Spatial and Sectoral Heterogeneity of Occupational Choice in Cameroon

Théophile Bougna
Pierre Nguimkeu
Abstract

This paper investigates the relationship between location, agglomeration, access to credit, informality, and productivity across cities and industries in Cameroon. Emphasizing the link between micro-foundations and the data, the paper develops and estimates a structural model of occupational choice in which heterogeneous agents choose between formal entrepreneurship, informal entrepreneurship, and non-entrepreneurial work. Their decision-making process is driven by institutional constraints such as entry costs, tax enforcement, and access to credit. The model predicts that agglomeration has a non-monotonic effect on formalization, and entrepreneurial profits increase with agglomeration effects. Estimating the model by the generalized method of moments, the paper finds that the returns to capital and labor are not uniform across sectors and cities. Manufacturing industries are highly constrained in capital and the elasticity of capital is higher in Yaoundé and Douala, whereas labor elasticity is higher in Kribi. Counterfactual simulations show that an increase in roads provision can have a substantial impact in terms of output, formalization, and productivity. A reduction in the current interest rate has a large and significant impact on formalization and no significant effect on business creation. Likewise, while the current tax rate is suboptimal for most cities, a tax reduction policy would have a much greater impact on formalization than on business creation. These effects differ substantially across cities and sectors, suggesting that those policy instruments could be implemented accordingly to support formalization and business creation.

This paper is a product of the Development Research Group, Development Economics. 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/research. The authors may be contacted at tbougna@worldbank.org and nnguimkeu@gsu.edu.
Spatial and Sectoral Heterogeneity of Occupational Choice in Cameroon*

Théophile Bougna†  Pierre Nguimkeu‡

JEL Classification: O12, O17, C51, C52, C54, R23, R11.

*We are grateful to Emma W. Hobson and two anonymous referees for very helpful comments and suggestions. Financial support from the Economic Performance of African Cities (NLTA) is gratefully acknowledged. The content is solely the responsibility of the authors.
†Development Research Group (DECRG), The World Bank, 1818 H Street, NW Washington, DC 20433, USA; Email: tbougna@worldbank.org
‡Department of Economics, Andrew Young School of Policy Studies, Georgia State University, 14 Mariette Street NW, Atlanta, GA 30303, USA; Email: nnguimkeu@gsu.edu.
1 Introduction

The informal sector forms an integral part of developing economies, employing more than two-thirds of the urban workforce and accounting for about one-third of the economy-wide output in Cameroon (INS, 2010). An extensive body of the literature tries to explain the emergence of informality. For instance, Djankov and others (2002) document the existence of lengthy and costly registration procedures that hamper the formalization process of firms. Also, several papers have emphasized tax burden as a constraint to formalization (see, e.g., Fortin and others, 1997; Amaral and Quentin 2006; de Paula and Scheinkman, 2011; Prado, 2011; and Ordonez 2014). Nguimkeu (2014, 2015) showed that skills and access to credit play a significant role in the productivity and formalization of informal firms as well as the creation of new formal firms. However, the extent to which these constraints vary according to the sector of activity and the location of firms has received very little attention in the literature. The informal economy is getting progressively more heterogeneous in terms of demographic groups entering the sector, the types of linkages (including many cross linkages to the formal sector) and in the variety of economic enterprises. Each group brings different knowledge and skill sets leading to an increasing variety of enterprises in type and size. As noted by the U.K. 2017 Cities Alliance report “any future research agenda directed towards policy measures to enable better management of, and support to the informal sector will need to take on board this growing complexity and heterogeneity.” Understanding the sectoral composition of the informal economy and micro-enterprises, spatial location, and the major constraints they face in their growth and formalization prospects, is critical in informing policy aiming at harnessing informality and the agglomeration potential within cities for greater productivity and job creation.

This paper examines the constraints hindering the creation of new formal firms, the formalization of informal firms and firm productivity in Cameroon, and how these constraints vary across sectors and spatial locations. We also analyze how these two dimensions (sector of activity and spatial location) interact with each other. Extending Nguimkeu (2014, 2015), we develop an occupational choice model in which individuals choose between formal entrepreneurship, informal entrepreneurship, and non-entrepreneurial work, and their decision-making process depends on entry costs, tax enforcement and access to credit all of which are heterogeneous in location and sector of activity. On the one hand, to join the formal sector, thereby enjoying better access to the capital market and a better technology, an entrepreneur must pay an upfront entry registration cost, and then pay taxes once in operation in the formal economy. On the other hand, the entrepreneur may choose informality, therefore saving the sunk cost, and avoiding taxes, unless caught and fined. However, the entrepreneur is bound
to remain credit-constrained and/or to operate an archaic technology in the informal economy. Our model allows, in particular, to trace various categories of informal firms, with respect to their productivity potentials and the level of constraints they face, a useful characterization in tailoring policy actions to address informality and promote entrepreneurship. The model predicts that agglomeration has a non-monotonic effect on formalization of informal firms, and entrepreneurial profits increase with agglomeration effects.

This paper also contributes to the literature looking at policies designed to bring informal businesses and individuals within the reach of the tax system. Evidence is nuanced and the growth benefits of formalization are positive but heterogeneous across firm types. The smallest firms usually benefit the least while mid-sized firms typically benefit the most. Much economic informality seems to be explained as rational sorting of firms and workers based on inherent productivity characteristics and so formalization per se may not address the fundamental determinants of productivity (Oviedo and others, 2009). In addition, corruption or poor public services may also reduce workers’ and firms’ willingness to formalize and pay taxes. A World Bank (2007) report shows a strong and significant relationship between informal behavior of the firm (measured as tax evasion), and the incidence of corruption (measured by extortion from public officials). Torgler and Schneider (2007) also look at how institutions affect individuals’ decisions to stay informal or to join the formal economy. In particular, they look at how tax morale—which is a measure of how justifiable people think it is to avoid paying taxes—is associated with informality. They found that this measure is highly correlated with institutional measures such as rule-of-law, and corruption. Their result highlights the role of individual’s trust in the ability of the public sector to serve the public interest when they decide to participate in the established social and economic system, versus instead deciding to stay informal. In a recent paper, Benhassine et al. (2018) used a randomized experiment in Benin, to document how formalization rates can be enhanced by policies that aimed at providing information, training, and support in opening a bank account to business. However, they found no significant gains in sales, profits, or welfare. The international evidence is cautionary against expecting a quick growth dividend from policies to encourage formalization. As rightly pointed out by Oviedo and others (2009), strengthening enforcement will work best when the appropriate incentives to formalize are created, so that informal agents have a viable transition path from informality to formality. Strengthening enforcement only may not be growth-enhancing. Policy packages can comprise a series of carrots and sticks that need to be adapted to the nature of informality in the country. In some countries, regulatory reform might be more relevant, while in others it could be regulation enforcement, administrative reform or tax reform. In this paper, we provide evidence of how investing in local infrastructure, increasing access to finance,
and performing tax reforms can affect entrepreneurship, formalization rate, output, and productivity.

We first test the implications of the model using reduced form estimations techniques, based on a cross-section of micro-enterprises from the Business census and the National Survey on Employment and the Informal Sector. This allows us to test the theory by investigating empirically the importance of location and sector of activity on the extent of informality and firm performance. Our cross-sectional results show that firm profits increase with agglomeration and informal agricultural activities are the least profitable, whereas the least profitable formal activities are in the service sector. In addition, informal entrepreneurs operating in manufacturing, commerce and services are more likely to formalize than their counterpart in the agriculture sector. Second, we perform specification tests and provide structural estimates of the model using the Generalized Method of Moments and micro-data (firm-level and household-level) from Cameroon. We find that the return to capital and labor is not uniform across sectors and cities in Cameroon. Manufacturing industries are highly constrained in capital and the elasticity of capital is higher in Yaoundé and Douala, while labor elasticity is higher in Kribi, reflecting a low availability of workforce. One of the main advantages of using a structural econometric approach is that it allows one to generate counterfactual samples on the basis of a well-tested theoretical model that is consistent with the data. The counterfactual policy simulations that we perform based on the estimated model are then used to quantify the effects of several policy measures on informality and productivity. In particular, by identifying prominent sectors of activities (and/or strategic locations) and constraints faced by agents operating in these sectors (and/or locations), we provide a quantitative assessment of the impact that a targeted policy design such as improved infrastructure, access to finance or adjusted tax rates would have on aggregate productivity, job creation, formalization of informal firms, and government revenues. The results show that an increase in roads provision can have a substantial impact in terms of output, formalization of informal firms and productivity. A reduction in the current interest rate has no effect on business creation, but large and significant impact on the formalization of informal firms. The current tax rate could be adjusted to decentralized local optimal rates that would yield higher rates of formalization and business creation as well as higher amounts of government revenues. All these effects are heterogeneous across cities and/or sectors.

The remainder of the paper is structured as follows. Section 2 documents our data, and briefly presents the stylized facts about informality, location and business constraints. It also provides a first look at the geographic concentration of industries, the specialization of cities, and the formal/informal entrepreneurship in Cameroon. In Section 3, we develop a structural model of occupational choice in which individuals choose between wagework, informal entrepreneurship and formal entrepreneurship.
In Section 4 we perform the estimation of the model by presenting an econometric strategy that uses differences in location entrepreneurship rates and earnings in Cameroon to identify the structural parameters of the theoretical model. We also present the structural estimates by cities and by sectors. In Section 5, we perform a set of counterfactual simulations to evaluate the impact of public policies on entrepreneurship and informality across cities and industries in Cameroon. Section 6 concludes while additional results and technical details are relegated to the appendix.

2 Facts about Firm Informality, Location and Constraints

Our structural and empirical analysis requires three key pieces of industry-level and city-level information: (i) measures of absolute geographic concentration; (ii) measure of specialization of cities (places), and (iii) measures of formal, and informal entrepreneurship. We discuss our data and these measures, and take a first look at Cameroon’s geographical concentration of industries, specialization of cities, firm productivity, the constraints hindering the creation of new firms (formal and informal), and how these constraints vary across sectors and spatial locations. This will allow us to uncover important features of the data used for this study and assess the empirical relevance of the model predictions for Cameroon’s cities and industrial sectors.

2.1 Data

Our primary data sources are the Cameroon National Survey of Employment and Informal Sector (henceforth EESI) and the General Enterprise Census (henceforth GEC). The EESI is at the household level and covers the years 2005 and 2010. This dataset will be used to measure and characterize formal and informal entrepreneurship. The General Enterprise Census is at the firm level and is only for the year 2009. We are using the GEC dataset to measure the absolute geographic concentration of industries, and the specialization of cities.

2.1.1 General Enterprise Census data

The GEC provides detailed micro-geographic information on all formal and informal establishments with fixed location, including the 4-digit NACE (Statistical classification of economic activities) classification, the number of employees of the establishment, turnover, as well as information on the business environment. Table 1 presents the 2009 distribution of formal and informal establishments and jobs in Cameroon (by region and by industries). As can be seen, the bulk of jobs (68 percent) and firms (85 percent) are in the commerce and services sectors. Manufacturing industries represents roughly 13 percent in terms of firm counts and 18 percent in terms of jobs. Economic activities
are strongly concentrated (in terms of employment) in Yaoundé and Douala and, more generally, along the region surrounding these economic and political capital cities, respectively. These two cities which represent less than 0.2 percent of Cameroon’s total surface account for more than 60 percent of employment. One of the key reasons for that strong overall clustering is plausibly the uneven distribution of the population arising from colonial settlement patterns. Geographical variations in industrialization are the primary factor affecting the spatial variations in incomes. We use the 2009 General Enterprise Census data to get a more precise picture of the extent and the strength of industrial localization of industries as well as the specialization of cities in Cameroon.

Table 1: Sectoral distribution of formal firms, informal firms and jobs in Cameroon

| Industries                        | Establishments | Jobs |
|-----------------------------------|----------------|------|
|                                   | Freq. | Percent | Freq. | Percent |
| Agriculture                       | 145   | 0.16    | 26,530 | 6.87    |
| Breeding                          | 86    | 0.10    | 549    | 0.14    |
| Forestry                          | 47    | 0.05    | 6,787  | 1.76    |
| Fishing and fish farming          | 6     | 0.01    | 39     | 0.01    |
| **Sub-total Primary**             | 284   | 0.32    | 33,905 | 8.78    |
| Mining                            | 25    | 0.03    | 1,162  | 0.30    |
| Food industry                     | 695   | 0.79    | 10,447 | 5.03    |
| Other Manufacturing               | 10,227 | 11.60  | 49,993 | 12.94   |
| Electricity water and gas         | 74    | 0.08    | 8,530  | 2.21    |
| Construction                      | 664   | 0.75    | 8,757  | 2.27    |
| **Sub-total Secondary**           | 11,685 | 13.26  | 87,889 | 22.75   |
| Commerce                          | 46,788 | 53.08  | 105,458 | 27.30   |
| Transport                         | 518   | 0.59    | 15,041 | 3.89    |
| Bank and insurance                | 596   | 0.68    | 12,584 | 3.26    |
| Other tertiary                    | 26,963 | 30.59  | 128,844 | 33.36   |
| **Sub-total Tertiary**            | 74,865 | 84.93  | 261,927 | 67.81   |
| Unknown                           | 1,310 | 1.49    | 2,542  | 0.66    |
| **Total**                         | 88,144 | 100.00 | 386,263 | 100.00  |

2.1.2 Cameroon National Surveys on Employment and Informal Sector

We use data from the Cameroon National Surveys on Employment and Informal Sector (EESI) to empirically investigate the characteristics of the Cameroon informal sector and its most salient features. The EESI surveys were conducted in 2005 and 2010 by the
National Institute of Statistics of Cameroon in partnership with the World Bank. These are nationwide cross-sectional and representative surveys that collect a comprehensive set of information about household characteristics and the characteristics of their economic activities. The surveys identify thousands of active household members aged 15 and above operating in the informal sector defined as “activities and production units that do not have written formal accounts and/or are not registered with the tax authorities.” We present the most salient features of the Cameroon informal sector using both these surveys.

