

DOCUMENTO DE TRABAJO

“Development Studies” Economic Expertise of Development Policies & Projects

Autores

Jade Lemoine



ALIANZA EFI
Economía Formal e Inclusiva



2020/2021

LEMOINE JADE

Internship Report

“Development Studies”

Economic Expertise of Development Policies & Projects

Informality in Colombia in the late 2010's: From the Use of Raw data to an Analytical Approach

1st of March 2021 - 1st of July 2021

UMR 201 Développement et Société (Nogent-sur-Marne)



Tutrice professionnelle :

Isabel Georges

Enseignant référent :

Alain Desdoigts



Acknowledgment

I would like to thank Mr. Desdoigts for his support throughout the master's programme, his availability and his attentiveness.

I would like to thank him for his understanding of our difficulties during our master's course, which was marked by distance learning.

I also thank him for this internship opportunity and the lessons I learned from it.

I would also like to thank Isabel Georges who allowed Atji and I to have access to her office in order to avoid teleworking.

A thought for the academics of Rosario in Bogota, Colombia who taught us a lot about informality and who were incredibly kind to us:

Andrés García-Suaza, Christina Fernandez, Fernando Jaramillo, Juan Gallego, Marlon Salazar Silva and Camilo Rios Bernal.

Table of Contents

Acknowledgment.....	3
I Introduction: Colombia: socio-economic context.....	6
II Internship’s scope	7
A. Presentation of the structure: The IRD	7
A.1) Historical background	7
A.2) Main goals	8
B. Purpose of the internship	9
B.1) Subject of the study.....	9
B.2) Principal objectives	9
III Studying informality: Literature review	10
A. A difficult terminology to define	10
A.1) Informality in grey literature	10
A.2) Informality according to international or national organizations	12
B. Characteristics of Informality	14
B.1) Firms	14
B.2) Type of worker.....	14
III A research internship	14
A. A database dedicated to micro businesses	14
A.1) EMICRON: General overview	14
A.2) Choices and definition of the variables of interest	17
B. Results	20
B.1) Descriptive statistic	21
B.2) Gender approach	22
B.2) Regression (See appendix 6)	26
B.3) Informality, wages, education.....	27

C. Has the challenge been met?.....	29
IV Conclusion and lessons learned	29
Appendix.....	31
Appendix 1: ILO’s definition of informality: Histogramms	31
Appendix 2: Informality as a continuum rather than a binary outcome (Shown by Christina Fernandez).....	33
Appendix 3: Infographics: database composition	33
Appendix 4: Survey coverage (in number of persons)	36
Appendix 5: Type of occupation according to gender (Informality with RUT criteria) ..	37
Appendix 6: <i>Log(wage) regression over independent variables in the extensive and the intensive margin of informality</i>	38
References.....	39

I Introduction: Colombia: socio-economic context¹



01/03/2021 - 01/07/2021

IRD - UMR D&S
(Jardin d'Agronomie Tropicale de Paris,
Nogent-sur-Marne, 94130)

Purpose : Replication of Ulyssea, 2018
with Colombian data.

According to the measurement of national accounts, in Colombia, in 2015, participation to unobserved economy represented 24.4% of total production and 29.9% of value added. These statistics were obtained thanks to the household survey (GEIH). With these results, the necessity to build a survey dedicated to small firms appears important. Indeed, GEIH results also allow to conclude that 67.7% of occupied persons in Colombia are in firms of maximum 10 persons². These firms represent an important share of the Colombian economy. Following CPC (*Consejo Privado de Competitividad*) in Colombia in 2019, 94.6% of functioning firms have less than nine employees. They correspond to 15.2% of labor workforce.

Colombian government has an intensive policy of informality reduction since the 2000's. According to International Monetary Fund, Laws have contributed to reduce the amount of informality in the economy (70% in 2007 to 62,3% in 2017). Policy measures have contributed to this decline (2006: Entrepreneurship Law, 2010: Formalization and Job Creation Law, 2012: Tax Reform).

The definition of informality used by DANE in Colombia considers: (i) the group of employees and employers work in firms with less than ten workers, (ii) workers are not covered by health and/or pension contribution. Following this approach, the question raised about the degree of informality of small firms appears relevant. Moreover, the second part of this definition raises the question of social security needs. Investigations over informal employment seek to fill the gap between two segments of the labour workforce: the formal ones who benefit from social security contribution, retirement contributions and paid leaves and those who do

¹ « Conformément aux orientations pour la rédaction du rapport de stage... »

² See below for a definition of informality. According to DANE, firm size (i.e less than 10 persons) is a criteria of informality.

not. It is a matter of interest because these issues depend on the objectives of decent work. Informal employment prevents its workforce from benefiting workers essential labour rights.

