



**SERIE  
DOCUMENTOS  
DE TRABAJO**

No. 283

Julio de 2022

# Informal versus Formal: Microfirms' Productivity Gaps

---

**Luis Hernando Gutiérrez**

**Paul Rodríguez-Lesmes**

# Informal versus Formal: Microfirms' Productivity Gaps

**Luis Hernando Gutierrez**

luis.gutierrez@urosario.edu.co

Universidad del Rosario

School of Economics

Calle 12C, No. 6-25, Bogotá,

Colombia

*Corresponding author*

**Paul Rodriguez-Lesmes**

paul.rodriguez@urosario.edu.co

Universidad del Rosario

School of Economics

Calle 12C, No. 6-25, Bogotá,

Colombia

## Abstract

Although evidence of a productivity gap between formal and informal firms is observed, this 'formality premium' is less explored for microfirms. The informality of microfirms is a central concern in low- and middle-income countries, and a crucial demand is noted for designing policies addressing this issue because they are the bulk of the economic tissue. We fill this void by estimating a productivity premium for the case of Colombia, considering two margins of informality: extensive, referring to business registration, and the intensive, which includes as well labor regulations. We use a unique longitudinal dataset from the Microenterprise Survey by the Colombian Statistics Department, which follows approximately 39,000 micro-establishments with up to 9 employees during 2012–2016. We utilize the transition into and out of formality to estimate the productivity premium (yearly sales per worker) between informal and formal firms, thereby exploring differences concerning initial productivity. We use a fixed-effects quantile regression to explore differential effects along the productivity distribution. We find evidence of a premium for both the extensive (20%) and intensive margins (6%), a gap that is decreasing along with the firm's productivity. The evidence of these premiums is related to two growth strategies of firms: an increase in capital investments for the extensive margin and an increase in human capital quality for the intensive margin. Further, we find the premium is notoriously wider for young firms (less than three years in the business) with a steeper gradient. We do not find systematic differences across sectors, gender of the owners, and motivation. These results are new evidence that supports the existence of a premium and the transition into and out of formality of microfirms in middle-income countries. Moreover, they suggest that microfirms' formalization and growth policies should be oriented toward promoting and enhancing formality's benefits.

Keywords: microfirms, firm informality, productivity premium, Colombia

# 1 Introduction

Micro-firms are a pervasive form of business organization that contribute substantially to the GDP and employment. In developing countries, the universe of those business units is integrated by formal and informal firms, both producing in the same industries regardless of their size. While a common argument for this was that these firms are essentially different (dualistic view), this theory was rebuffed by the coexistence of formal and informal micro-firms for similar size and productivity (Ulyssea, 2020a). Therefore, the motivations for the existence of informal micro-firms require a more detailed analysis. In particular, would an informal firm which becomes formal, obtain a productivity premium?

There is substantial evidence of the existence of a productivity gap between formal and informal firms (Allen et al, 2018) Hsieh and Olken 2014, Maloney 1999, Meghir, Narita and Robin 2015, Ulyssea 2018). The literature argues that potential reasons for the productivity advantage is because, among other things, formal firms are able to (i) invest in technology (Allen et al.); (ii) obtain better access to credit conditions (Hsieh and Olken 2014); and (iii) hire better human capital which is linked to higher managerial capital (Meghir et al, 2015). Yet, for micro-firms such evidence is scant. Microfirms, especially in developing countries, face important constraints that limit their ability to attain the potential benefits of formality. Access to capital markets, low human capital, and lack of managerial skills by owners of microfirms translate into low employment quality (La Porta and Schleifer, 2014, Perry et al., 2009). Also, costs of formality are proportionally higher, as it is the case of the (excessively) alleged tax burden. The high costs are, in some cases, supplemented with red-tapism, which do not allow firms to exploit their growth potential. Documenting if a productivity premium exists, and how it is related to characteristics linked to firm-growth potential, can help to provide elements for the design of successful formalization and productivity enhancement programs and policies, such

as training courses, the provisions of subsidies and credits, and other government initiatives.

This work aims to measure the productivity premium obtained by microfirms once they transit from informality to formality, allowing for differential effects according to their base productivity level. One of the main constraints to answering the questions above is the lack of suitable longitudinal data for formal and informal (micro) firms<sup>1</sup>. This paper uses a unique *longitudinal* survey from 2012 to 2016 for about 39000 micro-establishments. The sample survey includes only firms with less than 10 employees that operated in fixed premises, were located in urban areas and that produced in the retail, manufacturing and service industries. We estimate the premium using a quantile fixed effects regression (Powell, 2016, 2020). The data includes information about firms' status of formality and whether firms comply with the payment to their employees of social security regulations.

We calculate *two* margins of informality suggested by Ulyssea (2018) and used recently in the informality literature (Cisneros 2022, **Ponczek** and Ulyssea, 2020, Samaniego and Fernández 2020 and Ulyssea 2020b): (i) the *extensive margin* meaning that micro-firms have a business registry, a tax registry, and keep accounting records. That is, micro-firms paid all the entry costs; and (ii) the *intensive margin* in which firms that comply with (i) do not comply with labor regulations. That is, micro-firms did not pay social security benefits to their employees. The introduction of these two margins is an outstanding innovation. Most research on firm informality has generally used the extensive margin of informality measured with a binary decision to comply or not with (business) regulations<sup>2</sup>. However, the creation of the intensive margin goes beyond viewing informal firms as units that do not comply with the entry costs of operating but as a unit that operates as formal without incurring the cost related to paying social security benefits to their employees. We observe

---

<sup>1</sup> Longitudinal data exist for small, medium, and large firms in the USA and Colombia as documented in Eslava, Haltiwanger, and Pinzón (2022).

<sup>2</sup> See Rauch (1991), De Paula and Scheinkman, (2010).

that productivity (yearly sales per worker) distributions across sizes for both types of margins show an overlapping region between formal and informal microfirms, which goes in line with the coexistence of both types of firms documented elsewhere (Ulyssea, 2020).

In this study, the longitudinal nature of the dataset provides us with a crucial source of variation for the identification of the parameters determining the premium productivity dynamics: the *transitions into and out of formality* according to the two margins. We document that the extensive margin is a stepping-stone for the intensive margin. Also, most firms which became formal under the extensive margin remain formal for the rest of the study, while the intensive margin seems to be a more transitory state. We also observe substantial variation on workforce size, including transitions in and out self-employed status (one solo worker).

We find that for microfirms and using the *extensive* margin, a productivity premium on the mean of 18% that fades away after the 40th quantile being almost zero at the 95th quantile. For the *intensive* margin, the average premium is only about 6% but for several quantiles there is no evidence of a premium favoring formal microfirms. This is a crucial result that shows that for microfirms which transit into formality, being *formal* provides larger returns in productivity for those that are less productive. We also show that the premium is larger for new firms (less than 3 years old), and for necessity-motivated firms. The premium was stable across sectors, and gender ownership composition.

We show that the premium is attained under two distinctive growth strategies of firms. The extensive margin premium is linked to an expansion in capital investments (reflected on higher loans' uptake) and in workforce. In contrast, the intensive margin is linked to higher demand on human capital, as formalization is observed along with increases of the mean wage, no changes on loans' uptake, and even a reduction of workforce size.

The main contribution of this paper is to the literature on the existence of a microfirm's productivity premium. Recent research using both the extensive and

intensive margins has looked at, for instance how trade liberalization affects informality (Cisneros 2022, Ponczek and Ulyssea 2020b), and how labor inspections affect the costs of employing workers informally (Samaniego de la Parra and Fernández 2021). In those studies, the research focuses on analyzing how increasing the costs of informality affects the macro or micro level of informality. Several experimental and quasi-experimental studies<sup>3</sup> also aimed to measure the impact of becoming formal by varying benefits' perceptions and costs (see Floridi et al., 2020; Jessen & Kluge, 2021; Ulyssea, 2020a for recent reviews and meta-analysis). The common finding of those studies is that regardless of the type of (combined) interventions the effectiveness of the intervention in increasing microfirms' productivity has been limited and tends to disappear in the short run. We are measuring a different empirical object; rather than the average treatment effect of formalizing 'any' firm, we are measuring the premium attained by firms which self-select into formality. Given this, our contribution is threefold. First, in our study, by using the two margins of formality, we are able to document the existence of differentiated effects of the productivity premium. Second, we are able to explore differential effects according to ex-ante heterogeneity of microfirms, reflected on the existence of a wide productivity distribution, by using quantile regressions. Third, it allows us to measure what other contemporaneous decisions were undertaken by firms which decided to become formal, giving us the opportunity to explore potential mechanisms.

The rest of this paper is organized as follows. In section two of this paper, we present the dataset and Colombian context. Section three describes the empirical strategy of this study along with its limitations. In section four, we present the results and finish with the discussion and conclusions in sections five and six.

---

<sup>3</sup> Some types of the interventions are: i) provide information to treated firms on the benefits of formalization and how the process works, ii) provide information plus helps in registration, iii) provide information plus helps in registration plus reimbursement of direct costs, iv) registration and tax reduction or tax services, v) enforcement

## 2 Data and motivating facts

Firms in Colombia are classified as microfirms if they have less than 10 workers. As in most developing countries (Perry et al. 2009, La Porta and Shleifer 2008, Ulyssea 2018), they represent most of the Colombian firms. In 2017, 96.5% of all registered firms in the country are microbusiness. The 2020 Economic Census reports that in Colombia there were 5.8 millions of microfirms of which 857 thousands were businesses with fixed locations (micro-establishment).<sup>4</sup> By September 2017, in the main cities of the country, 43.9% of the working force was employed in a micro-establishment.<sup>5</sup>

The Micro-establishments Survey (MS) conducted annually from 2012 to 2016 by DANE, Colombia's official statistics agency, is a *longitudinal* study designed to study the evolution of the dynamics of a sample of 30.159 of microfirms that operated in fixed premises, which belong to the wholesale or retail trade (61%), services (29%), or industry (10%) sectors in the main 24 urban areas of the country.<sup>6</sup> The survey includes both self-employed workers (41.9%) and microfirms: two workers (29.4%), three to five workers (23.9%) and six to nine workers (4.8%). The firms should have been operating in fixed premises for at least a year prior to their inclusion in the study; they should be traced over time by the same commercial name, brand, or identification number; and were performing the same commercial activity.

### 2.1 Formality

---

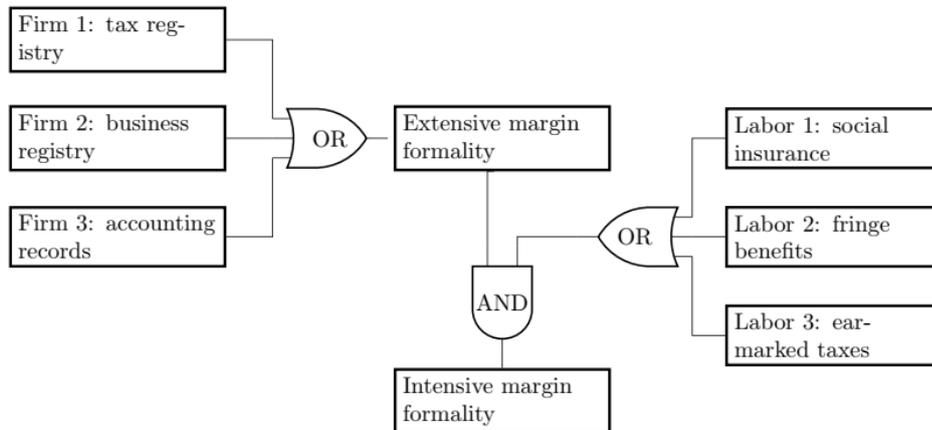
<sup>4</sup> Further details of the context of micro-establishments in Colombia can be found in Ortiz et al (2021).

