

**INFORMAL SECTOR HETEROGENEITY AND INCOME INEQUALITY:
EVIDENCE FROM THE DEMOCRATIC REPUBLIC OF CONGO***

FRANCK M. ADOHO AND DJENEBA DOUMBIA

The World Bank, USA

This paper empirically identifies three types of entrepreneurs in the Congolese informal sector, namely top-performers, constrained gazelles and survivalists. Based on logit and fixed effect OLS models, the paper finds that poverty and income inequality are more common among constrained gazelles and survivalists. Results also show that income inequality is explained mainly by educational disparities and lack of credit access among entrepreneurs. The outcomes of a Blinder-Oaxaca decomposition show that the performance of firms is a key factor in explaining differences in income. Moreover, the paper finds that human capital and managerial skills are important engines of performance.

Keywords: Informal Sector, Income Inequality, Poverty, Firm Performance.
JEL Classification: O17, O12, D21

1. INTRODUCTION

The urban informal sector represented 81.5 percent of employment in the Democratic Republic of Congo in 2012 (World Bank, 2016). This large proportion highlights the sector's critical importance for job creation, poverty reduction and income growth in the country. While the informal sector is defined in various ways in the

* The authors would like to thank Jean-Bernard Chatelain, Kathleen Beegle, Jean-Louis Combes, Chadi Bou Habib, Tidiane Kinda, Aly Sanoh, Clarence Tsimpo Nkengne, Pierella Paci, Philippe de Vreyer, Romeo Gansey as well as participants to the 12th Doctorissimes of University Paris 1 Pantheon-Sorbonne, the International Monetary Fund African Department Seminar, and the 57th Conference of the Canadian Association for Economic Science for their comments. The findings, interpretations and conclusions expressed in this paper do not necessarily reflect the views of the World Bank Group, its Board of Directors, or member states.

literature, it is generally said to encompass economic activities that are not subject to taxes and other regulations (De Soto, 1989; Loayza, 1996). This paper defines enterprises in the informal sector as enterprises without formal accounting books and national identification numbers as firms. This definition allows for comparability with studies in other developing countries.

Many aspects of the informal economy have been studied. Earlier studies have mainly focused on explaining the reasons underlying the emergence of the informal economy, the productivity of the informal sector, and why firms remain informal. Still, not enough attention has been paid to heterogeneity in the informal sector and even less to inequality within and between the various categories of informal firms. Understanding this heterogeneity and income inequality is crucial to the design of appropriate policies.

This paper contributes to the literature by analyzing heterogeneity in the informal sector in the Democratic Republic of Congo and the implications for the performance of firms and income inequality. Relying on a representative sample of informal firms derived from the 1-2-3 survey in the country, the paper identifies three distinct groups of informal entrepreneurs: top-performers, constrained gazelles, and survivalists. The first group of entrepreneurs is growth oriented and enjoys a greater access to capital. The second group includes entrepreneurs who share many characteristics, especially management skills, with the top-performers, but operate with less capital. The survivalists, entrepreneurs who are struggling to grow, represent the third group. Furthermore, the paper makes three other major contributions. First, it analyzes poverty and income inequality among these three categories of firms. Second, using a Blinder-Oaxaca decomposition, it explains income differences across the three groups. Third, it examines the drivers of performance and profitability among informal firms.

The results highlight a striking heterogeneity among informal firms. Poverty rates are higher among survivalists and constrained gazelles than among top-performers. Income inequality, which is more common among constrained gazelles and survivalists, is mainly explained by performance gaps, educational disparities, and a lack of credit access. The characteristics of entrepreneurs, such as age, educational attainment, and managerial skills, are important drivers of the quality of the performance of firms.

The rest of the paper is structured as follows. Section 2 reviews the literature on the informal sector. Section 3 describes the data and explains the identification strategy of the three groups of informal firms. Section 4 compares various characteristics associated with these groups. Section 5 investigates urban poverty and inequality among informal firms. Section 6 analyses the drivers of the performance of firms. Section 7 provides concluding remarks and policy recommendations.

2. LITERATURE REVIEW

The literature highlights a number of features of the heterogeneity in the informal

sector, including involuntary wage labor, queuing for formal employment, and voluntary self-employment. These are similar in many ways to the characteristics of the entrepreneurial small firm sector in developed countries (Grimm, Lay, Roubaud and Vaillant, 2011). For instance, Rogerson (1996) recognizes a conceptual distinction between two groups of informal enterprises. The first group consists of survivalist firms. These are involved in activities initiated by people unable to secure regular wage employment or access to an economic sector of their choice. These firms are characterized by poverty and must struggle to survive. The second group includes microenterprises or growth enterprises that are small. Fields (2004) differentiates between rationed upper-tier activities, which are based on methods of production and types of jobs that are similar to those of formal firms, and a lower tier, which is concentrated on easy entry activities and corresponding to the residual subsistence sector in the dualistic view (Bosch and Maloney, 2010). However, Cunningham and Maloney (2001), using data on Mexico, argue that there is strong heterogeneity among small firms that is of the same nature as the heterogeneity in developed countries, whereby small firms that have reached their optimal long-run size coexist with profitable starting firms and start-up firms that will fail.

Using a sample of informal entrepreneurs in seven West African countries, Grimm, Knorringa, and Lay (2012) in addition to the groups of top-performers and survivalists, empirically identify a third segment of informal entrepreneurs, constrained gazelles. These entrepreneurs show characteristics - managerial abilities and motivation - similar to those of growth-oriented entrepreneurs. The authors mostly find a larger share of constrained gazelles in more dynamic economies. Gindling and Newhouse (2014) investigate heterogeneity among the self-employed, using the international Income Distribution Database on almost 100 developing countries. Their results show that one-third of the unsuccessful entrepreneurs have the potential to become successful.

Using qualitative information, performance-based measures and employment growth, Diao et al. (2018) identified a group of firms in the Micro, Small and Middle-Sized Enterprises sector in Tanzania - the so-called "in-between". These firms operate mainly in the sectors of manufacturing and trade services, employ approximately the same number of employees as compared to those in the formal sector and have significant potential for future growth.

The informal sector has been studied in the literature, but few studies have specifically investigated inequality in the informal sector. Most that do have focused only on the causality between rising inequality and the size of the informal economy at the macroeconomic level (Chong and Gradstein, 2004; Winkelried, 2005; Lukiyanova, 2015).