II Internship's scope

A. Presentation of the structure: The IRD

A.1) Historical background

The IRD (« *Institut de Recherche pour le Développement* ») is a public scientific and technologic institution. It is placed under the supervision of Ministry of Higher Education, Research and Innovation and the Ministry of Europe and Foreign Affairs.

It has been created in 1943 and was originally called “*Office de la Recherche Coloniale*” and was under supervision of Secretary of State for the Navy and Colonies. At that time, its missions were to constitute a group of research workers specialized in the study of tropical regions. Its tasks evolved after the independency of African states in 1960s. The IRD's work took on a new form with the start of scientific cooperation with Third World countries (A. Sauvy). It entered in a research phase in order to contribute to their development. A fundamental reform took place in 1984, at that time, the IRD took came under the supervision of the Ministry of Research and the Ministry of Cooperation. Its role in the research for economic, social and cultural progress in developing countries gets more and more important. The institute grew up and divide in three scientific departments: environment, living resources, health and society. By 2006, it deploys a new policy with the creation of joint research unities (called UMR) in collaboration with universities and other public scientific and technological institution (EPST). IRD's headquarter moved in Marseille in 2008. In its objectives for 2011-2015 period, the IRD confirms its role in fundamental research towards the development of southern countries.

A.2) Main goals

Organisation



Figure1: IRD's presence around the world
Source: IRD's website

Its presence in in about fifty countries makes its network and its collaborations between Northern and Southern countries as fair as possible. IRD counts about 850 researchers and 1200 engineers and technical staff. It is about 100 different occupations. 30% of them work abroad. There are 75 research centres all over the world.

Network :

- 36 desks abroad
- 4 regional delegation in France
- 75 type of research centers : 1 team of labellised research (ERL), 8 associated unities (UA), 5 internaional miwed unities (UMI), 53 mixed research unities (UMR), 7 mixed unities of services (UMS), 1 service unity (US)
- 37 international laboratory (LMI)
- 40 teams associated to IRD
- 21 groups of southern-international research (GDRI sud)
- 8 interdisciplinary structural programs (PSIP)

Actions and Objectives

IRD's researches are focused on areas where most of ecological, social and political vulnerabilities are concentrated. Thus, the institute is a great contributor to approaches driven by SDGs.

The IRD's main objective is to strengthen collaboration between the so-called developed countries in the North and the developing countries in the South. By adopting a comprehensive approach to the objectives of sustainable development, IRD publications contribute to fundamental research on development and provide recommendations for public policies in the fields of health, economics, ecology, education and agriculture. The work of IRD staff is based on a multidisciplinary approach and contributes to the publication of original reports. Indeed, 62% of its publications are led with partners in southern countries.

B. Purpose of the internship

B.1) Subject of the study

My interest in research work in development economics motivated me to apply for a traineeship proposal from Mr Desdoigts. This internship was initially intended to focus on data collected by the AFD to deal with land tenure in Côte d'Ivoire. The unpublished character of the collected information created disagreements on the modalities of their exploitation, which delayed their delivery. We get out of this deadlock situation with the topic of informality. A conjoint work with academics of Rosario University in Bogota, Colombia was evoked. Their initial project was to replicate Ulyssea's 2018 paper on "Firms, Informality and Development: Theory and Evidence from Brazil" with Colombian dataset. This subject echoed previous work I made on this paper which had aroused my interest in the theme of labour and informality.

B.2) Principal objectives

During four months, I studied a Colombian survey with data about small businesses. The objective was to map informality in Colombia among small businesses. The aim of the internship was to gain efficiency in data mining and treatment data. At the beginning, I worked with R but the entry cost was too important and I realized it would take time before I could be efficient. Thus, the whole part of the project was carried out with STATA which fits more data analysis in the context of econometric problems.

We led this work in collaboration with some researchers of Rosario University of Bogota. As specialists of informality they expressed some special needs and told about interesting lines of research. We had weekly meetings during which we exchanged information on the progress of the research and on new elements to exploit. Part of my job was to respond to their requests to find specific information and to render it in a visual and comprehensive way.