<sup>5</sup> More precisely, micro-establishments employ 22.38% of the workforce as self-employed, 11.3% in firms of 2 to 3 workers, and 10.17% in firms of 4 to 10 workers. From the entire microfirms workforce, 71.9% of the workers are in micro-establishments; and for the specific case of self-employed, 62.7% of them constitute a micro-establishment. These are our own calculations using the work force survey (GEIH). In this survey, the questions for firm size do not allow to disentangle firms of 1-10 workers, and 1-9 workers (the official definition of a microfirm).

<sup>6</sup> Specifically firms belong either to a specific ISIC rev 4 category: (i) wholesale trade, retail trade and sale of motorcycles and their accessories, maintenance workshops and repair of motor vehicles; (ii) all service establishments not including financial, transportation, public education; or (iii) all micro-industry.

The formalization of microenterprises has been a recurring theme in research entrepreneurship theory and economic development theory perspectives (La Porta and Shleifer, 2014, Ulyssea, 2018, 2020a).<sup>7</sup> A persistent aspect of research on the informality of microenterprises has been to consider compliance by these productive units with legal requirements to operate. Some scholars have highlighted that the phenomenon of informality-formality is not a zero or one situation but rather a continuum wherein different degrees of (in)formality can coexist (Perry et al. 2007, Trebilcock 2004). We follow Ulyssea (2018, 2020a, 2020b) and distinguish between two margins of formality: (i) firms that comply with the firm-level legal requirements of entry (extensive margin); and ii) firms that comply with (i) and with labor regulations (intensive margin). Diagram in Figure 1 summarizes the definitions, which are explained below.

Figure 1: Lenient informality definitions



In the case of Colombia, with respect to firm-level factors, a firm can be considered *fully* formal if the following requirements are attained. The first is whether the productive micro-unit has a registry with the country's tax authority, which in Colombia consists of a single tax registry (RUT for its Spanish acronym). A second firm factor is whether the company has a business registry and if it had renewed it at the time of the survey. In Colombia, this registration is granted by the chamber of

<sup>7</sup> Further details on the formalization process are available at the official formalization guidelines (DNP, 2019).

commerce of each city. The last aspect that some investigations consider is the keeping of accounting records (Morón, Salgado, and Seminario, 2012). A firm is typically considered formal in the extensive margin (*lenient* definition) if at least *one* of these requirements is fulfilled. On top of this definition, we consider a *strict* version where the three requirements are fulfilled. In Figure 1, the "OR" operators will be replaced by "AND" of this *strict* version. According to the lenient [strict] definition, 76.5% [35.68%] of self-employed-year observations are formal. The figure is higher for microfirms: 93.32% [64.62%].

Panel A of Table 1 presents the mean of microfirms characteristics according to their formality status, considering the *lenient extensive margin* definition (Tables A1 and A2 in the appendix presents the other definitions). Among formal microfirms, the most common formality requirement is complying with the tax registry (85.7% self-employed; 91.9% microfirms), and next keeping accounting records (73.7%; 88.0%). Compliance with business registry is lower for both self-employed and micro (62.7%; 78.7%).

Concerning the labor regulations, on top of paying wages, these businesses must contribute a part of their payroll taxes to the social security system (health and pensions), mandatory fringe benefits and payments, and other earmarked payroll taxes. The low percentages are observed because some of these requirements are not applicable to certain types of labor contracts as well as because of the informality of labor contracts.<sup>8</sup> Colombian labor statistics show that of all employees in the country, 5 out of 10 do not have any type of contract, and only 2 have a written contract (López-Mera, 2021). Following Figure 1, a firm is lenient formal under the intensive

---

<sup>8</sup> The service-based contract was designed to hire solo-professionals that provide specific tasks given a deadline without the direction of the employer, and without a daily shift in specific premises. For instance, an accountant could be hired only to compute yearly taxes; a lawyer to handle a specific lawsuit; or a computer scientist to develop a function for a project. Under this idea, the professional is an independent contractor, self-employed (ILO, 2006). Yet, such a contract is used to conceal the true labor status of workers, even by the Government itself (Reyes, 2020).

margin if it is formal in the extensive margin, AND complies with at least one labor formality requirement. As above, there is a *strict* definition (*all* labor requirements).

Panel A of Table 1 also shows that among formal (extensive-wise) firms, the most common factor is to contribute to social insurance (17.7%; 16%). Compliance with the other two factors is half of it for self-employed (8.6%), and slightly less for microfirms (13%). These factors are uncommon for informal firms in the extensive margin (less than 3% for any factor). For this reason, establishing the extensive margin definition as the basis of the intensive margin is not particularly restrictive. Formality prevalence drops notoriously if we consider the intensive margin: it is only 13.71% [4.9%] for self-employed, and 15.28% [10.41%] for microfirms.

## **2.2 Formality and firms characteristics**

Panel B of Table 1 also presents the means of several firms' characteristics according to formality status.

Some stylized facts are observed. First, formal firms are more likely to be in the Services sector, less in Retail and even less in Manufacturing<sup>9</sup>. In this last type of activity, specific and more costly assets are needed and then may be a barrier to enter. Microfirms are more common than self-employed in Manufacturing. Second, regardless of the age group and their formality status, most of the business units, either self-employed or microfirms with employees, are old or with more than five years old. Furthermore, as businesses are younger, the share of microfirms becomes almost always smaller. This pattern suggests both a low entry and exit rate, and is different from the pattern found by Ulyssea in Brazil. He found that on average formal firms were older than informal implying a higher turnover in the informal sector. Third, regarding size groups, the greatest proportion of microfirms, either formal or informal, has two or less employees, although there exists a large difference in favor of informal ones. The group of firms with 6 to 9 workers is the smallest one not

---

<sup>9</sup> Ulyssea (2018) found that Brazilian microfirms were predominantly in Retail (49.1%) and Services (42.4%) and only 8.4% in Manufacturing.

reaching the 1% for informal microfirms and only 7% for formal. This pattern suggests a low rate of growth among microfirms, and hence a very low possibility they can become “gazelles”. Not surprisingly, Eslava et al (2022) found that for non-micro plants, “old (and frequently small) plants account for a larger fraction of employment in Colombia ”

Formal self-employed are more likely to be women, but formal microfirm are less likely to be owned by women. What is clear is that formal firms are almost twice more likely to be the result of a perceived investment opportunity, instead of other motives (ex. not finding a formal job or family inheritance).

The adoption of technologies follows an expected pattern. Formal firms are more likely to use information and communication technologies (both presence and ownership of devices), specially microfirms. However, if we consider usage of the financial services, there is a *striking* pattern. Informal firms are more likely to have requested and obtained loans, but this relationship is inverted if we consider loans obtained through the formal financial system.

A final important aspect to consider is the wage-level. Formal self-employed earn on average a 7% above the minimum wage, while informal ones earn almost 17% less than it. If we consider microfirms; extensive-margin formal wages represent 76.6% of the minimum wage, but for informal firms the proportion drops to 47%. These numbers far from the minimum wage are possible as firms can hire individuals for less than a year, but also because they can hire them informally, i.e., not paying the labor costs. The ability to reduce labor costs is usually argued as one of the main reasons firms do not formalize (Ulyssea, 2018). This wage bonus of formality (Meghir et al. 2015, Radchenko 2015) might be the result of a bonus in productivity, which we explore below.

Table 1: Proportion of firms according to characteristics and *extensive* margin lenient formality status

	Self-employed			Microfirm		
	Informal	Informal	Formal	All	Informal	Formal
<b>Panel A. Formality</b>						
Formal Extensive Lenient	0.765		1	0.933		1
Formal Extensive Strict	0.357		0.466	0.646		0.692
Firm-level factors of formality						
Tax registry	0.656		0.857	0.857		0.919
Business registry	0.48		0.627	0.735		0.787
Keeps accounting records	0.564		0.737	0.821		0.88
Formal Intensive Lenient	0.137		0.179	0.153		0.164
Formal Intensive Strict	0.049		0.064	0.104		0.112
Labor factors of formality						
Contributes to social insurance	0.142	0.027	0.177 *	0.15	0.009	0.160 *
Fringe benefits and payments	0.065	0.002	0.084 *	0.12	0.003	0.129 *
Pays firm-specific wage taxes	0.067	0.004	0.086 *	0.122	0.003	0.130 *
<b>Panel B. Firm characteristics</b>						
Economic sector						
Manufacture	0.063	0.073	0.059 *	0.128	0.098	0.130 *
Retail	0.705	0.726	0.698 *	0.551	0.62	0.546 *
Services	0.233	0.2	0.243 *	0.321	0.282	0.324 *
Age of the firm in years						
Less than 1	0.03	0.031	0.03	0.025	0.039	0.024 *
More than 1, less than 3	0.167	0.161	0.169 *	0.14	0.161	0.139 *
More than 3, less than 5	0.16	0.129	0.170 *	0.152	0.131	0.154 *
More than 5, less than 10	0.245	0.188	0.262 *	0.26	0.201	0.264 *
More than 10	0.398	0.492	0.369 *	0.423	0.468	0.419 *
Size of the firm						
2 workers				0.541	0.794	0.523 *
3 to 5 workers				0.39	0.199	0.403 *
6 to 9 workers				0.069	0.007	0.074 *
Female ownership	0.619	0.607	0.623 *	0.684	0.717	0.681 *
Firm based on a perceived opportunity	0.598	0.385	0.663 *	0.709	0.438	0.728 *
Loans						
Requested any	0.26	0.278	0.255 *	0.269	0.311	0.266 *
Requested formal	0.193	0.167	0.202 *	0.216	0.204	0.216 *
Obtained	0.245	0.262	0.239 *	0.252	0.297	0.249 *
Information and comm. technologies (ICT)						
Ownership of devices	0.179	0.048	0.220 *	0.425	0.073	0.450 *
Web presence	0.142	0.029	0.177 *	0.391	0.047	0.416 *
Relative♣ monthly wages	1.061	0.833	1.074 *	0.755	0.47	0.766 *
Relative♣ annual sales per worker	45.72	26.85	51.50 *	51.71	26.68	53.50 *

Notes: Own calculations using MS. Business registry: annual registration in the local chamber of commerce. Tax registry. Social insurance: employees and employers payroll taxes to fund health and retirement schemes. Firm-specific wage taxes: a proportion of the wage that the firm contributes to specific public sector institutions

(*parafiscales*). Female ownership: proportion of the owners who are women. Mandatory fringe benefits and payments include paid-holiday leaves, severance aid, service bonus, among others. ♣ Relative to the minimum wage of the year. The star (\*) represents the significance level at the 95% level of a *t*-test of equality of means between formal and informal firms.

