	0.159 [0.009]***	0.009 [0.009]
Number of countries	27	17	27	17	27	17
Observations	156,850	107,922	156,850	107,922	156,850	107,922
R-squared	0.218	0.318	0.384	0.384	0.075	0.137

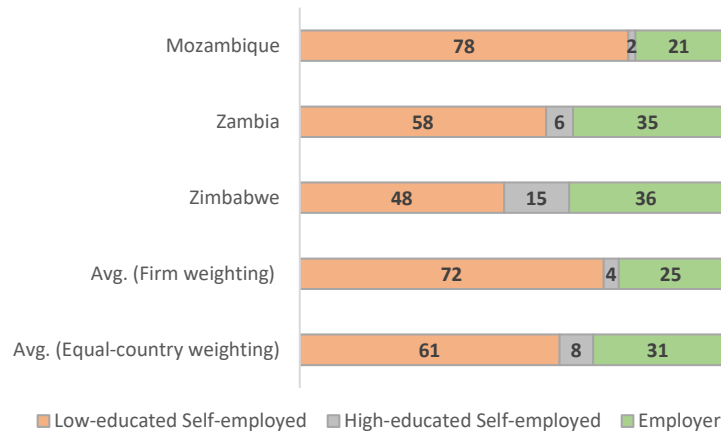
Source: Own estimations using GMD. Notes: Linear probability models estimated by Ordinary Least Squares. Omitted categories: 15-24 years old, no education, being household head, agriculture, low-skilled occupations. Robust standard errors in brackets. *** p<0.01, ** p<0.05, * p<0.

6. Profiling Urban Informal Firms

We now turn to the firm data to understand the nature of informal jobs from a firm perspective. We rely on the IES, keeping in mind that the harmonized data are from only three countries: Zambia, Zimbabwe, and Mozambique.

The IES sample differs from the GMD sample in a few important ways. The IES sample has fewer female firm owners in comparison to the share of women in our GMD sample, informal firm owners are more likely to be in the prime-age group than they were in the GMD sample, and informal firm owners are less educated than in our GMD sample. Though 72 percent of urban informal firms in the IES data are owned and operated by one person (self-employment) who has a low level of education, (Figure 12), the share of employers is much larger (25 percent) than in the GMD sample. This suggests that the findings based on worker analysis in the above section might underestimate the level of informality in urban Africa, since that analysis had assumed all employers were in the formal sector.

Figure 12. Distribution of Urban Informal Firms by Firm Type



Source: Own elaboration based on IES.

We consider two sets of variables to characterize urban informal firm types. The characteristics of the firms we measure are firm age, sector of activity, whether they operate outside household premises, the value of sales in the last month, the number of workers and productivity levels. We consider a set of variables related to financing: access to financial services and financial behavior including having a bank account, using banks or microfinance institutions for day-to-day operations, whether the accounts of the business are mixed with household expenses, and whether they use mobile money. We also separate “less vulnerable firms,” defined as those with productivity measures in the top 10 percent of the sample distribution, and “more vulnerable firms” to better understand the characteristics that are unique to higher productivity firms. We select the 10 percent cut-off following the informal firm heterogeneity analysis in the literature.⁵⁶ The productivity variable is calculated as the value of profits (sales minus total costs) per worker.

6.1 Distribution of Firm-Related Characteristics, by Informal Firm Type

A minority of informal firms manage to survive and grow. Less than 3 percent of firms are more than 20 years old, with 43.3 percent being start-ups (up to one year) and nearly 30 percent in existence for 2 to 5 years (Figure 13). This reflects high entry (and probably exit) rates of informal firms. When comparing different firm types, we find that HE-SE are younger than other informal firm types, with nearly half being 2 to 5 years old but a disproportionately small share with six or more years. The difference is statistically significant and robust even when controlling for owner and other firm characteristics (Table

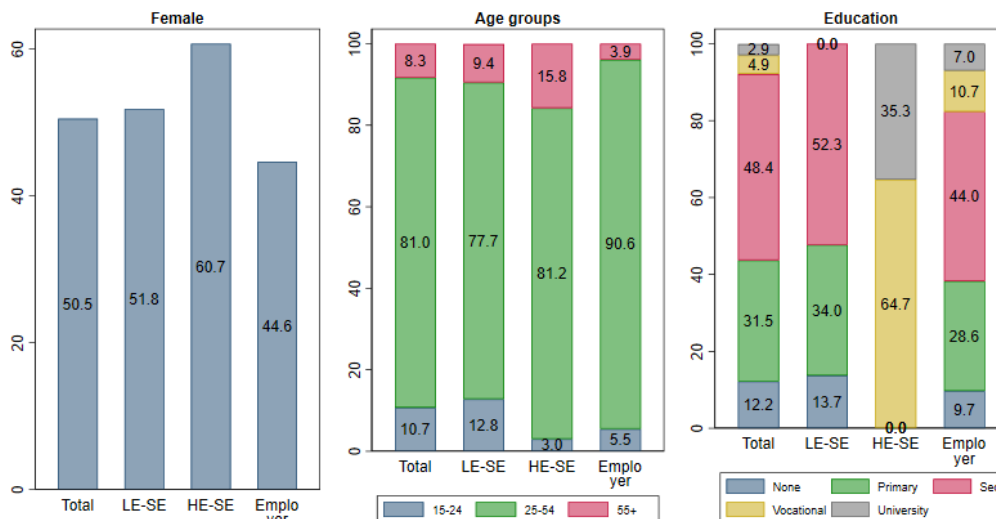
⁵⁶ Grimm et al. (2012), Adoho et al. (2018), Karlen et al (2022) and Aga, Campos, Conconi, Davies & Geginat (2021).

3). Though the three countries in our sample show the same pattern of most informal firms being young, intensity differs by country. These trends reflect the global literature.⁵⁷

Sales – a low value-added sector – is the dominant activity among informal firms, especially the self-employed. An estimated 74 percent of informal firms report that they sell food or other goods, with modest differences across countries (Figure 13) while 15.2 percent are engaged in manufacturing. Some heterogeneity emerges when comparing firm types. The share of firms that provide services is larger among employers (18.4 percent versus 10.8 percent for the average of all firm types) as is the share of manufacturing businesses (20.3 percent versus 15.2 percent for the average of all firm types), while the self-employed have a higher propensity for sales. The conditional estimates are only statistically significant at the 10 percent level (Table 3).

Most urban informal firms operate outside the household premises (59.1 percent on average), with small variation between firm types (Figure 13). The same pattern appears in the three countries included in the sample. The only exception is high-educated self-employed firms in Zambia, where only 32.5 percent reported operating outside household premises.

Figure 13. Characteristics of Urban Informal Firms



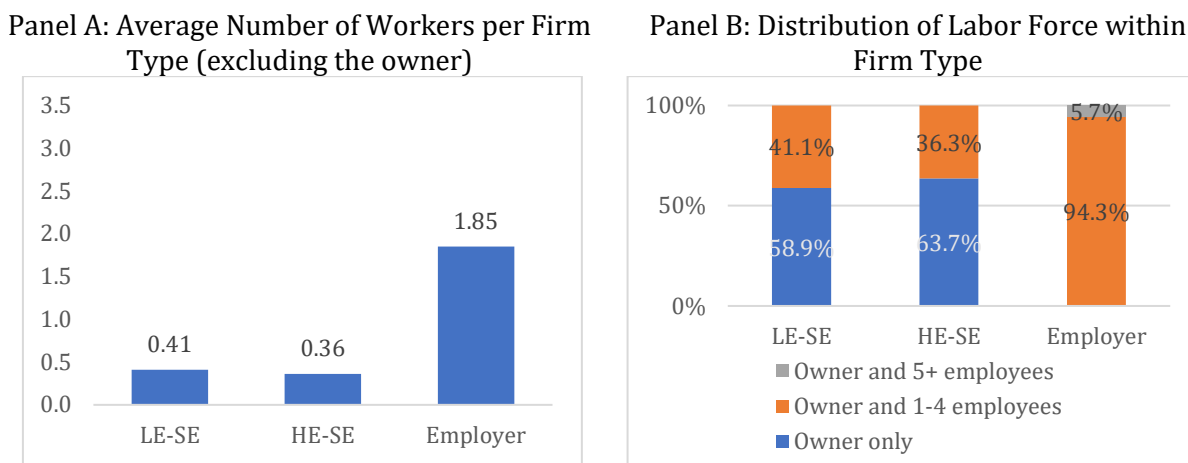
Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe.

Informal employers make a modest contribution toward employment creation, half of which is unpaid. Informal firms are small, primarily only employing their owners. An estimated 58.9 percent of LE-SE firms (63.7 percent of high-educated) only employ the

⁵⁷ See for example: Deléchat and Medina (2020) and La Porta and Schleifer (2014).

owner (Figure 14, Panel B). The rest employ an average of 0.41 (0.36) workers, all of whom are unpaid family member or apprentices (Figure 14, Panel A). By definition, employers hire paid workers, though they also employ a significant share of unpaid workers. The 94.3 percent of employers who are classified as micro firms (defined as firms that employ 1-4 employees), employ an average of 1.85 workers (in addition to the owners), where about 0.8 are unpaid workers. The 5.7 percent of employers who are classified as small firms (defined as firms having five or more employees), have an average of 3.7 paid workers and 2.5 unpaid workers.

Figure 14. Distribution of workers per firm, by informal firm type



Source: Own elaboration based on IES. Notes: Weighted average includes Mozambique, Zambia and Zimbabwe.

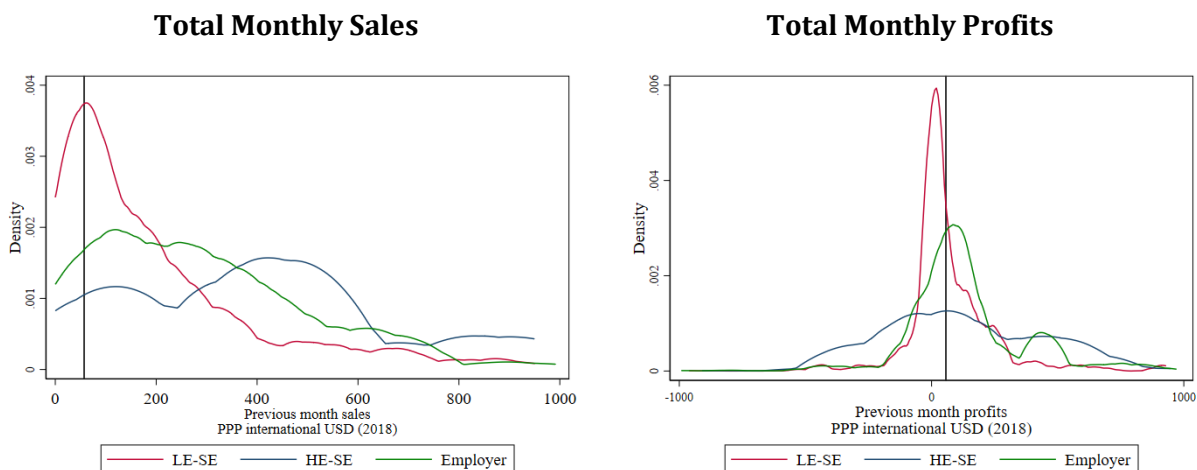
The difference in average sales and profits by firm types masks the high variance, and significant overlap, in the distributions. LE-SE firms have the lowest average monthly sales (250 USD PPP adjusted) as compared to HE-SE (600 USD) and employers (800 USD) (Figure 15, Panel A). However, an exploration of the left tail of the distribution shows that while half of LE-SE firms sold, on average, less than 200 USD in the previous month, a significant share of HE-SE (25 percent) and employers (30 percent) also had low sales. At the other extreme, 35 percent and 30 percent of HE-SE and employers, respectively, sold at least 600 USD in the last month on average while the share is 10 percent for LE-SE. When accounting for firm costs, average profits are particularly high for employers, earning an average of 302 USD monthly as compared to 31 USD for LE-SE and 42 USD for HE-SE. However, there is a bimodal distribution among employers, where profits for the larger sub-group is similar to average profits of the other firm types (Figure 15, Panel A).