The Cameroon labor market is characterized by a large informal sector. The informal sector employs around 90% of the occupied workforce in Cameroon. In 2010, there were approximately 2.5 million (90.5%) of Informal Production Units (UPI) in Cameroon. The regional distribution of households with informal businesses revealed that the highest number of informal businesses were in the Littoral region, including Douala (19.0%), followed by the Centre region, including Yaoundé (17.4%), the Far-North (14.4%), and the West (12.1%). The regions with the least number of informal businesses are respectively the South (2.1%), the Adamawa (4.3%), and the East (4.3%). The informal sector is largely concentrated in agricultural, and non-tradable services (e.g. retail trade, transportation, hotels or restaurants). Figure 1 shows the distribution of informal and formal firm by sector of activity in Cameroon. As can be seen, most informal businesses are primarily engaged in agricultural (48.5%) followed by services sector (23.1%), while formal firms operate mainly in services (74.2%). The sectoral distribution of informal businesses is heterogeneous across regions in Cameroon. While predominant activities in Douala and Yaoundé are related to services (about 50%), agricultural businesses are dominant in other regions. Manufacturing (26%) and hotels and restaurant (38%) businesses are more predominant in Douala (26%), while food processing (26.5%) is more common in the Far-North.

The Cameroon informal sector is a great contributor to its economy. This sector accounts for more than 30 percent of the GDP. During the first decade of the 2000s, there was an increase in the contribution of the informal sector to Cameroon’s GDP—from 37.3 percent in 2005 to 38.4 percent in 2010 (see Figure 2). Given its large contribution to the GDP, any diagnostic on Cameroon productivity and growth should include policy recommendations for the informal sector. The most important challenges are therefore how to alleviate constraints to business environment and improve productivity in the informal sector and incentivize them to transit into the formal sector. Indeed, despite this substantial contribution to the GDP and the labor market, the Cameroon informal sector consists of micro-units, often with one worker, very low-paid jobs and low labor productivity. The average size of informal units decreased from 1.5 employee in 2005 to 1.3 employee in 2010. This is driven by an increase in the share of
informal units that have only one employee (from 69.4% in 2005 to 86% in 2010), and a decrease in the share of informal units with more than 3 employees (from 4.5% in 2005 to 3.6% in 2010). The Cameroon National Institute of Statistics (INS) documents that labor productivity—as measured by output per worker-hour—in the informal sector seems to be declining over the period 2005-2010, and the optimal productivity is attained by a typical firm in the informal sector when it has at least three workers plus the entrepreneur. In addition, less than 1 percent of employment is a salaried job, and for those with a salary, more than 50 percent of them earn less than the statutory minimum wage (the SMIC)\(^1\). Compared to the formal sector, productivity per worker remains relatively low in the informal sector. One important factor of this low productivity level is the lack of skills and competence due in part to the mismatch between the education system and the demands of the labor market, as well as financial constraints due to credit market imperfections as discussed in Nguimkeu (2014).

To achieve structural transformation, Cameroon needs to boost its industrial sector. The share of informal workers in the Cameroon industry has decreased between 2005 and 2010. From nearly 15% in 2005, there were only 12.6% of occupied workforce, the majority of which was in the food industry. The predominance of agricultural activity in the Cameroon informal sector is slightly different from Uganda, where only 27 percent of informal businesses were engaged in agricultural activity in 2010. Manufacturing has been a key factor in many emerging Asian and Latin American countries, which have experienced rapid growth in productivity and employment creation, as well as technological upgrading. Since manufacturing has the potential for

\(^1\)Salaire Minimum Interprofessionnel Garanti.
value-addition of existing resources, it is therefore critical for Cameroon to expand its manufacturing production to foster economic growth.

Figure 2: Contribution of the informal sector to Cameroon GDP (%).

Figure 3: Informal Employment by Industry (in percentage of all employment) GDP (%).
2.2 Geographical Concentration, Specialization, and Entrepreneurship

Here, we discuss measures of geographical concentration, specialization and entrepreneurial activity in Cameroon.

2.2.1 Geographical Concentration of industries in Cameroon

The importance of agglomeration varies across industries, hence different industries have been found to exhibit surprisingly different levels of agglomeration (see Ellison and Glaeser, 1997). As pointed out by Kerr and Kominers (2015) “a complete understanding of the pathways by which agglomeration increases firms’ returns requires an effective agglomeration index, a measurement which identifies the concentration of industry and explains the relationship between firm location choice and industry concentration”. To get a precise measure and picture of the extent and strength of industrial localization in Cameroon, we compute the well-known Ellison and Glaeser (1997) discrete measure of localization using micro-geographic data from the GEC; see Appendix A.1 for details. This allows us to identifies the industries for which localization is potentially the strongest and to examine how this trend varies across different types of industries (formal vs informal, manufacturing, and new entrepreneur).

Table 2 reports our results, and several comments are in order. First, the mean of the EG index is 0.007 and the median is 0.012. Out of the 99 four-digit industries in Cameroon, 60 industries (60%) exhibit a positive value for the EG index. In other words, 60 percent of industries are weakly or strongly concentrated. Within those industries, roughly half of them are considered as strongly agglomerated. Second, formal firms are on average more geographically concentrated than informal firms. Third, new entrants tend to exhibit more geographical concentration than existing firms. Fourth, manufacturing industries are less concentrated than other industries in general. Fifth, few manufacturing and natural-resources related industries are among the most concentrated industries in Cameroon. Sixth, informal firms are on average less concentrated than formal firm in Cameroon. Our results show that around 50 percent of formal firm are geographically concentrated while less than 30% of informal industries are clustered in space. This suggests that informal firms do not tend to cluster in the same locations as formal firms and we should expect a very weak buyer-supplier and technology linkages between firms in these two sectors of the economy. Using data from India, Mukim (2015) show that buyer-seller and technology linkages explain much of formal-informal coagglomeration in India manufacturing. Our results

A comparison between our results and the different studies is not reasonable since the EG index should be sensitive to the size and shape of the underlying zoning system as pointed out by Briant et al. (2010).
Table 2: Summary statistics for the geographic concentration measures

| Statistics          | Region Level | Arrondissement Level |
|---------------------|--------------|----------------------|
|                     | All Informal | Formal               | New Entrants     |
| Mean                | 0.091        | 0.007                | -0.020           | -0.042          | -0.124          |
| Median              | 0.010        | 0.012                | 0.007            | 0.000           | 0.029           |
| Minimum             | -4.363       | -2.081               | -2.002           | -1.647          | -3.536          |
| Maximum             | 6.903        | 1.757                | 1.146            | 0.418           | 1.975           |
| Standard deviation  | 1.200        | 0.358                | 0.311            | 0.299           | 0.755           |
| Share < 0           | 42.424       | 39.394               | 41.414           | 39.190          | 28.790          |
| Share ∈ (0, 0.05]   | 24.242       | 31.313               | 32.323           | 12.160          | 13.630          |
| Share > 0.05        | 33.333       | 29.293               | 26.263           | 48.650          | 57.580          |
|                     | All Informal | Formal               | New Entrants     |
| Mean                | -0.138       | -0.078               | -0.075           | -0.162          | -0.120          |
| Median              | -0.015       | -0.001               | -0.016           | -0.044          | -0.006          |
| Minimum             | -4.927       | -1.272               | -1.225           | -1.201          | -0.941          |
| Maximum             | 4.481        | 0.387                | 0.387            | 0.143           | 0.182           |
| Standard deviation  | 1.211        | 0.292                | 0.286            | 0.300           | 0.284           |
| Share < 0           | 53.490       | 51.170               | 53.490           | 44.830          | 36.000          |
| Share ∈ (0, 0.05]   | 13.950       | 39.530               | 32.560           | 13.790          | 16.000          |
| Share > 0.05        | 32.560       | 9.300                | 13.950           | 41.380          | 48.000          |

Notes: Authors’ computations using gec data. Mean and median values are for 99 four-digit NACE classification of industries (43 manufacturing, 74 formal, and 66 new entrants). Share < 0 means ‘not clustered’; Share ∈ (0, 0.05] means ‘weakly clustered’; Share > 0.05 means ‘strongly clustered’. See Ellison and Glaeser (1997) for details.

show that almost 60 percent of new entrants are strongly concentrated. This result is consistent with Glaeser, Kerr, and Ponzetto (2010).

Looking at the geographical concentration of industries within cities, the following facts emerge. Douala and Bamenda exhibit significant geographical localization compared to the overall distribution of industries in Cameroon. Out of our 99 four-digit types of industries in Cameroon, 95 are operating in Douala, and 56 in Bamenda. The mean EG index of the overall industries in Cameroon is only 4.3 and 1.2 percent of that in Yaoundé and Bamenda respectively. Sugar manufacture, fabrics and stitches, and the growth of other industrial or export products are among the most localized industries in Douala. In Bamenda, the most localized industries are related to soap factories, perfume, detergents and cleaning products, the raising of other animals, and other wholesale trade. As expected, general public administration activities are the
second most concentrated sector in Yaoundé which is the country’s political capital. However, it is important to note that seven out of the ten most concentrated industries in Yaoundé are related to manufacturing activities (Manufacture of dairy products, pulp, paper, paperboard and paper products, cereal flour, soap, perfume, detergents and cleaning products, footwear, and other non-metallic products). The most concentrated industries in Kribi are related to the retail trade services, hotels and camps, restaurants and bars. This result is not surprising given that Kribi is one of the country’s loveliest cities endowed with natural touristic attribute.³

2.2.2 Regional Specialization in Cameroon

The previous analysis helps understand whether economic activities are localized or clustered in particular places in Cameroon. In this section, we ask a different but related question, which is whether cities or arrondissements specialized in particular activities. To answer this question, we need a relative specialization index that will help understand whether specific industries tend to account for a larger share in the economic activity of a city relative to their average share in all other cities. As its well known in the urban economic literature, cities manufacture different types of goods and, as such, production patterns differ across the urban hierarchy. Specialization and diversification are important for different types of entrepreneurial activities, and diversification is linked to more early-stage entrepreneurial activities (Fritsch and Storey, 2014). Using data on 23 urban areas in 12 European Union countries, they show that urban regions with high levels of economic growth and diversity of economic activities exhibit higher levels of opportunity-motivated entrepreneurial activity than their counterparts. Henderson, Lee and Lee (2001) showed that larger cities tend to have a more diverse industrial base, with providers of niche products and services who can find a market in a large agglomeration. On the other hand, smaller and medium size cities tend to be fairly specialized in food and beverage production, textiles, shoes, or pulp and paper products.

To get an accurate picture of the extent and the strength of cities specialization in Cameroon, we computed the Krugman’s specialization index at the 4-digit NACE level (see Appendix A.1 for details). A high value of the index (higher than 0.75) represents a highly specialized city, and a lower value of the index (below 0.35) represents a diverse one. As can be seen from Table 3 and Figure 4, localities are specialized on average (37 percent of cities are specialized in 2009). In other words, compared

³Kribi is a town located at the coastal lowlands of the southern region of Cameroon. In addition to its various touristic services such as hotels, bars, sporting complex, and cultural evenings for entertainment, the Chad-Cameroon pipeline project along with the seaport construction are expected to boost the industrial sector there in the coming years.
to the overall distribution of economic activities across cities, few industries tend to account for a large share in the economic activity of a city. Our results also show that specialization decreases with city size. In 2009, the average specialization index in large (0.271) and medium (0.379) agglomerations were respectively 61 and 86 percent of that in the small agglomerations (0.442). Figure 5 shows a significant negative relationship between locality specialization and population size. The corresponding regression also shows a negative and statistically significant relationship between agglomeration specialization and city size. We ran a simple corresponding regression for large cities and found statistically significantly negative relationship between specialization and city size. The slope is -0.31 with an $R$-square of 0.03. This result is consistent with Da Mata et. al (2005) who found that the specialization index in the largest agglomerations in Brazil was just 28 percent of that in the bottom agglomerations. As for Brazil’s large cities, large cities in Cameroon are on average more diversified than smaller ones. As can be seen from Figure 5, most arrondissements in Yaoundé and Douala are diversified, except Yaoundé II, while a small city like Kribi is specialized.

Natural-resources based localities are among the most specialized cities in Cameroon. In 2009, the most specialized localities were Niete and Dizangue—which are agro-industrial cities (see Table 12 in Appendix A.2). As can be seen from Figure 4 and Table 12 in Appendix A.2, most of the least specialized cities are located in the East, Far-North, and Adamawa regions—80% of the twenty least specialized localities are in these regions. The most specialized cities are in the South (Kribi, and Niete), the Centre, and the Littoral (Dizangue, Dibamba, and Melong) regions.

### 2.2.3 Entrepreneurship in Cameroon

There is an increasing interest on entrepreneurship especially at its regional dimension. Empirical research provides evidence that ‘space matter’ in entrepreneurship. In words, there are significant spatial differences in entrepreneurial activities across sub-national spatial units within a country. See, e.g., Reynolds (2004) and Armington and acs (2002) for the U.S.; Fritsch and Mueller (2004) for Germany, and Keeble and Walker (1995) for the U.K. However, there is little consensus about the most appropriate measure of entrepreneurship. Two main approaches to measure entrepreneurship

---

4We find similar results by using population quantiles to differentiate small and large cities. The average specialization index in the top quintile (largest agglomerations) is 66 percent of that in the bottom quintile (small agglomerations).

5We include only large cities i.e., cities with population above the median population. We also run similar regressions with all cities and found similar results i.e., a significantly negative relationship between specialization and city size, with a slope of -0.16 and an $R$-square of 0.03.

6Niete hosts the industrial rubber company *Hevea du Cameroun* (*HEVECAM*).
Figure 4: Specialization of Cities in Cameroon in 2009.