Classifying firms in the informal sector and recognizing the sector's role as a setting environment for profitable activity may contribute to economic growth (AfDB, 2013). Accounting for more than half the labor force in Sub-Saharan Africa, the informal sector represents prospects for better living standards among many people. Nonetheless, there is evidence that registering with tax authorities (formalization) increases the profits of

firms with two to five workers and the middle third of capital stock (the midsize group), while leading to losses among smaller and larger informal firms (McKenzie and Sakho, 2010). Determining how to reduce inequality and poverty between and within the three distinct categories of informal firms - top-performers, constrained gazelles, and survivalists - could help enhance the productivity of firms and in the coordination of strategies and policies to support the formalization of the sector.

3. DATA AND METHODOLOGY

3.1. 1-2-3 Survey

The empirical investigation reported in this paper relies on survey data from the 1-2-3 survey collected in the Democratic Republic of Congo in 2012. The survey was conducted by the National Institute of Statistics, in partnership with other actors, including Afristat and the World Bank. The 1-2-3 survey is a representative, multilayer survey that covers three nested surveys, three phases involving separate statistical populations: individuals, informal production units, and households.

Phase 1 provides detailed information on employment, unemployment, and household and individual sociodemographic characteristics. It includes a household module, an individual questionnaire for individuals ages 5 or above, and a community questionnaire. Phase 2 is a specific survey covering the heads of informal production units. This informal sector survey collects information on the characteristics of firm owners, production units, economic performance, business requirements (costs, investment), and employees. This phase is important as it allows measuring the Congolese informal economy. Phase 3 is a survey on household expenditures that involves interviews among a subsample of the informal production units identified in phase 1. The paper focuses on a representative subsample of 4,504 informal firms. Table A1 in the appendix defines the variables used in the paper.

3.2. Identification Strategy: Informal Firms

3.2.1. *Defining a Top-Performer*

The strategy to identify the three categories of informal entrepreneurs involves three steps.¹ The first step consists in defining growth-oriented entrepreneurs or top-performers in terms of physical capital and value added. Physical capital is defined as

¹ The paper uses different definitions of top-performers, and different methodologies in addition to those used in Grimm, Knorringa, and Lay, 2012.

the replacement value of all business-related assets that have been used in the operation of the enterprise in the previous year, including the business establishment, machinery, furniture, vehicles, and utilities. In the second step, owner and firm characteristics that are correlated with physical capital and value added, are identified and the empirical probability that a firm will be highly productive is assessed based on these correlations. In the third step, entrepreneurs are classified into the three groups - top-performers, constrained gazelles, and survivalists - conditional on the actual status as a top-performer and the predicted probability of this status.

The paper defines top-performers as those entrepreneurs in the sample who are the most productive (the top 50 percent in terms of capital productivity, that is, value added per physical capital) among the top 40 percent of the capital distribution. Because capital accumulation is part of the economic growth process, it is considered an accurate factor in identifying a growth-oriented firm.

3.2.2. *Sample Selection Bias*

Because of missing values in the variable of capital and because the definition of the top-performers is based mainly on this variable, the Heckman (1979) two-step estimation is viewed as a reliable way to test for selection bias. The first step consists in analyzing the determinants of capital possession. To do so, a probit model is estimated, as the follows:

$$\Pr(\text{Capital} = 1) = \theta(\alpha_0 + \alpha_1 X_i + \omega_i), \quad (1)$$

where capital is a dummy variable taking 1 when a firm owns capital and 0 otherwise; θ is the cumulative standard distribution function; X_i is a vector of the following factors: business association, financial knowledge that is measured by a variable equal to 1 if the firm's owner has knowledge about microfinance institutions and 0 otherwise, family tradition taking 1 if the owner started a business because of family tradition, the age of the owner, the age squared; a gender dummy equal to 1 if the owner is a man and 0 otherwise, and firm age. Dummies are included for sectors. The equation also includes regions fixed effects.

The results show that membership in a business association, the possession of financial knowledge, and being a man increase the probability of owning capital (Table 1). They also support the evidence that industry is a highly capital-intensive sector. Entrepreneurs in the industry sector are more likely (around 6.7 percent) than entrepreneurs in other sectors to have capital (Table 2). Entrepreneurs operating in hotels and restaurants are also likely to own capital. Entrepreneurs in retail trade and in repair services show less likelihood of capital ownership.

In the next step, the paper derives the predicted probabilities from the estimation of (1) to calculate the inverse Mills ratio, that is, the ratio of the probability density

function to the cumulative distribution function, as follows²:

$$\widetilde{MR} = \frac{\phi(\alpha_1 X_i / \widehat{\sigma_\omega})}{\Phi(\alpha_1 X_i / \widehat{\sigma_\omega})}. \quad (2)$$

Table 1. Step 1: Determinants of Capital Ownership among Firms

	(1) Coefficients	(2) Marginal effects
Business association	0.2090** (0.0972)	0.0548
Financial knowledge	0.1510*** (0.0585)	0.0397
Family tradition	0.1040 (0.0896)	0.0273
Male (=1)	0.0887* (0.0522)	0.0233
Age firm	0.0057 (0.0130)	0.0015
Age owner	0.0037 (0.0081)	0.0010
Age squared owner	-0.0000 (0.0000)	-0.0000
Constant	0.8020*** (0.1840)	
Sector fixed effects	yes	
Region fixed effects	yes	
Pseudo-R ²	0.1138	
Observations	4410	4410

Notes: Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

The paper then includes this ratio in estimating the following equation:

$$\Pr(Y_i^{Top} = 1) = \theta(\beta_0 + \beta_1 X_i + \mu \widetilde{MR} + \omega_i), \quad (3)$$

where Y_i^{Top} is the binary variable indicating a firm that is a top-performer; θ is the cumulative distribution function of the logistic distribution; β_1 is the vector of coefficients that defines the link between owner and firm characteristics and the probability that the firm is a top-performer; and X_i is the vector of owner and firm characteristics, different from the vector in Equation (1). The vector X_i encompasses the predetermined factors that are correlated with identity as a top-performer. To address

² Box A1 provides a detailed explanation about the inverse Mills ratio, the probability density and the cumulative distribution functions.

the issue of endogeneity, predetermined variables are included, as follows: age, age squared, gender, educational attainment, and the motivation of the entrepreneur to set up the business, which is measured by a dummy variable taking the value of 1 if the entrepreneur created the firm because s/he could not find a paid job in a large company. Besides these predetermined variables which allow to control for the firm's age and avoid selection against younger firms with a high potential, a sector dummy is included.