On a per worker basis, HE-SE stand out as having higher and more concentrated profits, a proxy for productivity, than the other firm types. The average monthly value of sales per worker is 464 USD in HE-SE firms, as compared to 224 USD and 276 USD for LE-SE and employers, respectively (Figure 15, Panel B). LE-SE and employers also have a similar distribution of sales per worker. The similar pattern between these two firm types is because LE-SE combine a low value of sales with a workforce of small size. Employers, on the other

hand, sell more than LE-SE but they employ more people. HE-SE have a more disperse distribution of sales than the other firm types. The value of profits per worker follows a similar pattern as sales.

Figure 15. Sales and sales per worker in previous month. In PPP 2018 USD.

Panel A. Total Month Sales and Profits



Panel B. Per Worker Monthly Sales and Profits

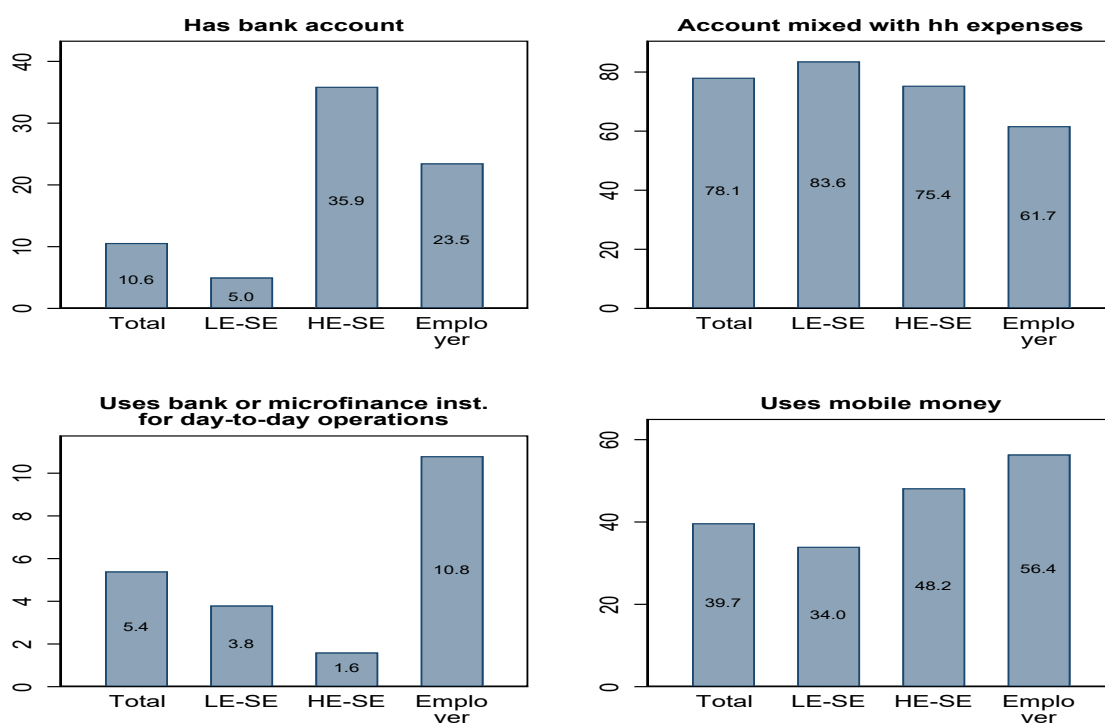


Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe. The figure shows the value of sales, sales per workers, profits, and profits per worker in last month expressed in PPP USD of 2018. Figures are restricted to the (-1000,1000) range to facilitate visualization. The vertical line indicates the monthly value of the 1.9 USD-a-day poverty line.

6.2 Distribution of Financial Tools and Management Practices, by Informal Firm Type

Informal firms are financially excluded, use few and poor financial management practices and tools, and while LE-SE have the worst practices, all three types of informal firms are lagging. On average, only 10.6 percent of informal firms have a bank account (Figure 16). Among LE-SE, the share is only 5 percent, while it reaches 35.9 percent for HE-SE. The conditional correlations confirm that having a bank account is negatively associated with being LE-SE (Table 3).⁵⁸ An estimated 5.4 percent of the sample regularly uses banks or microfinance institutions, with employers having the highest use rate at 10.8 percent. Instead, nearly 40 percent of informal firms use mobile money in their regular operations with higher rates of use by employers and HE-SE. In terms of business practices, most firms mix business accounts with household expenses –almost 80 percent on average. The percentage is largest among LE-SE (83.6 percent) in comparison to HE-SE (75.4 percent) and employers (61.7 percent).

Figure 16. Access to Financial Services and Financial Management



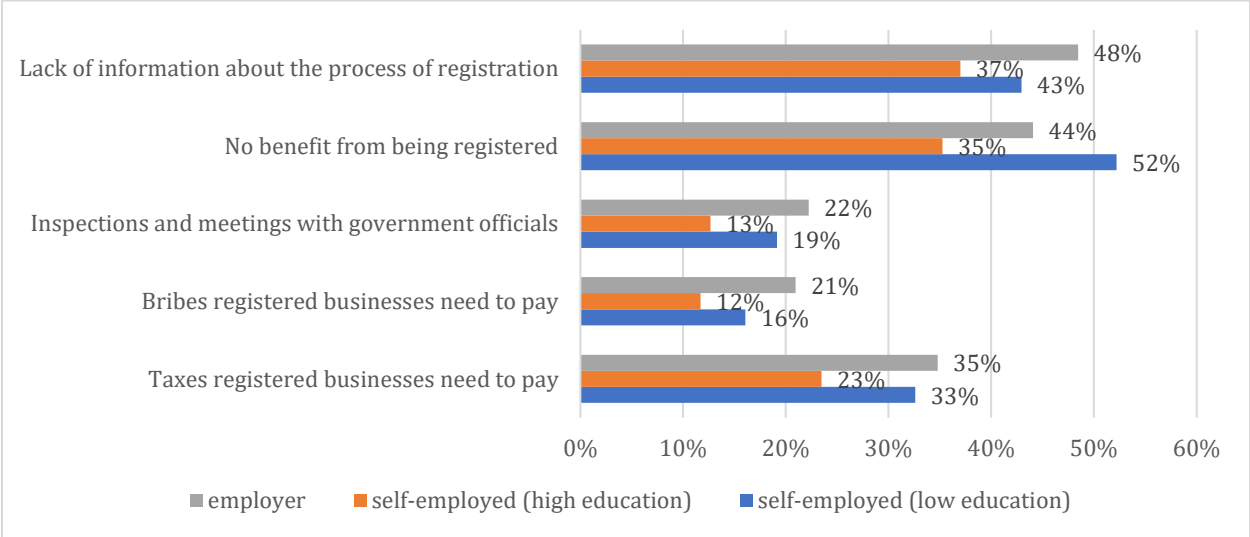
Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe except for mobile money which is not available for Zimbabwe.

All three informal firm types identify similar motivation for not being registered with authorities, though the top rationale differs by firm type. Employers mostly point to the lack of information about how to register as an impediment, while also citing low benefits from

⁵⁸ For HE-SE we do not find a statistically significant association. This is due to the education regressor, which is likely positive correlated with having a bank account. When the education dummies are dropped, we find a positive relationship between having a bank account and HE-SE.

registration (Figure 17). LE-SE either do not see the benefit from registering (52 percent) or do not have information about how to register (43 percent). HE-SE also identify a lack of information (37 percent) and seem to focus on the costs and benefits, citing that they do not see the benefit of registering (35 percent). A smaller share (12 to 22 percent) of respondents cite inspections and the need to pay bribes as a reason for not registering. The findings are aligned with the idea that formality can be considered a factor of production that business owners will “purchase” when the benefits outweigh the costs.⁵⁹

Figure 17: Reason for not Registering with Authorities, by Informal Firm Type



Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe.

⁵⁹ For example, Levenson and Maloney (1998) find that firms become formal when the perceived benefits that registration buys (access to finance, lower risk of fines, etc) exceeds the monetary and transaction costs of registering. De Mel et al. (2010) finds that it is the recurrent costs of being registered, rather than the start-up costs, that employers consider when making the choice to register.

Table 3. Determinants of Urban Informal Types of Firms

VARIABLES	Low educated self-employed			High-educated self-employed			Employer		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
Main owner, 25-54 years old	-0.097 [0.054]*		-0.061 [0.060]	-0.011 [0.007]		-0.012 [0.009]	0.107 [0.054]**		0.073 [0.059]
Main owner, 55 years or older	0.073 [0.078]		0.109 [0.087]	-0.006 [0.031]		-0.003 [0.031]	-0.067 [0.082]		-0.106 [0.091]
Main owner is female	0.013 [0.047]		-0.018 [0.051]			0.022 [0.011]**	0.019 [0.010]*		-0.001 [0.052]
Primary education	-0.005 [0.099]		0.022 [0.097]	-0.001 [0.013]		-0.001 [0.013]	0.006 [0.100]		-0.022 [0.098]
Secondary education	0.041 [0.096]		0.068 [0.092]	-0.002 [0.014]		0.001 [0.014]	-0.039 [0.097]		-0.069 [0.093]
Vocational training	-0.752 [0.086]***		-0.645 [0.103]***	0.435 [0.111]***		0.523 [0.103]***	0.317 [0.140]**		0.122 [0.145]
University training	-0.777 [0.086]***		-0.715 [0.093]***	0.600 [0.205]***		0.547 [0.194]***	0.177 [0.207]		0.167 [0.194]
Firm age ==2-5 years		-0.032 [0.051]	-0.040 [0.056]		0.023 [0.019]	0.033 [0.014]**		0.009 [0.050]	0.007 [0.058]
Firm age ==6-20 years		-0.033 [0.062]	-0.027 [0.068]		-0.011 [0.012]	0.008 [0.010]		0.044 [0.063]	0.019 [0.068]
Firm age ==More than 20 years		0.049 [0.066]	-0.006 [0.086]		-0.008 [0.023]	0.010 [0.009]		-0.041 [0.065]	-0.003 [0.089]
Retail trade (food)		0.159 [0.085]*	0.085 [0.084]		0.002 [0.016]	-0.007 [0.008]		-0.161 [0.084]*	-0.078 [0.084]
Selling other goods		0.153 [0.093]*	0.103 [0.097]		0.016 [0.018]	-0.002 [0.007]		-0.169 [0.092]*	-0.101 [0.096]
Provide services		0.010 [0.101]	-0.045 [0.114]		-0.011 [0.021]	-0.010 [0.012]		0.001 [0.103]	0.056 [0.114]
Outside household premises		0.010 [0.050]	0.038 [0.053]		0.020 [0.013]	0.001 [0.010]		-0.030 [0.050]	-0.039 [0.054]
Has a bank account		-0.395 [0.087]***	-0.191 [0.104]*		0.102 [0.051]**	-0.019 [0.031]		0.294 [0.088]***	0.210 [0.107]**
Constant	0.869 [0.103]***	0.721 [0.098]***	0.780 [0.133]***	-0.003 [0.014]	-0.012 [0.016]	-0.007 [0.017]	0.134 [0.104]	0.291 [0.097]***	0.227 [0.132]*
Number of countries	3	3	2	3	3	2	3	3	2
Observations	1,151	1,793	1,129	1,151	1,793	1,129	1,151	1,793	1,129
R-squared	0.218	0.144	0.232	0.482	0.072	0.515	0.069	0.099	0.104

Source: Own estimations using IES. Notes: Linear probability models estimated by Ordinary Least Squares. Omitted categories: owner 15-24 years old, owner with no education, firm 0-1 years old, manufacturing. Robust standard errors in brackets. * significant at 10%, ** 5%, *** 1%

6.3 Characteristics of Informal Firms by Productivity Level

In parallel with the above exercise that considered low- versus high-welfare workers, we profile low versus high productivity firms. Using Grimm et. al. (2012) as a guide,⁶⁰ we divide the sample into two groups. Firms in the 90th percentile or above of profits per worker – a proxy for productivity - are labelled “high productivity” while firms with profits per worker below the 90th percentile are considered “low productivity.”⁶¹

Men are disproportionately owners of the most productive informal firms. Figure 18 shows that for all three types of informal firms, owners of low-productivity firms are more likely to be women as compared to owners of high productivity firms. The story with regard to age is more nuanced. For the highly educated self-employed, virtually all of the most productive firms have owners aged 25-54, while only about 80 percent of the low-productivity firms do. However, approximately 20 percent of highly productive employers are 55 and up, while almost none of the low-productivity employers are.