Note: Cities with missing data are in gray. The specialization index range between 0 and 2. Highly specialized (greater than 0.75), weakly specialized (between 0.35 and 0.75), Not specialized (less than 0.35).
Table 3: Summary statistics the Krugman specialization measure

| Statistics          | All   | City size |
|---------------------|-------|-----------|
|                     | Mean  | Small     | Medium   | Large   |
| Mean                | 0.408 | 0.442     | 0.379    | 0.271   |
| Median              | 0.232 | 0.249     | 0.265    | 0.138   |
| Minimum             | 0.000 | 0.000     | 0.000    | 0.020   |
| Maximum             | 1.958 | 1.958     | 1.347    | 1.021   |
| Standard deviation  | 0.426 | 0.461     | 0.359    | 0.285   |
| Number of cities    | 300   | 199       | 66       | 35      |

| Share of specialized cities | All   | Small     | Medium   | Large   |
|------------------------------|-------|-----------|----------|---------|
| Highly specialized (ksi > 0.75) | 21.00 | 25.13     | 15.15    | 8.57    |
| Weakly specialized (ksi ∈ (0.35, 0.75]) | 20.33 | 18.09     | 28.79    | 17.14   |
| Not specialized (ksi < 0.35)      | 58.67 | 56.78     | 56.06    | 74.29   |
| Total                          | 100.00| 100.00    | 100.00   | 100.00  |

Notes: Authors’ computations using gec data. Mean and median values of the Krugman Specialization Index (ksi) for 300 arrondissements in Cameroon. Small, medium, and large cities are respectively cities with less than 50,000, between 50,000 and less than 100,000, and more than 100,000 inhabitants in 2011. Population data are from Cameroon Census Bureau (BUCREP), and the National Institute of Statistics (INS).

have been extensively used in the literature (see Glaeser and Kerr, 2009). The first approach consisted of using (i) the share of startups within a single industry, (ii) the new product introductions, (iii) the venture capital placement, and (iv) the creation of new firms. All these metrics allow to capture the dynamic nature of entrepreneurship. However, obtaining these data and information is difficult, especially in the developing context and for confidentiality reasons. The second approach consists in associating entrepreneurship with the number of people leading independent enterprises. To this end, relevant metrics are (i) self-employment, and (i) average firm size. Hence, in this paper, we define as an informal entrepreneur a self-employed in the informal sector who owns a business that has at least one employee. The definition of formal entrepreneur is the traditional one (that is a self-employed operating in the formal sector)

Despite significant improvement in the government willingness to improve the business climate in Cameroon, more than 50% of entrepreneurs are not satisfied with the overall business environment. Table 4 reports, for each 2-digit industry the weighted share of firms that identify a given factor as one of the top five most challenging con-
Figure 5: Cameroon City Specialization and population size in 2009.

Constraints for doing business in Cameroon. As can be seen from this table, corruption, high taxes and complex tax regime, credit constraints, excessive bureaucracy, unfair competition, financing cost, infrastructure, and insufficient public-private dialogue to promote collective action are the most challenging obstacles. Overall, these eight obstacles represent around 85 percent of the reported complaints. Table 4 and Figure 6 also show how the reported obstacles vary across sector of activities. Taxation, corruption, and access to credit are the major challenges in almost all industries, though it is important to note some disparities. For example, administrative delays and the lack of skills and training are major constraints in sectors such as energy, transport, and banking and insurance. Meanwhile, high financing costs are reported to be more important in the food, and construction industries, compared to other industries. These obstacles plague the business environment in Cameroon and undermine its ability to compete in global markets. A promising avenue to improve the business environment is to promote local economic policies that will help alleviate these obstacles. For example, a strong local level public-private dialogue can complement the existing dialogue taking place at the national level. The main advantage of such approach is that government
can better address or respond to the constraints that local prominent industries are facing.

Table 4: Entrepreneurship obstacles by sub-sector of activities in 2009

| Obstacles                       | Overall | Food | Other Mfg | Utilities | Construction | Commerce | Transport | Banking | Other tertiary |
|--------------------------------|---------|------|-----------|-----------|--------------|----------|-----------|---------|----------------|
| Tax system                     | 24.8%   | 24.9%| 24.2%     | 22.2%     | 21.0%        | 24.8%    | 22.5%     | 21.8%   | 24.8%          |
| Corruption                     | 16.5%   | 17.6%| 16.8%     | 14.4%     | 14.4%        | 16.8%    | 15.8%     | 15.9%   | 16.1%          |
| Access to credit               | 13.1%   | 9.8% | 12.7%     | 5.6%      | 12.6%        | 13.1%    | 18.0%     | 16.1%   | 13.2%          |
| Bureaucracy                    | 9.1%    | 8.8% | 9.4%      | 12.2%     | 10.1%        | 9.0%     | 11.1%     | 9.9%    | 9.2%           |
| Unfair competition             | 6.7%    | 2.4% | 6.2%      | 8.9%      | 6.5%         | 6.5%     | 5.5%      | 4.3%    | 6.9%           |
| Costs of financing             | 5.3%    | 6.2% | 5.7%      | 4.4%      | 7.1%         | 5.2%     | 5.3%      | 4.4%    | 5.3%           |
| Infrastructure                 | 5.6%    | 4.9% | 5.7%      | 3.3%      | 3.6%         | 4.9%     | 3.0%      | 3.6%    | 4.9%           |
| Insufficient public-private Dialogue | 3.6% | 6.0% | 3.9%      | 4.4%      | 5.1%         | 3.6%     | 5.1%      | 3.9%    | 3.7%           |
| Water and Energy               | 2.6%    | 4.1% | 2.8%      | 2.2%      | 2.8%         | 2.9%     | 1.8%      | 1.7%    | 3.0%           |
| Transportation                 | 2.6%    | 4.1% | 2.2%      | 0.0%      | 2.8%         | 2.8%     | 0.4%      | 2.8%    | 2.4%           |
| Skills and training            | 2.5%    | 0.8% | 2.8%      | 8.9%      | 3.1%         | 2.4%     | 3.3%      | 4.8%    | 2.3%           |
| Justice system                 | 2.1%    | 2.8% | 2.4%      | 3.6%      | 3.9%         | 2.5%     | 1.8%      | 1.7%    | 2.3%           |
| Labor legislation              | 0.3%    | 1.9% | 0.0%      | 0.7%      | 1.7%         | 2.0%     | 1.4%      | 1.6%    | 1.7%           |
| Supply of raw materials        | 1.6%    | 0.0% | 1.4%      | 2.2%      | 0.8%         | 1.6%     | 2.2%      | 2.1%    | 1.6%           |
| Market opportunities           | 1.2%    | 2.1% | 1.0%      | 3.3%      | 2.5%         | 1.2%     | 2.2%      | 1.7%    | 1.3%           |
| Preferential schemes           | 0.3%    | 0.0% | 0.2%      | 2.2%      | 0.3%         | 0.3%     | 0.0%      | 0.4%    | 0.3%           |
| No obstacles                   | 0.2%    | 0.0% | 0.5%      | 0.0%      | 0.0%         | 0.1%     | 0.0%      | 0.0%    | 0.2%           |
| Others                         | 0.8%    | 2.1% | 0.6%      | 0.0%      | 0.8%         | 0.8%     | 0.9%      | 0.7%    | 0.9%           |
| Observations                   | 39      | 742  | 9         | 75        | 3950         | 45       | 73        | 2066    | 6660           |

Note: Firms were asked to report the five (in decreasing order) most important obstacles to entrepreneurship. We used these rankings to compute a weighted average score by allowing a decreasing weighting scheme (0.5 for the top and 0.1 for the bottom). The percentage represents the average score of the weighted sum. Other Mfg gathers all non-food manufacturing activities. Utilities include Gas, electricity, and Water.

2.3 Regression analysis: Entrepreneurial activities and Infrastructure

In this subsection, we use simple OLS regression analysis to test the relationship between entrepreneurial profits (from formal and informal activity) with population density and regional infrastructure in Cameroon. Specifically, we implement an OLS model where we regress entrepreneurs net income as a function of their individual characteristics such as education, age, wealth and city common factors such as population density.\(^7\) We include location dummies for the following six cities: Douala, Yaoundé, Garoua, Bamenda, Bafoussam, and Kribi. Transport infrastructure is an important factor for the development of economic activity, especially in developing countries, where it plays an essential role by reducing the trade and transportation costs, providing economic and social opportunities and benefits that result in accessibility to markets and employment opportunities. The size or density of local population also captures an important aspect of agglomeration. This is standard in the literature since the pioneering work by Ciccone and Hall (1996).

\(^7\)We use principal component analysis to compute the wealth index using household belonging. The description of the procedure as well as variables taken from the questionnaire are given in the appendix.
We assume that the total income $\pi_{ij}$ of an entrepreneur operating in industry $i$ within city $j$ is given by the following equation:

$$\pi_{ij} = \alpha X_{ij} + \beta Z_j + \eta_i + \eta_j + \epsilon_{i,j}$$ (1)

where $\pi$ is the reported profit of individual $i$ in city $j$. The vector $X$ is the set of individual characteristics such as the formality status, entrepreneurial versus workers status, age, and education; $Z$ is the set of city characteristics such as population density and paved road density\(^8\), and $\epsilon$ is the error term. We also include in the regression locations and industries dummies. In addition, we consider alternative binary outcome regression specifications where the dependent variables are indicators of entrepreneurship, and formality. This is to estimate the probability of entrepreneurship and the probability of formalizing, respectively. Summary statistics of these variables are given in Table 5. Average monthly income is the highest in Douala and Yaoundé,

---

\(^8\)Figure 16 illustrate the spatial distribution of paved roads across cities in Cameroon. As can be seen, the Eastern region is the least connected and there is a lot of variability across cities. We are going to use this variability to investigate how a change in roads provision can impact entrepreneurship and formalization rate across cities and industries.
which is reasonable, given that these cities are also the most expensive in the country. Interestingly, our wealth index shows that wageworkers in Yaoundé are wealthier than those in Douala, although those in Douala have higher wages.

Table 6 reports the results of the regression equation given by (1). Columns (1) and (2) provide profit regression results for informal entrepreneurs and Columns (3) and (4) give results for formal entrepreneurs. In specifications (2) and (4), we add industries and city dummies as controls. These results show that firm profits increase with agglomeration, captured here by the population density. Moreover, in the informal sector, agriculture is the least profitable activity whereas the least profitable formal activities are in the services sector. As expected, age has a nonlinear effect on profits for all our specification. As entrepreneurs earn more profit as they get older, plausibly due to more experience. However, at a certain age threshold, aging becomes an handicap to profit. Education is also positively associated with informal firms’ profits, while wealth is insignificant. These effects are similar for formal entrepreneurs except that wealth is an important factor of productivity in the formal sector, essentially because it is used as a collateral credit borrowing. Finally, Column (5) and (6) report the probability of formalizing. Although most of the coefficients are insignificant, their sign seem consistent with our theory, as we discuss below. Results also show that informal entrepreneurs operating in manufacturing, commerce, and services are more likely to formalize than their agriculture counterpart. This could be explained by the fact that agricultural firms are usually located in the suburbs making it difficult for the tax authority to reach them.

Table 5: Summary statistics of the variables by specific cities

| Variables                      | Douala | Yaoundé | Garoua | Bamenda | Bafoussam | Kribi | Average |
|--------------------------------|--------|---------|--------|---------|-----------|-------|---------|
| Share of worker in agriculture | 0.05   | 0.05    | 0.36   | 0.23    | 0.32      | 0.02  | 0.11    |
| Share of worker in manufacture | 0.28   | 0.17    | 0.21   | 0.24    | 0.20      | 0.34  | 0.23    |
| Share of worker in commerce    | 0.24   | 0.21    | 0.13   | 0.17    | 0.22      | 0.18  | 0.21    |
| Share of worker in service     | 0.54   | 0.66    | 0.37   | 0.49    | 0.44      | 0.57  | 0.55    |
| Share of formal worker         | 0.31   | 0.36    | 0.23   | 0.19    | 0.22      | 0.66  | 0.31    |
| Share of non-entrepreneur      | 0.78   | 0.84    | 0.80   | 0.74    | 0.77      | 0.85  | 0.80    |
| Share of formal entrepreneur   | 0.03   | 0.02    | 0.02   | 0.02    | 0.02      | 0.07  | 0.02    |
| Share of informal entrepreneur | 0.19   | 0.14    | 0.17   | 0.25    | 0.22      | 0.08  | 0.18    |
| Average monthly income         | 141.12 | 132.14  | 86.60  | 91.40   | 82.72     | 104.72| 124.31  |
| Average monthly wages          | 117.26 | 108.30  | 77.55  | 80.87   | 74.25     | 94.96 | 104.18  |
| Education                      | 8.75   | 8.99    | 5.10   | 7.93    | 7.14      | 7.75  | 8.15    |
| Age                            | 29.43  | 28.44   | 28.17  | 28.89   | 26.61     | 29.16 | 28.79   |
| Wealth                         | 149.12 | 131.44  | 120.92 | 126.84  | 21.87     | 7.85  | 117.38  |
| Paved road density             | 0.08   | 0.15    | 0.10   | 0.12    | 0.13      | 0.07  | 0.11    |
3 Model Economy

The economy comprises $J$ cities (or locations) and a mobile population of size $N_j$ in each city $j$. Individuals choose between wagework, informal entrepreneurship and entrepreneurship. We assume as in Evans and Jovanovic (1989), Paulson et al (2006) and others, that each individual is endowed with an initial wealth $z$ and an intrinsic level of entrepreneurial ability $\theta \in [0, +\infty)$, drawn from a fixed, continuously differentiable, city distribution $G_j(\theta)$. There is a large number of industries (or sectors), indexed by $i$. Each sector produces a homogeneous final consumption good. Production of each good requires labor and capital, both of which are freely mobile across cities. Workers are hired locally and paid city-specific wages, whereas capital is owned globally and is rented at the same price everywhere. We assume that the total output $y_{ij}$ of an entrepreneur with ability $\theta_{ij}$, operating in industry $i$ within city $j$ is given by:

$$y_{ij} = \theta_{ij} A_{ij} k_{ij}^{\alpha} l_{ij}^{\beta}$$  \hspace{1cm} (2)

\begin{table}[h!]
\centering
\caption{Regression analysis of firm profits and formalization}  
\begin{tabular}{l|cccc|cc}
\hline
Variables & \multicolumn{4}{c|}{Firm Profits} & \multicolumn{2}{c}{Formalization} \\
 & (1) & (2) & (3) & (4) & (5) & (6) \\
\hline
Education & 6.517*** & 6.279*** & 5.930** & 5.310* & 0.127*** & 0.142*** \\
 & (0.723) & (0.726) & (2.891) & (2.960) & (0.019) & (0.021) \\
Age & 3.133** & 3.502*** & 6.328 & 5.022 & 0.012 & 0.046 \\
 & (1.019) & (1.026) & (5.436) & (5.403) & (0.033) & (0.034) \\
Age$^2$ & -0.027** & -0.028** & -0.032 & -0.014 & 0.000 & 0.000 \\
 & (0.012) & (0.012) & (0.069) & (0.069) & (0.000) & (0.000) \\
Wealth & -0.015 & -0.018 & 14.40*** & 14.73*** & -0.001 & -0.002 \\
 & (0.049) & (0.028) & (4.408) & (4.352) & (0.004) & (0.004) \\
PopDensity & 1.251** & 0.843*** & 1.561 & 2.388 & -0.018 & -0.070 \\
 & (0.615) & (0.626) & (2.944) & (2.923) & (0.017) & (0.062) \\
PopDensity$^2$ & 0.002 & 0.002 & & & & \\
 & & & & & (0.004) & (0.004) \\
Manufacturing & 18.59*** & 51.85** & & & 1.669*** & \\
 & (7.087) & (26.22) & & & (0.181) & \\
Commerce & 22.50*** & & -0.944 & & 1.331*** & \\
 & (6.829) & & (29.03) & & (0.189) & \\
Services & 119.8*** & & 45.77 & & 2.191*** & \\
 & (6.591) & & (31.96) & & (0.191) & \\
City Dummies & NO & YES & NO & YES & NO & NO \\
R$^2$ / Pseudo R$^2$ & 0.062 & 0.055 & 0.136 & 0.118 & 0.039 & 0.153 \\
Obs & 1897 & 1807 & 247 & 247 & 2143 & 2143 \\
\hline
\end{tabular}
\end{table}

Notes: Our unit of observation is individual. Standard errors are in parentheses. Column (1) and (2) and profit regressions for informal firms; Column (3) and (4) are profit regression for formal firms; Column (5) and (6) are probit regressions for the probability of formalizing. Significance codes: ‘*’ 10%, ‘**’ 5%, and ‘***’ 1%.
where \( k_{ij} \) and \( l_{ij} \) denote the capital and labor inputs, with economy-wide capital and labor shares \( \alpha_i, \beta_i \in (0, 1) \), respectively; and \( A_{ij} \) is a location and sector specific agglomeration externality (that includes both road infrastructure and firm/household agglomeration). We also have \( \gamma_i = \alpha_i + \beta_i < 1 \), implying diminishing returns to scale in variable factors at the establishment level, as in Lucas (1978). As suggested by Rosenthal and Strange (2003) we allow the agglomeration externality \( A_{ij} \) to act as a multiplier on the production function.