Table 2 displays the estimation results. It highlights that men have a greater chance (5.2 percent) than women of being top-performers. In addition, the probability of inclusion as a top-performer rises with age, albeit at a decreasing pace. Vocational training is a key determinant of being a top-performer. Individuals who have started businesses because they could not find jobs in large firms exhibit a higher probability (approximately 4.2 percent) of inclusion among the top-performers. Moreover, the inverse Mills ratio is not significant, that is, there is no selection bias in the model estimation.

Table 2. Step 2: Probability of being a Top-performer

	(1) Coefficients	(2) Marginal effects
Age owner	0.1170*** (0.0287)	0.0088
Age squared owner	-0.0014*** (0.0004)	-0.0001
Male (=1)	0.7010*** (0.1560)	
Age firm	0.0239 (0.0426)	0.0018
No diploma	Ref.	
Primary completed	0.1280 (0.3990)	0.0074
Some secondary	0.4390 (0.3110)	0.0288
Post-secondary	0.5990 (0.3740)	0.0419
Other post primary	1.0570** (0.4320)	0.0886
No opportunity in large firms	0.5640 (0.1830)	0.0424
Family tradition	0.1710 (0.2660)	0.0129
Inverse Mills ratio	-2.1290 (1.3410)	-0.1600
Constant	-4.7690*** (0.9240)	
Sector fixed effects	yes	
Region fixed effects	yes	
Pseudo-R ²	0.0800	
Observations	3401	3401

Notes: Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

3.2.3. Identification of constrained gazelles and survivalists

Based on regression (3), the paper predicts, for each entrepreneur in the country, the statistical probability of being a top-performer. This probability is conditional on the estimated parameters β_0 and β_1 and on the vector of observed variables X . The following model is performed:

$$\widehat{\Pr}(Y_i^{Top} = 1) = \theta(\widehat{\beta}_0 + \widehat{\beta}_1 X_i), \quad (4)$$

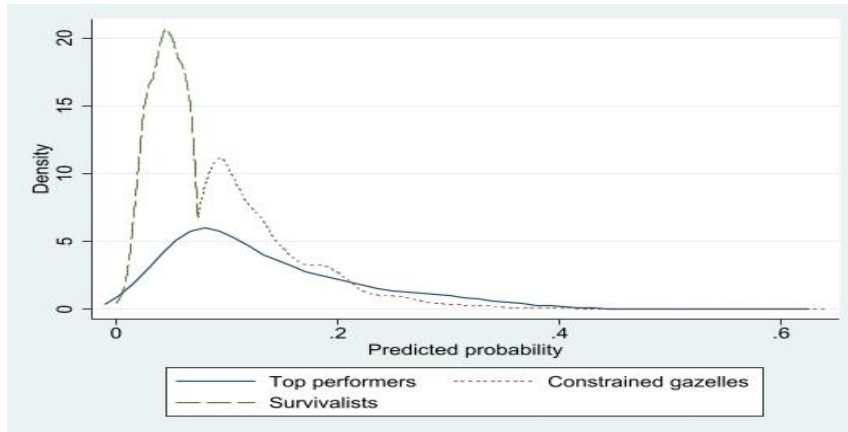
where $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are the estimated coefficients. To separate the sample of informal entrepreneurs into three categories: top-performers, constrained gazelles and survivalists the paper uses these predicted probabilities.

To determine the shares of the other segments, the cut-off point is selected for $\widehat{\Pr}(Y_i^{Top} = 1)$ between constrained gazelles - a group of entrepreneurs who have business skills that resemble in many respects those of the top-performers but are clearly different from those skills characterizing survivalists - and survivalists such that the mean of this predicted probability is similar in the group of top-performers and the group of constrained gazelles. This suggests that the distribution of the observable factors is equal in both categories and that on average constrained gazelles should be as likely as the actual top-performers to be top-performers. The survivalists are defined as all entrepreneurs for whom $\widehat{\Pr}(Y_i^{Top} = 1)$ is below the threshold.³

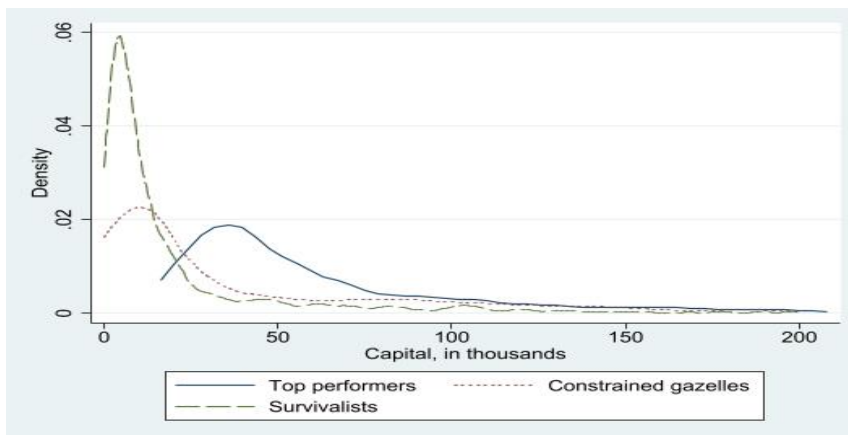
Estimates based on Equation (3) illustrate that constrained gazelles and top-performers share some characteristics. Table A3 in appendix illustrates the determinants of inclusion among the three distinct groups of informal firms that have been identified. The results show that firms managed by men, people who created businesses because they could not find jobs as wage earners in large firms, more well educated, and older people have a higher probability of being constrained gazelles. However, these determinants also decrease the probability of being survivalists.

To check the consistency of the method of determining the three groups of entrepreneurs, the distribution of predicted probabilities is plotted for the top-performers, survivalists, and constrained gazelles. Figure 1 shows that the distribution among survivalists is dissimilar relative to the distribution within the two other groups. It also displays the distribution of capital: top-performers and constrained gazelles possess more capital than survivalists. This confirms the results of the analysis

³ As robustness check, the paper relies on full information maximum likelihood (FIML) model to estimate Equation (3). The FIML model handles missing values problems well by using all the information available. The results remain unchanged (Table A2). The two groups (constrained gazelles and survivalists) are then reclassified using the predicted probabilities based on the model regressions. The groups of constrained gazelles and survivalists using Heckman two step and FIML estimations match at about 96 percent.



a. Distribution of predicted probability of being a top-performer



b. Distribution of capital

Figure 1. Distributions Across Informal Firms

4. HETEROGENEITY IN THE INFORMAL SECTOR

Assessing differences among top-performers, constrained gazelles, and survivalists helps identify the best policy design for each category. This section considers three features - individual entrepreneurial characteristics, choice of sector, and firm characteristics. Constrained gazelles and top-performers might be expected to exhibit similar basic management characteristics that aligns with how this analysis defined the two groups. Table 3 shows comparison results.