The sector, firm size, and location of the firm is not uniformly correlated with firm productivity; instead, differences emerge by firm type. Among self-employed firms, a disproportionate share of non-food retail is among the 10 percent most productive firms, while employers who provide services are particularly productive. The smallest LE-SE firms are disproportionately among the 10 percent most productive LE-SE firms, as are small (2-5 employees) HE-SE firms (relative to all HE-SE firms) and larger (more than 6 employees) employers. Notably, HE-SE and employers with one employee or less are far over-represented among low-productivity firms. The results for employers could be the result of self-selection: the top 10% is represented by those very productive informal firms that have more workers (6+) and are in sub-sectors that are more typically high-growth (services) rather than simply trading. Finally, LE-SE who work outside the home are particularly productive.

⁶⁰ Grimm et al. (2012) defines firms that are in the top 90 percent of a composite score composed of profits per worker and value of capita stock, as “top performers” in a study of eight African countries. Adoho et al. (2018) follows the same methodology when studying small firms in the Democratic Republic of Congo. This 10 percent is exogenously set by the researchers. A more recent study (Karlen 2023) allows for the top performers to endogenously emerge from data from Togo through a cluster analysis, finding 18 percent are top performers. A similar exercise using data from Côte d’Ivoire finds that 11 percent are in the top tier (Karlen and Johanssen 2023). In Aga, Campos, Conconi, Davies & Geginat (2021) 8% of informal firms have similar characteristics and performance than formal firms (hence they have higher productivity levels).

⁶¹ As is common in firm data, a large share of firms reports negative profits and the distribution has long left and right tails. The data do not provide sufficient information to determine if these outliers are caused by measurement error or if they reflect the potential of all firm types to earn very high and very low profits. The distribution of profits per capita heavily overlaps between firm types.

Figure 18. Profile of Urban Informal Firms by Owner Characteristics and Productivity Levels

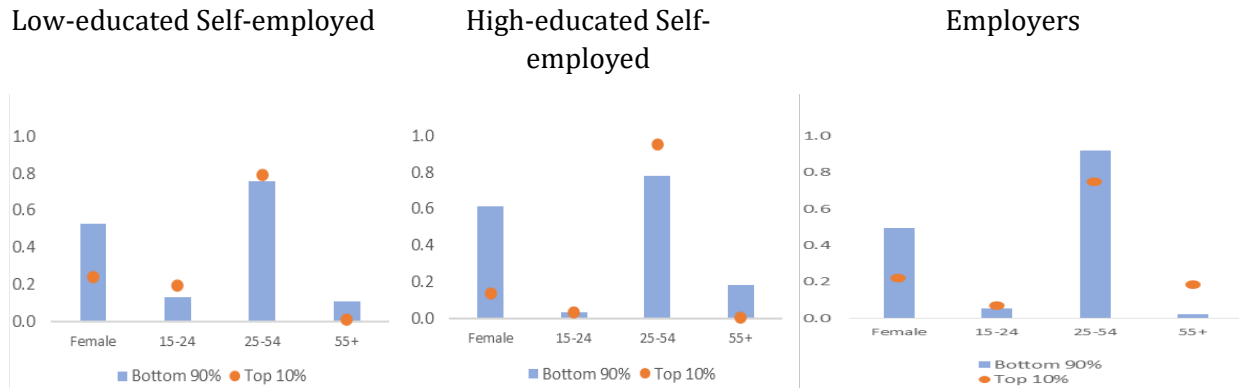
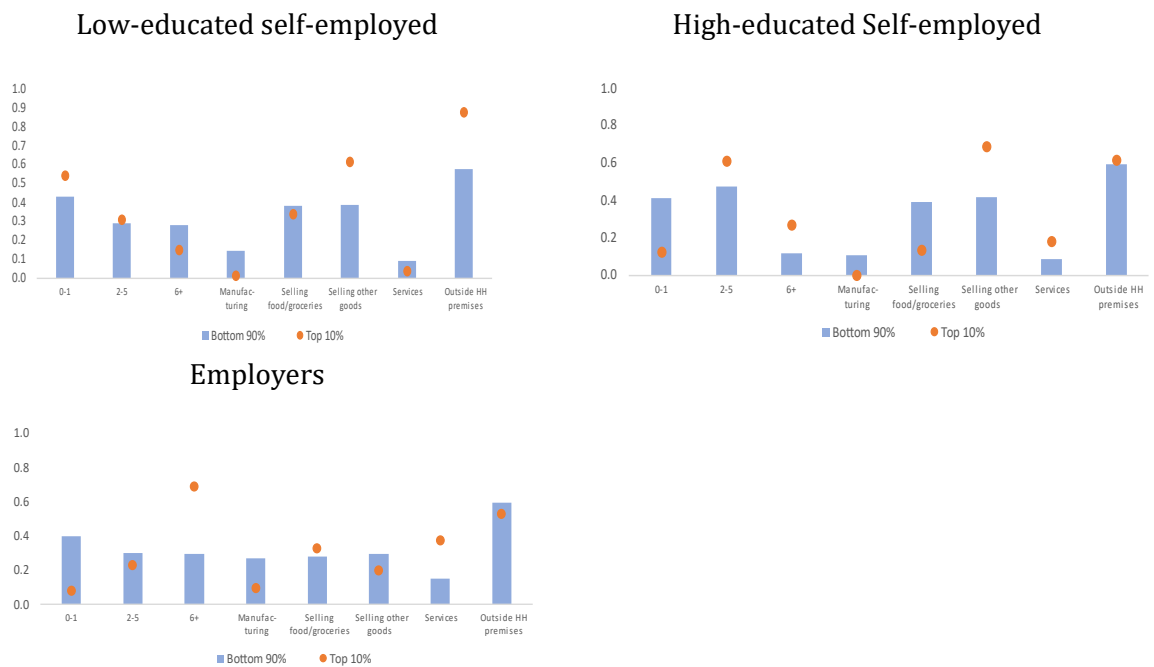


Figure 19. Profile of Urban Informal Firms by Firm Characteristics and Productivity Levels

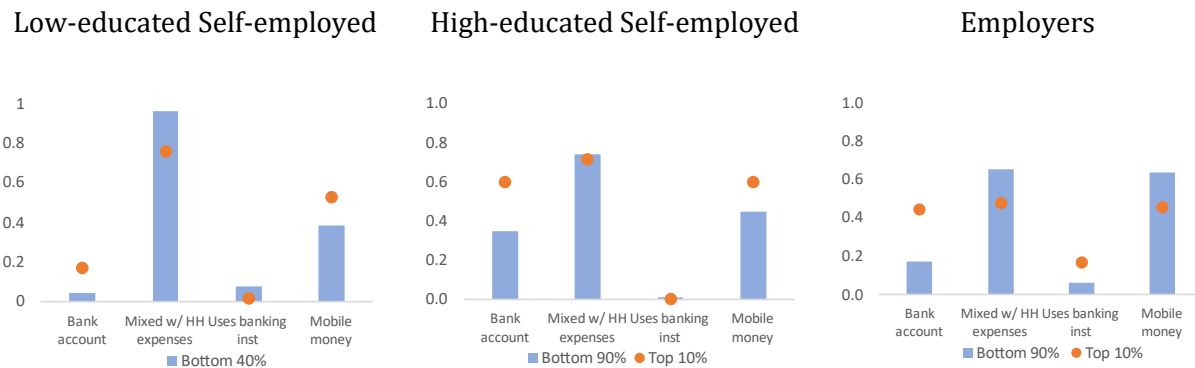


Source: Own elaboration based on IES. Notes: Cross-country averages using firm weights. The number of countries used for each variable is the same as the one indicated in previous figures. The first three bars are firm size, excluding the owner: 0-1, 2-5, and 6+. The next three bars are industry. The last bar indicates that the firm operates outside of the owner's home.

More productive firms are more likely to be banked and use management practices, though mobile money accounts may provide productivity-enhancements to informal self-employed firms. Figure 20 shows that among all three firm types, those that have bank accounts are also disproportionately high productive firms. However, both LE-SE and HE-SE who use mobile money are also among the most productive. Firms that mix business and

household expenses, a bad management practice, are disproportionately low productivity, cutting across all three firm types.

Figure 20: Profile of Urban Informal Firms by Financial Services and Management



Source: Own elaboration based on IES. Notes: Cross-country averages using firm weights. The number of countries used for each variable is the same as the one indicated in previous figures.

7. Urban

To assess how urban spaces, infrastructure, and services interact with the labor market, we consider which urban factors are correlated with different job types. We use the GMD to explore correlations between job type and the urban variables including access to electricity, water, improved sanitation, internet, and mobile phones. The results from the GMD should be interpreted with caution since the survey presents information about public services at the respondent’s home, which is not necessarily the location of the business.⁶² The IES, on the other hand, provides insights into which urban goods and services informal firms need and use at their workplace. Finally, we use big data to proxy the correlation between urban spaces and informal job types in urban Mozambique, including poverty, urban density and cell tower density. Again, the urban variables observed where workers live, which may not be where they are working.

7.1. Access to Urban Infrastructure and Services

A small share of owners of informal firms identifies a business need for public utilities, with significant differences by firm type. Less than 20 percent of firm owners sampled in the IES state that they need water or electricity services (Figure 21). However, low average electricity needs are driven by the large LE-SE sector; 40 percent of the other two firm types

⁶² The three country IES sample found that 60% of informal businesses operated outside the household premises. We do not have similar data for the 26-country GMD.

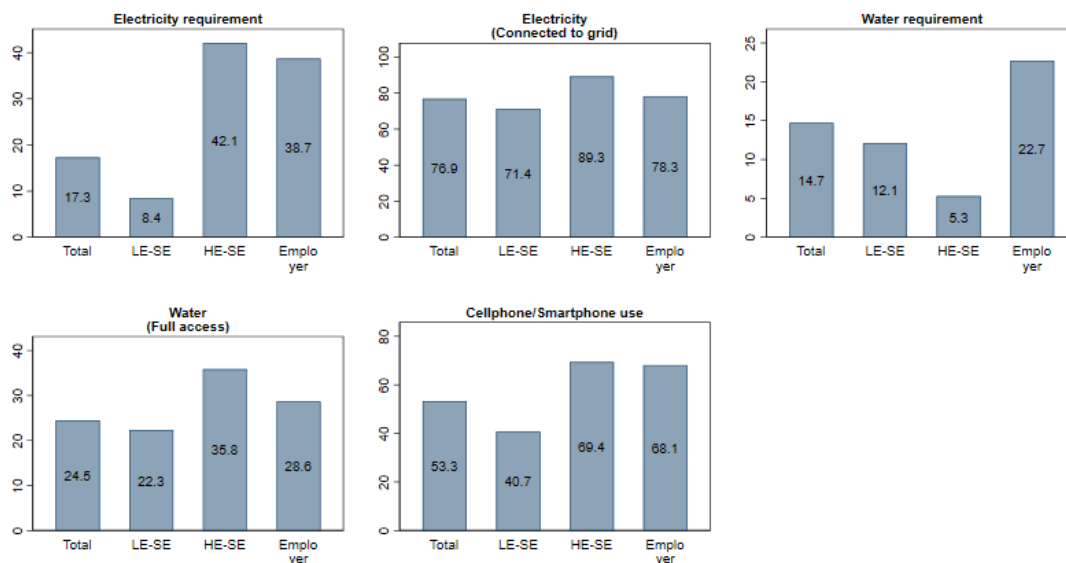
state that they need electricity to run their businesses. In contrast, the HE-SE have little need for water while employers particularly need this service.