III Studying informality: Literature review

A. A difficult terminology to define

A.1) Informality in grey literature

Informality is not a binary concept. It usually refers to legal activities and can represent employment opportunities for low skilled people. Defining the extent to which firms are formal or informal depends on the degree they adhere to institutional rules.

But, there is no general agreement on the definition of informality. This lack of consensus makes it difficult to study informality.

For **Lipton (1984)**: “The informal sector concept has become discredited on account of three alleged deficiencies: misplaced dualism, misplaced isolation and confusion”.

- Misplaced dualism: there is no clear split between formal and informal sector. There’s more like a continuum. Thus, we cannot define informality on a sectoral criterion.
- Misplaced isolation: the interaction between informal and formal economy needs further investigations
- Related to confusion is the fact that the definition of the informal sector is not quite clear and need more clarity.

These conclusions are still valid. Indeed, La Porta and Shleifer (2014) consider a dual view of informality. To them formal and informal sectors are highly disconnected and informal firms exist without productivity growth. They conclude that the difference of productivity between formal and informal firms is due to the fact that informal firms are small, run by poorly educated people and are inefficient. This lack of productivity prevents them from entering in the formal sector. To them, as formal and informal firms are different, they do not work in the same activities. Thus, informal firms are not a source of disloyal competitiveness.

From Ulyseas (2018) we are able to distinguish two margins of informality. This paper breaks the dual view and builds a bridge between formal and informal sectors. The author mentions the extensive and the intensive margins of informality:

“Two margins of informality: (i) not register their business, the extensive margin; and (ii) hire workers “off the books”, the intensive margin.”

He considers an approach where there exist two types of informality. His approach is innovative because it allows the existence of informal workers within formal firms to be considered. That

is to say, workers without a work permit - giving them access to social security rights - working in formal firms that have chosen to declare only some of their employees. Until now, only informal firms and specificities of informal workforce within the informal sector have been studied.

He develops a general equilibrium model and finds out that lower informality does not necessarily leads to higher production. Indeed, as informality allows less productive firms to compete with more productive firms, it leads to misallocation of resources but also allows less educated persons to be part of the workforce. If the legislation ensure that the formal sector shrinks, usually, informal workers would be kicked out of the labor market. So, sometimes, less informality is not desirable. In his paper, Ulyssea proposes a graphic showing the different views of informality:

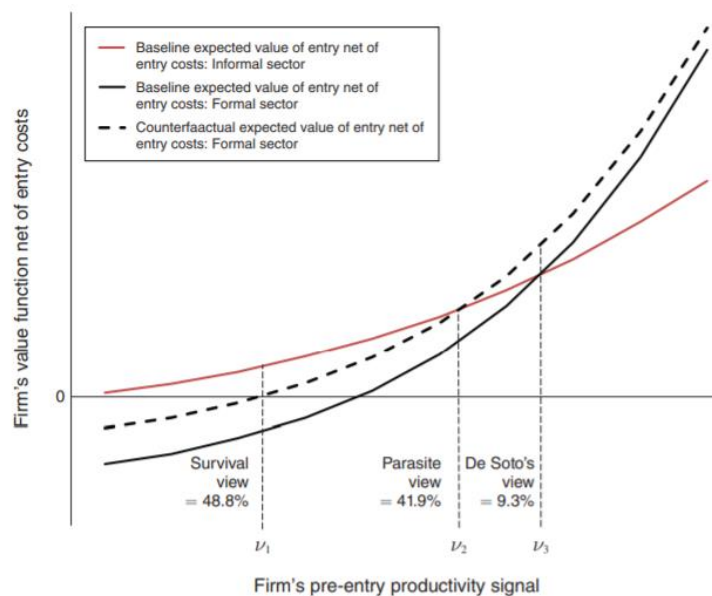


FIGURE 6. THE DISTRIBUTION OF INFORMAL FIRMS TYPES IN THE DATA

Notes: The figure shows, for each productivity level, firms' expected value function net of entry costs in the formal and informal sectors, $V_f^f(\nu) - E_f$ and $V_f^i(\nu) - E_i$, respectively. The third curve displays the net expected value of entering the formal sector in a scenario where its entry costs are equalized to informal sector's ($E_f = E_i$): $V_f^{f,c}(\nu) - E_i$.