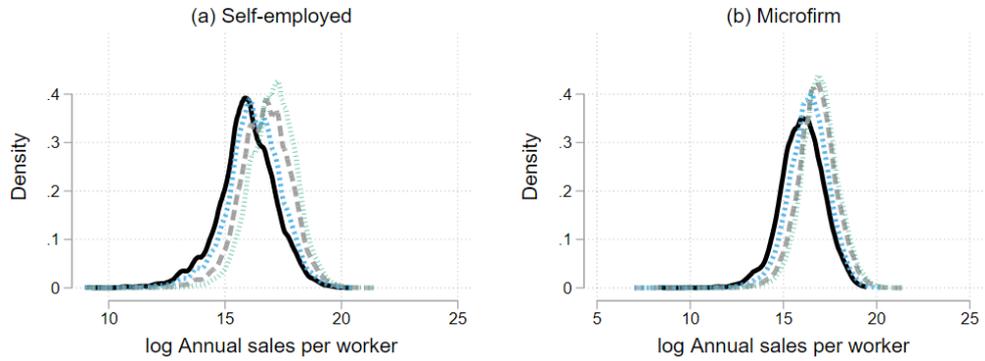
### 2.3 Productivity

As a proxy of labor productivity, we use the ratio of a firm's sales to total workers. Unfortunately, we could not account for firms' costs, and thus, any effects of firms' efficiency that can be captured through the estimation of a raw measure of the value added are outside the scope of this paper.

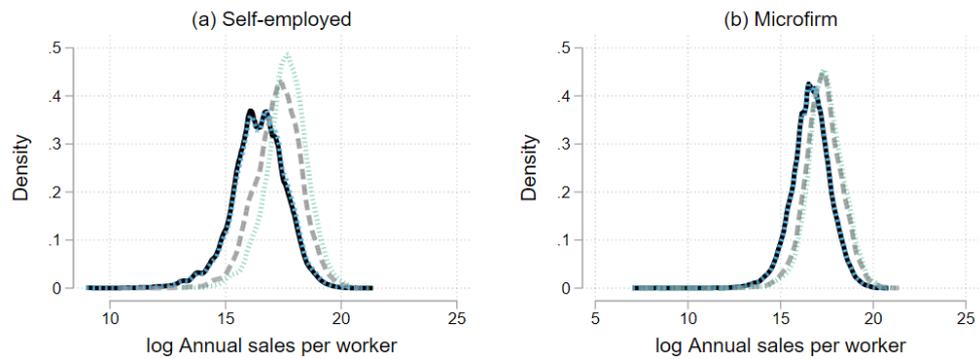
The last row of Table 1 presents this variable standardized by the monthly minimum wage. Formal sales mean productivity is twice for formal than informal firms, for both self-employed and microfirms. While this might suggest the presence of duality, Figure 2 presents the distribution of the logarithm of the measure according to formality status, in both the lenient and strict versions, and for both the extensive and intensive margin. Three facts are observed in the Figure. First, in all cases there is a large overlapping region between formal and informal microfirm size and productivity distributions. In other words, for most of the productivity domain, there are both formal and informal firms. Second, productivity distributions of formal firms are in all cases shifted to the right. Last, the mean is higher the stricter the formality definition is. These results corroborate the findings by Allen et al, 2018, Almeida and Carneiro 2012, Meghir, Narita and Robin 2015, and Ulyseia 2018 all for Brazil. Yet, all this does not tell us how different the productivity for a firm would be if it becomes formal, *after* being informal and vice versa. As our aim is to measure this, we study such transitions next.

Figure 2. Distribution of productivity by formality status

I. Extensive margin definition



II. Intensive margin definition



— Informal - Lenient    - - - Formal - Lenient  
 ··· Informal - Strict    ····· Formal - Strict

Notes: The graph presents the densities of log annual sales, a proxy of productivity, according to their formality status.

**2.4 Microfirm dynamics**

The present dataset allows us to explore microfirm dynamics. In particular, we explore transitions out of and into formality, and firm size in terms of the number of workers.

First, we explore formality transitions. For this analysis, we integrate the extensive and intensive margin definitions into a single variable.<sup>10</sup> Table 2 shows the transition matrices between formality and informality and vice versa according to three categories: 1) informal extensive, that is, firms that do not comply with any business regulations, 2) formal extensive -

<sup>10</sup> Tables A3 and A4 in the appendix presents the transition matrices by the extensive and intensive margins separately. The transition informal to formal is less common under the strict margin definition (Table A3). Transitions from informality to formality are larger for firms with two or more workers (microfirms) at t-1 (Table A4 in the appendix).

informal intensive, or the *extensive* margin defined by Ulyssea, and 3) formal intensive, that is firms that comply with both business registration regulation and labor regulations. The table illustrates stayers and switchers forwards and backward for the three constructs. The principal (primary) diagonal shows the percentage of firms that were stayers. A 61.8% of informal extensive microfirms were stayers, and a sizable portion of the remaining switched to a higher formality status, i.e., they started to comply with business regulations while not incurring labor costs mandated by regulations. Regarding formal extensive - informal intensive, a large percentage of firms, 81.1%, did not switch their status.

Additionally, the number of microfirms in this group was by far the largest (77 thousand). Unusually, 7.2% of firms transitioned to a *lower* formality status, i.e., informal - extensive. Again a non-negligible proportion of microfirms upgrade their status to the next one, i.e., formal intensive. Last, there are the microfirms that, at t-1, were formal - intensive. About 40.6% were stayers. The remaining percentage switched to a lower formality status, most (57.6%) to the extensive margin defined by Ulyssea. Table 3A in the appendix presents a supplementary view of the transitions between formality status. The transitions of microfirms from informal to formal are less common (about 5%) under the strict margin definition. Also, some 6270 firms that were formal in a year switched their status to informal under that margin.

Characteristics associated with the transition into formality are explored in table A5 in the appendix. Using a linear probability model, we explore which of the firm characteristics are associated with the transition into formality of previously informal firms. As expected, a higher probability of transition to formality is observed for more productive firms, younger, larger (in terms of employees), constituted in response to a perceived opportunity, users of ICT, which pay higher wages.

A second main characteristic of the analysis to consider is size. In some of the analyses in the previous section above, we separated the data between self-employed and microfirms. Yet, firm size is an endogenous variable. Table 3 shows that there are substantial transitions between firm size categories.

We observe that 22.27% of the self-employed became microfirms with two workers in a year, and 4.91%, became microfirms with three workers or more.

Table 2: Formality transition matrices

		Status at <i>t</i>		
		[1]	[2]	[3]
Status at <i>t</i> - 1	[1]. Informal <i>extensive</i>	10,834 <b>61.76%</b>	6,383 36.39%	325 1.85%
	[2]. Formal <i>extensive</i> , informal <i>intensive</i>	6,906 7.23%	77,466 <b>81.09%</b>	11,154 11.68%
	[3]. Formal <i>intensive</i>	361 1.82%	11,452 57.60%	8,068 <b>40.58%</b>

Notes: Own calculations using DANE Micro-enterprise survey 2012-2016. Results are based on the strict definition (a firm is formal if it meets all requirements of being legally established).

Table 3: Firm size transitions

Size at <i>t</i> -1		Size at <i>t</i>			
		1	2	3-5	6-9
Size at <i>t</i> -1	1 (Self-employed)	37,882 <b>72.82%</b>	11,584 22.27%	2,430 4.67%	124 0.24%
	2	12,938 28.99%	23,309 <b>52.23%</b>	8,091 18.13%	287 0.64%
	3 to 5	2,641 8.43%	7,517 24.01%	18,820 <b>60.10%</b>	2,334 7.45%
	6 to 9	151 3.02%	314 6.29%	1,909 38.24%	2,618 <b>52.44%</b>

Notes: The survey excludes firms if they have 10 workers or more.

Furthermore, 29% of firms with two workers became self-employed in a year. 47.5% of those with 6 to 9 workers turned into either firms with 2 or 3 to 5 workers; and even self-employed (3.02%). Last, most microfirms with three or more workers keep their size or transit to a greater size.

A limitation of this dataset is that we cannot determine transitions from microfirm into small-firm or above (10 or more employees). The implications of this censoring will be discussed in the last section of the document.

### 3 Methods

#### 3.1 Empirical model

To estimate the impact of formality on our proxy of productivity, we exploit the variation over time on each firm, in other words, their transition between informal and formal status. The description of the data in the firm's dynamics section above show that these transitions are common. As a result, we will measure how much the productivity of a firm grows [shrinks] when a microfirm decides to become formal [informal], by comparing it with similar microfirms but which stay informal [formal].

We estimate the premium using a fixed-effects ordinary least squares model (FEOLS):

$$y_{it} = x'_{it}\beta + \gamma I_{it} + \alpha_i + \varepsilon_{it} \quad (1)$$

where  $I_{it}$  is an indicator of the formality status of the firm  $i$  in year  $t$ , and  $y_{it}$  is the logarithm of the annual sales per worker. Control variables correspond to vector  $x_{it}$ , the number of employees, a dummy indicating if at least one woman is a part of the owners of the firm, ICT hardware and usage dummies, ICT and web presence, age of the firm dummies (1-3, 3-5, more than 5 years), sector dummies (manufacturing, services, commerce), and year and region (*departamento*) level dummies. Dummies  $\alpha_i$  correspond to the fixed effects that capture both observable and unobservable firm

characteristics, which are time invariant. Therefore, coefficient  $\gamma$  corresponds to the productivity premium of formal vs. informal firms.

Following Nordman et al. (2016), the role of formality across the productivity distribution may exhibit crucial differences. Therefore, we implement a fixed-effects quantile regression (FEQR) that allows us to identify impacts across the distribution but under a similar identification strategy (within variation) as in equation 1. In this case, the dependent variable would be  $q_Q(y_{it})$ , which corresponds to the  $Q^{\text{th}}$  conditional quantile  $Q \in [0,1]$ . Therefore, a set of coefficients  $\gamma(Q)$  are estimated for 5<sup>th</sup> to 95<sup>th</sup> percentiles with steps of 5. We chose the estimator by Powell (2016, 2020), which considers nonadditive fixed effects. The advantage of this estimator is that the coefficients  $\gamma(Q)$  can be interpreted similarly to  $\gamma$  obtained with the FEOLS (i.e., the impact of formality on the  $Q^{\text{th}}$  quantile of productivity). The estimator is implemented using Baker, Powell, and Smith (2016) algorithm `qregpd` in Stata 16. Alternative estimators, such as Koenker (2004), which are based on additive fixed effects deliver estimates that their interpretation is based on the quantile of the outcome net of the individual fixed effect. Both distributions (and therefore results for the quantiles) might differ substantially if the individual effects are correlated with the productivity level.

To provide insights on why there is a premium (if any) for becoming formal, we undertake two additional sets of exercises. These exercises will allow us to explore potential channels explaining the productivity premium.

First, we assess if there are differences across firms according to characteristics typically linked to growth expectations. There exist many different types of firms with different growth patterns (Delmar & Wiklund, 2008; Coad, 2009; Coad and Tamvada, 2012). Delmar and Wiklund (2008, p. 196) argue that the lack of consensus in firm growth studies is failing to account for four primary firm features: firm size, firm age, type of industry, and firm governance.

Second, we explore whether firms' size and loan use change contemporaneously with the variation of formality status. Hermans et al. (2015), in their review of firm growth studies, mention that firms' expectations of growth or growth aspirations are referred to as the introduction of innovation, increase in the number of employees, and internationalization. While apart from workforce size, we do not observe these variables in detail, investments are proxy by loans uptake.

The analyses above are performed under both the extensive and intensive definitions. Yet, it is important to consider that identification for the extensive margin comes mainly from transitions *into* formality - the most common transition as seen in section 2-, and for the intensive margin from transitions *out-of* formality. Also, as extensive margin transitions are more common, standard errors are expected to be smaller than those of the intensive margin.