### 3.1 Entrepreneurship

In addition to renting capital at rate \( r_j \) and paying wages at rate \( w_j \), there are two regulation costs that formal firms pay. In order to formalize, they need to pay a registration cost, \( c_j \), that represents administrative costs such as license fees as well as bribes and the opportunity cost of the days spent for the registration procedure. They also pay a profit tax \( \tau \) at the end of each period. Moreover, at the end of the period, the government levies a tax \( \tau \) on the profit of formal enterprises. Their formal status however provides them with a better access to credit from financial institutions so that capital is assumed to be freely optimized (see Bruhn 2011, Laporta and Shleifer 2014).

Denoting by \( \pi^F_{ij} \) the formal firm’s one period profit in location \( j \) and industry \( i \), the (static) profit maximization problem of the firm is:

\[
\max_{k_{ij} \geq 0, l_{ij} \geq 0} \left\{ (1 - \tau) \left[ \theta_{ij} A_{ij} k_{ij}^{\alpha_i} l_{ij}^{\beta_i} - w_j l_{ij} - r_j k_{ij} \right] - r_j c_j \right\}
\]

The interior solutions of the entrepreneurs maximization problem are:

\[
k_{ij}^F = \left( \theta_{ij} A_{ij} \right)^{\frac{1}{1 - \beta_i}} \left( \frac{\alpha_i}{r_j} \right)^{\frac{1 - \beta_i}{\gamma_i}} \left( \frac{\beta_i}{w_j} \right)^{\frac{\beta_i}{\gamma_i}} \quad \text{and} \quad l_{ij}^F = \left( \theta_{ij} A_{ij} \right)^{\frac{1}{1 - \beta_i}} \left( \frac{\alpha_i}{r_j} \right)^{\frac{1 - \beta_i}{\gamma_i}} \left( \frac{\beta_i}{w_j} \right)^{\frac{\beta_i}{\gamma_i}}
\]  

(3)

The optimal formal entrepreneur’s profit can therefore be expressed as follows:

\[
\pi^F(\theta_{ij}, A_{ij}) = (1 - \tau)(1 - \gamma_i)(\theta_{ij} A_{ij})^{\frac{1}{1 - \beta_i}} \left( \frac{\alpha_i}{r_j} \right)^{\frac{1 - \beta_i}{\gamma_i}} \left( \frac{\beta_i}{w_j} \right)^{\frac{\beta_i}{\gamma_i}} - r_j c_j.
\]  

(4)

Unlike the formal entrepreneur, the informal entrepreneur pays no taxes, but faces a capital cost of \( \lambda r_j \), where \( \lambda \geq 1 \). This means that informal entrepreneurs have less access to formal credit, since their cost of capital is higher than that of the formal

---

9 Notice that while the agent pays the one-time entry cost \( c_j \) to formalize his business, his periodic formal profit accounts for an amount of \( r_j c_j \) instead. This is because \( c_j \) can be seen as the present value of a sum of periodic payments of \( r_j c_j \) across the lifetime of the firm, i.e. \( c_j = \sum_{t=1}^{\infty} \frac{r_j c_j}{(1 + \tau)^t} \), assuming payments begin at the end of the current period.
entrepreneurs. Moreover, the informal entrepreneur also faces a probability of being detected by the authorities in which case all profit is forfeited. We assume that the detection probability increases with the “visibility” of the firm. It is therefore denoted $p(l_{ij})$ and depends monotonically on the size of the firm measured by its labor force, as follows:\textsuperscript{10}

\[ p(l_{ij}) = \begin{cases} 1 & \text{if } l_{ij} \leq \bar{l}_j \\ 0 & \text{otherwise} \end{cases} \tag{5} \]

In other words, an informal firm in sector $i$ and location $j$ cannot employ more than $\bar{l}_j$ units of labor without being caught by the authorities, but will not suffer any penalty when $l_{ij} \leq \bar{l}_j$.

The informal entrepreneur’s optimal investment capital and labor then solves the expected profit maximization problem:

\[
\max_{k_{ij} \geq 0, l_{ij} \geq 0} \left\{ (1 - p(l_{ij})) \left[ \theta_{ij} A_{ij} k_{ij}^p l_{ij}^p - w_j l_{ij} - \lambda r_j k_{ij} \right] \right\}
\]

The optimal choices for the informal entrepreneur are given by:

\[
k_{ij}^* = (\theta_{ij} A_{ij})^{\frac{1}{1-\gamma_i}} \left( \frac{\alpha_i}{\lambda r_j} \right)^{\frac{1-\beta_i}{1-\gamma_i}} \beta_i \left( \frac{\beta_i}{w_j} \right)^{\frac{1}{1-\gamma_i}} \quad \text{and} \quad l_{ij}^* = (\theta_{ij} A_{ij})^{\frac{1}{1-\gamma_i}} \left( \frac{\alpha_i}{\lambda r_j} \right)^{\frac{1-\beta_i}{1-\gamma_i}} \beta_i \left( \frac{\beta_i}{w_j} \right)^{\frac{1}{1-\gamma_i}} \tag{6}\]

This solution is feasible only when $l_{ij}^* \leq \bar{l}_j$, which, with the monotonicity of the demand function with respect to entrepreneurial ability, ensures the existence of a unique ability threshold $\theta_{ij}^c(A_{ij})$ above which the informal entrepreneur is constrained on the labor market. This means all informal entrepreneurs with $\theta_{ij} \leq \theta_{ij}^c$ operate unconstrained with labor $l_{ij} = l_{ij}^* \leq \bar{l}_j$, while all those with $\theta_{ij} > \theta_{ij}^c$ operate constrained with $l_{ij} = \bar{l}_j$. In the former case, the operational capital is $k_{ij} = k_{ij}^*$ whereas in the latter it is given by:

\[
k_{ij}^f = \left( \frac{\alpha_i \theta_{ij} A_{ij}}{\lambda r_j} \right)^{\frac{1}{1-\gamma_i}} \beta_i \left( \frac{\beta_i}{w_j} \right)^{\frac{1}{1-\gamma_i}} .
\]

The optimal informal entrepreneur’s profits can therefore be expressed as follows:

\[
\pi^f(\theta_{ij}, A_{ij}) = \begin{cases} (1 - \gamma_i)(\theta_{ij} A_{ij})^{\frac{1}{1-\gamma_i}} \left( \frac{\alpha_i}{\lambda r_j} \right)^{\frac{1-\beta_i}{1-\gamma_i}} \beta_i \left( \frac{\beta_i}{w_j} \right)^{\frac{1}{1-\gamma_i}} & \text{if } \theta_{ij} \leq \theta_{ij}^c(A_{ij}) \\ (1 - \alpha_i)(\theta_{ij} A_{ij})^{\frac{1}{1-\alpha_i}} \left( \frac{\alpha_i}{\lambda r_j} \right)^{\frac{1-\beta_i}{1-\alpha_i}} \beta_i \left( \frac{\beta_i}{w_j} \right)^{\frac{1}{1-\gamma_i}} - w_j \bar{l}_j & \text{otherwise.} \end{cases} \tag{7}\]

\textsuperscript{10}Note that this functional form is adopted for simplicity. A more general form for the function $p(\cdot)$ can be assumed and our qualitative results would be unchanged.
Although this paper has a similar broad objective of investigating occupational choice in the Cameroon context as in Nguimkeu (2015), this framework substantially differs from the latter in at least three distinct ways. First, it extends the model in Nguimkeu (2015) to allow for agglomeration externalities and examine how they affect the sorting of talents across cities and occupations. Second, unlike in Nguimkeu (2016) where the detection probability in the informal economy was constant and exogenously fixed, we endogeneize it in the current paper so that it monotonically vary with the size of the firm. This better fits the reality since in real economies, the likelihood of getting in trouble with the tax authorities often increases with the degree of visibility of the enterprise (see De Paula and Sheinkman 2011). Finally, we allow for the cost of capital to differ with the formality status of the firms in order to reflect the perception that informal firms face a higher cost of capital compare to their formal counterparts. These differences allow to obtain additional insights and to uncover potentially subtle empirical patterns that may help design more refined and targeted policies aiming at enhancing the economic performance of Cameroon cities in particular, and African cities in general.

3.2 Occupational Choice and Agglomeration

Each agent in location $j$ also has the choice of becoming a worker (or non-entrepreneur) and receiving the local market wage $w_j$. In this class of models, usual arguments (continuity, monotonicity, differentiability) guarantee the existence of unique ability cutoff points $\theta^W_{ij}(A_{ij})$ and $\theta^F_{ij}(A_{ij})$ such that:

- If $\theta_{ij} < \theta^W_{ij}$, the agent is a worker,
- if $\theta^W_{ij} \leq \theta_{ij} < \theta^F_{ij}$, the agent is an informal entrepreneur,
- if $\theta_{ij} \geq \theta^F_{ij}$, the agent is a formal entrepreneur.

These cutoff points are implicitly defined by:

$$\max\{I(A_{ij}, \theta^W_{ij}), F(A_{ij}, \theta^F_{ij})\} = w_j,$$

and

$$I(A_{ij}, \theta^F_{ij}) = F(A_{ij}, \theta^F_{ij}).$$

The expressions of $\theta^W_{ij}(A_{ij})$, $\theta^W_{ij}(A_{ij})$ and $\theta^F_{ij}(A_{ij})$ are given in the Appendix. Notice that if $1 - \tau < \beta = \frac{1}{\lambda}$, that is, taxes are too high compared to the capital cost wedge provided by the formal sector, then unconstrained informal entrepreneurs in
sector $i$ will not formalize, regardless of the registration cost. We assume that this condition holds in the model; it assures that firms whose detection probability is zero will always prefer to stay informal. The equilibrium in the local labor market requires $w_j$ to satisfy:

$$
\sum_i \int_{\theta_{ij}^H(w_j)}^{\theta_{ij}^N(w_j)} l_{ij}^I(\theta, w_j) dG_j(\theta) + \sum_i \int_{\theta_{ij}^H(w_j)}^{\theta_{ij}^N(w_j)} l_{ij}^F(\theta, w_j) dG_j(\theta) = \sum_i \int_0^{\theta_{ij}^H(w_j)} dG_j(\theta)
$$

\[
\text{Demand for labor} - \text{Supply for labor}
\]

where $l_{ij}^I(\theta, w_j) = l_{ij}^I(\theta, w_j)1[\theta < \theta_{ij}^I(w_j)] + l_{ij}^I[\theta \geq \theta_{ij}^I(w_j)]$, and the arguments emphasize the dependence of the cutoffs and labor demand on the level of wages. The unicity of the local equilibrium wage $w_j$ follows from usual arguments. The left hand side of Equation (8) is decreasing in $w_j$ and exceeds the right hand side when $w_j$ tends to zero, while the right hand side is increasing in $w_j$ and exceeds the left hand side when $w_j$ tends to $\infty$. The existence and unicity of the local equilibrium wage then follows from the continuity of these functions. This local equilibrium wage is sensitive to agglomeration as well as policies that directly or indirectly affect the labor supply and labor demand curves.

Figure 7: The Effects of Agglomeration on Profits

Agglomeration economies is a large concept that includes any effect that increases firms’ and workers’ income when the size of the local economy grows. Cities play an important role in the economy since economic proximity makes for more efficient production. These efficiency gains typically arise because of agglomeration externalities. Theoretical and empirical evidence shows that agglomeration effects are important to understand firm dynamics. Duranton and Puga (2001) study the the effect of agglomeration externalities in innovation and the development of production processes, while
Dumais, Ellison, and Glaeser (2002) examine the effect of firm dynamics (entry, exit, expansion, and contraction) on the concentration of economic activity. Therefore, agglomeration plays a fundamental role in shaping occupational choice and associated payoffs, and all the key qualities derived above—ability cutoffs, labor demands, capital demands, and the equilibrium wages—depend on the specific location and sector in which the agent operates. Note that these key qualities can also depend on other factors like natural advantages. However, Ellison and Glaeser (1999) find that natural advantage account for only a small fraction of industrial concentration in the United States, and Combes, Duranton, and Gobillon (2008) conclude that sorting and agglomeration account for the bulk of spatial wage disparities in France.

As explain in the previous sections, agglomeration externalities regroup interactions within and between sectors that give rise to various sorts of complementarities and indivisibilities, such as matching, sharing, and learning externalities (Duranton and Puga, 2004). Many studies (e.g. Behrens, Duranton, and Robert-Nicoud 2014) find that stronger agglomeration effects lead to larger cities and firms are more productive, on average, in larger cities (see Combes et al, 2012.). We therefore have the following results.

**Proposition 1.** We have the following results, ceteris paribus.

1. Entrepreneurial profits increase with agglomeration effects.
2. Stronger agglomeration effects increase the demand for labor.
3. Agglomeration has a non-monotonic (U-shaped) effect on the formalization of informal firms.