Electricity needs are generally met across firm types while water needs are not. Again turning to the IES survey, among those informal firms needing electricity, 76.9 percent are connected to the grid, with HE-SE having greater connectivity than the others (Figure 21). The conditional estimates confirm that those who are connected to the grid are more likely to be employers and less likely to be LE-SE. Water access is a different story. Among those needing water, only 24.5 percent have full access, again, with HE-SE having more access than the others. The conditional estimates find that LE-SE have particular water deficiencies (Table 4). This in contrast to water and electricity access in the homes of workers. If we consider all workers (rather than firms) surveyed in the GMD, regardless of whether they need water or electricity, we find high access to water and electricity in their homes, about 80 percent access for each (Figure 22), with few differences between formal and informal workers. Access to improved sanitation is markedly lower for informal workers than formal sector workers.

Informal workers and firms' absolute and relative access to the internet is low, with significant variation across employment types. Only 40.6 percent of informal workers have access to the internet, with particularly low connectivity among the unpaid and LE-SE (Figure 22). The conditional correlates, controlling for poverty level, confirm that those workers with access to the internet are less likely to be unpaid or LE-SE as compared to all other job types (Table 5). The HE-SE have internet access on par with formal wage employees (58 percent) (Figure 22).

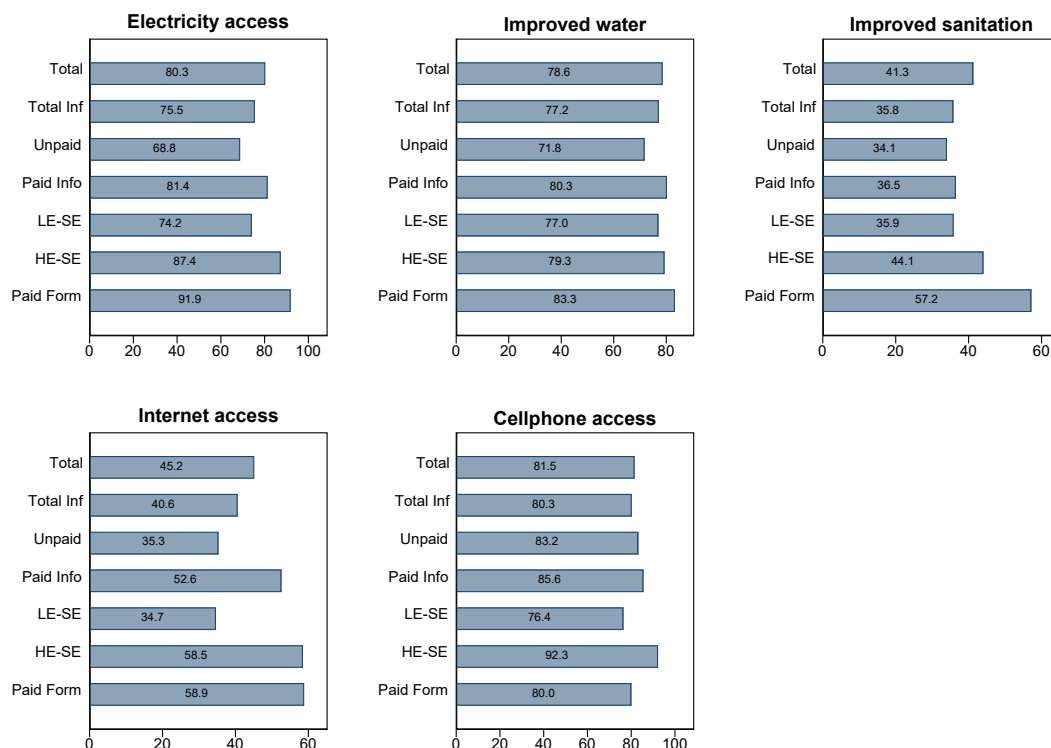
Cell phone access is much higher than internet access and relatively equal across employment types. It exceeds 80 percent for all employment types except LE-SE (76.4 percent), with the HE-SE having the most access (Figure 22). The IES reports that informal firms similarly report high usage (rather than simply access) of cell phones, with an average of 53.3 percent, with the LE-SE having lower usage rates than the other two categories of informal firms (Figure 21). The high rates of cellphone or smartphone usage could reflect ways of running the business, e.g., contact with customers and providers by phone or by social media, and also the possibility of using mobile money. In fact, the different rate of cellphone/smartphone use between firm types reflects the differential use of mobile money reported above.

Figure 21. Access and Use of Infrastructure, Services and Technology, by informal firm type



Source: Own elaboration based on IES. Notes: Weighted average including Mozambique, Zambia and Zimbabwe for electricity and water; and including Zambia and Zimbabwe for cellphone/smartphone.

Figure 22. Access to urban infrastructure and services, by employment type



Source: Own elaboration based on GMD. Notes: Cross-country averages using population weights. Countries included in electricity, water, sanitation, and cellphone averages: Angola, Burundi, Benin, Burkina Faso, Chad, Cote d'Ivoire, Cameroon, Gabon, Ghana, Guinea, Guinea Bissau, Kenya, Liberia, Mali, Malawi, Niger, Nigeria, Senegal, Sierra Leone, Sao Tomé and Príncipe, Eswatini, Togo, Uganda, Zambia and Zimbabwe. Internet average includes Cameroon, Benin, Burkina Faso, Chad, Cote d'Ivoire, Gabon, Ghana, Guinea, Guinea Bissau, Kenya, Mali, Senegal, Sierra Leone, Sao Tomé and Príncipe, Togo and Zambia.

Table 4. Urban Infrastructure and Services Correlates with Urban Informal Firms

VARIABLES	Low self-employed		High self-employed		Employer	
	(1)	(2)	(1)	(2)	(1)	(2)
Has grid electricity	-0.405 [0.084]***	-0.242 [0.112]**	0.050 [0.026]*	0.01 [0.016]	0.355 [0.085]***	0.232 [0.115]**
Has non-grid electricity	-0.234 [0.174]	-0.306 [0.203]	0.013 [0.024]	-0.005 [0.027]	0.221 [0.180]	0.311 [0.204]
Full water access	0.105 [0.126]	0.01 [0.154]	-0.064 [0.029]**	-0.084 [0.058]	-0.041 [0.128]	0.074 [0.168]
insufficient water access	0.026 [0.080]	0.195 [0.085]**	-0.040 [0.015]***	-0.024 [0.015]	0.014 [0.082]	-0.171 [0.084]**
Use computer	-0.371 [0.074]***	-0.371 [0.157]**	-0.049 [0.025]*	0.036 [0.036]	0.421 [0.077]***	0.335 [0.164]**
Demographic controls	no	yes	no	yes	no	yes
Firm controls	no	yes	no	yes	no	yes
Constant	0.823	0.634	0.017	-0.016	0.160	0.382

	[0.033]***	[0.173]***	[0.010]*	[0.030]	[0.032]***	[0.173]**
Observations	1120	669	1120	669	1120	669
R-squared	0.170	0.371	0.032	0.605	0.150	0.247

Source: Own estimations using IES Mozambique and Zambia. Notes: Linear probability models estimated by Ordinary Least Squares. Demographic controls include: owner age dummies, gender, and education dummies. Firm controls include: firm age dummies, sector, location of business, and use of bank account, and if the firm is among the top 10% in productivity. Robust standard errors in brackets. * significant at 10%, ** 5%, *** 1%.

Table 5: Urban Infrastructure and Services Correlates with Urban Informal Types of Workers

VARIABLES	Unpaid workers		Low-educated Self-employed		Informal Wage employees	
	(1)	(2)	(1)	(2)	(1)	(2)
Access to electricity	-0.056 [0.005]***	-0.001 [0.006]	-0.061 [0.008]***	-0.024 [0.007]***	0.010 [0.007]	0.013 [0.007]*
Has access to internet	-0.010 [0.004]***	-0.007 [0.005]	-0.112 [0.006]***	-0.024 [0.006]***	0.014 [0.006]**	0.020 [0.007]***
Use computer	-0.021 [0.004]***	-0.015 [0.006]***	-0.177 [0.007]***	-0.009 [0.007]	-0.042 [0.008]***	-0.041 [0.009]***
Access to water	-0.003 [0.004]	-0.005 [0.004]	-0.002 [0.006]	0.004 [0.006]	-0.010 [0.006]*	-0.004 [0.006]
Non-poor	-0.102 [0.005]***	-0.035 [0.005]***	-0.050 [0.007]***	-0.009 [0.007]	0.049 [0.006]***	0.026 [0.006]***
Demographic controls	no	yes	no	yes	no	yes
Jobs controls	no	yes	no	yes	no	yes
Constant	0.244 [0.045]***	0.477 [0.014]***	0.621 [0.118]***	0.398 [0.018]***	0.105 [0.019]***	0.005 [0.021]
Number of countries	17	14	17	14	17	14
Observations	87,008	53,958	87,008	53,958	83,510	53,958
R-squared	0.054	0.34	0.087	0.379	0.032	0.162

Source: Own estimations using GMD.

Notes: Linear probability models estimated by Ordinary Least Squares. Demographic controls are: age dummies, education dummies, gender, and relationship to household head. Job controls are: sector of employment, occupational skill level. Robust standard errors in brackets. significant at *10%, ** 5%, *** 1%.

7.2 The Urban Environment and Employment Types

To explore how the urban environment correlates with job types, we use big data and the Mozambique Census to explore how urban density, poverty, and cell service correlate with the spatial distribution of job types across five urban areas. A visual depiction of the spatial distribution of the variables is provided in Annex 2. The multivariate estimates are presented in Table 6.

Table 6: Multivariate estimates, by job type

Variables	LE-SE	HE-SE	Unpaid Worker	Employer	Informal Employee	Formal Employee
Log Building Density	0.029 (0.018)	0.001 (0.003)	0.003 (0.014)	-0.012*** (0.004)	0.010 (0.008)	0.019 (0.021)
Log Population	0.001 (0.011)	-0.0001 (0.002)	-0.001 (0.011)	-0.008 (0.007)	0.016** (0.008)	-0.041*** (0.014)
Log Luminosity	-0.146** (0.069)	-0.006 (0.015)	-0.131* (0.067)	0.024 (0.018)	0.043 (0.052)	0.235*** (0.118)
Avg Electrification Likelihood	-0.068 (0.058)	0.048** (0.014)	-0.179*** (0.038)	0.007 (0.020)	0.198*** (0.046)	0.067 (0.087)
Proportion of Urban Area	-0.083 (0.057)	0.019** (0.009)	-0.038 (0.056)	0.055*** (0.014)	-0.012 (0.043)	-0.067 (0.063)
Log Cell Tower Count per sqkm	-0.015 (0.009)	0.004 (0.003)	0.007 (0.007)	0.002 (0.003)	-0.006 (0.007)	0.012 (0.012)
R-squared	0.813	0.801	0.368	0.883	0.706	0.849

Notes: Linear Models of employment-type probability. Control variables for all regressions were poverty proxies and city fixed effects. Each model is estimated with the full set of urban variables. Each column presents the results of a single regression, where only the coefficient estimate and standard error for the urban variables in the regression are reported. Sample size for each estimate is 212.