## 4 Results

### 4.1 General

Figure 3 presents the coefficients of both the FEOLS and FEQR with 95% confidence intervals (CI). Each graph shows, on the horizontal axis, the quantile of the dependent variable (productivity proxy: log annual sales per worker) and the value of the coefficients on the vertical axis. For the FEQR estimates, each dot corresponds to a  $\gamma(Q)$  estimate, and the gray area corresponds to the CI (diamonds in the case of the strict definition of informality). For the FEOLS coefficient, a solid horizontal line is presented with its corresponding CI in dashed lines: on the left, Figure 3a shows a premium of 18.9% (SE=0.00975) which corresponds to the lenient extensive margin definition; on the right, Figure 3b presents a premium of 11.0% (SE=0.00683), for the lenient intensive margin. Some of the points are not presented if the value could not be estimated because of the lack of observations or if the standard errors were excessively large for the figure. For example, the first point to the left (5th percentile) in Figure 3a shows that (extensive) formal microbusinesses had a productivity about

22.1% higher than informal microfirms of *similar* productivity levels in the previous year. These differences are identified by comparing changes in the outcome of firms that changed their formality status against those, which did not change it depending on experiencing similar time trends in terms of ownership, usage of ICTs, time in business, activity, and region. Table A6 in the appendix presents the coefficients behind these two graphs.

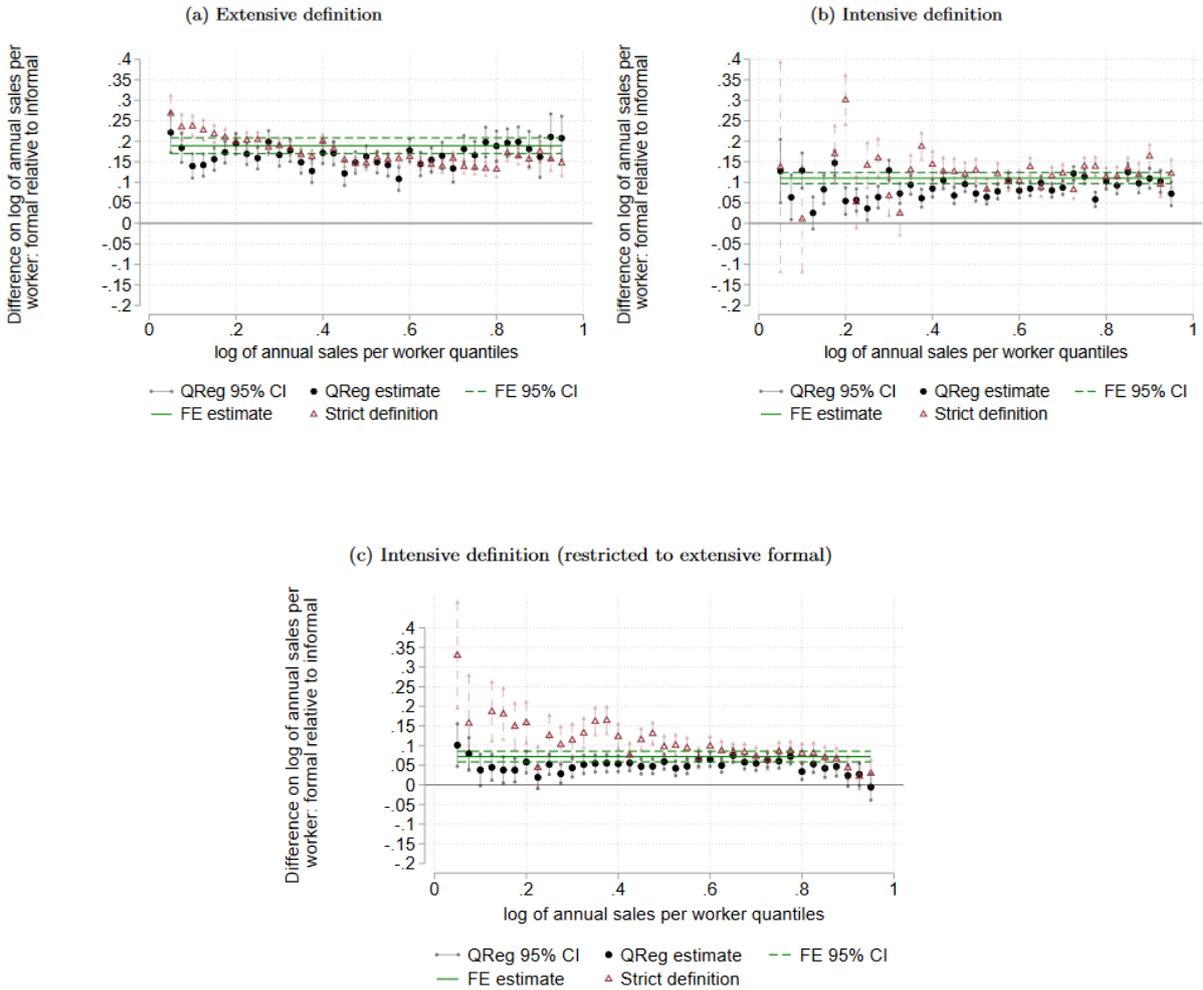
In figure 3a, we observed that under the lenient extensive formality definition, there is no gradient in terms of the productivity premium. While for some quantiles the point estimate is below the fixed effects confidence interval, there is no systematic pattern. However, for the strict definition, there is a decreasing pattern: firms with relatively lower productivity levels, are the ones that obtain larger benefits for their transition into formality. These results contrast with the findings of Nordman et al. (2016), who analyzed the earning gaps among categories of informal wage workers versus formal wage workers. When comparing formal wage workers with informal self-employed, Nordman et al. find a premium of about 12% in favor of the informal self-employed, which the researchers argue is due to their better unobserved entrepreneurial skills. They find a steep gradient along the earning distribution, which is not as clear in our exercise when considering firms.

In figure 3b, for the intensive margin definition, there is also no evidence of a gradient across productivity. Moreover, for the first quantiles, estimates are noisy, for both the lenient and strict definitions, and several of the estimates cannot be rejected to be equal to zero. Therefore, we perform a further analysis which consists in restricting the sample to extensive formal microfirms (Figure 3c). Then, we calculated the productivity premium between informal and formal microfirms using the intensive margin. The results confirm that the marginal gain of the intensive is much smaller compared to the case when there is no restriction for the extensive formal microfirms. Furthermore, with the strict definition, the results are more precise than those presented in Figure 3b: there exists a decreasing trend along with the productivity distribution, as the one found for the extensive margin.

The results from these analyses show clear evidence of the existence of a premium for becoming formal in terms of productivity. The result is expected as the returns of the investments that (extensive margin) formality allow are potentially larger: accessing credit in better conditions (amount and rates), becoming a provider of large firms and the government, or being a part of accelerators. We further explore this channel in section 4.3. The productivity gradient on the premium for the (strict) extensive margin might be related to the characteristics linked to the potential-growth of the firm. We explore this in section 4.2.

For the intensive margin, the premium is close to half of the one obtained for the extensive margin. To understand this, we should consider that prior to becoming formal in the intensive margin, firms typically become formal in the extensive margin (as discussed in section 2.4), thus, they have already obtained some of the benefits. A firm being registered with the local authorities and the tax system, including its employees in the social security system, provide a small marginal advantage in terms of productivity than to the benefits attained for registering the firm in the first place.

Figure 3: Informality premium on annual sales



Notes: The graph presents the coefficients of the informality premium and their 95% confidence interval after a fixed-effects quantile regression. Controls: the proportion of female ownership, ownership of ICT devices and web presence, age of the firm dummies, economics sector dummies, year dummies, and city fixed effects. See Table 2 for further details on these variables.

## 4.2 Heterogeneity

In this section we explore heterogeneity on the lenient extensive margin definition of formality. As we use subgroups of firms, for the other definitions the estimates become highly imprecise and do not show radically different patterns.

First, we consider the age of the firm. We expect the formality premium to be more significant for the younger firms undertaking the investments required to grow

their businesses. Figure 4a shows that for microfirms that are less than three years old in the market, the premium is around 40% by the 20th percentile compared to the 15% premium for older firms until the median. The premium gap between younger and older microfirms continues until the 70th percentile. Afterward, the difference disappears.

Second, we explore results according to the motivation for the business if they found an opportunity in the market. The selection into formality between opportunity and necessity entrepreneurs might be substantially different (Reynolds et al., 2001). Opportunity-driven individuals are more likely to take more risks that can materialize in requesting bank loans to aid their growth (Block et al. 2015). Given this, at first sight, there is an unexpected result in Figure 4b. For microentrepreneurs who self-declared to have started their businesses as a means of providing subsistence income<sup>11</sup>, the premium was consistently larger than for those micro-owners who declared to have searched for a business opportunity.

Necessity microfirms are more likely to be informal, 23.9% of them vs 8.1% of opportunity microfirms. Moreover, informal opportunity firms are more likely to transit into formality (44.8% of them in one year) than necessity ones (31.2%), and more likely to stay formal (95.4% vs 89.9%). We conjecture that necessity firms are (i) more likely to be dependent on the informality status for the viability of the business, (ii) less likely to undertake investments (due to risk aversion), and (iii) less likely to quantify the cost and benefits of formalizing their business. Thus, those which transit into formality are firms for which there is a substantial expected benefit.

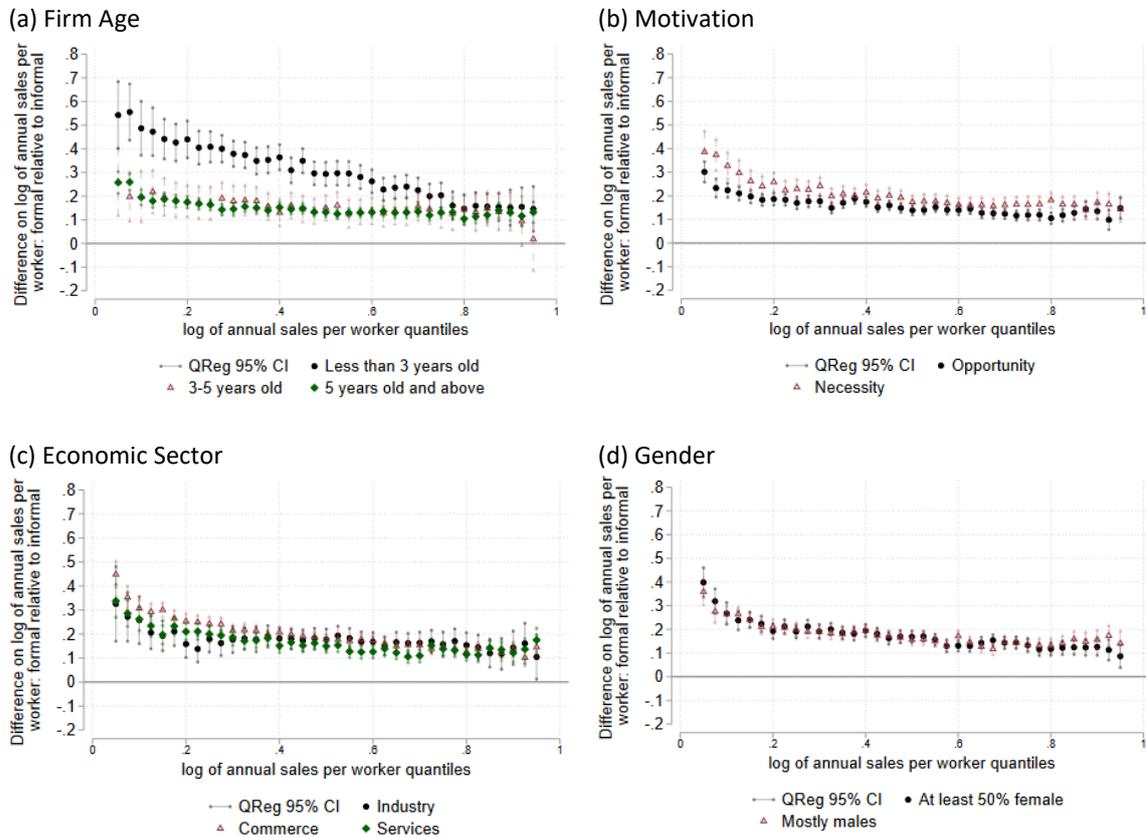
Next, we consider the economic sector (Figure 4c) and the gender of the owners (Figure 4d). Despite the known differences in growth according to characteristics such

---

<sup>11</sup> Includes other motivations: family tradition, time constraints, and loss of employment.

as economic sector and gender of the owner (Delmar & Wiklund, 2008), we did not find that the informality premium depends markedly on such characteristics.<sup>12</sup>

Figure 4: Informality premium according to firm characteristics



**Notes:** The graph presents the coefficients of the informality premium and their 95% confidence interval after a fixed-effects quantile regression. Controls: number of workers, proportion of female ownership, ownership of ICT devices and web presence, age of the firm dummies, economics sector dummies, year dummies, and city fixed effects. See Table 1 for further details on these variables.