Item (1) of the proposition implies that more talented workers, who stand a higher chance of becoming entrepreneur, sort themselves to larger cities. However, it is also possible that individuals living in larger cities have the opportunity to obtain more and better education, making them more talented. The result given by Item (2) implies that stronger agglomeration effects may translate into higher wages, because labor demand might increase while labor supply is unchanged. Item (3) means that the likelihood of formalizing decreases with agglomeration effects up to a certain agglomeration threshold after which it becomes increasing beneficial for firms to formalize. Essentially, Firms decide to formalize only when their expected profits as formal entrepreneurs net of their expected profits as informal entrepreneurs is non-negative. In cities with low levels of agglomeration, the gains from formality are low so for the same level of talent it is more profitable to remain informal; This is especially true for unconstrained informal entrepreneurs who have no incentive to formalize. However, above a certain agglomeration threshold formality becomes increasingly attractive. This is because agglomeration is a scaling factor on entrepreneurial talent and hence entrepreneurs with
highly scaled talent would want to operate at a larger scale but would be constrained in the informal sector.

Figure 8: The Effects of Agglomeration on Transition Probabilities

While this model is similar to Nguimkeu (2015) in many ways, there are important differences that are worth pointing out. In particular, unlike in Nguimkeu (2015) where formal enterprises are not financially constrained, both formal and informal firms are constrained in this model while still allowing for the magnitude of the constraint to differ with formality status. This makes it much closer to reality and is expected to better fit the data. Likewise, allowing for the detection probability to vary with the size of the firm is more realistic than having a fixed detection probability across the board for all categories of informal firms as in Nguimkeu (2015), and this is also expected to improve the fit. Adding agglomeration externalities as we do in this model allows, as we initially stated, to account for spatial heterogeneity which is the main focus of this paper. While these additions do not qualitatively change the model predictions, they allow to get a better fit when we take the model to the data as well as allowing to derive new insights pertaining to spatial heterogeneity.

4 Structural Estimation

One important implication of the model developed above is the effect of sector and location heterogeneities with regard to determining the supply of entrepreneurs in a region (city) and industry. In this section, we present an econometric strategy that uses differences in location entrepreneurship rates and earnings in Cameroon to identify the structural parameters of the theoretical model presented in the previous section.

To arrive at specific functional forms, we follow the literature (e.g. Evans and Jovanovic 1989, Paulson et al 2006) and assume a lognormal distribution function for the entrepreneurial ability:
where \( x_{ij} \) are years of education, and sector and location fixed effects are included to capture productivity shocks shared by all firms in a sector and all firms in a location, respectively. In other words, we specify the covariance matrix of \( \varepsilon = (\varepsilon_{ij}) \) to allow for spatial correlation between locations \( j \) and across sectors \( i \).

\[
\ln \theta_{ij} = \delta_0 + \delta_1 x_{ij} + \varepsilon_{ij}, \quad \varepsilon_{ij} \sim N(0, \sigma_{ij}) \tag{9}
\]

Hence,

\[
E[\varepsilon_{it} \varepsilon_{js} | \cdot] = \sigma_{ij} \text{ if } t = s, \text{ and } 0 \text{ otherwise}
\]

Note that assuming log-normality for the talent distribution is not restrictive; other distributions such as the Pareto distribution can be used. However, as shown by Combes et al. (2012), the productivity distribution is usually better approximated by a log-normal than by a Pareto.

The model estimation also requires to determine both the agglomeration variable \( A_{ij} \), and the minimum visible size of informal firms \( \bar{l}_j \). For the agglomeration, we assume that \( A_{ij} = Z_{ij} \rho_{ij} R_{ij} \), where \( Z_{ij} \) is an agglomeration index, \( R_{ij} \) is infrastructure index, and \( \kappa_j, \rho_j \geq 0 \) are elasticity parameters to be estimated. This construction of these indices are describec in the appendix.

For the estimation procedure, we have a total sample of \( N \) households, \( i = 1 \ldots N \) and \( J \) cities, \( j = 1, \ldots, J \), with data on their initial wealth, \( z_{ij} \), years of schooling of the principal earner, \( x_{ij} \) and occupational status \( E_{ij} \) (with \( E_{ij} = 1 \) if the household runs a formal business and zero otherwise), \( F_{ij} \) (with \( F_{ij} = 1 \) if the household runs an informal business and zero otherwise), \( W_{ij} \) (with \( W_{ij} = 1 \) if the household does not run a business and zero otherwise). Institutional parameters such as the borrowing interest rate and the tax rate on formal firms income are set to the national averages taken from the 2010 Doing Business report for Cameroon.

### 4.1 Estimation Approach: GMM

Denote by \( \psi \) the parameter vector of the structural model to be estimated. These structural parameters are related to technology, credit and labor market access as well as the distributional parameters of entrepreneurial ability. This parameter vector is given
by \( \psi = (\alpha_j, \beta_j, \delta_j, \kappa_j, \sigma_{s,j})' \), where \( \alpha_j, \beta_j, \kappa_j, \sigma_{s,j} \) and \( \delta_j, j = 1, \ldots, M \), are the elasticities of capital, labor, agglomeration factors and fixed-effects for each location, and \( \sigma_{s,j}, s = 1, \ldots, S \), is the spatial correlation between locations and sectors.

We adopt a generalized method of moments (GMM) approach to estimate the parameters of our model that treats factor inputs as a latent variables. The estimation strategy relies on the idea that the structural model should replicate the observed joint empirical distribution function of location choice, earnings, informality status, and employment conditional on occupational choice. The GMM method therefore allows us to match a list of entrepreneurship probabilities and income moments predicted by the model to their data counterparts, for the observed personal characteristics \( x_{ij} \) and agglomeration index \( Z_{ij} \) and infrastructure index \( R_{ij} \).

In our baseline specification we match a list of moments corresponding to the probabilities (proportions) of formal business ownership and informal business ownership in different cities and subsamples defined based on the terciles of years of schooling and initial wealth. These set of moments includes static and conditional moments on occupational outcomes, location, formality status, education, earnings, and labour market outcomes. Tables 7 lists all moments which are used to identify the model. The parameters in the first column are identified by one or several corresponding moments given in the second column, and the number of moments is given in column 3.

If we combine all moments used in the estimation procedure into one vector \( m_N \) and denote by \( m(\psi) \) their theoretical counterparts in the model, then the orthogonality conditions are given by

\[
\dot{h}_N(\psi) = m(\psi) - m_N
\]

Following Hansen (1982), \( \psi \) can be estimated using the following moments estimator:

\[
\hat{\psi} = \arg \min_{\psi} \dot{h}_N(\psi)' \hat{\Omega} \dot{h}_N(\psi)
\]

for some positive semi definite matrix \( \hat{\Omega} \) which converges in probability to \( \Omega \). The estimator \( \hat{\psi} \) is consistent and

\[
\sqrt{N}(\hat{\psi} - \psi) \xrightarrow{d} N \left(0, (D'\Omega D)^{-1}D'\Omega V\Omega D(D'\Omega D)^{-1}\right),
\]

where \( D = \mathbb{E}[\partial h(\psi)/\partial \psi] \), and \( V \) is the asymptotic covariance matrix of the vector of sample moments, and denotes the parameter of the data generating process. The asymptotically efficient estimator is obtained by setting \( \hat{\Omega} = \hat{V}^{-1} \), so that

\[
\sqrt{N}(\hat{\psi} - \psi) \xrightarrow{d} N \left(0, (D'\hat{V}^{-1}D)^{-1}\right).
\]
Moreover, standard $J_N$-statistics can be used for hypothesis and specification tests. The statistic $J_N$ has an asymptotic $\chi^2$ distribution with degrees of freedom equal to the number of moments minus the number of parameters.

Table 7: Model Moments and Sample Analogs

| Moment | Model $M_f(X_i, \psi)$ | Sample $m_i$ |
|--------|------------------------|--------------|
| 1. Prob. formal | $\frac{1}{n_j} \sum_{i=1}^{n_j} \Pr[F_{ij} = 1|\cdot]$ | $\frac{1}{n_j} \sum_{i=1}^{n_j} F_{ij}$ |
| 2. Prob. formal, $x < x_m$ | $\frac{\sum_{i=1}^{n_j} \Pr[F_{ij} = 1|\cdot][I_{ij} = 1|x < z_m]}{\sum_{i=1}^{n_j} [I_{ij} = 1|x < z_m]}$ | $\frac{\sum_{i=1}^{n_j} F_{ij} [I_{ij} = 1|x < z_m]}{\sum_{i=1}^{n_j} I_{ij}}$ |
| 3. Prob. informal | $\frac{1}{n_j} \sum_{i=1}^{n_j} \Pr[I_{ij} = 1|\cdot]$ | $\frac{1}{n_j} \sum_{i=1}^{n_j} I_{ij}$ |
| 4. Prob. informal, $x < x_m$ | $\frac{\sum_{i=1}^{n_j} \Pr[I_{ij} = 1|\cdot][I_{ij} = 1|x < x_m]}{\sum_{i=1}^{n_j} [I_{ij} = 1|x < x_m]}$ | $\frac{\sum_{i=1}^{n_j} I_{ij} [I_{ij} = 1|x < x_m]}{\sum_{i=1}^{n_j} I_{ij}}$ |
| 5. Income formal | $\frac{\sum_{i=1}^{n_j} \mathbb{E}[y_{ij}|F_{ij} = 1] \Pr[F_{ij} = 1|\cdot]}{\sum_{i=1}^{n_j} \Pr[F_{ij} = 1|\cdot]}$ | $\frac{\sum_{i=1}^{n_j} y_{ij} F_{ij}}{\sum_{i=1}^{n_j} F_{ij}}$ |
| 6. Income formal, $x < x_m$ | $\frac{\sum_{i=1}^{n_j} \mathbb{E}[y_{ij}|F_{ij} = 1] \Pr[F_{ij} = 1|\cdot][I_{ij} = 1|x < x_m]}{\sum_{i=1}^{n_j} \Pr[F_{ij} = 1|\cdot][I_{ij} = 1|x < x_m]}$ | $\frac{\sum_{i=1}^{n_j} y_{ij} I_{ij} F_{ij}}{\sum_{i=1}^{n_j} I_{ij} F_{ij}}$ |
| 7. Income informal | $\frac{\sum_{i=1}^{n_j} \mathbb{E}[y_{ij}|I_{ij} = 1] \Pr[I_{ij} = 1|\cdot]}{\sum_{i=1}^{n_j} \Pr[I_{ij} = 1|\cdot]}$ | $\frac{\sum_{i=1}^{n_j} y_{ij} I_{ij}}{\sum_{i=1}^{n_j} I_{ij}}$ |
| 8. Income informal, $x < x_m$ | $\frac{\sum_{i=1}^{n_j} \mathbb{E}[y_{ij}|I_{ij} = 1] \Pr[I_{ij} = 1|\cdot][I_{ij} = 1|x < x_m]}{\sum_{i=1}^{n_j} \Pr[I_{ij} = 1|\cdot][I_{ij} = 1|x < x_m]}$ | $\frac{\sum_{i=1}^{n_j} y_{ij} I_{ij} I_{ij}}{\sum_{i=1}^{n_j} I_{ij} I_{ij}}$ |
| 9. Income non-entrep | $\frac{\sum_{i=1}^{n_j} \mathbb{E}[y_{ij}|W_{ij} = 1] \Pr[W_{ij} = 1|\cdot]}{\sum_{i=1}^{n_j} \Pr[W_{ij} = 1|\cdot]}$ | $\frac{\sum_{i=1}^{n_j} y_{ij} W_{ij}}{\sum_{i=1}^{n_j} W_{ij}}$ |

Notes: The table reports averages of the described quantities. The derivation of the predicted model moments are given in Nguimeu (2015). $x_m$ is the median years of schooling.

4.2 Estimation results

The results of the GMM parameter estimates are reported in Tables 8 and 9. Table 8 presents the structural estimates according to sectors. These estimates are obtained for Agriculture, Manufacturing, Commerce and Services. Pairwise residual correlations among these industries are not reported in the Table but are available from the authors. The results show that average expected monthly wages are not significantly different across sector and are roughly the same as the national average of 104,000 FCFA. The estimates of $\delta_1$ show that education has a differential effect on productivity, depending on the sector of activity in which one operates. At the national level, a
10% increase in years of education is associated with a 5% increase in Yaoundé entrepreneurial productivity on average, whereas this effect is larger in Agriculture and lower in Manufacturing. The elasticities of capital and labor, $\alpha$ and $\beta$, are 0.23 and 0.49 respectively at the national level. It means that a 10 percent increase in the capital devoted to a business leads to a 2.3% increase in output, while a 10 percent increase in hired labor increases output by 4.9%, respectively. However, when we look at the sectoral level, we see that the elasticity of capital is higher in Agriculture and lower in Manufacturing. Since returns to capital are usually high in the informal sector (e.g., Udry and Anagol 2006, De Mel et al. 2008) and low in the formal sector in Africa (e.g., Alby, Auriol and Nguimkeu 2015), this results may reflect the fact that compared to the manufacturing sector, the agricultural sector is highly informal (roughly 95% in 2010) so that the capital elasticity estimate is driven by firms from the informal sector. The borrowing constraints is measured by $\lambda$ and is estimated at 2.27 for the whole sample, 2.69 for Agriculture, 2.86 for Manufacturing, 1.61 for Commerce and 1.92 for Services. This implies that entrepreneurs from the Commerce and Services industries are less constrained than those from the Agriculture and Manufacturing industries. This is presumably because the capital needed to start a business in the former is usually lower than in the later, especially in the informal sector.

Table 9 presents the structural estimates according to the main cities of Cameroon. These cities include Douala, Yaoundé, Bamenda, Kribi, Bafoussam and Garoua. Unlike above stratification where wages are uniform across industries, in this case wages significantly vary by city. The average expected monthly wages are higher in Douala and Yaoundé, estimated at 117,300 FCFA and 108,300 FCFA respectively. This is consistent with the fact that these two metropolitan cities gather more than 70% of the countries economic activities so that workers outside options are higher. The elasticity of capital is higher in Douala and Yaoundé and lower in Bafoussam and Garoua. In contrast, the elasticity of labor is the highest in Kribi, reflecting a low availability of workforce in this city as given by the summary statistics.