Crowded neighborhoods are also those with less formal work arrangements. The results are robust to two measures of urban density. Population density is positively and significantly correlated with the density of informal employees but negatively correlated with formal employees. Similarly, building density is negatively correlated with being an employer. The findings from the multivariate regression are consistent with literature that finds that urban density is characterized by greater congestion and thus less conducive to operating a formal sector firm (Grover and Maloney 2022).

Less poor neighborhoods are home to formal and more structured types of work. The two measures of neighborhood economic development – luminosity and electrification – tell slightly different stories.⁶³ Luminosity is positively and significantly correlated with formal employees but negatively correlated with LE-SE and unpaid family workers. Electrification

⁶³ Henderson et al. (2012) find that night-time lights are a proxy for the density of poverty.

shows a slightly different pattern, where greater electrification is correlated with more structured forms of work, namely HE-SE and both formal and informal employees (though not statistically significant).

The number of cell towers in a location is not statistically significant for any job type. The estimated signs of the coefficient do not reveal a pattern since the correlation is negative for LE-SE yet positive for unpaid workers. This is perhaps not surprising since cell phone access may reach far beyond the neighborhood where the towers are located, barring physical barriers. Further, as shown above, cell phone coverage is high across all job types.

8. Discussion

Combining all the previous information, we conclude that the profile of urban informal workers for the average of the 26 countries under analysis could be summarized as follows:

Unpaid workers are loosely linked to the urban labor market. They are generally characterized as married women (58.1 percent) or youth (61.8 percent), having limited education (80 percent), and poorer than all other type of informal and formal urban workers (18.9 percent of them are extreme poor). Though living in urban areas, nearly half work in the agriculture sector and in low-skilled occupations (45 percent), especially workers with lower levels of welfare, while the other half work in middle-skilled occupations, especially workers with high-level of welfare. Their access to technology and sanitation is low (35.3 percent to internet and 34.1 percent to improved sanitation), lower than for other informal types, though 83.2 percent have mobile phones. They are more common in less economically developed neighborhoods.

Low-educated self-employed is the semi-skilled service sector dominated by women. They are mainly female (60 percent), even when controlling for their marital status, adults (72.7 percent), low-educated (100 percent, by definition), and poorer than formal workers (16 percent of them are extreme poor). They work in middle-skilled occupations (73.3 percent) in the services sector (54.4 percent), especially those defined as high-welfare levels, and have a similar sectoral distribution of employment as high-educated self-employed workers. Their access to technology and sanitation is below that of all other formal and informal employment types (34.7 percent have access to the internet and 35.9 percent to improved sanitation), which may limit their business potential. Three-quarters have access to a mobile phone. They live in denser urban neighborhoods, but also those that are a bit more economically developed.

Informal wage employment is the domain of vulnerable men. This sector is mostly men (64.8 percent), prime-age adults (72.3 percent), low-educated (71.5 percent) although their educational level is higher than for other informal types, and they are poorer than formal workers (9.2 percent of them are extreme poor). They work in middle-skilled occupations (55.8 percent) across sectors, though a higher share should be in these occupations given their other characteristics. Their access to technology and sanitation is

low, but better than for the other informal types, suggesting that although they do not enjoy the benefits associated with a formal wage job, their earnings are such that they can access more goods and services than those working in other informal employment types. More than 85 percent have access to a mobile phone.

Looking within the informal firm sectors, some differences arise:

Low-educated self-employed firms are simple retail operations with low profits, job creation, or modern business practices. Most owners are adults, despite a larger share of young owners in comparison to other firm types. Half of them are women with a larger female share among low-productivity firms. Most firms engage in retail trade, operating outside household premises, and 41 percent include unpaid employees. On average, LE-SE have the lowest profits, equal to the monthly poverty line, though some have profits in-line with employers. Access to financial services is very low, with less than one-twentieth of firms having a bank account or using banks or microfinance institutions for their regular operations. More than half report that they do not see the benefit of formalization.

High-educated self-employed enterprises are young and productive retail businesses. Most HE-SE owners are adults and more than half of the sample are women, especially among high-productivity firms. HE-SE are younger than other firms, with 90 percent operating for no more than five years. Their profits are double those of other business types, their profits per worker, a proxy for labor productivity, is higher and more concentrated than the other firm types. They are mostly in retail trade and operate outside the household premises, operating in less poor neighborhoods. Though using few modern business practices, they are more connected to the financial sector than other firm types. In addition to not seeing the benefit of formalization, HE-SE are not registered due to lack of information on how to register.

Employers are older firms that managed to grow and create jobs though firm productivity is modest. By definition, the key defining characteristic of employers is that they employ an average of 2.5 paid workers, as compared to no paid workers in other firm types. A majority of employers are low-educated adult men (55.5 percent). They are more active in services or manufacturing than are the self-employed. They have the lowest use of good business practices, such as not mixing business accounts with household expenses, but slighter better use of financial services, including mobile money, than the self-employed. They tend to live in less dense areas that are accessible to other urban neighborhoods. While employers have the highest total revenues and profits, their per capital revenues and profits are on par with LE-SE enterprises.

9. Conclusions

Informality strongly shapes Sub-Saharan Africa's labor markets. This paper uses data from 26 harmonized household and labor surveys (GMD) and three Informal Enterprises (IES) to sketch a profile of the urban informal sector in Sub-Saharan Africa.

The paper finds that urban informality is large and persistent in SSA. We estimate that 56 to 65 percent of the urban labor force can be classified as working in the informal sector, rising to 70-76 percent if we omit Nigeria and South Africa from the sample. Half of those are low-educated self-employed, which we equate with the informal self-employed. Another third are wage employees who do not receive social benefits as a part of their remuneration package. The balance are unpaid (mostly family) workers. These estimates are similar orders of magnitude as those in ILO (2018). Back of the envelop calculations for a subset of countries suggest that the share of urban informal employment can fluctuate significantly in the short-run, with expansion in some countries and contraction in others. The negative correlation between the share of urban informal workers and GDP growth rates suggests that economic growth may bring formalization in the long-run, with significant noise in the short run. Despite the strong correlation, the direction of causality is not necessarily obvious.

These regional averages mask significant heterogeneity across countries. While most countries follow the regional trends, orders of magnitude may differ significantly. For example, the share of the urban labor force that can be classified as informally employed ranges from 20 percent of South Africa's labor force to 92 percent in Benin. The biggest share of informal sector employment in most countries is in LE-SE, ranging from 24-45 percent. Thus, no single country profile can be a proxy for the region and regional proxies can only approximate country-specific profiles.

Urban informal employment can be characterized as prime-aged workers with incomplete secondary education, slightly more female, non-poor, and working in middle-skilled jobs in the service sector. The informal sector is more heavily female than the formal sector, largely due to female under-representation in the latter as women crowd into self-employment and unpaid labor. It is younger, particularly in unpaid jobs. Informal sector employees have less education, principally due to an absence of tertiary workers, though all other education levels are found in the informal sector. It is more comprised of middle-skilled workers while the formal sector is balanced between middle- and high-skilled. While informal sector workers are poorer than formal sector workers, the difference in average poverty rates is only 8 percentage points.

Sub-profiles emerge within the informal sector, principally driven by life-cycle and gender. Unpaid workers are those who are least connected to the labor force, being young, female spouses, low-skilled and over-represented in the low-skilled agriculture sector. Prime-aged women dominate informal self-employment while prime-aged men dominate informal wage employment. While both informal employment types are characterized by middle-skilled occupations in the service sector, informal wage employees are less poor and more engaged in high-skilled and "other sector" jobs. From the perspective of welfare, highly-vulnerable workers look the same as low-vulnerable workers in terms of demographics, but the latter are more in services and in medium- and high-skilled occupations.

The profiles of the three types of informal urban enterprises are rather similar, in terms of owner characteristics, sector, and motivation for informality, with most differences emerging in earnings, business practices, and access to financial tools. Informal enterprises are primarily led by prime-aged men and women (gender balanced) with a secondary education or less. Seventy percent of firms are younger than five years, reflecting challenges in firm survival, though firms with paid employees are a little older and report higher earnings, suggesting that firm growth is possible. Most firms are in the retail sector, which is also a disproportionately productive sector. These firms use few modern financial tools and practice poor business practices. Half do not see the value of formalizing, do not have information on how to formalize, or do not want to pay the taxes associated with formalization. More productive firms are led by prime-aged adults, are young, and are in the retail sector. HE-SE enterprises have the highest productivity while profits are highest among employers.

Urban variables were weakly correlated with job type. A minority of urban firms express a need for urban public services (electricity, water). Among those who need these services, HE-SE and employers have high access while LE-SE do not, even when controlling for firm productivity level. All express a need for, and access to cell phones, with sufficient cell tower across urban areas. More crowded neighborhoods are home to informal business owners while less poor neighborhoods house more structured work arrangements.

The findings of this study are suggestive due to data limitations. Although the study covers a wide range of African countries across the continent, the GMD only include data for 26 of 48 countries in Sub-Saharan Africa, equivalent to 61 percent of the region's population. The IES only includes three countries for which we have recent data. The worker and firm data are not strictly comparable due to differences in the number of countries included in the analysis, the national versus main cities representativeness, and the informality focus of the firm level data. The worker and firm data have some overlap – namely among LE-SE, HE-SE, and employers – though comparisons across surveys on this subset are limited due to the IES sample frame limitations by region and industrial sector.

The paper findings offer some general policy guidance. First, informal employment and informal firms dominate the urban space, suggesting an immediate need for greater policy attention to these segments of the urban labor force and enterprise landscape. Second, the rate of formalization does not seem to have increased systematically in the last decade, despite economic growth and poverty reduction, pointing to a need for policies to support the urban informal sector to be more productive and resilient rather than assuming that SSA will “grow out of” informality in the short term. Third, increasing the availability of data on informal firms should become a priority especially for African countries where most firms are in the informal sector; this would allow for better policy making. Fourth, policy will need to take into consideration both heterogeneity across the informal sector in terms of job types and the different characteristics of those working in each job type. Gender and age are particularly relevant variables for informal self-employed (former) and unpaid (latter)

workers. Fifth, the three types of informal enterprises are rather similar, but all would benefit from greater access to mobile money and improved financial and management tools; digital innovations can be a way to acquire these productivity-enhancing practices. Last, conditional on worker characteristics, the presence of public services such as electricity and water is not strongly associated with formality. That said, mobile phone coverage is important for all types of informal workers, highlighting the need to prioritize investments in the sector to support the urban informal workers. Since urban density and poverty have some effect on the prevalence of employment type, encouraging greater urban development in a less crowded manner could affect the shape of the labor market.