### 1.1 What else changes with formality?

Firms which become formal reveal that for them, there is a net benefit of such a decision (Perry et al. 2009). For this reason we explore other margins of response that firms have. We consider hiring new staff, loans acquisition, and having large firms as clients as some of the most important dimensions which motivate formality transition (Perry et al 2009, Williams et al. 2017).

<sup>12</sup> We explored another dimension: the size of the city. Formality benefits might be larger in big cities (metropolitan areas of Bogotá, Medellín, Cali and Barranquilla; all of them above 2 million inhabitants) than smaller ones. We did not find any differential pattern.

Figure 5 presents, for both margins of formality (Panels A and B), the size profile through the observation window, for firms which were classified as informal in 2012. The graph is presented according to the formality status of the firms in each year, and by firm's age at the start of the observation window (2012). Age is considered for two groups of microfirms: one group composed of firms aged less than three years (left), a second group composed of microfirms older than 10 years old (right). Several stylized facts can be observed. The first is the opposite positioning of extensive-intensive margins profiles of microfirms. By the extensive margin and for *both* age-groups of microfirms, formal (extensive-lenient) outpaced informal formal peers. However, by the intensive margin, informal microfirms exceed formal microfirms. Ulyssea (2020b) estimated a model with simulated microdata to analyze what firm growth looks like under both margins. When he estimated firm growth using the intensive margin, he found a slightly flatter age-size profile for Brazilian (micro) firms. A very similar pattern for Colombian microfirms can be seen in Panel B of Figure 5.

The profiles above suggest two different strategies. We complement the analysis above by estimating equation (1) with four alternative outcomes: mean wages, firm size, obtaining a loan (from any source), obtaining a loan from a formal source (banks and other financial institutions).

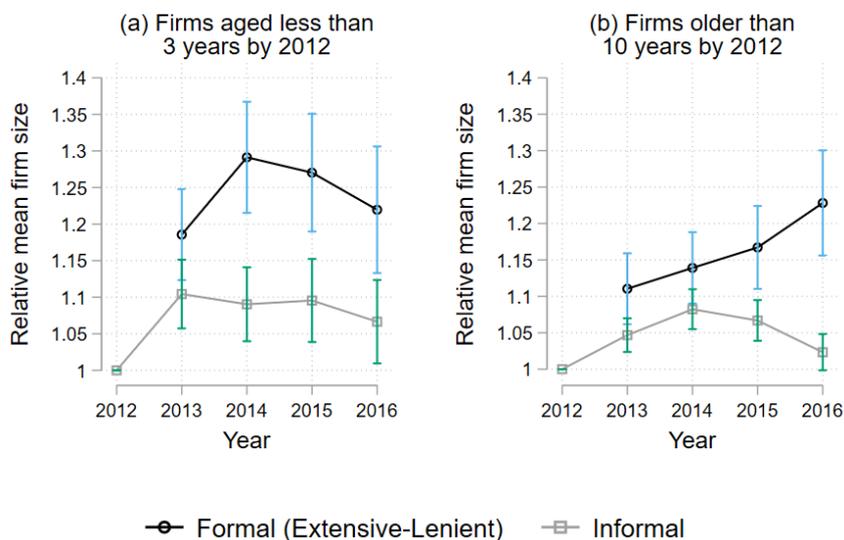
The first strategy consists of firms growing in their number of employees as they become formal (extensive margin). Therefore we expect these firms to undertake investments that are possible (or at least cheaper) for a formal than an informal firm. Panel A of Table 3 shows that on average, these firms are obtaining more loans, in particular formal loans. These investments are complementary to the labor force of the firm. Thus, the capital injection is the main reason for the increase in productivity. Firms which formerly had higher levels of productivity, are able to increase their workforce in larger numbers (Table 3, Panel B).

Second, firms which formalize their working force (intensive margin), do not grow in the number of employees relative to those firms which did not become formal in such margin. If these firms are older than 10 years, they actually shrink their workforce. Table 3 shows that these firms reduce their workforce (0.38 workers less) but the remaining workers are better paid (32% increase on the mean wage). Thus, these firms are looking to attract (or retain) skilled workers, which not only require better wages but also the benefits of a formal job. Interestingly, these firms do not require more capital

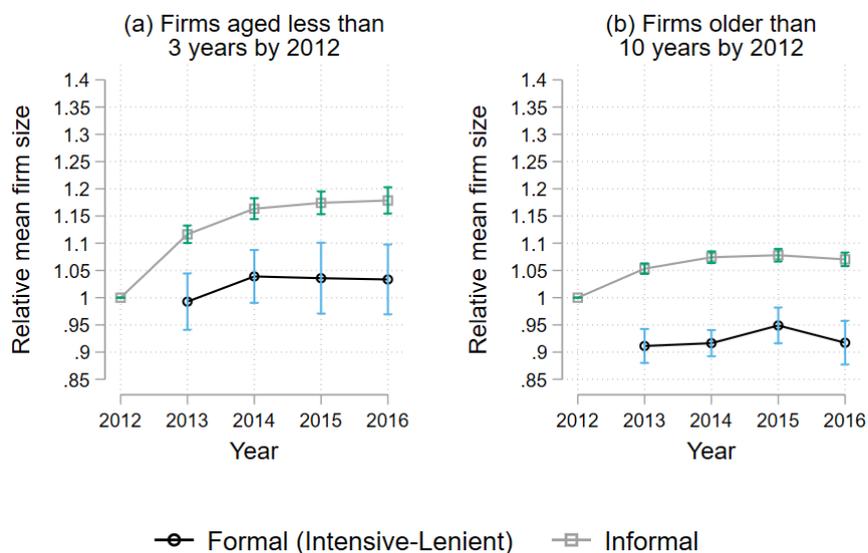
(represented in the loans) to improve their productivity. These margins of response do not seem to be dependent on prior productivity levels (Table 3, Panel C).

Figure 5. Size profile by informality status and firm's age, for informal firms by 2012

A. Extensive margin formality



B. Intensive margin formality



Notes: Confidence intervals correspond to the 95% level, based on the standard error of the mean.

Table 3. Other response margins to formalization

	(1)	(2)	(3)	(4)
	Log-wages	N employees	Loan any (=1)	Loan formal (=1)
<b>Panel A. No tertile interactions</b>				
Formal Extensive Lenient	0.0690*** (0.022)	0.122*** (0.008)	0.0210*** (0.007)	0.0249*** (0.006)
Formal Intensive Lenient	0.322*** (0.007)	0.385*** (0.009)	0.00807*** (0.005)	-0.000402 (0.005)
<b>Panel B. Extensive Lenient, tertile interactions</b>				
Formal Extensive Lenient X Low Productivity	0.0859*** (0.025)	0.0989*** (0.010)	0.0158*** (0.007)	0.0194*** (0.006)
Formal Extensive Lenient X Mid Productivity	0.0761*** (0.024)	0.129*** (0.011)	0.0253*** (0.008)	0.0302*** (0.007)
Formal Extensive Lenient X High Productivity	0.0679*** (0.024)	0.186*** (0.012)	0.0301*** (0.008)	0.0325*** (0.007)
<b>Panel C. Intensive Lenient, tertile interactions</b>				
Formal Intensive Lenient X Low Productivity	0.343*** (0.019)	0.392*** (0.019)	-0.0157 (0.010)	-0.00229 (0.009)
Formal Intensive Lenient X Mid Productivity	0.325*** (0.010)	0.422*** (0.015)	-0.000894 (0.008)	0.00191 (0.007)
Formal Intensive Lenient X High Productivity	0.321*** (0.008)	0.419*** (0.014)	-0.00449 (0.007)	0.00207 (0.007)

Notes: Each row corresponds to a separate regression. Own calculations using DANE Micro-enterprise survey 2012-2016. Lenient definition: The informality dummy takes the value of 1 if the firm has none of the 3 respective requirements of being legally established.

## 2 Discussion

Results above present robust evidence of the existence of a premium for microfirms for becoming formal in terms of productivity. This premium is measured for microfirms which transit into (and out of) formality, while comparing with comparable firms which did not follow such transition. Such firms undertook such decisions

because there were clear benefits of becoming formal, observed either by increasing capital investments and/or human capital expansion. This is a scenario where we observe substantial transitions of microfirms across the formality status, workforce size, and substantial overlapping over productivity levels for them.

Furthermore, our results show that the premium size highly depends on the formality margin type, and slightly from the initial productivity level. There is an important increase for microfirms that start to comply with business registrations - the extensive margin-, and after this, there is a smaller one for microfirms that start to comply with labor regulations, the intensive margin. Moreover, and under the strict definitions, there is evidence of a gradient on the premium according to the initial productivity level: gains are higher for firms which are less productive to begin with.

The existence of a gradient in the premium constitutes evidence of higher barriers for formalization for the less productive firms: for these microfirms, there is a need for higher benefits in order to overcome relatively higher costs of formalization. A first factor behind this is size: most microfirms, as reported above, are very small, with no more than three employees. Part of this is an optimal size plant according to the owners' interest: Hurst and Pugsley (2011, p.73) report that "many small businesses have little desire to grow big or to innovate in any observable way." This can be the case for many Colombian microfirms. However, for those that have an intention to grow, being too small precludes them from overcoming the costs of transitioning to a higher formality status. A second factor could be the low levels of human capital and managerial skills that owners of these business units have, limiting their ability to assess the true returns of formalization (Jovanovic, 1982). Unfortunately, the dataset does not provide information to contrast each potential factor separately.

We have shown that Microfirms opportunistically take advantage of the formality/informality status. First, they transit into the extensive margin formality in order to obtain, at least, better conditions in terms of loans, for instance. Second, they transit into the intensive margin formality once they need more productivity workers

(reflected in higher mean wages). In fact, it is common that firms transition into informality in order to hire more workers. These results are consistent with results from the literature which shows that reducing the fixed costs of becoming formal does not increase formality, what has been found to produce such an increase is to increase the relative benefits of formality (Ulyssea, 2020b). This behavior is possible due to a weak enforcement of labor regulations which are not coordinated at the national level and result in fines that are often not paid (DNP, 2019).