Except for Bafoussam, borrowing constraints, captured by the parameter $\lambda$ are also lower in these two cities, compared to the rest of the country (since their estimates of $\lambda$ are 2.37 and 2.51, which are relatively lower). The only exception is Bafoussam where $\lambda$ is estimated at 1.14 and is the lowest in the country. This is presumably because many entrepreneurs in this part of the country rely more on informal finance such as kinship networks and rotating saving and credit associations (rosca) to finance their business (see, Baland et al 2017), rather than collateral-based lending as in large cities. At the city level, we also provide structural estimates for output elasticities with respect to agglomeration (captured by population density index) and infrastructure (captured by paved road density index), $\kappa$ and $\psi$, respectively. We estimated $\kappa$ at 0.163
for the whole sample, implying that a 10 percent increase in the population density increases firm productivity by 1.63%. For Douala and Yaoundé, these values are 0.065 and 0.91 which are lower than the national average for the obvious reason that the population density in these cities is already quite high. In other less populated cities such as Kribi and Bamenda, this elasticity is higher, estimated at 0.28 and 0.26 respectively. We estimated $\varrho$ at 0.27 for the whole sample, implying that a 10 percent increase in paved road density increases firm productivity by 2.7%. This elasticity is higher in Garoua and Bamenda compared to the rest of the cities.

Before we perform counterfactual simulations, it is important to assess how well the model fits the data. We do this by comparing the model predicted moments and their sample analogs. Table 10 reports the model fit for the nine chosen moments, of which five were targeted in the gmm estimation as defined in Table 7 above, and four were not targeted in the gmm estimation. Apart from the proportion of formal entrepreneurs which is matched within a 20% discrepancy, all the remaining moments are matched within 10% or lower of their counterpart values in the data. A good fit in these moments can be seen as a validation of the model with data that have or have not been used directly in the estimation.

### Table 8: Structural Estimates - Industries

| Parameter | Name   | Whole   | Agri    | Manuf   | Com    | Serv   |
|-----------|--------|---------|---------|---------|--------|--------|
| Average wage | $w$     | 104.2   | 103.64  | 104.63  | 104.7  | 104.6  |
|            |        | (15.9)  | (16.1)  | (15.4)  | (15.8) | (15.7) |
| Ability - educ | $\delta_1$ | 0.51    | 0.62    | 0.35    | 0.52   | 0.45   |
|            |        | (0.079) | (0.118) | (0.042) | (0.076)| (0.17) |
| Capital share | $\alpha$ | 0.230   | 0.321   | 0.186   | 0.292  | 0.213  |
|            |        | (0.057) | (0.023) | (0.073) | (0.019)| (0.021) |
| Labor share | $\beta$ | 0.499   | 0.519   | 0.617   | 0.696  | 0.561  |
|            |        | (0.059) | (0.023) | (0.074) | (0.019)| (0.024) |
| Borrow. constr. | $\lambda$ | 2.27    | 2.69    | 2.86    | 1.61   | 1.92   |
|            |        | (0.314) | (0.320) | (1.164) | (0.569)| (0.245) |
| Number of Obs. |        | 10709   | 1150    | 2454    | 2268   | 5911   |

Asymptotic standard errors in parenthesis
Parameters $\kappa$ and $\varrho$ are not included because geographical factors may not vary across industries.
Table 9: Structural Estimates - Cities

| Parameter      | Name | All     | Dla     | Yde     | Bda     | Krb     | Baf     | Gar     |
|----------------|------|---------|---------|---------|---------|---------|---------|---------|
| Avg wage       | $w$  | 104.2   | 117.3   | 108.3   | 80.97   | 94.96   | 74.25   | 77.55   |
|                |      | (15.9)  | (12.1)  | (11.2)  | (9.27)  | (9.35)  | (7.12)  | (7.36)  |
| Ability - educ | $\delta_1$ | 0.51    | 0.48    | 0.49    | 0.52    | 0.56    | 0.41    | 0.63    |
|                |      | (0.04)  | (0.02)  | (0.03)  | (0.01)  | (0.01)  | (0.13)  | (0.03)  |
| Capital share  | $\alpha$ | 0.23    | 0.28    | 0.27    | 0.20    | 0.21    | 0.18    | 0.17    |
|                |      | (0.05)  | (0.02)  | (0.07)  | (0.01)  | (0.03)  | (0.08)  | (0.06)  |
| Labor share    | $\beta$ | 0.499   | 0.523   | 0.531   | 0.471   | 0.601   | 0.432   | 0.425   |
|                |      | (0.06)  | (0.03)  | (0.07)  | (0.02)  | (0.01)  | (0.08)  | (0.06)  |
| Borrow. constr.| $\lambda$ | 2.27    | 2.37    | 2.51    | 2.83    | 2.79    | 1.14    | 2.60    |
|                |      | (0.31)  | (0.51)  | (0.45)  | (0.79)  | (1.33)  | (0.32)  | (0.98)  |
| Elast - aggro  | $\kappa$ | 0.163   | 0.065   | 0.091   | 0.257   | 0.275   | 0.211   | 0.195   |
|                |      | (0.031) | (0.055) | (0.044) | (0.069) | (0.032) | (0.027) | (0.05)  |
| Elast - infrast| $\rho$  | 0.266   | 0.201   | 0.214   | 0.292   | 0.197   | 0.258   | 0.299   |
|                |      | (0.107) | (0.093) | (0.094) | (0.096) | (0.085) | (0.102) | (0.121) |
| Number of Obs. |      | 10709   | 4334    | 3647    | 653     | 87      | 850     | 1138    |

Asymptotic standard errors in parenthesis.

5 Policy Simulations

The informal sector can generate negative externalities (Loayza, 2007). For example, by using and congesting public infrastructure without paying a contribution, and by taking formal business market shares. The latter externality is however theoretically ambiguous and there is little empirical evidence that informal firms unfairly take market share from law-abiding formal firms. However, while informality helps firms avoid certain costs, it may also preclude access to certain opportunities, including greater access to cheaper credit, increased opportunities to engage with larger firms and government contracts, reduced harassment by police, and access to broader training and support programmes. In this section, we perform a set of counterfactual simulations to evaluate the impact of public policies on entrepreneurship and informality across cities and industries in Cameroon. The estimated model represents the current state and the purpose is to evaluate how this equilibrium would change as a result of possible changes in some quantities in the model. For each policy change, an equilibrium wage rate is computed assuming the same distribution of wealth and the same exoge-
Table 10: Matched and Unmatched Moments at the GMM estimates

| Description                                      | Model | Data | Discrepancy (%) |
|--------------------------------------------------|-------|------|-----------------|
| 1. Prob. of formal entrep (%)                    | 2.78  | 2.31 | 20.1            |
| 2. Prob. of formal entrep, $z < z_m$ (%)         | 0.65  | 0.55 | 18.2            |
| 3. Prob. of informal entrep (%)                  | 18.5  | 17.71| 10.2            |
| 4. Prob. of informal entrep, $z < z_m$ (%)       | 12.8  | 11.70| 9.40            |
| 5. Income of formal entrep                       | 179.8 | 189.12| -10.4          |
| 6. Income of formal entrep, $z < z_m$            | 116.8 | 106.13| 10.1           |
| 7. Income of informal entrep                     | 80.7  | 85.60| -5.72           |
| 8. Income of informal entrep, $z < z_m$          | 59.8  | 66.93| -10.6           |
| 9. Income of non-entrep                          | 104.2 | 104.2| 0.00            |

Notes. The table reports averages. Income is in thousands of FCFA (§1 ~ 500 FCFA) $z_m$ is the median initial wealth in the data.

nously fixed interest rate, and agents make their choices based on these factors. Our focus is on the impact of factors such as infrastructure and borrowing constraints. We provide graphical representations that focus on how various policy options lead to differences within cities or industries.

5.1 Impact of Infrastructure: Paved road density

This counterfactual experiment consists in examining how the equilibrium would change should there be a change in the country infrastructure level. Our measure of infrastructure index is the paved road density. This policy could be a project of road construction, a bridge construction or the paving of an existing road. We assume a change in the infrastructure index $R_{ij}$ defined by

$$R'_{ij} = R_{ij} + \eta$$

where $\eta$ represents the increment in infrastructure quality implied by the policy.

Impact on Industries

The impact of this infrastructure policy with the Cameroon data within the context of the model is quantified for a range of relative quantity increments $\eta$, starting from $\eta = 0$, the current state as produced by the structural estimates, to $\eta = 5 \bar{R}$, where $\bar{R}$ is
Figure 9: Impact of infrastructure on entrepreneurship and informality: industries

(a) Agriculture  
(b) Manufacturing  
(c) Commerce  
(d) Services

Notes: In these figures the horizontal axis is the increment in infrastructure measured in proportions of the average road density index. The vertical axis are transition probabilities. The average infrastructure index in the current equilibrium taken as initial state.

Figure 9 shows the variation in the fraction of new enterprise creation and the fraction of informal firms that become formal across the four main industries in the country. As the density of paved road increases, both the fractions of new enterprises and the fraction of formal entrepreneurs increase. The increase in paved road density in the country has a similar high effect on entrepreneurial choice in both the Agricultural and Commerce sectors, but a lower effect in Manufacturing and Services. A 15% increase in paved roads would generate the creation of twice as much new firms in the former, but very insignificant fractions in the latter. In contrast, relatively higher fractions of informal firms would formalize in the Manufacturing and Services sectors.
Figure 10: Impact of infrastructure on output and productivity: industries

Note: In these figures the horizontal axis is the increment in infrastructure measured in proportions of the average road density index. The left vertical axis is output in thousands of CFA, and the right vertical axis is productivity in thousands of CFA.

compared to the Agriculture and Commerce sectors as a result of increased infrastructure. Figure 10 depicts variations in enterprise output and productivity across the four main industries resulting from an increase in the density of paved roads. In general, there are stark differences in productivities between industries. For increased quantities of paved roads, the increase in output seems to follow a convex path (left vertical axis). A 10% or lower increase in infrastructure only has a modest effect on output. More significant increase are obtained for increases of 20% and above. The most striking result is from the Commerce sector. The change in output is insignificant for increment lower than 50%, but drastically increases after this threshold. In contrast, the path for industrial productivity is seemingly concave and increases at a
relatively constant pace across the four sectors (right vertical axis). On average, a 20% increase in infrastructure more than double the initial productivity in all four sectors of activities.

**Impact on Cities**

We now look at the impact of infrastructure on cities. Figure 11 shows the variation in the fraction of new enterprise creation and on the fraction of informal firms that become formal in six major cities of Cameroon. While Douala and Yaoundé present similar patterns of formalization for increasing quantities of paved road density, the patterns of creation of new firms are different between these two metropolitan cities. While a small increase in paved road density has an immediate impact on entrepreneurship in Yaoundé, it is only when this density increases above 15% that entrepreneurship kicks off in Douala. This may appear quite surprising because paved road density is initially lower in Douala than in Yaoundé (as shown by the summary statistics given in Table 5). Moreover, above an increment threshold of about 30%, additional paved roads do not seem to generate any additional new firm creation in Douala. In contrast, the relationship between paved road density and firm creation seems to be steadily increasing in Yaoundé. This suggests that unlike in Yaoundé, paved roads are not necessarily a crucial ingredient for business creation in Douala. The reason may be found in the difference in the nature of landscape between these two cities. While Yaoundé has a hilly and mountainous landscape which may be a serious obstacle to entrepreneurship (e.g., for commerce and transportation businesses) if roads are unpaved, Douala, on the other hand, has a low-lying plain landscape, making it more practicable even when unpaved.

Another striking pattern of business creation depicted in Figure 11 is the counterfactual impact for the city of Kribi. Perhaps for similar reasons as in Douala (that is, low-lying plain landscape), the relationship between paved road density and business creation appear to be piecewise increasing (or piecewise horizontal). The remaining cities, that is, Bamenda, Bafoussam and Garoua, exhibit similar and regular upward sloping patterns of formalization for increasing availability of paved roads. Likewise, the patterns of business creation are also upward sloping in these three cities, although these patterns seem to be steeper in Bamenda and Garoua compared to Bafoussam.

That the patterns of formalization are, in general, upward sloping across all cities for increasing quantities of paved roads suggests that the availability of paved roads encourages informal firms to formalize. This could be either because paved roads make it easier to do business by increasing productivity (as we discuss below) so that
informal entrepreneurs quickly grow and graduate to the formal sector,\(^\text{11}\), or because paved roads provide more access to tax inspectors to otherwise remote areas, which discourages informality as informal entrepreneurs become easily detectable.

Figure 12 depicts variations in enterprise output (left vertical axis) and productivity (right vertical axis) across the six major Cameroon cities resulting from an increase in the density of paved roads. The patterns of productivity are all similar across the cities: for increasing availability of paved roads, firms are more productive. This suggests that paved roads make it easier to do business in general. As for the industry, there are stark differences in productivity effects between cities. This important role of infrastructures on firm performance is consistent with earlier work on Cameroon in the context retail firms (see Nguimkeu 2016). Likewise, output appear to be similarly increasing across the cities for increasing availability of paved roads. Interestingly, these patterns seem to be pairwise comparable among cities. For example, Yaoundé is quite similar to Kribi, whereas Bafoussam is similar to Garoua, and Douala is similar to Bamenda. In addition, these patterns are all convex, meaning that the impact of paved roads on enterprise output increases at an increasing rate in the major Cameroon cities considered.

In summary, a policy that improves the level of infrastructure has heterogeneous impacts across industries and cities. Depending on their comparative advantage in terms of having more or less hospitable landscape, economic activities in some cities would be more or less sensitive to road pavement than in others. Likewise, some industries for which road infrastructure are key ingredient for their well-functioning would benefit more from a road improvement policy than others. We now turn to the counterfactual impact of a policy aiming at relaxing borrowing constraints for potential entrepreneurs.

### 5.2 Impact of Borrowing constraint: Interest rate

We analyze a credit access policy by evaluating the counterfactual impact of a reduction of the borrowing interest rate.\(^\text{12}\) We assume an exogenous change in interest rate defined by

\[
r' = r - b, \quad 0 \leq b < r
\]

\(^{11}\)Indeed, according to ILO (2002), the informal sector is an incubator for business potential and a stepping stone for accessibility and graduation to the informal sector.

\(^{12}\)Nguimkeu (2014) simulated the impact of an exogenous policy change in collateral requirement through microfinance in Cameroon and found a positive impact on entrepreneurship and earnings. Here, we take another route by looking at an exogenous policy change in the price of capital.
Figure 11: Impact of infrastructure on entrepreneurship and informality: cities

Notes: In these graphs, the horizontal axis is the increment in infrastructure in proportions of the average road density index. The vertical axis are transition probabilities.
Figure 12: Impact of infrastructure on output and productivity: cities

(a) Douala (b) Yaoundé

(c) Bamenda (d) Kribi

(e) Bafoussam (f) Garoua

Note: The horizontal axis is the increment in infrastructure in proportions of the average road density index. The left vertical axis is output and the right vertical axis is productivity.
where $b$ represents the decrement in the gross interest rate implied by the policy.