References

- Abreha, K.G., Kassa, W., Lartey, E., Mengistae, T., Owusu, S., & Zeufack, A. (2021). *Industrialization in Sub-Saharan Africa: Seizing Opportunities in Global Value Chains*. Washington, DC: World Bank.
- Adoho, F. & Doumbia, D. (2018). Informal sector heterogeneity and income inequality: Evidence from the Democratic Republic of Congo. *World Bank Policy Research Working Paper*, (8328).
- Aga, G., Campos, F., Conconi, A., Davies, E. & Gegina, C. (2021). "Informal Firms in Mozambique: Status and Potential." World Bank Policy Research Working Paper #9712.
- Amit, R. & Muller, E. (2013). ' "Push" And "Pull" Entrepreneurship.' *Journal of Small Business and Entrepreneurship*, 12(4).
- Bandiera, O., Elsayed, A.H., Heil, A., & Smurra, A. (2022). "Economic Development and the Organisation of Labour: Evidence from the Jobs of the World Project." *Journal of the European Economic Association*, 20(6).
- Basu, A.K., Chau, N.H., Fields, G.S. & Kanbur, R. (2018). 'Job Creation in a Multi-sector Labour Market Model for Developing Economies'. *Oxford Economic Papers*, 71(1): 119–44.
- Bosch, M. & Maloney, W. F. (2008). "Cyclical Movements in Unemployment and Informality in Developing Countries." World Bank Policy Research Working Paper No. 4648.
- Brewer, J. & Gibson, S. (2014). *Necessity Entrepreneurs: Microenterprise Education and Economic Development*. Edward Elger Publishers, Cheltenham UK.
- Calderon, C. (2022). *Boosting Productivity in Sub-Saharan Africa: Policies and Institutions to Promote Efficiency*. Washington, DC: World Bank.
- Coulibaly, S., W. Kassa; A. Zeufack; A. Mattoo. (2022). *Africa in the New Trade Environment: Market Access in Troubled Times*. Washington, DC: World Bank.
- Cunningham, W., Ngarachu, M., & Kuddo, A. (2022). *Benchmarking Labor Legislation and Enforcement in South Africa*. World Bank Group: Washington, DC.
- Cunningham, W., McGinnis, L., García Verdú, R., Tesliuc, C. & Verner, D. (2008). *Youth at Risk in Latin America and the Caribbean: Understanding the Causes, Realizing the Potential*. Directions in Development; Human Development. Washington, DC: World Bank
- Danquah, M., Schotte, S. & Sen, K. (2019). "Informal work in sub-Saharan Africa Dead end or steppingstone?" WIDER Working Paper 2019/107.
- Deléchat, C. & Medina, L. (2020) "What is the Informal Economy?" Finance and Development, 54-55. International Monetary Fund, Washington, DC.
<https://www.imf.org/en/Publications/fandd/issues/2020/12/what-is-the-informal-economy-basics>

- De Mel, S., McKenzie, D., & Woodruff, C. (2010). "Who Are the Microenterprise Owners? Evidence from Sri Lanka." in J. Lerner and A. Schoar (eds), *International Differences in Entrepreneurship*. Chicago, IL: University of Chicago Press.
- De Vreyer, P., & Roubaud, F. (eds) (2013). *Urban Labor Markets in sub-Saharan Africa*. Washington, DC: World Bank.
- Dhillon, N. & Youssef, T. (2009). *Generation in Waiting: The Unfulfilled Promise of Young People in the Middle East*. Brookings Institute: Washington, DC.
- Fajnzylber, P., Maloney, W., & Rojas, G.M. (2006). Microenterprise dynamics in developing countries: How similar are they to those in the industrialized world? Evidence from Mexico. *The World Bank Economic Review*, 20(3), 389-419.
- Fields, G.S. (2005). 'A Guide to Multisector Labor Market Models'. Social Protection Discussion Paper Series 0505. Washington, DC: World Bank.
- Fiess, N., Fugazza, M. & Maloney, W. F. (2010). "Informal self-employment and Macroeconomic Fluctuations," *Journal of Development Economics*, 91(2), 211-226.
- Filmer, D., & Fox, L. (2014). *Youth Employment in Sub-Saharan Africa*. World Bank Publications.
- Fox, L., & Sohnesen, T. P. (2012). Household enterprises in Sub-Saharan Africa: Why they matter for growth, jobs, and livelihoods. *World Bank Policy Research Working Paper #6184*. World Bank: Washington, DC.
- Gindling, T. H., & Newhouse, D. (2014). Self-employment in the developing world. *World Development*, 56, 313-331.
- Grimm, M., P. Knorringa, & J. Lay (2012). "Constrained Gazelles: High Potentials in West Africa's Informal Economy." *World Development*, 40(7): 1352-68.
- Grover, A. & Maloney, W.F. (2022). "Proximity without Productivity: Agglomeration Effects with Plant-level Output and Price Data." Policy Research Working Paper Series #9977. World Bank: Washington, DC.
- Henderson, J.V., Storeygard, A. & Weil, D.N. (2012). "Measuring Economic Growth from Outer Space." *American Economic Review*, 102 (2): 994-1028.
- ILO (2018). *Women and Men in the Informal Economy: A Statistical Picture*. (Third edition). Geneva: International Labour Office.
- Kanbur, R. (2017). "Informality, Causes, Consequences, and Policy Responses." *Review of Development Economics*, 21. 939-961.

- Karlen, R., Rougeaux, S., Johansson de Silva, S., & Simon, B. (2023). *Resilience and productivity: Keys to a better future for urban informal sector workers in Côte d'Ivoire*. World Bank, Washington, DC.
- Karlen, R. & Rother, F. (2023). "Togo Jobs Diagnostic—Confronting Challenges and Creating Opportunities for More Good Quality Jobs for All." World Bank, Washington, DC.
- Kinda, T. & Loening, J. L. (2010). Small enterprise growth and the rural investment climate: Evidence from Tanzania. *African Development Review*, 22(1), 173-207.
- Kiaga, A. and Leung, V. (2020). "The Transition from the Informal to the Formal Economy in Africa." In ILO (eds) *The transition to formality: Comparing policy approaches in Africa, Asia and Latin America*." ILO, Geneva.
- Lall, S., Henderson, J.V. & Venables, A. (2017). "Africa's Cities: Opening Doors to the World." World Bank, Washington, DC.
- Lam, D., Leibbrandt, M., & Allen, J. (2019). "The Demography of the Labor Force in Sub-Saharan Africa: Challenges and Opportunities." GLM/LIC Synthesis Paper #10. IZA, Bonn, Germany. https://g2lm-lic.iza.org/wp-content/uploads/2019/11/glmlic_sp010.pdf
- La Porta, R.& Shleifer, A. (2014). Informality and development. *Journal of economic perspectives*, 28(3), 109-26.
- Levenson, A. R. & Maloney, W. F. (1998). "The informal sector, firm dynamics, and institutional participation." Policy Research Working Paper #1988. The World Bank, Washington, DC.
- Loening, J. & Mikael Imru, L. (2009). "Ethiopia: diversifying the rural economy. An assessment of the investment climate for small and informal enterprises." MPRA Paper No. 23278. https://mpra.ub.uni-muenchen.de/23278/1/MPRA_paper_23278.
- Loening, J. L., Rijkers, B., & Söderbom, M. (2008). Nonfarm microenterprise performance and the investment climate: evidence from rural Ethiopia. *World Bank Policy Research Working Paper*, (4577).
- Maloney, W.F. (2004). "Informality Revisited." *World Development*, 32(7), 1159-1178.
- McCaig B., Pavcnik N. 2021. "Entry And Exit of Informal Firms and Development". Working Paper 28986: <http://www.nber.org/papers/w28986>.
- Mehrotra, S. (2020). From Informal to Formal: A Meta-Analysis of What Triggers the Conversion in Asia. Global Employment Policy Review Background paper No 2. ILO, Geneva.
- Merotto, D., Weber, M., & Aterido, R. (2018). Pathways to Better Jobs in IDA Countries. Washington DC, World Bank.

Nguimkeu, P. (2014). "A Structural Econometric Analysis of Informal Sector Heterogeneity." *Journal of Development Economics*, 107, 175-191.

Peeters, P., Cunningham, W., Acharya, G., & Van Adams, A. (2009). *Youth Employment in Sierra Leone : Sustainable Livelihood Opportunities in a Post-conflict Setting*. World Bank Publications, Washington DC..

Perry, G., Maloney, W.F., Arias, O., Fajnzylber, P., Mason, A. & Saavedra-Chanduvi, J. (2007). *Informality: Exit and Exclusion*. World Bank Latin American and Caribbean Studies. Washington, DC: World Bank.

Schoar, A. (2010). The divide between subsistence and transformational entrepreneurship. *Innovation policy and the economy*, 10(1), 57-81.

Ulysea, G. (2020). "Informality: Causes and Consequences for Development." *Annual Review of Economics*, 12. 525-46.

United Nations Population Division. 2018. *World Urbanization Prospects: 2018 Revision*. United Nations, New York USA.

Appendix

Annex 1: Data

Table A1. Surveys and year used

Workers data: Global Monitoring Database			
Country	Country Code	Source Survey	Year of Data Collection
Angola	AGO	IDREA	2018
Burundi	BDI	ECVMB	2013
Benin	BEN	EHCVM	2018
Burkina Faso	BFA	EHCVM	2018
Côte d'Ivoire	CIV	EHCVM	2018
Cameroon	CMR	ECAM-IV	2014
Gabon	GAB	EGEP	2017
Ghana	GHA	GLSS-VII	2016
Guinea	GIN	EHCVM	2018
Guinea-Bissau	GNB	EHCVM	2018
Kenya	KEN	IHBS	2015
Liberia	LBR	HIES	2016
Mali	MLI	EHCVM	2018
Malawi	MWI	IHS-V	2016
Niger	NER	EHCVM	2018
Nigeria	NGA	GHSP-4	2018
Senegal	SEN	EHCVM	2018
Sierra Leone	SLE	SLIHS	2018
São Tomé and Príncipe	STP	IOF	2017
Eswatini	SWZ	HIES	2016
Chad	TCD	EHCVM	2018
Togo	TGO	EHCVM	2018
Uganda	UGA	UNHS	2016
South Africa	ZAF	QLFS	2019
Zambia	ZMB	LCMS-VII	2015
Zimbabwe	ZWE	PICES	2017
Firms database: Informal Enterprise Survey			
Mozambique	MOZ		2018
Zambia	ZMB		2019
Zimbabwe	ZWE		2017

Source: Own elaboration based on GMD and IES.

Annex 2: Geospatial data

Using the 2017 Mozambique Population Census (a 10% extract of relevant variables was made available to the research team) and big data from various sources, we generate the set of urban variables and the distribution of job types by neighborhood (*bairro*).

Urban variables

We introduce publicly available geospatial big data to proxy for six characteristics of the urban environment that cannot be captured through survey data. We use a three-step process to generate the variables, drawing from different databases. First, we create a 1km-by-1km tessellation (grid) of the entire country. Second, we pull the geospatial indicator and use the data to estimate statistics for each grid. Third, we aggregate each geospatial grid-level indicator to the *bairro* (neighborhood) level.⁶⁴ Therefore, each variable is estimated for each neighborhood in per square kilometer terms. By so doing, we make variables comparable between neighborhoods.

We generate six variables at the *bairro* (neighborhood level): nighttime lights, electrification access, population density, building density, urban density, and ICT density. These variables were selected because they are a proxy for variables that characterize the urban environment: population density, poverty, urban density and ICT density.⁶⁵ The six variables were generated as follows:

1. *Night-time lights*: This variable measures the intensity of luminosity at night. The luminosity may be due to electrification or any other light-emitting source (such as fires, stray lights and other light sources that can be picked up from space that are not electricity). We rely on the Visible Infrared Imaging Radiometer Suite (VIIRS) satellite (~463 meter resolution) data for a measure of annual nighttime lights brightness composites (per sqkm) within each *bairro*. We extract the total radiance values from the raster dataset provided by NASA's Earth Observatory database into a 1km-by-1km square tessellation of the study areas.⁶⁶ To clean the data, we employ the straylight corrected band composite which extends the visible areas closer to the poles and improves dynamic range. The result is the total luminosity (in nanoWatts/cm²/sr) for each square km in the five-city study area. We estimate average values for each *bairro* within each city. Figure 1 shows significant variation in night-time luminosity across each urban area, with lighter areas being those with

⁶⁴ This is the "5th administrative division", where the 9 provinces are the 1st administrative division, the districts are the 2nd, the *postos* are the 3rd, the *localidades* are the 4th and the *bairros* are the 5th.

⁶⁵ A larger set of potential variables was available, including various measures of build coverage. We analyzed the full set of variables and dropped those that proxied similar concepts. We then selected the six final variables based on their usage in published literature.

⁶⁶ A raster is an image file. It is a picture where each position in the image is an x-y coordinate and the intensity of the color is a functional mapping to a specific value. They are also called GeoTIFF (Tagged Image File Format) Raster or tif files.

greater luminosity. Night-time lights are used to proxy density of poverty (Henderson et al. 2012).