We have also found that the premium is higher for necessity-driven businesses. As above, this tells us that these firms need higher benefits to decide to formalize. Such results motivate entrepreneurship (Psychological) training, i.e., business practices (McKenzie & Woodruff, 2017; Frese et al., 2016), which can occur at early stages, for instance at the school and universities that enhance entrepreneurial intention (Nabi et al., 2017). Consequently, it would produce business ideas that do not depend on being informal in the market. Moreover a non-negligible 27% of self-employed turned to microfirms with some employees. It shows that there are self-employees that, irrespective of their relatively low productivity, might get large benefits for becoming formal as is found in Leal and Chacón (2017). This would require focused efforts in the area of business training, empowerment and the like.

Additionally, we also obtained evidence that microfirms in their first years are those which attain the highest productivity gains, but also the steeper gradient observed highlight that they face even greater barriers to formalization. These findings suggest that public policies must strengthen programs like accelerators, seed capital, and similar ones that focus on nascent firms (Coad et al. 2022, Gonzalez and Leatherbee 2018, and González and Reyes 2021). These programs should stress to participants, from the start, the benefits (and costs) of adopting both business and labor regulations. With regard to this last, showing that paying labor regulations can be a guarantee to attract (or at least retain) better human capital.

Finally, we observed that several categorizations that are known to be relevant for the growth-potential of firms, like gender composition of the owners or economic

sector, do not exhibit differential premiums. Thus, policy approaches intended to help microfirms grow might not need to discriminate by these factors. At least, if the sole goal is to enhance productivity.

It is important to highlight a key restriction to the interpretation of our results. Since our goal is to describe the observed premium in productivity for firms that *change* their formality status over time, the estimates presented here cannot be interpreted as an average treatment effect, i.e., the average premium in the productivity of a randomly selected firm to be formal. Results are closer to the average productivity gain of firms that decided to become formal. They could also be interpreted as the size of the gains needed for those firms to transit into formality. Therefore, it cannot be inferred that a program that motive or force firms into formality, would necessarily result in the gains measured in this article. Such gains are likely to be smaller as firms that transit into formality tend to be better prepared in terms of factors that promote growth.

Further research in this topic is needed, especially using more detailed measures of productivity, growth, and benefits. This article is based on self-reports, which result in estimates that are imprecise due to a measurement error, and small premiums cannot be properly detected. Also, questions in this survey do not provide detailed information on the timing of the transition into formality; thus, our results might be downward biased as we do not have firms that were exposed for an entire year to the benefits of formalization. Moreover, the design of the dataset censors microfirms when they report having ten or more employees, and this situation induces further downward bias as the highest achievers were removed from the dataset by construction.

### 3 Conclusions

We have used a rich and unique longitudinal dataset of Colombian microfirms over the period 2012-2015 to analyze the existence of a productivity premium between informal and formal microfirms according to the two margins proposed by Ulyssea (2018). To this end, we first proceed to distinguish microfirms that comply with business regulations, i.e., paying the mandated registration costs from those microfirms that incur the registration costs, but that hire, or contract informal workers to avoid incurring the costs mandated by labor regulations. This separation allows us to disentangle heterogeneous results from the dynamics of microfirms transitioning out and into formality.

In terms of public policy, the OECD recommends four lines towards formalization of labor: to lower costs, visibilize the benefits of formalization, enhance enforcement, and promote skills development (OECD, 2019). When considering microfirms, our results show that these guidelines are equally valid.

## 4 References

Allen, J., Nataraj, S., & Schipper, T., (2018) Strict duality and overlapping productivity distributions between formal and informal firms. *Journal of Development Economics*, 135, 534-554.

Almeida, R. and Carneiro, P. (2012). Enforcement of Labor Regulation and Informality. *American Economic Journal: Applied Economics*, 4, 64-89.

Baker, M., Powell, D., & Smith, T. (2016). Qregpd: Stata module to perform quantile regression for panel data. *Statistical Software Components, S458157, Boston College Department of Economics*.

Blau, D. M. (1985). Self-employment and self-selection in developing country labor markets. *Southern Economic Journal*, 52 (2), 351–363.

Block, J., Kohn, H. K., Miller, D., & Ullrich, K. (2015). Necessity entrepreneurship and competitive strategy. *Small Business Economics*, 44(1), 37–54.

Campos, F., Frese, M., Goldstein, M., Iacovone, L., Johnson, H. C., McKenzie, D., & Mensmann, M. (2017). Teaching personal initiative beats traditional training in boosting small businesses in West Africa. *Science*, 357 (6357), 1287–1290.

Cisneros, C. (2022). Unfolding Trade Effect in Two Margins of Informality. The Peruvian Case. *The World Bank Economic Review*, 36, 141–170

Coad, A. (2009). *The growth of firms: A survey of theories and empirical evidence*. Edward Elgar Publishing.

Coad, A., & Tamvada, J. P. (2012). Firm growth and barriers to growth among small firms in India. *Small Business Economics*, 39 (2), 383–400.

Coad, A., Harasztosi, P., Pál, R., & Teruel, M.. (2022). Policy Instruments for High-Growth, in K. Wennberg & C. Sandström. Editors Questioning the Entrepreneurial State. Status-quo, Pitfalls, and the Need for Credible Innovation Policy. Springer.

Departamento Nacional de Planeación [DNP] (2019). Política de Formalización Empresarial. Documento CONPES 3956.

De Mel, S., McKenzie, D., & Woodruff, C. (2013). The demand for, and consequences of, formalization among informal firms in Sri Lanka. *American Economic Journal: Applied Economics*, 5 (2), 122–50.

de Soto, H. (1989). *The other path: the invisible revolution in the third world*. Harper Collins.

Delmar, F., & Wiklund, J. (2008). The effect of small business managers' growth motivation on firm growth: A longitudinal study. *Entrepreneurship Theory and Practice*, 32 (3), 437–457.

Eslava, M., Haltiwanger, J., & Pinzón, Á. (2022). Job Creation in Colombia Versus the USA: 'Up-or-out Dynamics' Meet 'The Life Cycle of Plants'. *Economica*, 89: 511-539. <https://doi.org/10.1111/ecca.12418>.

Farrell, D. (2006). Tackling the informal economy. *Business Week*, 10, 262–276.

Floridi, A., Demena, B., and Wagner, N. (2020). Shedding light on the shadows of informality: A meta-analysis of formalization interventions targeted at informal firms. *Labour Economics*, 67, 101925..

Frese, M., Gielnik, M. M., & Mensmann, M. (2016). Psychological training for entrepreneurs to take action: Contributing to poverty reduction in developing countries. *Current Directions in Psychological Science*, 25 (3), 196–202.

González, J. and Leatherbee, M. (2018) The effects of business accelerators on venture performance: Evidence from start-up Chile. *The Review of Financial Studies* 31, 1566-1603.

González, J. and Reyes, S. (2021). Identifying and boosting "Gazelles": Evidence from business accelerators. *Journal of Financial Economics*, 139, 260-287

Hermans, J., Vanderstraeten, J., Van Witteloostuijn, A., Dejardin, M., Ramdani, D., and Stam, E. (2015). Ambitious entrepreneurship: A review of growth aspirations, intentions, and expectations. *Entrepreneurial Growth: Individual, Firm, and Region*.

Hsieh, C., and Olken, B., (2014). The Missing "Missing Middle". *Journal of Economic Perspectives*, 28, 89-108:

Hurst, E. and Pugsley, B. (2011). What Do Small Businesses Do? *Brooking Papers on Economic Activity*. Fall.

International Labor Organization. (2006). Employment Relationship Recommendation, 2006 (No. 198) Recommendation concerning the employment relationship. R198

Jessen, J., & Kluge, J. (2021). The effectiveness of interventions to reduce informality in low- and middle-income countries. *World Development*, 138, 105256.

Jovanovic, B. (1982). Selection and the evolution of industry. *Econometrica*, 50(3), 649–670.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis*, 91 (1), 74–89.

La Porta, R., & Shleifer, A. (2008). The unofficial economy and economic development. *Brookings Papers on Economic Activity*, 2008, 275–352.

La Porta, R., & Shleifer, A. (2014). Informality and development. *Journal of Economic Perspectives* 28 (4), 109–126.

Leal Calderon, Z., & Chacón, N. (2017). Self-employment in the Andean countries: Motivations and the link to productivity. *Inter-American Development Bank*.

López-Mera, S. F. (2021). ¿Qué tan lejos está el ODS# 8 para Colombia? una década de medición del trabajo decente. *Sociedad y Economía*, (43).

Maloney, W. (2004). Informality revisited. *World Development*, 32 (7), 1159–1178.

McKenzie, D., & Woodruff, C. (2017). Business practices in small firms in developing countries. *Management Science*, 63 (9), 2773–3145.

Meghir, C, Narita, R, & Robin, J. (2015). Wages and Informality in Developing Countries. *American Economic Review*, 105, 1509-1546.

Morón, E., E. Salgado, and C. Seminario (2012). Financial dependence, formal credit and firm informality: Evidence from Peruvian household data. Technical report, IDB Working Paper Series; 288.

Nabi, G., Liñán, F., Fayolle, A., Krueger, N., & Walmsley, A. (2017). The impact of entrepreneurship education in higher education: A systematic review and research agenda. *Academy of Management Learning & Education*, 16 (2), 277–299.

Nordman, C. J., F. Rakotomanana, and F. Roubaud (2016). Informal versus formal: A panel data analysis of earnings gaps in Madagascar. *World Development*, 86, 1–17.

OECD (2019). Chapter 2. Fostering high-quality jobs for all in Colombia, in *OECD Economic Surveys: Colombia 2019*. OECD Publishing, Paris, <https://doi.org/10.1787/e4c64889-en>.

Ortiz, A. F., Medina Rojas, I. D., Rodríguez Lesmes, P. A. y Gutiérrez, L. H. (comp.) (2021). La economía de las tiendas de barrio en Colombia. Editorial Universidad del Rosario. <https://doi.org/10.12804/urosario9789587847284>

Perry, G., Maloney, W., Arias, O., Fajnzylber, P., Mason, A., Saavedra-Chanduvi, J., (2009). Informality : exit and exclusion. *World Bank Publications*.

Ponczek, V. and Ulyssea, G. (2022). Enforcement of Labour Regulation and the Labour Market Effects of Trade: Evidence from Brazil. *The Economic Journal*, 132, 361–390,

Powell, D. (2016). Quantile regression with nonadditive fixed effects. Quantile treatment effects. *Unpublished manuscript, available at: https://works.bepress.com/david\_powell/1*.

Powell, D. (2020). Quantile treatment effects in the presence of covariates. *Review of Economics and Statistics*, 102 (5), 994–1005.

Radchenko, N. (2014). Heterogeneity in Informal Salaried Employment: Evidence from the Egyptian Labor Market Survey. *World Development*, 62: 169-188.

Rauch, J. E. (1991). Modelling the informal sector formally. *Journal of Development Economics*, 35 (1), 33–47.

Reynolds, P., Camp, S., Bygrave, W., Autio, E., & Hay, M. (2001). *Global Entrepreneurship Monitor, 2001 Executive Report*. Kansas City, MO: Kauffman Center for Entrepreneurial Leadership at the Ewing Marion Kauffman Foundation.