The impact of this policy with the Cameroon data within the context of the model is depicted in Figure 13. The impact is quantified for a range of interest rate decrements, starting from 0, the current state as produced by the structural estimates, to 20 percentage points (that is, increasing values on the horizontal axis in the graphs represent corresponding reductions in interest rates). The effects of these decrements on the creation of new firms, formalization of informal firms and enterprise profits are evaluated across the different industries and pictured in the Figure 13 (The left vertical axis are transition probabilities and the right vertical axis are profits). The results show that reducing the current interest rate has no effect on business creation in the Manufacturing and Services sectors, and has a moderate effect on commerce when the reduction is above 15 percentage point. In the Agriculture sector, however, the impact on business creation is significant although moderate for low values of the decrements and larger when these decrements increase by 15 percentage point or above.

On the other hand, reducing the interest rate has a much larger and significant impact on the formalization of informal firms across sectors. The patterns of formalization in the Agriculture and Commerce sectors are similar and upward sloping in a smooth and steady way. In the Manufacturing and Service sectors, formalization is inelastic to interest rate for low levels of decrements. It is only for interest rates reductions under a certain threshold (7 percentage point for Manufacturing and 12 percentage points for Services) that entrepreneurs in these sectors find it beneficial to formalize.

A reduction of interest rates has, as expected, a positive impact on firm profits in general. The Manufacturing and Service sectors exhibit similar increasing and smooth patterns of profit when the interest rates go down. The pattern of profits in Agriculture is also increasing, although the impact is quite moderate for decrements lower than 10 percentage points. The most striking case is in the Commerce sector. Reductions of less than 18 percentage points have no impact on firm profits. However, for interest rate reductions above this threshold, firm profits increase dramatically.

5.3 Impact of Taxation

In this counterfactual experiment, we investigate the impact of tax reforms on entrepreneurship and informality across cities. Formally, we assume reductions in tax rates of magnitude $d$, such that

$$
\tau' = \tau - d, \quad 0 \leq d < \tau
$$
Figure 13: Impact of credit access on entrepreneurship, informality and profits

Notes: The horizontal axis is the decrement in borrowing interest rates. The left vertical axis are fractions of entrepreneurs, and the right vertical axis are firm profits.

The reduction in tax rate ranges from 0, the current state, to about 0.4, representing an 85% tax reduction. Figure 14 depicts the effects of variations in tax rates on the creation of new firms and the formalization of informal firms across cities.

Tax reduction has different incentive effects on the creation of new firms and on the formalization of informal firms across cities. Apart from Kribi, a continuous reduction in taxes would have continuous and substantial effects on the formalization on informal firms. In Kribi, the incentive to formalize does not take up until the tax reduction is at least 7% less the current rate. We also notice that tax reduction in the ranges of 15% to about 24%, yield the same effect on formalization. This behavior is
Figure 14: Impact of tax reduction on entrepreneurship and informality: cities

Note: The horizontal axis is the decrement in the tax rate on formal firms profit. The left vertical axis are transition probabilities.
consistent with those found earlier where counterfactual policy only have piecewise effects.

On the other hand, a reduction of taxes only has minor effect on the creation of new firms. For example, in all the major cities, the incentive to create new firms as a reaction to a decrease in taxes does not take up unless the tax decrement is at least 7% of firm profits (Bamenda, Yaoundé, and Bafoussam). In other cities such as Douala, Kribi and Garoua, it is quite high and takes a 24%, 13% and 14% decrements in taxes, respectively, to see any incentive to start a formal firm in those cities. This might be due to fact that wages are relatively higher in these cities, or the cost of registration (such as the number of days for the registration procedure and the foregone income incurred during the days spent in the registration office for these procedures) are relatively higher.

Figure 15 depicts the effects of variations in tax rates on firm profits and tax revenues across cities. Note that the tax revenues reported here include both government revenues in terms of taxes income collected on firm profits as well as registration fees collected from firms in the formalization procedures. As one would obviously expect, a decrease in taxes increases the aggregate profit of firms, since it increases the profit of formal firms and has no effect on informal firms profits. The more interesting pattern, however, is the impact of taxes on tax revenues across cities. The counterfactual results show that taxation rates have the well-known Laffer’s inverted-U pattern on government revenues; there exists an optimal tax rate for the government that generates a maximum tax revenue gains. Interestingly, this optimal tax rate is not the same across cities. In Yaoundé and Bafoussam, the optimal tax decrement is around 15%, implying an optimal tax rate of about 33%, lower than the current rate of 48%. In Bamenda and Garoua, the optimal tax decrement is about 8%, implying an optimal government tax rate of 40%. The two extreme cases are Douala and Kribi. In Douala, the optimal tax decrement is the highest in the country, evaluated at 40%, implying an optimal government tax rate of about 8%. In Kribi, the optimal tax decrement is the lowest in the country, evaluated at 0%, implying that the current tax rate is optimal for the government.
Figure 15: Impact of tax reduction on firm profit and tax revenues: cities

Note: The horizontal axis is the decrement in the tax rate on formal firms profit. The left vertical axis are firm profits, and the right vertical axis are tax revenues.
6 Conclusion

We use data from the Cameroon National Surveys on Employment and Informal Sector (eesi), and the General Enterprises Census to empirically investigate the characteristics of the informal sector in Cameroon and its most salient features. Cameroon’s labor market is characterized by a large informal sector that employs around 90% of the occupied workforce, is largely concentrated in agricultural, and non-tradable services (e.g. retail trade, transportation, hotels or restaurants), and accounts for more than 30% of the GDP. Despite this substantial contribution to the GDP and the labor market the informal sector is made of micro-units, often with one worker, very low-paid jobs and low labor productivity. We document that economic activities are strongly concentrated in Yaoundé and Douala and 60% of Cameroon’s industries are geographically concentrated. Informal firms are on average less concentrated than formal firms, and cities are less specialized. Our results show that only 16 percent of these firms were highly specialized in 2009, and the most specialized cities are natural-resources based cities.

To further investigate the relationship that exists between location, agglomeration, access to credit, informality and productivity across cities and industries, we develop an occupational choice model where individuals decide between formal entrepreneurship, informal entrepreneurship and non-entrepreneurial work. Our model predicts that while entrepreneurial profits increase with agglomeration effects, agglomeration has a non-monotonic (U-shaped) effect on formalization of informal firms. Using data from eesi, we provide structural estimates of the model test some of its salient implications using the generalized method of moments, disaggregated along the main cities and industries of the country. We found that certain sectors such as Manufacturing are highly constrained in capital, compared to other sectors such as Commerce. We also found that the return to capital is not uniform across sectors since some of them are more populated by informal firms compared to others.

The estimated model is used in counterfactual policy experiments to simulate the effect of infrastructure on entrepreneurship, informality and productivity. We found that increased availability of paved road, better access to finance, and tax reforms can have a substantial impact on the economy in terms of entrepreneurship, output, formalization of informal firms and productivity. This is particularly true in sectors such as Agriculture and Services which are likely to be benefited the most from such policy (e.g., improving roads has an obvious positive impact on the transportation of goods and production factors). As countries and regions strive to remain competitive in the face of globalization, local and national governments seek increasingly to support competitive cities and industries. Three key determinants of cities’ wealth are related
to human capital, location, and size. These advantages are the outcome of long term investments in human capital, infrastructure and institutions. Our findings highlight three implications for policy makers interested in supporting formalization, business creation and productivity. These policy instruments include: (i) increase infrastructure provision by investing in local infrastructure; (ii) increase access to finance, (iii) local tax reforms, and (iv) increase the public-private dialogue at local level to better understand the other heterogenous constraints that firms face.

References

[1] Alby, P., Auriol, E., and Nguimkeu, P., (2014). “Does Social Pressure Hinder Entrepreneurship in Africa? The Forced Mutual Help Hypothesis”, Working Paper.

[2] Amaral, P. and E. Quentin, (2006). “A Competitive Model Of The Informal Sector”, Journal of Monetary Economics, 53: 1541-1553.

[3] Ana Maria Oviedo., Mark R. Thomas., Kamer Karakurum-Ozdemir, (2009). “Economic Informality : Causes, Costs, and Policies - A Literature Survey”, World Bank Publications, The World Bank, 5917.

[4] Armington, C. and Acs, Z., (2002). “The determinants of regional variation in new firm formation”. Regional Studies, 36: 33-45.

[5] Baland RM, Yseboodt M, Rivoldini A, Van Hoolst T., (2017). “Obliquity of Mercury: Influence of the precession of the pericenter and of tides”, Icarus. 291, 136-159. doi: 10.1016/j.icarus.2017.03.020.

[6] Behrens, K., Duranton, G., Robert-Nicoud, F., (2014). “Productive Cities: Sorting, Selection, and Agglomeration”, Journal of Political Economy, 122(3): 507-553.

[7] Benhassine, N, D McKenzie, M Santini and V. Pouliquen, (2018). “Does Inducing Informal Firms to Formalize Make Sense? Experimental Evidence from Benin?”, Journal of Public Economics, 157: 1-14.

[8] Briant, A., P-P. Combes, and M. Lafourcade, (2010). “Dots to Boxes: Do the Size and Shape of Spatial Units Jeopardize Economic Geography Estimations?”, Journal of Urban Economics, 67(3): 287-302.

[9] Bruhn, M., (2011). “License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico”, Review of Economics and Statistics, 93(1): 382-386.

[10] Ciccone, A., and R. E. Hall, (1996). Productivity and the density of economic activity. American Economic Review, 86(1): 54-70.
[11] Combes, P.P., Duranton, G., Gobillon, L., Roux, S., (2010). “Estimating Agglomeration Economies with History, Geology, and Worker Effects”, Chapter in NBER book Agglomeration Economics (2010), Edward L. Glaeser, editor, 15 - 66.

[12] Combes, P.P., Duranton, G., Gobillon, L., Roux, S., (2008). “Spatial Wage Disparities: Sorting Matters!”, Journal of Urban Economic, 63(2): 723-742.

[13] Combes, P.P., Duranton, G., Gobillon, L., Roux, S., (2012). “Sorting and local wage and skill distributions in France”, Regional Science and Urban Economics, 42(6): 913-930.

[14] da Mata, U., Deichmann, U., Henderson, J.V., Lall, S.V., H.G. Wang, (2005). “Determinants of city growth in Brazil”, Journal of Urban Economics, 62(2): 252-272.

[15] De Mel, S., McKenzie, D., and, C., (2008). “Returns to Capital in micro-enterprises: Evidence from a Field Experiment”, Quarterly Journal of Economics, 123(4): 1329-1372.

[16] de Paula, A. and J. Scheinkman, (2011) “The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil”, Review of Income and Wealth, Series 57, Special Issue.

[17] Djankov, S., R. LaPorta, F. Lopez-De Silanes, and A. Shleifer, (2002). “The Regulation of Entry”, Quarterly Journal of Economics, 117: 1-37.

[18] Duranton, G., Puega, D., (2004). “Micro-foundations of urban agglomeration economies”, In Handbook of Regional and Urban Economics, vol. 4, edited by J. Vernon Henderson and Jacques-Francois Thisse, 2063-2117. Amsterdam: North-Holland.

[19] Ellison, G. and E. Glaeser, (1997). “Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach,’ Journal of Political Economy, 105(5): 889-927.

[20] Ellison, Glenn D., and Edward L. Glaeser, (1999). “The Geographic Concentration of Industry: Does Natural Advantage Explain Agglomeration?? American Economic Review, Papers and Proc, 89(2): 311-316.

[21] Ellison, G., Glaeser, E. and W. Kerr, (2010). What Causes Industry Agglomeration? Evidence from coagglomeration patterns, American Economic Review, 100(3): 1195-1213.

[22] Evans, David S. and Boyan Jovanovic, (1989). “An Estimated Model of Entrepreneurial Choice under Liquidity Constraints.” Journal of Political Economy, 97: 808-827.
[23] Fortin, B., N. Marceau, and L. Savard, (1997). “Taxation, Wage Controls and the Informal Sector”, *Journal of Public Economics*, 66: 293-312.

[24] Fritsch, Michael, and David J. Storey., (2014) “Entrepreneurship in a Regional Context: Historical Roots, Recent Developments and Future Challenges’, *Regional Studies*, 48(6): 939-954.

[25] Fritsch, M., AND Mueller, P., (2004). “The effects of new business formation on regional development over time”. *Regional Studies*, 38: 961-975.

[26] Glaeser, E., Kerr, W., Ponzetto, G., (2010), “Clusters of entrepreneurship” *Journal of Urban Economics*, 67(1): 150-168

[27] Henderson, V., Lee, T., Y.J., Lee, (2001). “Scale Externalities in Korea”, *Journal of Urban Economics* 49(3): 479-504

[28] International Labour Organization, (2002). “Decent Work and the Informal Economy”, International Labor Conference, Geneva.

[29] Institut National de la Statistique, (2005). *Enquête sur L’emploi et le secteur informel au Cameroun en 2005*, Rapport principal, Yaoundé, Cameroon.

[30] Keeble, David, and Sheila Walker, (1994). “New Firms, Small Firms and Dead Firms: Spatial Patterns and Determinants in the United Kingdom’. *Journal of Regional Studies*, 28(4): 411-427.

[31] Kerr, William R. and Scott D. Kominers, (2015). “Agglomerative forces and cluster shapes.” *Review of Economics and Statistics*, 97(4): 877–899.

[32] Krugman P., (1991). “Increasing returns and economic geography”, *Journal of Political Economy*, 99:483–499.

[33] Krugman P., (1995). “Development, Geography and Economic Theory”, *MIT Press, Cambridge, MA*.

[34] Krugman P., (2009). “The new economic geography, Now Middle-aged”, *Regional Studies*, 45:1–7.

[35] Laporta, R. and A. Schleifer, (2014). “Informality and Development”, *Journal of Economic Perspectives*, 28(3):109-126.

[36] Loayza, Norman and Jamele Rigolini, (2006). “Informality Trends and Cycles”, *World Bank Policy Research Working Paper*, 4078.

[37] Lucas Jr., R.E., (1978). “On the size distribution of business firms.” *Bell Journal of Economics*, 9: 508-523.
[38] Megha, Mukim, (2015). “Coagglomeration of formal and informal industry: evidence from India”, Journal of Economic Geography, 15(2): 329-351.

[39] Nguimkeu, P., (2014). “A structural econometric analysis of the informal sector heterogeneity”, Journal of Development Economics, 107:175-191.

[40] Nguimkeu, P., (2015). “An Estimated Model of Informality with Constrained Entrepreneurship”, Working Paper.