2. *Electrification access.* An additional measure of urban poverty is electrification access. It differs from the luminosity variable in two ways. First, it attributes night lights to electricity, cleaning the data of other sources of light. Second, it does not account for the brightness of the light. The electrification access measure is created by the University of Michigan's High Resolution Electricity Access (HREA) data. The HREA project leverages the VIIRS high resolution satellite data and uses machine learning techniques to estimate electrification access between 2012 and present. This is tracked against new population estimates based on computer vision techniques identifying all human-built settlement structures. For each 30m by 30m grid, the algorithm estimates the likelihood that the observed light is due to electrification. It assigns each grid a value of 1 if the light is likely due to electricity and 0 otherwise. We aggregate this data to the *bairro* level. The distribution of electrification for the *bairros* in the five cities is presented in Figure 2. A comparison of Figures 1 and 2 find high access to electricity in *bairros* with moderate (or low) levels of luminosity, and the luminosity shows greater spatial variance than the electrification access variable.⁶⁷
3. *Population density:* We utilize the WorldPop population rasters to estimate population and population density at the *bairro* level. The WorldPop geotiff rasters have a spatial resolution of approximately 100m x 100m (3 arc seconds). Figure 3 shows the population density for each of the five cities. The denser areas are generally those with more informal settlements. More population density is correlated with more electrification or luminosity.⁶⁸
4. *Building density:* An alternative measure of population density, we again use the WorldPop building rasters to estimate building density at the *bairro* level. The WorldPop building rasters contain information on the count, density, total area, mean area, mean length and total length of the buildings. We generate average building density per square kilometer and then aggregate the data to the *bairro* level (Figure 4). This variable is highly correlated with the population density variable (0.95 correlation coefficient).
5. *Urban density.* The WorldPop raster contains a rural/urban classification for each 100m grid cell. It uses information about building density, building length, other measurements of buildings in the grid cell, and similar variables in surrounding grid cells to assess how densely built an area is. If the algorithm estimate exceeds a pre-set threshold, the 100m grid cell is assigned a value of 1 and given a value of 0 otherwise. We use these data to generate the proportion of a 1k square area that is urban, and then aggregate to the level of the *bairro* (Figure 5). Though based on the built environment, this variable correlates highly with electrification (0.87), with a correlation 0.7 – 0.8 with the other urban variables. However, as seen in Figure 5, the

⁶⁷ The correlation coefficient is 0.73 (Annex 1).

⁶⁸ The correlation coefficients are 0.58 and 0.81, respectively (Annex 1).

general patterns are similar to the other maps, with greater urban density and greater population density or poverty variable co-occurring.

6. *ICT density*: Finally, to study access to ICT, we queried the Open CellID database for locations of all crowd-sourced cellphone towers in Mozambique. We estimate the number of cell phone towers per square kilometer within each neighborhood (as of April 9, 2022). Figure 6 shows that very low coverage across all cities, with higher concentration in one area of each city. The variable somewhat coincides with luminosity but has low correlation (around 0.25-0.4) with the other urban variables.

Poverty controls: We also use a 10% Census extract to control for neighborhood poverty in our analysis. The Census includes geo-spatial information on the *bairro* of residence for each respondent, which we can use to match respondent to the grid. We use variables that are directly generated in the 10% census extract: individual ownership for radio, landline phone, laptop, iron, electric stove, refrigerators, and car. Then, using the information on *bairro* of the respondent, we generate neighborhood rates of ownership (see Annex). We use the Census weights so that the data are representative at the *bairro* level.

Figure 1: log luminosity for each *bairro* in 5 urban areas

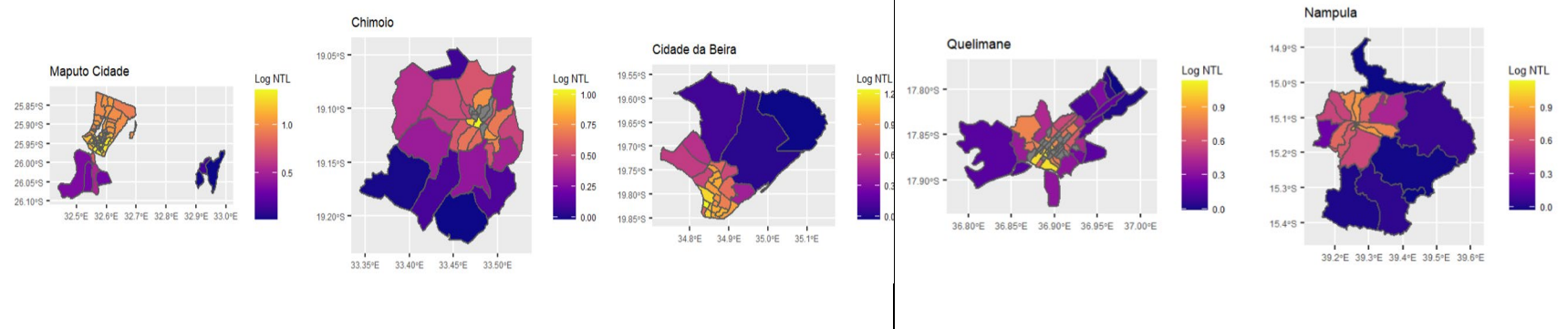


Figure 2: electrification access for each *bairro* in 5 urban areas

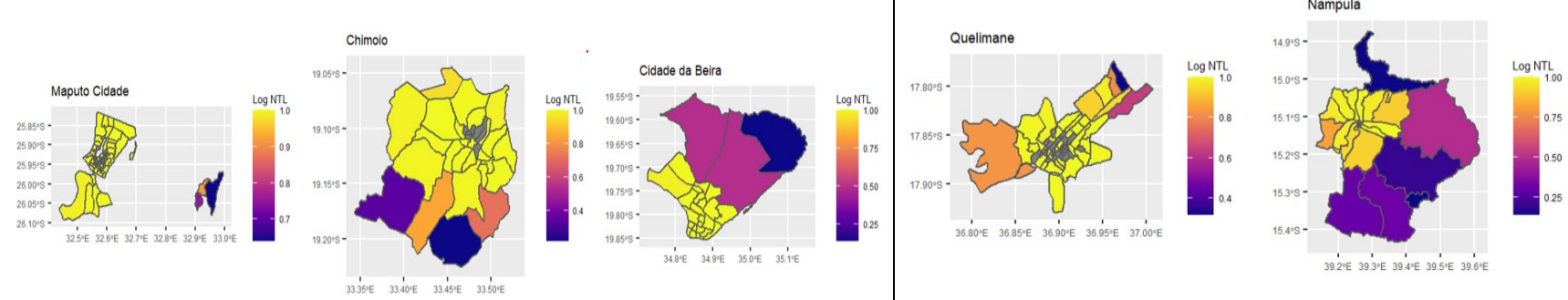


Figure 3: Population density by *bairro* in 5 urban areas

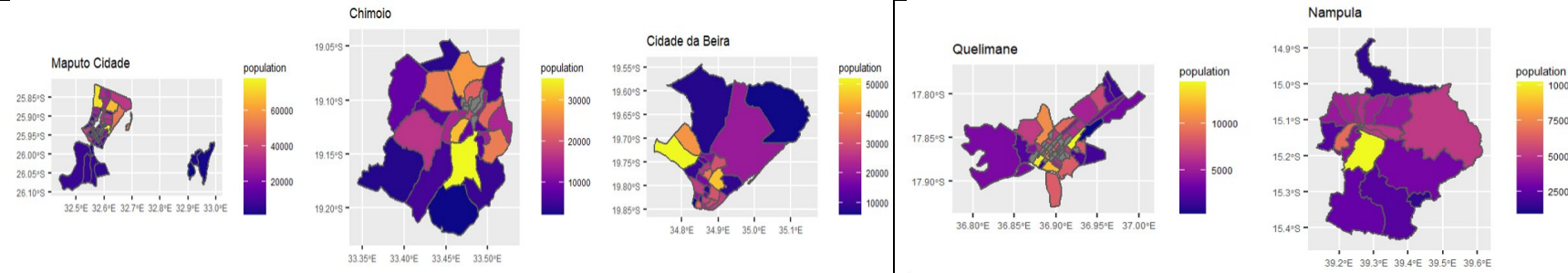


Figure 4: Building density by *bairro* in 5 urban areas

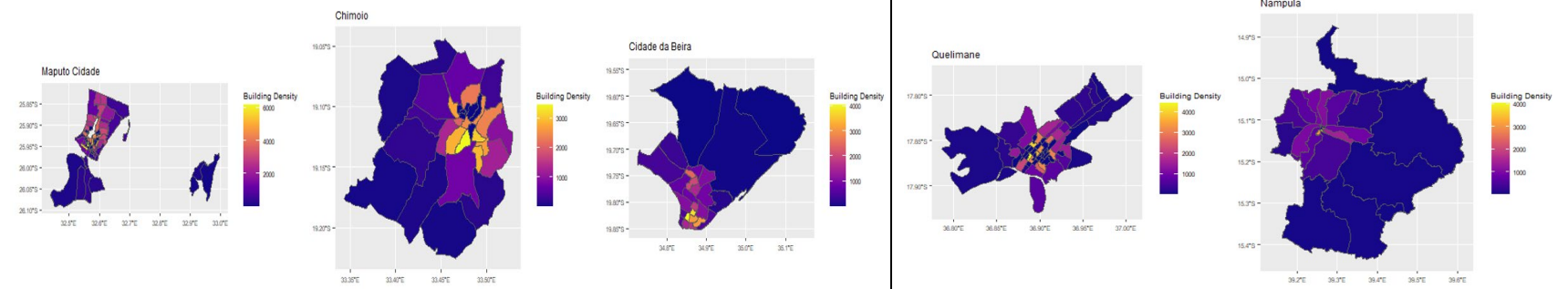


Figure 5: Proportion of the *bairro* that is urban

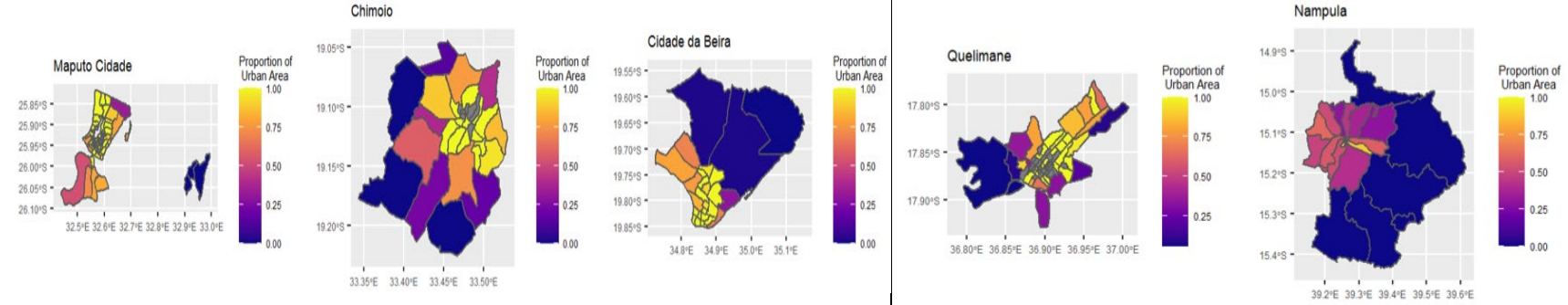
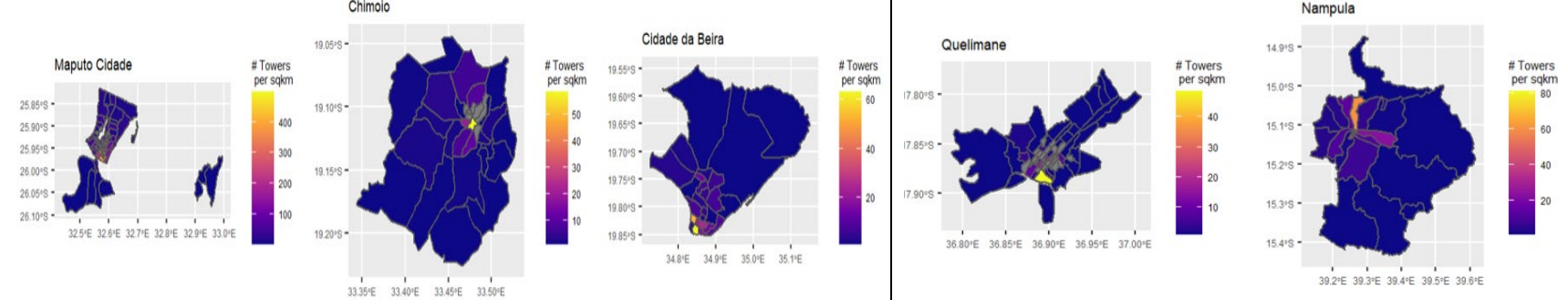