Table 3. Comparison across Top-performers, Survivalists, and Constrained Gazelles

	(1) Top performer	(2) Constrained Gazelles	(3) Survivalists	(4) t- Test (1 vs 2)	(5) t- Test (1 vs 3)	(6) t- Test (2 vs 3)
Basic owner characteristics						
Age owner	37.753	38.132	36.582	0.561	0.199	0.001***
Male (= 1)	0.580	0.712	0.132	0.000***	0.000***	0.000***
No diploma	0.046	0.034	0.119	0.298	0.000***	0.000***
Primary completed	0.110	0.098	0.231	0.526	0.000***	0.000***
Some secondary	0.658	0.666	0.561	0.797	0.001***	0.001**
Post-secondary	0.140	0.158	0.082	0.448	0.001***	0.000***
Other post-primary	0.043	0.042	0.005	0.961	0.000***	0.000***
No opportunity in large firms	0.201	0.195	0.026	0.832	0.000***	0.000***
No opportunity in small firms	0.110	0.148	0.133	0.091*	0.285	0.222
Poor	0.568	0.644	0.733	0.014**	0.000***	0.000***
Firms characteristics and Sector						
Size of firm	1.756	1.496	1.191	0.001***	0.000* **	0.000***
Paid workers	0.450	0.2178	0.057	0.0003***	0.000* **	0.000***
Age Of firm	2.263	2.295	2.030	0.766	0.021**	0.000***
Value added	7,755,333	1,789,103	1,049,235	0.000***	0.000***	0.000***
Capital [#]	64,342.53	57,903.35	28,567.01	0.214	0.000***	0.000***
Turnover	31,400,000	5,315,526	3,893,473	0.000***	0.000***	0.000***
Gross operating profit	7,028,248	1,533,952	996,596.9	0.000***	0.000***	0.000***
Food/Agribusiness	0.230	0.184	0.532	0.073*	0.000***	0.000***
Clothing and apparel	0.083	0.104	0.054	0.2726	0.047**	0.000***
Industry	0.156	0.162	0.147	0.811	0.675	0.250
Construction	0.030	0.033	0.001	0.787	0.000***	0.001*
Transport	0.026	0.033	0.025	0.528	0.910	0.172
Hotels and restaurants	0.080	0.049	0.000	0.037**	0.000***	0.000***
Wholesale	0.020	0.018	0.005	0.889	0.006***	0.000***
Primary and mining	0.046	0.048	0.014	0.874	0.000***	0.000***
Retail and repair services	0.266	0.290	0.172	0.405	0.000***	0.000***
Other activities	0.060	0.073	0.046	0.407	0.323	0.001***
Structural and behavioural factors						
Firm set up with other person	0.256	0.216	0.115	0.131	0.000***	0.000***
By family tradition	0.087	0.090	0.077	0.800	0.574	0.155
Firm has a location	0.116	0.081	0.034	0.049**	0.000***	0.000***
Mobile location	0.146	0.152	0.165	0.799	0.405	0.313
Fixed location on the street	0.176	0.170	0.126	0.800	0.018**	0.000***
Firm located at home	0.313	0.349	0.423	0.235	0.000***	0.000***
Firm located at clients' home	0.033	0.048	0.017	0.247	0.065*	0.000***
Fixed location on market	0.033	0.033	0.015	0.981	0.027**	0.000**
Improvised location on market	0.183	0.160	0.210	0.342	0.273	0.000***
Shop/restaurant	0.030	0.027	0.013	0.778	0.027**	0.004***
Other location	0.083	0.057	0.027	0.099*	0.000***	0.000***
Water connection	0.083	0.068	0.049	0.366	0.016**	0.023**
Electricity connection	0.133	0.142	0.053	0.690	0.000***	0.000***
Telecommunication	0.200	0.202	0.077	0.925	0.000***	0.000***
Self-employment / Firm is staffed by one person	0.656	0.752	0.878	0.000***	0.000***	0.000***
Trade register	0.016	0.024	0.006	0.401	0.067*	0.000***
Financial knowledge	0.216	0.228	0.143	0.652	0.001***	0.000***
Member of a business association	0.113	0.101	0.051	0.547	0.000***	0.000***
Active reaction to demand shocks	0.786	0.735	0.744	0.081*	0.136	0.602
Risk aversion	0.164	0.124	0.114	0.065*	0.015**	0.406
Observations	300	1331	1839			

Notes: Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1. # Few outliers have been dropped.

4.1. Individual Entrepreneur Characteristics

The incidence of poverty in the Congolese informal sector is high. Among Congolese entrepreneurs, 61 percent are poor. However, the incidence of poverty is lower within top-performers compared to the two other groups. About 57 percent of top-performers are poor whereas respectively 64 percent and 73 percent of constrained gazelles and survivalists are poor. Policies that aim to improve the well-being of poor informal workers should thus target the survivalists and constrained gazelles.

Gender-based disparity in the informal sector is more noticeable by type of firm. Survivalists informal firms are mostly woman-owned (87 percent). Women own 42 percent and 29 percent, respectively, of top-performers and constrained gazelles (see Table 3).

Educational differences are pronounced, with top-performers and constrained gazelles more highly educated than survivalists. While the difference in educational attainment between top-performers and constrained gazelles is not statistically significant, the education gap between top-performers and survivalists is noticeable. While 14 percent of top-performers have a postsecondary education, only 8 percent of survivalists have attained this level. Survivalists have the lowest overall educational attainment.

Vocational training correlates positively with the performance of informal firms. This correlation may provide a useful link between education and firm performance. Through better education and training, entrepreneurs may develop greater capability, acquire new management and marketing skills, and learn efficient production processes.

4.2. Firm Typology and The Choice of Sector

More than 50 percent of survivalists are engaged in food trade and agribusiness compared with only 23 percent of top-performers and 18.4 percent of constrained gazelles. This sector comprises for instance retail trades of palm oil, tomatoes and other agricultural products. The primary sector of top-performers and constrained gazelles (26.6 percent and 29.0 percent, respectively) is repair services and the retail trade, which includes retail trade in coal, fuel, and kitchen appliances and the repair of bicycles, motorcycles, and household items. Our results are similar to those of Gindling and Newhouse (2014) showing that the successful self-employed are more likely to operate in retail and services compared to unsuccessful self-employed.