Figure 2: Different views of informality
Source: Ulyssea (2018), Page 2037

The survival view seems to represent the major part of informal firms, it echoes La Porta's vision (2014). The Parasite view corresponds to firms productive enough to enter the formal market but who choose do to not in order to avoid regulation. In De Soto's view, informality is a transitional state. These are firms which have the productivity potential to enter the formal

market but decide to wait a productivity signal high enough to allow them being performant enough when they leave the informal market.

A.2) Informality according to international or national organizations

ILO’s approach (See Appendix 1)

The informal sector plays an important role in job creation, production and growth revenues in developing countries. It’s a major issue because it creates difficulties in identifying social security needs. It also represents important fiscal loss. That is why statistics on informality are essential in order to gain knowledge about employment quality.

Work is supposed informal if it is not submitted to national legislation, worker does not pay income tax, has no access to social security or advantages linked to employment (paid holidays, work accident insurance...).

Informal Employment	Informal Firms
Informal firm	No identity distinct from its owners
Work for proper consumption (subsistence economy)	Controlled and owned by one or more members of the household
Casual or short-term work Jobs or hours worked are lower than a certain threshold	Doesn’t have any balance sheet
Operational criteria: No social security No paid holidays or sick leave No written labor contract	Some criteria depending on the country of residence: number of employees, registration of the firm, employees are not declared

Production units by type	Status in employment (ICSE-93)									
	Employers		Own-account workers		Members of producers’ cooperatives		Employees		Contributing family workers	
	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal	Informal	Formal
Formal sector enterprise	Dark Grey	Light Blue	Dark Grey	Light Blue	Dark Grey	Light Blue	Dark Red	Light Blue	Dark Red	Dark Grey
Informal sector enterprise	Dark Red	Dark Grey	Dark Red	Dark Grey	Dark Red	Dark Grey	Dark Red	Light Blue	Dark Red	Dark Grey
Households ^(a)	Dark Grey	Dark Grey	Dark Red	Dark Grey	Dark Grey	Dark Grey	Dark Red	Light Blue	Dark Grey	Dark Grey

Source: ILO, 2003a, annex.

(a) Household enterprises producing exclusively for own final use and households employing paid domestic workers.

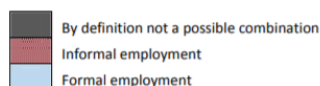


Figure 3: Intensive and Extensive margins of Informality according to ILO

DANE's definition

Through the document CONPES 3956, the DANE define informality via different steps of firm's life:

- **Entry formality:** Registration criteria
- **Inputs and production factor formality:** Declared workers, lands, capital and intermediate consumptions
- **Production formality associate to production process and commercialisation:** Respect of sanitary norms, technical rules and norms
- **Tax formality:** Declare revenue and workers and pay taxes on production

Also, one should note that according to the DANE, informality is a urban concept.

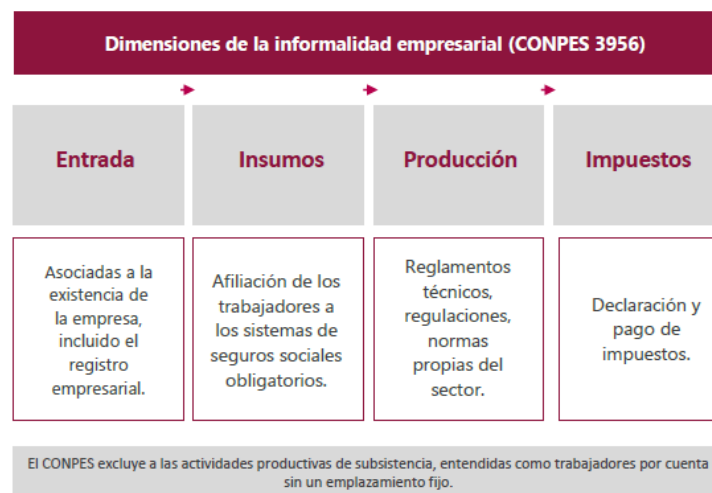


Figure 4: Dimensions of firms formality.
Source: CONPES 3956

Others

In a 2009 paper for the Colombian Ministry of Labor and Planning Office (DNP), Bernal calculates 23 potential measures of informality for Colombia and concludes that considering Health and Pension benefits as criteria for worker informality is the more suitable in the case of Colombia. Because it is highly correlated with other measures of informality, it makes is a good indicator that the individual has the entire package of benefits associated with formal employment.