Reyes Torres, C. E (2020). Encubrimiento del contrato de prestación de servicios al contrato realidad: consecuencias para la administración pública en el respeto y garantía de los derechos laborales. Master thesis. Universidad Militar Nueva Granada.

Reynolds, P., Camp, S., Bygrave, W., Autio, E., & Hay, M. (2001). *Global Entrepreneurship Monitor, 2001 Executive Report*. Kansas City, MO: Kauffman Center for Entrepreneurial Leadership at the Ewing Marion Kauffman Foundation.

Samanaiego de la Parra, B. and Fernández, L. . (2020). increasing the cost of informal workers: Evidence from Mexico. Working Papers No. 2020-19. banco de México, Ciudad de México.

Trebilcock, A. (2004). Decent work and the informal economy. *WIDER Discussion Paper No. 2005/04*, 1–37.

Ulyssea, G. (2018). Firms, informality, and development: Theory and evidence from Brazil. *American Economic Review*, 108 (8), 2015–2047.

Ulyssea, G. (2020a). Informality: Causes and consequences for development. *Annual Review of Economics*, 12 (1), 525–546.

Ulyssea, G. (2020b). Formal and Informal Firms Dynamics. Mimeo, January 16, 2020.

van der Zwan, P., Thurik, R., Verheul, I., & Hessels, J. (2016). Factors influencing the entrepreneurial engagement of opportunity and necessity entrepreneurs. *Eurasian Business Review*, 6 (3), 273–295.

Williams, C.C., Martínez, A., Kedir, A. (2017). Informal entrepreneurship in developing economies: The impacts of starting up unregistered on firm performance, *Entrepreneurship: Theory and Practice*, 41(5), 773-799.

# **Informal versus Formal: Microfirms' Productivity Gaps**

## **Appendix**

Table A1. Proportion of firms according to characteristics and formality status; lenient definitions

	I. Lenient, Extensive						II. Lenient, Intensive					
	Self-employed			Microfirm			Self-employed			Microfirm		
	(1) All	(2) Informal	(3) Formal	(4) All	(5) Informal	(6) Formal	(7) All	(8) Informal	(9) Formal	(10) All	(11) Informal	(12) Formal
Formal Extensive Lenient	0.765	0	1.000 ***	0.933	0	1.000 ***	0.765	0.728	1.000 ***	0.933	0.921	1.000 ***
Formal Extensive Strict	0.357	0	0.466 ***	0.646	0	0.692 ***	0.357	0.301	0.705 ***	0.646	0.603	0.884 ***
Firm-level factors of formality												
Tax registry	0.656	0	0.857 ***	0.857	0	0.919 ***	0.656	0.609	0.949 ***	0.857	0.835	0.981 ***
Business registry	0.48	0	0.627 ***	0.735	0	0.787 ***	0.48	0.43	0.791 ***	0.735	0.703	0.910 ***
Keeps accounting records	0.564	0	0.737 ***	0.821	0	0.880 ***	0.564	0.517	0.863 ***	0.821	0.794	0.973 ***
Formal Intensive Lenient	0.137	0	0.179 ***	0.153	0	0.164 ***	0.137	0	1.000 ***	0.153	0	1.000 ***
Formal Intensive Strict	0.049	0	0.064 ***	0.104	0	0.112 ***	0.049	0	0.358 ***	0.104	0	0.682 ***
Labor factors of formality												
Contributes to social insurance	0.142	0.027	0.177 ***	0.15	0.009	0.160 ***	0.142	0.007	0.986 ***	0.15	0.001	0.976 ***
Fringe benefits and payments	0.065	0.002	0.084 ***	0.12	0.003	0.129 ***	0.065	0.001	0.469 ***	0.12	0	0.787 ***
Pays firm-specific wage taxes	0.067	0.004	0.086 ***	0.122	0.003	0.130 ***	0.067	0.001	0.479 ***	0.122	0	0.797 ***
Economic sector												
Manufacture	0.063	0.073	0.059 ***	0.128	0.098	0.130 ***	0.063	0.065	0.044 ***	0.128	0.131	0.113 ***
Retail	0.705	0.726	0.698 ***	0.551	0.62	0.546 ***	0.705	0.699	0.740 ***	0.551	0.544	0.590 ***
Services	0.233	0.2	0.243 ***	0.321	0.282	0.324 ***	0.233	0.235	0.215 ***	0.321	0.325	0.297 ***
Age of the firm in years												
Less than 1	0.03	0.031	0.03	0.025	0.039	0.024 ***	0.03	0.029	0.036 ***	0.025	0.025	0.024
More than 1, less than 3	0.167	0.161	0.169 ***	0.14	0.161	0.139 ***	0.167	0.165	0.179 ***	0.14	0.141	0.134 ***
More than 3, less than 5	0.16	0.129	0.170 ***	0.152	0.131	0.154 ***	0.16	0.158	0.175 ***	0.152	0.153	0.15
More than 5, less than 10	0.245	0.188	0.262 ***	0.26	0.201	0.264 ***	0.245	0.242	0.259 ***	0.26	0.261	0.253 ***
More than 10	0.398	0.492	0.369 ***	0.423	0.468	0.419 ***	0.398	0.405	0.350 ***	0.423	0.42	0.439 ***
Size of the firm												
2 workers	0	0	0	0.541	0.794	0.523 ***	0	0	0	0.541	0.559	0.438 ***
3 to 5 workers	0	0	0	0.39	0.199	0.403 ***	0	0	0	0.39	0.38	0.443 ***
6 to 9 workers	0	0	0	0.069	0.007	0.074 ***	0	0	0	0.069	0.06	0.119 ***
Female ownership	0.619	0.607	0.623 ***	0.684	0.717	0.681 ***	0.619	0.594	0.781 ***	0.684	0.64	0.924 ***
Firm based on a perceived opportunity	0.598	0.385	0.663 ***	0.709	0.438	0.728 ***	0.598	0.569	0.783 ***	0.709	0.687	0.838 ***
Loans												
Requested any	0.26	0.278	0.255 ***	0.269	0.311	0.266 ***	0.26	0.269	0.212 ***	0.269	0.28	0.213 ***
Requested formal	0.193	0.167	0.202 ***	0.216	0.204	0.216 **	0.193	0.197	0.173 ***	0.216	0.223	0.176 ***
Obtained	0.245	0.262	0.239 ***	0.252	0.297	0.249 ***	0.245	0.252	0.200 ***	0.252	0.262	0.198 ***
Information and comm. technologies (ICT)												
Ownership of devices	0.179	0.048	0.220 ***	0.425	0.073	0.450 ***	0.179	0.156	0.327 ***	0.425	0.385	0.644 ***
Web presence	0.142	0.029	0.177 ***	0.391	0.047	0.416 ***	0.142	0.118	0.292 ***	0.391	0.347	0.637 ***
Relative♣ monthly wages	1.061	0.833	1.074 ***	0.755	0.47	0.766 ***	1.061	0.992	1.119 ***	0.755	0.656	1.110 ***
Relative♣ annual sales per worker	45.715	26.855	51.504 ***	51.71	26.68	53.502 ***	45.715	40.149	80.754 ***	51.71	45.595	85.619 ***

Table A2. Proportion of firms according to characteristics and formality status; strict definitions

	III. Strict, Extensive						IV. Strict, Intensive					
	Self-employed			Microfirm			Self-employed			Microfirm		
	(13) All	(14) Informal	(15) Formal	(16) All	(17) Informal	(18) Formal	(19) All	(20) Informal	(21) Formal	(22) All	(23) Informal	(24) Formal
Formal Extensive Lenient	0.765	0.635	1.000 ***	0.933	0.811	1.000 ***	0.765	0.753	1.000 ***	0.933	0.925	1.000 ***
Formal Extensive Strict	0.357	0	1.000 ***	0.646	0	1.000 ***	0.357	0.324	1.000 ***	0.646	0.605	1.000 ***
Firm-level factors of formality												
Tax registry	0.656	0.465	1.000 ***	0.857	0.597	1.000 ***	0.656	0.638	1.000 ***	0.857	0.841	1.000 ***
Business registry	0.48	0.191	1.000 ***	0.735	0.25	1.000 ***	0.48	0.453	1.000 ***	0.735	0.704	1.000 ***
Keeps accounting records	0.564	0.323	1.000 ***	0.821	0.495	1.000 ***	0.564	0.542	1.000 ***	0.821	0.8	1.000 ***
Formal Intensive Lenient	0.137	0.063	0.271 ***	0.153	0.05	0.209 ***	0.137	0.093	1.000 ***	0.153	0.054	1.000 ***
Formal Intensive Strict	0.049	0	0.137 ***	0.104	0	0.161 ***	0.049	0	1.000 ***	0.104	0	1.000 ***
Labor factors of formality												
Contributes to social insurance	0.142	0.072	0.267 ***	0.15	0.05	0.204 ***	0.142	0.097	1.000 ***	0.15	0.051	1.000 ***
Fringe benefits and payments	0.065	0.016	0.152 ***	0.12	0.029	0.170 ***	0.065	0.017	1.000 ***	0.12	0.018	1.000 ***
Pays firm-specific wage taxes	0.067	0.02	0.151 ***	0.122	0.032	0.171 ***	0.067	0.019	1.000 ***	0.122	0.02	1.000 ***
Economic sector												
Manufacture	0.063	0.071	0.048 ***	0.128	0.121	0.132 ***	0.063	0.064	0.025 ***	0.128	0.13	0.110 ***
Retail	0.705	0.679	0.751 ***	0.551	0.516	0.570 ***	0.705	0.699	0.823 ***	0.551	0.542	0.621 ***
Services	0.233	0.25	0.201 ***	0.321	0.363	0.298 ***	0.233	0.237	0.152 ***	0.321	0.327	0.269 ***
Age of the firm in years												
Less than 1	0.03	0.032	0.028 ***	0.025	0.031	0.021 ***	0.03	0.03	0.039 ***	0.025	0.025	0.023
More than 1, less than 3	0.167	0.167	0.168	0.14	0.15	0.134 ***	0.167	0.166	0.200 ***	0.14	0.14	0.137
More than 3, less than 5	0.16	0.148	0.182 ***	0.152	0.15	0.154	0.16	0.158	0.194 ***	0.152	0.153	0.15
More than 5, less than 10	0.245	0.226	0.278 ***	0.26	0.24	0.271 ***	0.245	0.244	0.257 *	0.26	0.262	0.246 ***
More than 10	0.398	0.427	0.344 ***	0.423	0.428	0.419 ***	0.398	0.402	0.309 ***	0.423	0.42	0.443 ***
Size of the firm												
2 workers	0	0	0	0.541	0.686	0.462 ***	0	0	0	0.541	0.56	0.375 ***
3 to 5 workers	0	0	0	0.39	0.292	0.443 ***	0	0	0	0.39	0.379	0.479 ***
6 to 9 workers	0	0	0	0.069	0.022	0.095 ***	0	0	0	0.069	0.06	0.146 ***
Female ownership	0.619	0.587	0.678 ***	0.684	0.656	0.699 ***	0.619	0.6	0.988 ***	0.684	0.649	0.982 ***
Firm based on a perceived opportunity	0.598	0.512	0.750 ***	0.709	0.593	0.772 ***	0.598	0.585	0.849 ***	0.709	0.692	0.863 ***
Loans												
Requested any	0.26	0.266	0.250 ***	0.269	0.28	0.263 ***	0.26	0.267	0.125 ***	0.269	0.278	0.201 ***
Requested formal	0.193	0.188	0.204 ***	0.216	0.212	0.217	0.193	0.198	0.101 ***	0.216	0.222	0.166 ***
Obtained	0.245	0.249	0.236 ***	0.252	0.263	0.246 ***	0.245	0.251	0.115 ***	0.252	0.26	0.187 ***
Information and comm. technologies (ICT)												
Ownership of devices	0.179	0.135	0.259 ***	0.425	0.266	0.511 ***	0.179	0.169	0.377 ***	0.425	0.395	0.683 ***
Web presence	0.142	0.096	0.224 ***	0.391	0.22	0.485 ***	0.142	0.131	0.361 ***	0.391	0.356	0.696 ***
Relative monthly wages	1.061	1.053	1.065	0.755	0.634	0.801 ***	1.061	1.054	1.077	0.755	0.681	1.145 ***