Top-performing firms are generally larger than constrained gazelles and survivalists. Top performing firms are approximately 20 percent larger than constrained gazelles and almost 50 percent larger than survivalist firms. Firm size is measured as the total number of paid and non-paid workers. In addition, the number of paid workers is significantly higher in top-performing firms than in the two other types of firms. This suggests that constrained gazelles and survivalists tend to employ more non-paid workers (mainly family members) as compared to the top-performers.

Constrained gazelles are generally older than both top-performers and survivalists, and survivalists are, on average, the youngest of the three groups. This age dimension may reflect accumulated business experience. Firms may also experience both positive and negative shocks during their existence. The learning-by-doing channel enables firms to learn and to perform better as they age, making them more likely to survive (Baker and Kennedy, 2002). However, older firms in developing countries may suffer from productivity losses, for example, if their owners become too old to adopt new technologies (Nichter and Goldmark, 2009).

Other firm characteristics tend to be significantly different when top-performing firms, constrained gazelles and survivalists are compared. Top-performers generate significantly higher value added, sales turnover and gross operating profit as compared to constrained gazelles and survivalists. For instance, sales turnover averages more than CGF 31 million. This number compares with only CGF 5 million for constrained gazelles and CGF 3 million for survivalists. This suggests that the three types of firms may differ in terms of their behavior in managing their firms.

4.3. Structural and Behavioural Factors

This section examines structural and behavioural factors that could explain why top-performers, constrained gazelles and survivalists differ.

Setting up an informal firm with more than one owner may reduce credit market constraints. Enterprises with multiple owners are more likely to be top-performers or constrained gazelles. About 26 percent of the top-performers have founded their businesses as joint ventures. About 22 percent of constrained gazelles have done so while only 11 percent of survivalists are joint ventures. About 9 percent of top-performing firms and constrained gazelles have been created by family tradition while this is the case for 7 percent of survivalists.

Top-performers and constrained gazelles tend to operate from fixed locations (streets, markets, shops), while survivalists tend to operate from home and mobile selling points. The transitory locations of survivalists may explain the greater constraints they face in taking advantage of basic infrastructure. Constrained gazelles and survivalists are more constrained in terms of location than top-performers. However, top-performers and constrained gazelles are close in terms of access to basic infrastructure while they are clearly different from survivalists.

Survivalists possess less financial knowledge, are less likely to register their businesses and to be involved in a business association than either of the other two groups. Top-performers and constrained gazelles are not significantly different in these characteristics. This highlights that top-performers and constrained gazelles share some similar management skills and professional networks. Yet, the owners of the two types of firms differ in terms of other management skills - their owners' active reaction to demand shocks and their aversion to risks. The chapter constructs two dummy variables to assess (1) the active reaction of entrepreneurs to demand shocks and (2) the risk

aversion of entrepreneurs. The first dummy variable equals 1 if in response to an important decline in demand, the entrepreneur reduces his/her profit, increases and/or diversifies the quality of products. The second dummy variable equals 1 if the motivation of the entrepreneur to take up his/her current activity is - the assurance of more stable receipts than in other products. Results show that top-performers tend to be more reactive to demand shocks, and less averse to risks as compared to constrained gazelles and survivalists.

5. URBAN POVERTY AND INCOME INEQUALITY IN THE INFORMAL SECTOR

Poverty and income inequality are particularly significant in the informal sector. This section examines these two issues among the three distinct groups of informal firms.

Table 4 shows the distribution of quintile of income by informal firms. The top-performers are better off than the other categories of informal firms. While more than 17 percent of survivalists are in the bottom quintile of revenue only respectively 13 percent of constrained gazelles and 9 percent of top-performers are found in this quintile. Moreover, the logit estimates of the probability of being poor confirm that survivalists and constrained gazelles are more likely to be poor than top-performers (Table 5).⁴

Table 4. Quintiles of Per Adult Equivalent Consumption, Informal Firms (%)

Quintiles of consumption	Top performers	Constrained Gazelles	Survivalists
1	8.94	12.85	17.55
2	18.43	21.17	26.23
3	21.37	17.17	21.94
4	23.04	21.47	19.66
5	28.23	27.35	14.63
Total	100	100	100

Sources: Calculations using 1-2-3 survey data

Notes: Percentages are weighted using sampling weights.

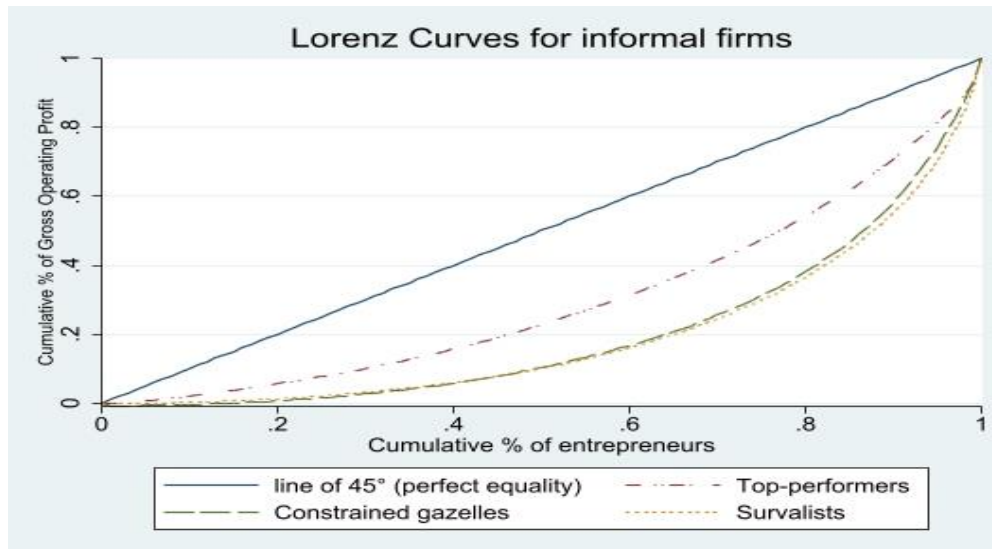
⁴ These results are robust to two other classification methods (using the mean in capital and in value added per hour worked) for the three groups of informal firms.