Figure 6: Cell Tower Locations by City



Job type variables

Our dependent variable is the density of an employment type within a *bairro*. To generate the employment type variable, we use the 10% Census extract, and estimate the number of individuals who are working in each of six classifications of types of employment in each *bairro*, as well as the share of the employed in each of those job types living in a *bairro*. The employment types are defined based on education and employment type since the Census does not provide information on more standard indicators of informal employment, such as social security payments or firm registries. The groups are:

- Employer, defined as “self-employed with employees”
- Formal Self-Employed, defined as “self-employed without employees” and having completed, at least, a secondary education
- Informal Self-Employed, defined as “self-employed without employees” and having less than a completed secondary education
- Formal Employee, defined as working in the public, private, or non-profit sector and having completed, at least, a secondary education
- Informal Employee, defined as working in a private home or in the public, private, or non-profit sector and having less than a completed secondary education
- Unpaid Workers, defined as being an unpaid worker in a family business.

Figure 8: City Maps of Employer Prevalence (neighborhood level disaggregation)

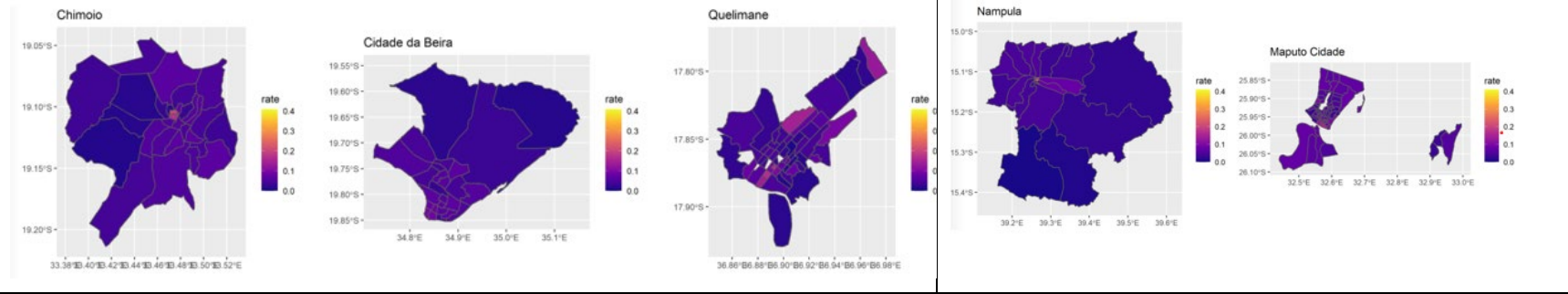


Figure 9: City Maps of Formal Self-employed Prevalence (neighborhood level disaggregation)

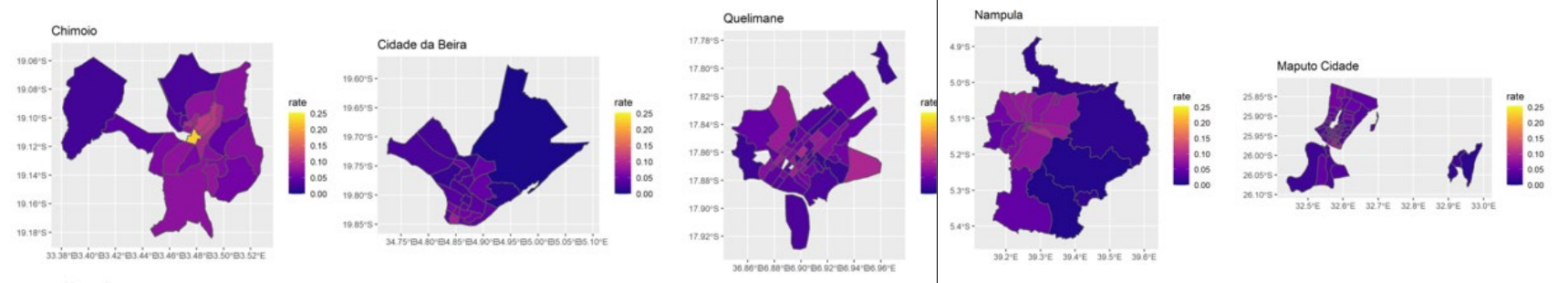


Figure 10: City Maps of Formal Wage-employee Prevalence (neighborhood level disaggregation)

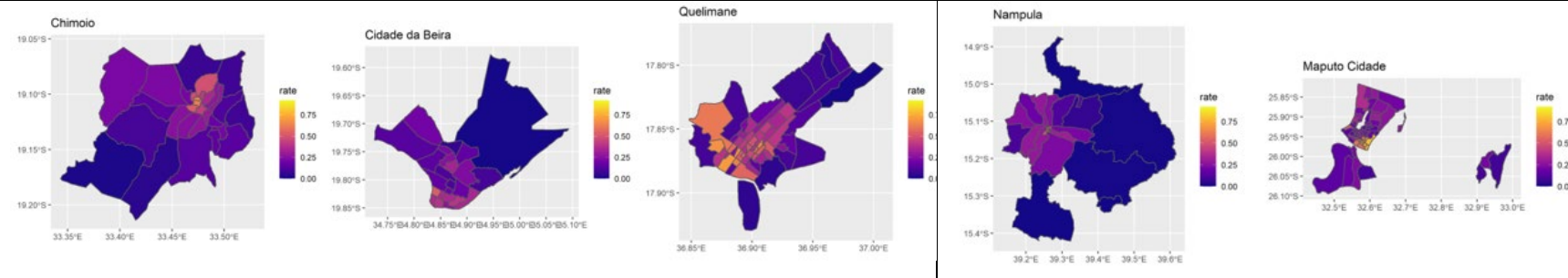


Figure 11: City Maps of Informal Wage-employee Prevalence (neighborhood level disaggregation)

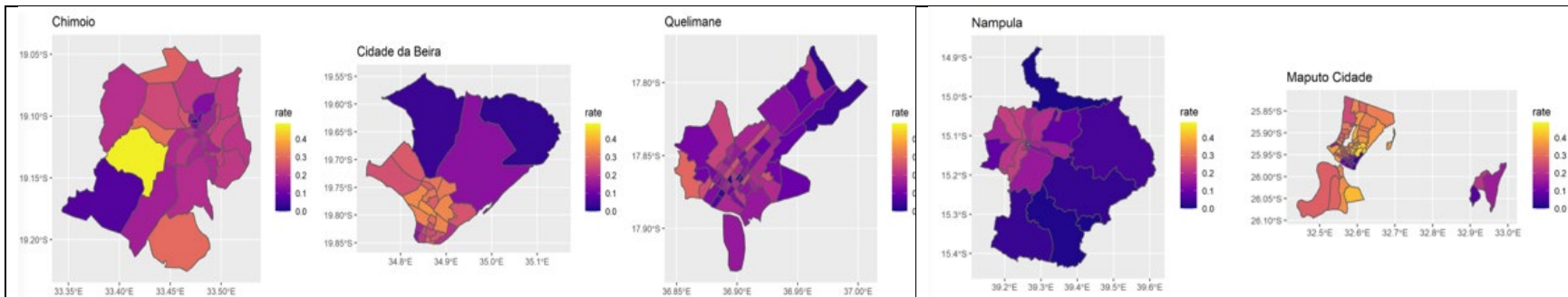


Figure 12: City Maps of *Informal Self-employed* Prevalence (neighborhood level disaggregation)

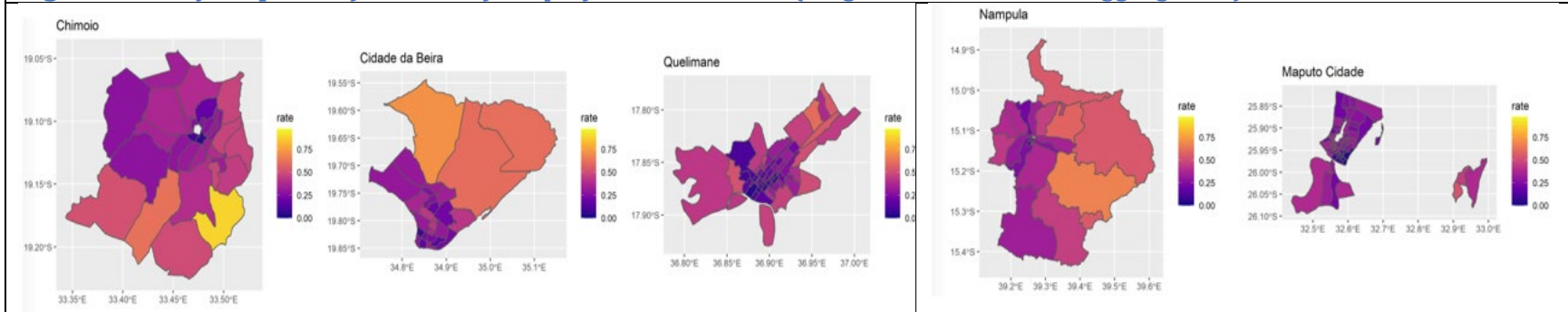


Figure 13: City Maps of *Unpaid Employee* Prevalence (neighborhood level disaggregation)

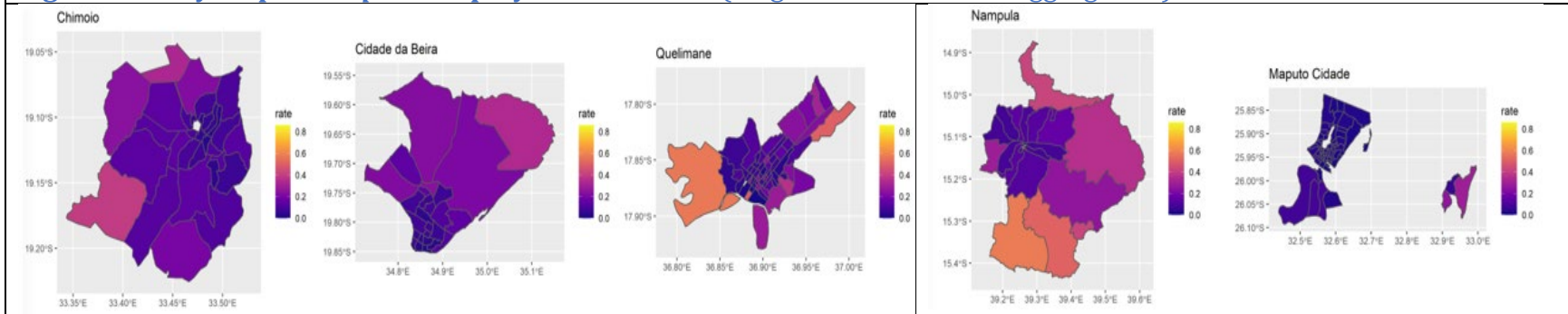


Figure 14: City Maps of *Informal Self-employed* Prevalence (city level disaggregation)