[41] Nguimkeu, P., (2016). “Some Effects of Business Environment on Retail Firms”, Applied Economics 48(18): 1647-1654

[42] Ordoñez, J.C., (2014). “Tax collection, the informal sector, and productivity”, Review of Economic Dynamics, 17(2): 262-286.

[43] Oviedo, A. M.; Thomas, M. R.; Karakurum-Ozdemir, K., (2009). “Economic Informality: Causes, Costs, and Policies - A Literature Survey”, World Bank Working Paper no. 167

[44] Paulson, A., R. Townsend and A. Karaivanov, (2006). “Distinguishing Limited Commitment from Moral Hazard in a Model of Entrepreneurship”, Journal of Political Economy, 144(1): 100-144.

[45] Reynolds, P., Bosma, N., Autio, E., Hunt, S., De Bono, N., Servais, I., Lopez-Garcia, P. and Chin, N., (2005). “Global entrepreneurship monitor: data collection design and implementation, 1998–2003”, Small Business Economics, 24(3): 205-231

[46] Rosentha, S., Strange, W.C., (2003). “Evidence on the Nature and Sources of Agglomeration Economies”, Handbook of Regional and Urban Economics, 4: 2119-2171

[47] Prado, M., (2011). “Government policy in the formal and informal sectors”, European Economic Review 55(8): 1120-1136.

[48] Torgler, Benno and Friedrich Schneider, (2007). “Shadow Economy, Tax Morale, Governance and Institutional Quality: A Panel Analysis”, IZA, dp No. 2563.

[49] Udry, C. and Anagol S., (2006). “The return to capital in Ghana”, American Economic Review, 96(2): 388-393.

[50] World Bank, (2005). “Doing Business in 2005: Removing Obstacles to Growth”, World Bank report, Washington, DC.

[51] World Bank, (2007). “Economic performance in Latin America and the Caribbean: A Microeconomic Perspective”, World Bank report, No. 40171-LAC.
A Appendix

A.1 Measures and Additional Results

Ellison and Glaeser geographic concentration index

The Ellison-Glaeser index (Ellison and Glaeser, 1997) defines concentration as agglomeration above and beyond what we would observe if plants simply chose locations randomly (as opposed to a uniform spatial distribution). This measure provides an unbiased estimate of agglomerative forces independently of their source. It can be interpreted as the probability that a firm choosing its location follows the prior firm rather than locating randomly. The Ellison-Glaeser index is given by the following formula:

\[ \gamma_i = \frac{G_i - (1 - \sum_r x_r^2) H_i}{(1 - \sum_r x_r^2)(1 - H_i)}, \]

where:

- \( G_i \equiv \sum_r (s_{ri} - x_r)^2 \) is the spatial Gini coefficient of industry \( i \);
- \( x_r \) is the share of total employment in each locality \( r \);
- \( s_{ri} \) is the share of employment of locality \( r \) in industry \( i \);
- \( H_i \equiv \sum_j z_{ji}^2 \) is the Herfindahl index of the plant size distribution of industry \( i \);
- \( z_{ji} \) represent the employment share of a particular firm \( j \) in industry \( i \).

**Interpretation:** Following Ellison and Glaeser (1997), an industry is strongly concentrated if \( \gamma_i > 0.05 \), weakly concentrated if \( \gamma_i \in (0, 0.05] \), and not concentrated if \( \gamma_i < 0 \).

Krugman index

The Krugman Specialization index (\( \text{ksi} \)) is a widely-used specialization measure. It measures the standard error of industry shares, by computing the share of employment which would have to be relocated to achieve an industry structure equivalent to the average structure of the reference group. Technically, the Krugman relative specialization index for an arrondissement, region or city is the sum over industries of the absolute differences between each industry’s share of regional or city employment and its corresponding share of national employment. The Index can take values in between zero (identical territorial structures) and two (totally different structures). In our case, we have 300 arrondissements, 146 industries, and we will measure the relative employment specialization at arrondissement level.
\[ KSI_a = \sum_{i=1}^{\text{arrondissements}} |S_{a,i} - \bar{S_i}| \]  

(11)

Where \( S_{a,i} \) is the output or employment share of industry \( i \) in arrondissement \( a \), and \( \bar{S_i} \) is the average share of industry \( i \) in the total employment across all arrondissements in Cameroon. Interpretation: If the Krugman relative specialization measure is zero, then the economic structure of an arrondissement is identical to the economic structure of the overall economy. The higher the index, the more the economic structure differs or deviates from the overall country and the more an arrondissement is considered to be specialized.

**Household Wealth index**

A Principal Component Analysis is used to create the aggregate wealth index used as household initial wealth in the empirical analysis. We have data on \( p \) variables representing quantities of household belongings. The \( i \)th principal component (or principal factor) is a linear combination of the \( p \) variables \( X_1, X_2, \ldots, X_p \),

\[ Z_i = a_{i1}X_1 + a_{i2}X_2 + \ldots + a_{ip}X_p \]

that has the largest possible variance, subject to the condition that \( a_{i1}^2 + a_{i2}^2 + \ldots + a_{ip}^2 = 1 \).

The principal factor can be extracted from this method through the following steps:

1. Define the matrix \( X \) (whose columns are the \( X_i, i = 1, \ldots, p \)).
2. Subtract the mean of each column of \( X \).
3. Calculate the sample covariance matrix \( \Sigma \) of \( X \).
4. Find the eigenvalues \( \lambda_1, \lambda_2, \ldots, \lambda_p \) and the corresponding eigenvectors \( a_1, a_2, \ldots, a_p \).

The eigenvectors should be scaled (by their euclidean norm) so that they are unit vectors. The coefficients of the \( i \)th principal component are then given by \( a_i \) while \( \lambda_i \) is its variance.

5. Assuming that the eigenvalues are ordered as \( \lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p \), choose the eigenvector that is associated with the highest eigenvalue, that is \( a_1 \). The principal component then corresponds to the eigenvector \( a_1 \),

\[ Z_1 = a_{11}X_1 + a_{12}X_2 + \ldots + a_{1p}X_p \]

In particular, \( \text{Var}[Z_1] = \lambda_1 \) which is the highest possible variance. To construct this single index, we use household items in the questionnaire that relate to the number of durable goods in the household: number of vehicles, number of TVs, number of radios, number of DVD/Video-CD, number of fridges, number of freezers, number of gas cookers, number of fans, number of sewing machines, a dummy variable that is equal to 1 if the household has air conditioning, number of mobile phones, number
of computers, number of electric irons, number of houses owned by the household. We consider only items that were acquired by the households prior their entry to their current activity. The resulting index computed from the data explains 45% of the variance in households durables.

A.2 Additional Results:

Table 11: Ten Most and Least localized NACE 4-digit industries in 2009 (EG index)

| NACE | Industry description                                      | EG index |
|------|----------------------------------------------------------|----------|
| 202  | Growing of coffee                                        | 1.757    |
| 1004 | Manufacture of sugar                                     | 1.244    |
| 1402 | Manufacture of other alcoholic beverages (except beer and malt) | 0.523    |
| 600  | Extraction of crude petroleum                            | 0.491    |
| 902  | Preparation of rice                                      | 0.386    |
| 3405 | Other transport (air, water)                             | 0.283    |
| 3406 | Auxiliary and auxiliary transport activities              | 0.203    |
| 2901 | Electric Power Generation, Collection, Transmission and supply | 0.186    |
| 3602 | Insurance activities (except social security)            | 0.165    |
| 103  | Growing of Oilseed crops (except palm)                    | 0.144    |

Ten Least localized industries

| NACE | Industry description                                      | EG index |
|------|----------------------------------------------------------|----------|
| 2702 | Manufacture of other transport equipment                  | -2.081   |
| 205  | Growing of palm                                          | -1.363   |
| 2301 | Manufacture of cement                                     | -0.533   |
| 1500 | Manufacture of tobacco products                           | -0.355   |
| 2902 | Water collection, treatment and supply                     | -0.271   |
| 1202 | Mfg of macaroni, noodles, couscous and similar farinaceous products | -0.237   |
| 4103 | Social work activities                                    | -0.230   |
| 2202 | Manufacture of rubber products                            | -0.189   |
| 3402 | Transport by taxi                                         | -0.183   |
| 702  | Mining of non-ferrous metal ores                          | -0.162   |

Notes: This ranking is based on author’s computation of the Ellison and Glaeser inde at arrondissement level.
Table 12: Twenty Most and Least Specialized cities in 2009 (Krugman index)

| Code | region        | Label         | ksi  |
|------|---------------|---------------|------|
| 355  | South         | NIETE         | 1.958|
| 158  | Littoral      | DIZANGUE      | 1.897|
| 325  | Centre        | MENGUEME      | 1.782|
| 154  | Littoral      | DIBAMBA       | 1.685|
| 59   | Centre        | NGOUMOU       | 1.542|
| 38   | Centre        | ELIG-MFOMO    | 1.454|
| 101  | Far-North     | PETTE         | 1.428|
| 120  | Far-North     | KAR-HAY       | 1.374|
| 94   | East          | DIANG         | 1.349|
| 102  | Far-North     | BOGO          | 1.347|
| 60   | Centre        | AKONO         | 1.340|
| 311  | Centre        | BIBEY         | 1.328|
| 138  | Far-North     | MINDIF        | 1.310|
| 64   | Centre        | MATOMB        | 1.303|
| 106  | Far-North     | WAZA          | 1.297|
| 61   | Centre        | BIKOK         | 1.256|
| 351  | West          | BANWA         | 1.256|
| 272  | South         | KRIBI II      | 1.220|
| 149  | Littoral      | MELONG        | 1.217|
| 207  | North-West    | NGIE          | 1.213|

Ten Least specialized cities

| Code | region        | Label         | ksi  |
|------|---------------|---------------|------|
| 68   | Centre        | NGOG-MAPUBI   | 0.00 |
| 111  | Far-North     | HILE-HALIFA   | 0.00 |
| 112  | Far-North     | DARAK         | 0.00 |
| 350  | North         | MADINGRING    | 0.00 |
| 337  | East          | MBOTORO       | 0.00 |
| 137  | Far-North     | TAIBONG       | 0.00 |
| 25   | Adamawa       | MAYO-BALEO    | 0.00 |
| 257  | South         | OVENG         | 0.00 |
| 23   | Adamawa       | DIR           | 0.00 |
| 129  | Far-North     | MAYO-MASKOTA  | 0.00 |
| 67   | Centre        | MESSONDO      | 0.00 |
| 280  | South-West    | LIMBE II      | 0.00 |
| 117  | Far-North     | VELE          | 0.00 |
| 51   | Centre        | NGORO         | 0.00 |
| 105  | Far-North     | LOGONE-BIRNI  | 0.00 |
| 167  | North         | MAYO HOURNA   | 0.00 |
| 139  | Far-North     | MOULVOUDAYE   | 0.00 |
| 136  | Far-North     | PORHI         | 0.00 |
| 307  | Adamawa       | NGAOUI        | 0.002|
| 77   | East          | GARI-GOMBO    | 0.011|

Notes: This ranking is based on author’s computation of the Krugman specialization index at arrondissement level, using 4-digit classification.
Figure 16: Paved roads density in Cameroon.

Note: Cities with missing data are in gray. This metric is the ratio of the length of the country’s total road network to the country’s land area km.
B Appendix: Ability Cutoffs

- The ability cutoff to be a constrained informal firm in sector $i$ and location $j$, $\theta_{ij}^c$, is defined by $l_{ij}^c(\theta_{ij}^c(A_{ij}), A_{ij}) = \bar{l}_j$, implying:

$$\theta_{ij}^c(A_{ij}) = \frac{l_{ij}^{1-\gamma}}{A_{ij}} \left( \frac{\lambda r_j}{\alpha_i} \right)^{\alpha_i} \left( \frac{w_j}{\beta_i} \right)^{1-\alpha_i} \quad (12)$$

- The ability cutoff to become an entrepreneur in sector $i$ within city $j$, $\theta_{ij}^w(A_{ij})$ is the solution to the equation $\max\{\pi^l(\theta_{ij}, A_{ij}), \pi^F(\theta_{ij}, A_{ij})\} = w_j$. Denoting by $\theta_{ij}^{W,1}$ and $\theta_{ij}^{W,2}$ the unique solutions of the equations $w_j = \pi^F(\theta_{ij}, A_{ij})$ and $w_j = \pi^l(\theta, A_{ij})$ respectively, we have $\theta_{ij}^{W,1}(A_{ij}) = \frac{1}{A_{ij}} \left( \frac{w_j + r_j c_j}{(1-\gamma_i)(1-\tau)} \right)^{1-\gamma_i} \left( \frac{r_j}{\alpha_i} \right)^{\alpha_i} \left( \frac{w_j}{\beta_i} \right)^{\beta_i}$, and

$$\theta_{ij}^{W,2}(A_{ij}) = \begin{cases} 
\frac{(1+\bar{l}_j)^{1-\alpha_i} \bar{l}_j^{-\beta_i}}{A_{ij}} \left( \frac{w_j}{1-\alpha_i} \right)^{1-\alpha_i} \left( \frac{\lambda r_j}{\alpha_i} \right)^{\alpha_i} & \text{if } \bar{l}_j \leq \frac{\beta_i}{1-\gamma_i} \\
\frac{1}{A_{ij}} \left( \frac{w_j}{\beta_i} \right)^{1-\alpha_i} \left( \frac{\lambda r_j}{\alpha_i} \right)^{\alpha_i} \left( \frac{\beta_i}{1-\gamma_i} \right)^{1-\gamma_i} & \text{otherwise}
\end{cases}$$

The desired cutoff is then defined by $\theta_{ij}^w(A_{ij}) = \min\{\theta_{ij}^{W,1}(A_{ij}), \theta_{ij}^{W,2}(A_{ij})\}$.

- The ability cutoff to become formal entrepreneur in sector $i$ within city $j$, $\theta_{ij}^f(A_{ij})$ is the solution to $\pi^F(\theta_{ij}) = \pi^l(\theta_{ij})$. Denote $V^{l,F}(\theta_{ij}) = \pi^F(\theta_{ij}) - \pi^l(\theta_{ij})$, and notice that for $V^{l,F}(\theta_{ij})$ is continuous in $\theta$, $V^{l,F}(\theta_{ij}) < 0$ for $\theta_{ij} < \theta_{ij}^F$ and $V^{l,F}(\theta_{ij}) \to +\infty$ as $\theta_{ij} \to +\infty$. Hence $\theta_{ij}^F(A_{ij})$ exists and is unique.