Table 5. Poverty between Informal Firms, Logit Estimates

	(1) Coefficients Poor = 1	(2) Marginal effects	(3) Coefficients Poor = 1	(4) Marginal effects
Top-performers	Ref.		Ref.	
Constrained gazelles	0.320** (0.146)	0.068** (0.0309)	0.409*** (0.154)	0.072*** (0.027)
Survivalists	0.736*** (0.132)	0.157*** (0.0277)	0.582*** (0.154)	0.102 *** (0.027)
Constant	0.276** (0.124)		1.339*** (0.193)	
Sector fixed effects			yes	
Region fixed effects			yes	
Pseudo-R2	0.011		0.152	
Observations	3398	3398	3398	3398

Notes: Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

A Lorenz curve illustrates inequality across the three groups of informal firms (Figure 2). Profit is more equally distributed across top-performers than across the other two groups of informal firms. The poorest 25 percent of top-performers share about 8 percent of the profits, whereas the poorest 25 percent of the constrained gazelles possess only 1.5 percent of the profits. The Gini coefficient varies across informal firms. The top-performing firms are not only wealthier, but also the income of the group is distributed more equally across the group.

**Figure 2.** Lorenz Curves for Informal Firms

An ordered logit regression that highlights the factors explaining income inequality for all the firms together (all sample) and for each of the three segments of firms separately is then performed, as follows:

$$Y_{ij}^* = \alpha_0 + \beta_0 \text{Gross_profit}_{ij} + \beta_1 X_{ij} + \beta_2 H_{ij} + \beta_3 M_{ij} + \beta_4 S_{ij} + \beta_5 R_{ij} + \varepsilon_{ij}, \quad (5)$$

where Y_{ij}^* represents the ordered quintiles of income per adult equivalent for entrepreneur i in segment j . The segments refer to the three distinct groups of informal firms: top-performers, constrained gazelles and survivalists. Gross_profit_i is the gross operating surplus of firm i . If an increase in the gross operating surplus raises the income of the less well-off entrepreneurs, it means that inequality narrows. The vector X_i includes standard characteristics of the firm's owner, that is, age, age squared, gender and the age of the firm. The vector H_i includes educational attainment, financial knowledge, reading knowledge and a dummy taking the value 1 if the owner is member of a business association. The vector M_i represents proxies for the firm's management skills. It includes lack of customers, credit access and equipment, management problems, firm location and employees' discipline issues. Finally, dummies for sector (S_i) and region (R_i) are included.

Results show that an increase of Congolese Franc 1,000 in the gross operating surplus, the odds of being in the top quintile versus the (combined) other categories are 1.001 times greater (all sample) (Table A4 in Appendix). An increase in profits is more likely to positively impact the income of the richest entrepreneurs than the others' incomes. When considering each category of firms, the constrained gazelles and survivalists show the same pattern. However, among the top-performers there is an equal chance of being in the top quintile of income, but the estimated coefficient is not significant.

Human capital, including educational attainment and the ability to read in any language is a key factor, that is, it increases the probability of being in the top quintile for all types of informal firms. Still, the impact of postsecondary educational attainment among top-performers is the greatest. Attainment of other post-primary education, including vocational training, increases the odds of being in the top income quintile among constrained gazelles, though this is not the case for the two other groups. Moreover, firms with less access to credit are less likely to be in the top income quintile. The impact of the lack of credit access on income inequality is greater among constrained gazelles and survivalists.

Firm size and the income of the owners appear to be independent among top-performers and survivalists. However, among constrained gazelles, there is a significant negative relationship between firm size and the probability of being in the highest income quintile. Regarding the factors underlying firm organization, firms constrained by credit access are less likely to be in the top income quintile.

These results highlight that pro-poor policies aimed at improving education and providing access to finance are the most favorable to increasing income among the

poorest among informal firms.

The results of inequality regressions are robust to two other classification methods of the three groups of informal firms and to two other definitions of top-performers. First, the groups of top-performers and constrained gazelles are reclassified based on the median value of capital (Table A5 in Appendix). Second, the top-performers are redefined as those entrepreneurs in the top 10 percent of the distribution of value added per hour worked (Table A6). The paper also tests whether the classification changes according to the definition of top-performers. First, top-performers are defined as those entrepreneurs in the sample who are in the top 40 percent in terms of capital productivity among the top 40 percent of the capital distribution. Second, the paper defines top-performers as those entrepreneurs in the sample who are in the top 40 percent in terms of capital productivity among the top 50 percent of the capital distribution. Then, the paper classifies the two other groups, the constrained gazelles and survivalists. Tables A7 and A8 present the results.⁵

6. DRIVERS OF THE PERFORMANCE OF INFORMAL FIRMS

The findings from the previous section show that policies aimed at reducing the performance gap between top-performers and the two other groups, namely, constrained gazelles and survivalists, may also help lift entrepreneurs in the latter two groups out of poverty and reduce income inequality. This section first corroborates these results based on a Blinder-Oaxaca decomposition and then examines the factors that make one informal firm more successful than another.

The Blinder-Oaxaca decomposition (Blinder, 1973; Oaxaca, 1973) explains the gap in the means of an outcome variable - the logarithm of per adult equivalent revenue in this case - between two groups. This paper compares top-performers with constrained gazelles and top-performers with survivalists.

The formalization of income differences is as follows:

$$\ln(\text{Income})_{Tt} - \ln(\text{Income})_{it} = (X_{Tt} - X_{it})\beta_{Tt} + (\beta_{Tt} - \beta_{it})X_{it}, \quad (6)$$

$$i \in \{\text{Constrained gazelles}, \text{Survivalists}\},$$

where $\ln \text{Income}_{Tt}$ and $\ln \text{Income}_{it}$ represent the mean of the logarithm of per adult equivalent income for respectively top-performers and constrained gazelles or survivalists in time t . X is the vector of control characteristics for individuals. The first

⁵ We performed an endogeneity test to assess the potential endogeneity of our profit variable (gross operating surplus). Using the sector-region average of value added per hour worked as an instrument for gross operating profit, the test did not find evidence of endogeneity.

term is the explained component, and the second term represents the unexplained component as described above.

Results show that differences in firms' performance explain more than half a percent of the revenue gap between top-performers and constrained gazelles (Table 6). Differences in other factors including the individual characteristics of entrepreneurs such as age and human capital, and other managerial characteristics such as management and employee discipline issues, explain more than 20 percent of the revenue gap even though the coefficient is not significant. Nonetheless a gap of more than 25 percent of the gap remains unexplained.

Overall, the performance of firms is an important factor in explaining income differences between top-performers and constrained gazelles while it explains less than 15 percent of income differences between top-performers and survivalists (Tables 6 and 7). The results of the Blinder-Oaxaca decomposition are in line with the previous findings that well performing informal firms (the top-performers) are less poor and exhibit less income inequality relative to constrained gazelles and survivalists.