Relative annual sales per worker	45.715	34.579	65.793 ***	51.71	35.138	60.785 ***	45.715	42.484	108.406 ***	51.71	46.523	96.335 ***
----------------------------------	--------	--------	------------	-------	--------	------------	--------	--------	-------------	-------	--------	------------

---

Table A3: Formality transition matrices

<u>Panel A. Lenient definition</u>		Status at t			
		(a) Extensive margin		(b) Intensive margin	
		Informal	Formal	Informal	Formal
Status at t – 1	Informal	10,834 61.76%	6,708 38.24%	101,589 89.85%	11,479 10.15%
	Formal	7,267 6.30%	108,140 93.70%	11,813 59.42%	8,068 40.58%
<u>Panel B. Strict definition</u>		Status at t			
		(a) Extensive margin		(b) Intensive margin	
		Informal	Formal	Informal	Formal
Status at t – 1	Informal	47,063 76.20%	14,699 23.80%	115,858 94.95%	6,164 5.05%
	Formal	15,795 22.19%	55,392 77.81%	6,270 57.38%	4,657 42.62%

Notes: Own calculations using DANE Micro-enterprise survey 2012-2016. Results are based on the lenient definition (a firm is formal if it meets at least *one* of the respective requirements of being legally established), and the strict definition (a firm is formal if it meets *all* requirements of being legally established).

Table A4: Formality transition matrices by firm size at t-1

<b>Panel (A) Self-employed at t-1</b>		Status at t			
		(a) Extensive margin		(b) Intensive margin	
		Informal	Formal	Informal	Formal
Status at t – 1	Informal	7,995 66.06%	4,108 33.94%	41,408 92.77%	3,228 7.23%
	Formal	4,123 10.33%	35,794 89.67%	4,797 64.96%	2,587 35.04%
<b>Panel (B) Microfirm at t-1</b>		(a) Extensive margin		(b) Intensive margin	
		Informal	Formal	Informal	Formal
Status at t – 1	Informal	2,839 52.2%	2,600 47.8%	60,181 87.94%	8,251 12.06%
	Formal	3,144 4.16%	72,346 95.84%	7,016 56.14%	5,481 43.86%

Notes: Own calculations using DANE Micro-enterprise survey 2012-2016. Results are based on the lenient definition (a firm is formal if it meets at least one of the respective requirements of being legally established).

Table A5. Determinants of formalization

VARIABLES measured at time t	Formality status at time t+1			
	Extensive		Intensive	
	(1) Lenient	(2) Strict	(3) Lenient	(4) Strict
log of annual sales	0.0261* (0.0152)	0.0574*** (0.00602)	0.0340*** (0.00284)	0.0275*** (0.00223)
Economic Sector				
Manufacture [Base]	-	-	-	-
Commerce	-0.0380 (0.0585)	-0.0469** (0.0196)	0.0457*** (0.00785)	0.0406*** (0.00619)
Services	-0.0528 (0.0601)	-0.129*** (0.0196)	0.0217*** (0.00775)	0.00420 (0.00596)
Age of the firm in years				
Less than 1				
More than 1, less than 3	0.0461 (0.110)	-0.0675 (0.0464)	-0.00359 (0.0242)	-0.0393* (0.0215)
More than 3, less than 5	0.0202 (0.108)	-0.0406 (0.0453)	-0.0182 (0.0235)	-0.0544*** (0.0211)
More than 5, less than 10	0.0207 (0.104)	-0.0807* (0.0443)	-0.0246 (0.0230)	-0.0543*** (0.0207)
More than 10	-0.0955 (0.101)	-0.148*** (0.0435)	-0.0336 (0.0227)	-0.0595*** (0.0206)
Size of the firm				
Self-employed = 0,				
2 workers	0.124** (0.0530)	0.000867 (0.0186)	-0.00706 (0.0109)	-0.00184 (0.00717)
3 to 5 workers	0.144** (0.0610)	0.0503** (0.0204)	-0.0134 (0.0113)	0.0152** (0.00762)
6 to 9 workers	0.212 (0.199)	0.130*** (0.0416)	0.00389 (0.0165)	0.0505*** (0.0130)
Female ownership	0.0670* (0.0394)	0.0750*** (0.0134)	-0.00193 (0.00575)	0.0134*** (0.00456)
Firm based on a perceived opportunity	0.0795** (0.0333)	0.0988*** (0.0122)	0.0321*** (0.00558)	0.0224*** (0.00431)
Information and comm. Technologies (ICT)				
Ownership of devices	0.181*** (0.0629)	0.0570*** (0.0164)	0.00907 (0.00672)	0.00467 (0.00529)
Web presence	0.159* (0.0938)	0.0500*** (0.0177)	0.0427*** (0.00727)	0.0330*** (0.00579)
Requested loans	-0.0402 (0.0351)	-0.0186 (0.0127)	-0.0406*** (0.00549)	-0.0384*** (0.00420)
Relative♣ monthly wages	0.126*** (0.0340)	-0.00474 (0.00743)	0.0288*** (0.00866)	0.00827** (0.00372)
Constant	-0.126 (0.245)	-0.598*** (0.103)	-0.511*** (0.0486)	-0.411*** (0.0393)
Observations	869	6,518	15,916	18,061
R-squared	0.096	0.067	0.036	0.039

Notes: Estimates from linear probability models, conditional on being informal at time t. Robust standard errors in parentheses. Significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

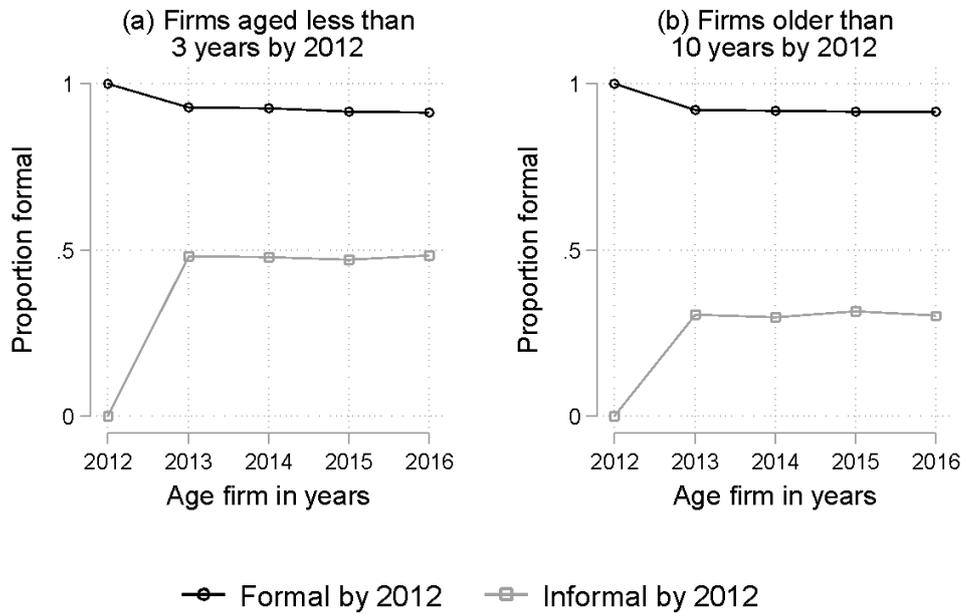
Table A6. Main estimates coefficients

	FEOLS	Coefficients associated to quantiles of sales per worker (productivity measure)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Mean	.1	.2	.3	0.4	0.5	0.6	0.7	0.8	0.9
<b>Extensive lenient</b>	0.189*** (0.00975 )	0.140*** (0.0152)	0.194*** (0.0129)	0.167*** (0.0140)	0.172*** (0.0133)	0.163*** (0.0130)	0.178*** (0.0137)	0.134*** (0.0172)	0.188*** (0.0187)	0.162*** (0.0255)
<b>Extensive strict</b>	0.183*** (0.00590 )	0.237*** (0.0129)	0.200*** (0.00997)	0.189*** (0.00946)	0.200*** (0.00840)	0.146*** (0.00818)	0.163*** (0.00794)	0.158*** (0.00810)	0.132*** (0.00937)	0.175*** (0.0131)
<b>Intensive Lenient</b>	0.110*** (0.00683 )	0.129*** (0.0218)	0.0544*** (0.0166)	0.129*** (0.0127)	0.0848*** (0.0108)	0.0727*** (0.00957)	0.0797*** (0.00915)	0.0870*** (0.00873)	0.102*** (0.00952)	0.109*** (0.0123)
<b>Intensive Strict</b>	0.151*** (0.00963 )	0.0104 (0.0658)	0.300*** (0.0307)	0.0666*** (0.0247)	0.144*** (0.0163)	0.129*** (0.0141)	0.104*** (0.0123)	0.122*** (0.0109)	0.112*** (0.0120)	0.163*** (0.0144)

Notes. Column (1) presents the estimates for the formality coefficient for each of the definitions (each one comes from a separate regression). Columns (2) to (10) presents the estimates for a subset of the estimated quantiles (deciles) on the productivity distribution. All regressions include 52,159 firms. Robust standard errors are presented in parentheses. Significance: \*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Figure A1. Informality status profile according to the 2012 lenient informality

Panel A. Extensive margin



Panel B. Intensive margin

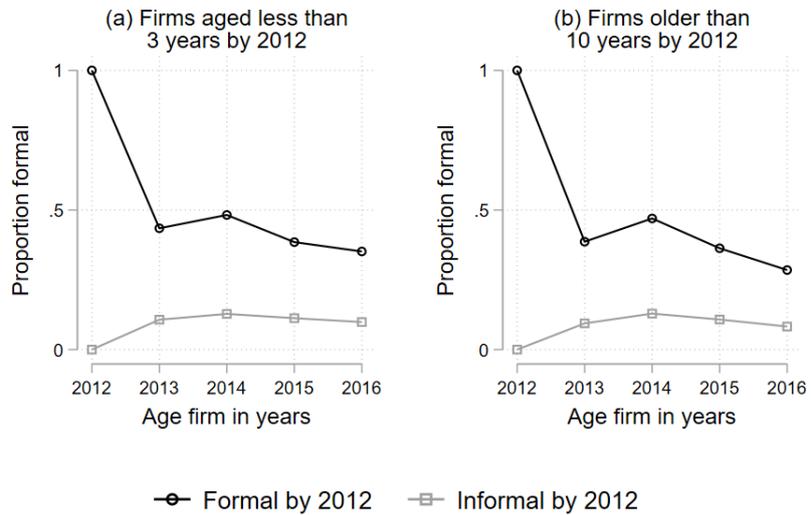


Figure A2. Self-employed status profile according to the 2012 lenient extensive margin informality