Table 6. Blinder-Oaxaca Decomposition of Per Adult Equivalent Income, Top-performers versus Constrained Gazelles

Per adult equivalent consumption, Log	Coefficients	Std. Err.	P > z
Differential			
Mean prediction (top-performers)	13.180	0.044	0.000
Mean prediction (gazelles)	13.050	0.019	0.000
Difference	0.120	0.048	0.009
Explained			
Value added per hour worked, Log	0.065	0.017	0.000
Other factors	0.026	0.018	0.154
Total	0.091	0.024	0.000
Unexplained	0.033	0.046	0.468
Observations = 1427			

Sources: Calculations using 1-2-3 survey data, 2012.

Table 7. Blinder-Oaxaca Decomposition of Per Adult Equivalent Income, Top-performers versus Survivalists

Per adult equivalent consumption, Log	Coefficients	Std. Err.	P > z
Differential			
Mean prediction (top-performers)	13.180	0.044	0.000
Mean prediction (survivalists)	12.890	0.014	0.000
Difference	0.290	0.046	0.009
Explained			
Value added per hour worked, Log	0.039	0.016	0.018
Other factors	0.117	0.026	0.000
Total	0.157	0.029	0.000
Unexplained	0.133	0.049	0.007
Observations = 1889			

Sources: Calculations using 1-2-3 survey data, 2012.

In the next step, the drivers of the performance of firms are analyzed. The following regression explains the value added per hour worked by a number of characteristics of the entrepreneurs such as human capital including educational attainment, financial knowledge, a proxy for social networks and proxies for management skills. The Equation 6 is as follows:

$$\begin{aligned} \ln vah_i = & \alpha_0 + \alpha_1 \ln K_i + \alpha_2 \text{nocapital}_i + \alpha_3 X_i + \alpha_4 H_i + \alpha_5 M_i + \alpha_6 S_i \\ & + \alpha_7 R_i + \varepsilon_i, \end{aligned} \quad (7)$$

where the explained variable $\ln vah_i$ is the log of the ratio of value added with the total hours worked in firm i . $\ln K_i$ stands for the log of physical capital. Given that in our sample some firms do not possess any capital, a dummy variable nocapital_i that equals 1 when the firm has no capital is also introduced. This allows us to avoid obtaining unbiased estimates of the returns to capital when some firms have zero observations of capital (Grimm, Knorringa and Lay, 2012). The vector X_i includes standard characteristics of the firm's owner, (age squared, gender and firm age). The vector H_i includes educational attainment, financial knowledge, reading knowledge and a dummy taking the value 1 if the owner is member of a business association. The vector M_i includes proxies for firm's management skills such as lack of customers, lack of credit access, management issues, employees' discipline problems, and problems of location and lack of equipment. The equation also includes sector dummies S_i and region fixed effects R_i .

Equation (7) explains the performance drivers of the three groups of firms. Figure A.1 in appendix shows that the higher a firm's value added per hour worked, the better the firm performs. The results are displayed in Table 8. The first column shows the results when we estimate Equation (7) with only some basic characteristics of the firm's owner, including education. The second column includes the variables estimated by vectors X_i and H_i . And the third column displays the results if one estimates the equation with all the variables.

The estimated capital elasticity is about 0.115 (column 3). The results also show that the older the firm's owner, the better is the firm's, but at a narrowing rate. Human capital, including secondary educational attainment, vocational training, and reading knowledge, is significantly and positively correlated with the performance of a firm. For instance, some other postprimary educational attainment, such as vocational training, raises the performance of firms by around 41 percent (column 1).

Membership in a business association is another important factor in the success of firms. The availability of a professional network seems to be a key advantage that can help entrepreneurs gain access to resources such as information and credit.

⁶ The regressions on the drivers of firm performance rely on Cobb-Douglas production function.

Table 8. Drivers of Performance

Dependence variable: Log of value added per hour worked	(1)	(2)	(3)
Capital, log	0.112*** (0.019)	0.105*** (0.017)	0.115*** (0.016)
No capital	0.943* (0.486)	0.888* (0.511)	1.268*** (0.471)
Age	0.0585*** (0.011)	0.0571*** (0.010)	0.0579*** (0.009)
Age squared	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Male (=1)	0.0497 (0.0693)	0.0459 (0.0626)	0.0553 (0.0661)
N0 diploma	Ref.	Ref.	Ref.
Primary completed	0.037 (0.124)	0.034 (0.117)	0.005 (0.123)
Some secondary	0.227** (0.116)	0.223** (0.110)	0.209* -0.121
Post-secondary	0.222 (0.138)	0.211 (0.131)	0.180 (0.144)
Other post primary	0.344* (0.191)	0.325 (0.224)	0.261 (0.262)
Can read	0.331*** (0.087)	0.320*** (0.092)	0.300*** (0.097)
Finl age	0.036** (0.018)	0.0354** (0.0179)	0.043** (0.019)
Financial Knowledge		0.147* (0.0783)	0.171** (0.075)
Business association		0.140* (0.081)	0.169* (0.090)
Lack of customers			-0.134** (0.060)
Lack of credit access			-0.037 (0.067)
Management problems			-0.109 (0.069)
Location problem			-0.106 (0.075)
Lack of equipment			0.169** (-0.074)
Discipline problema			-0.430*** (0.125)
Constant			5.544*** (0.232)
Sector fixed effects			yes
Region fixed effects			yes
R-squared			0.106
Observations			3,119

Notes: Standard errors in parentheses *** p < 0.01, ** p < 0.05, * p < 0.1.

All the coefficients associated with the factors used as proxies for management skills have the expected signs except the variable for the lack of equipment, which may not be a great problem for those entrepreneurs providing services such as hairdressers or luggage carriers. In any case, building performance-enhancing skills is always central to enhancing productivity and reducing inequality and poverty among informal firms. Entrepreneurs in the informal sector, especially survivalists, need formal (secondary) education. Including constrained gazelles and top-performers, they also require specific vocational training programs to upgrade their business skills.

Employees' discipline issues and the lack of customers are the most relevant challenges to the good performance of informal firms in the Democratic Republic of Congo. Several sectors exhibit high growth potential. Besides construction and other household service activities (including hairdressing), apparel, and mining and other primary activities in production and the retrieval of raw materials are the sectors in which entrepreneurs are most likely to increase their productivity. Moreover, the analysis shows that, relative to survivalists, more top-performers and constrained gazelles operate in these sectors. These results are robust to three other measures of firm performance, namely, value added per worker, labor productivity and real labor productivity (Table A9).⁷

7. CONCLUSION AND POLICY RECOMMENDATIONS

Using survey data, this paper contributes to the empirical literature in three main areas. First, the paper examines poverty and income inequality within three distinct categories of firms that were empirically identified. Second, using a Blinder-Oaxaca decomposition, it explains income differences across the three groups. Third, the paper analyzes the drivers of performance and profitability among informal firms.

The paper shows that income inequality among informal firms is explained primarily by disparities in educational attainment and infrastructure, including the lack of access to credit. Vocational training raises the odds of being in the top quintile among constrained gazelles. Access to finance, including financial knowledge and access to credit, is a more important factor among constrained gazelles and survivalists than among top-performers. The results also show that performance is a key factor in explaining income differences between informal firms. Additionally, the paper finds that human capital and managerial skills are important drivers of firm performance.

Policies to reduce poverty in the country should be tailored to each of the three groups of firms in the informal sector. For the survivalists, the priorities should involve

⁷ Labor productivity is the ratio of production to the number of workers. Real labor productivity is the ratio of turnover to the number of workers.

providing vocational training to improve technical and managerial skills and raise the sense of entrepreneurship. Improving financial literacy and easing the access to finance are also essential for this group. Unlocking the potential of the constrained gazelles would require facilitating the access to credit and supplying training programs to enhance managerial skills. The assistance for top-performers should seek to strengthen managerial skills and improve the access to infrastructure, enabling these firms to expand markets.

APPENDIX

Box A1. Probit Model with Selection Bias – Heckman Two-Step Procedure

The paper used the Heckman procedure to correct for sample bias. The first step consists in estimating the probit model (Equation 1) i.e. calculating the probability that a firm has capital and then obtaining the linear predictors from the model. The second step consists in calculating the inverse Mills ratio and including it in Equation (1). The inverse Mills ratio is the ratio of the probability density function and the cumulative distribution function divided by the standard error of the probit estimation. This ratio is calculated as follows:

$$\widetilde{MR} = \frac{\phi(X'_i \alpha_1 / \widehat{\sigma}_\omega)}{\Phi(X'_i \alpha_1 / \widehat{\sigma}_\omega)}$$

The probability density and cumulative distribution functions are such that:

$$\Pr(\text{Capital} = 1 | X) = \int_{-\infty}^{X \alpha_1} \phi(t) dt = \Phi(X \alpha_1).$$

Other Tables and Figures of the Appendix are available online at:
https://jed.cau.ac.kr/archives/47-4/47-4-3_appendix.pdf

REFERENCES

AfDB - African Development Bank (2013), “Recognizing Africa’s Informal Sector Championing Inclusive Growth across Africa” (blog), March 27. <https://www.afdb.org/en/blogs/afdb-championing-inclusive-growth-across-africa/post/recognizing-africas-informal-sector-11645/>.

- Baker, G.P. and R.E. Kennedy (2002), "Survivorship and the Economic Grim Reaper," *Journal of Law, Economics, and Organization*, 18(2), 324-361.
- Blinder, A.S. (1973), "Wage Discrimination: Reduced Form, and Structural Estimates," *Journal of Human Resources*, 8(4), 436-455.
- Bosch, M. and W.F. Maloney (2010), "Comparative Analysis of Labor Market Dynamics Using Markov Processes: An Application to Informality," *Labour Economics*, 17(4), 621-631.
- Chong, A. and M. Gradstein (2004), "Inequality, Institutions, and Informality," Research Department Working Paper 516 (September), Inter-American Development Bank, Washington, DC.
- Cunningham, W.V. and W.F. Maloney (2001), "Heterogeneity among Mexico's Microenterprises: An Application of Factor and Cluster Analysis," *Economic Development and Cultural Change*, 50(1), 131-156.
- De Soto, H. (1989), *The Other Path: The Invisible Revolution in the Third World*, New York: Harper and Row.
- Diao, X., J. Kweka and M. McMillan (2018), "Small Firms, Structural Change and Labor Productivity Growth in Africa: Evidence from Tanzania," *World Development*, 105(C), 400-415.
- Fields, G. S. (2004), "Dualism in the Labor Market: A Perspective on the Lewis Model after Half a Century," *Manchester School*, 72(6), 724-735.
- Gindling, T.H. and D. Newhouse (2014), "Self-Employment in the Developing World," *World Development*, 56(C), 313-331.
- Grimm, M., P. Knorringa and J. Lay (2012), "Constrained Gazelles: High Potentials in West Africa's Informal Economy," *World Development*, 40(7), 1352-1368.
- Grimm, M, J. Krüger and J. Lay (2011), "Barriers to Entry and Returns to Capital in Informal Activities: Evidence from Sub-Saharan Africa," *Review of Income and Wealth*, 57(s1), S27-S53.
- Grimm, M., J. Lay, F. Roubaud and J. Vaillant (2011), "Informal Sector Dynamics in Times of Fragile Growth: The case of Madagascar," Washington, DC: World Bank. <http://documents.worldbank.org/curated/en/357541468337172457/Informal-sector-dynamics-in-times-of-fragile-growth-the-case-of-Madagascar>
- Heckman, J. J. (1979), "Sample Selection Bias as a Specification Error," *Econometrica*, 47(1), 153-161.
- Loayza, N.V. (1996), "The Economics of the Informal Sector: A Simple Model and Some Empirical Evidence from Latin America," *Carnegie-Rochester Conference Series on Public Policy*, 45(1), 129-162.
- Lukiyanova, A. (2015), "Earnings Inequality and Informal Employment in Russia," *Economics of Transition*, 23(2), 469-515.
- McKenzie, D. and Y.S. Sakho (2010), "Does It Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability," *Journal of Development Economics*, 91(1), 15-24.

- Nichter, S. and L. Goldmark (2009), "Small Firm Growth in Developing Countries," *World Development*, 37(9), 1453-1464.
- Oaxaca, R.L. (1973), "Male-Female Wage Differentials in Urban Labor Markets," *International Economic Review*, 14(3), 693-709.
- Rogerson, C.M. (1996), "Urban Poverty and the Informal Economy in South Africa's Economic Heartland," *Environment and Urbanization*, 8(1), 167-179.
- Winkelried, D. (2005), "Income Distribution and the Size of the Informal Sector," Working paper, St. John's College, University of Cambridge, Cambridge, UK.
- World Bank (2016), "The Democratic Republic of Congo Poverty Assessment," World Bank, Washington. DC.

Mailing Address: Djeneba Doumbia, The World Bank, 701 18th St NW, Washington DC 20006, USA, Email: ddoumbia@worldbank.org

Received April 23, 2016, Revised March 15, 2022, Accepted June 27, 2022.