B. Characteristics of Informality

B.1) Firms

Informality is characterised by its flexibility and its mobility. In Colombia, informality is mainly characterised by self-employment. Following Bhowmik (2005), informal workers are usually sellers who have no fixed shop or permanent structure where they can exercise their economic activity. The number of informal firms is countercyclical: when there is a crisis, the number of formal firms tends to shrink to the benefit of the informal sector. Workers whose jobs have not been safeguarded also move into the informal sector, which appears to be a refuge.

B.2) Type of worker

Fields (1990) develops an approach where less educated people tend to have informal jobs. He highlights that the dualism between formal and informal market can also be seen in the informal market. Less skilled people are more likely to have unpaid jobs.

A cross-sectional analysis conducted by Bernal (2009) through the 23 largest Colombian cities shows that higher score in higher education (Index by Ministry of Labor) explains between 44-49% of variation in informality. Data show that informality decreases with higher level of education as only 9% of people without high level of education are formal.

III A research internship

A. A database dedicated to micro businesses

A.1) EMICRON: General overview

Purpose of the database

Microfirm (DANE): Economic unit with a maximum of nine occupied persons and developing an activity of production of goods or services, with the aim of earning an income, and owning or renting their production facilities.

The survey is produced by DANE (Departamento Administrativo de Nacional Estadística), it is the equivalent of French INSEE. DANE produces national statistics over the whole Colombian economy. It contributes in shaping national policies.

EMICRON (*Encuesta de Micronegocios*) is a database constructed thanks to the Colombian household survey (GEIH – *Gran Encuesta Integrada de Hogares*)³. Its objective is to provide some statistical information on micro firms, that is to say, firms up to 9 employees. The survey is representative at national level and the agricultural, manufacturing, retailing and services sectors are surveyed.

These activities are gathered under two variables:

- **GRUPOS4:** Agriculture, Industry, Rental, Services
- **GRUPOS12:** Different activities are specified under the four sectors of the economy.

For the DANE, the necessity to build a survey specific for micro firms is justified by the fact that they constitute the bigger share of functioning firms in Colombia. EMICRON aims at reaching information about their activities and people involved in these activities. It also aims at providing complementary information to other surveys such as: Annual Manufacture Survey (EAM – *Encuesta Anual Manufacturera*), Annual Business Survey (EAC – *Encuesta Anual de Comercio*) and Annual Services Survey (EAS – *Encuesta Anual de Servicios*).

The objective of exhaustivity is reached through eleven modules of investigation. The database gives information over the type of activity, firm borrowings, its place of activity, characteristics of employees and variables of formalization, characteristics of the firm and variables of formalization, use of technologies (TIC), access to financial services, firm's costs and sells and information over social capital.

A photograph of Colombian micro firms⁴

In EMICRON 2019 there is an amount of 86 969 firms. When we compute this amount with a factor of expansion⁵ we can say that there are 5 874 177 micro-firms of less than nine workers in Colombia. 62.4% of them are run by men (3 666 041) and 36.7% of them are run by women (2 208 136). When we look at their status, 12.4% of them are said “*Patron*” and 87.6% of them are “*Trabajadores por cuenta propia*”. It is a status that depends on the size of the firm run by the owner. About 75% of them work alone.

Sectors of activity (share of firms with factor of expansion):

³ See below for further details

⁴ For a visual representation of the database please refer to appendix 3

⁵ The EMICRON database is intended to be nationally representative. The questionnaire was not submitted to every entrepreneur and some regions were better covered than others. The terminology expansion factor expresses the fact that a coefficient is applied to each data item so that the product obtained reflects the real value in the Colombian economy.

Agriculture (65% rural)	22.5%	1 319 240
Industry	11.6%	683 778
Retail (34% urban)	27.9%	1 638 724
Services (48% urban)	38%	2 232 435

70.6% of Colombian micro firms are urban and the other 30% are rural. In urban areas, 31% of activities take place in owner's house, among them only 19.2% have a space dedicated for the activity. Among all micro firms only 19.2% operate in a permanent shop. 20% are said itinerant and take place in street or are door to door activities (such as informal hairdressers).

Agricultural activities are mainly market gardening activities which do not need previous investment: 51.2% of activities take place in land fields and 20% in owner's place. Main activities in urban areas correspond to car repair and sales. There are also services activities such as taxis or Uber. Other activities of services are hairdresser or barber. Concerning retail activities most of them are street seller and itinerant shops.

42.4% of firms cannot be seen from the street. They are hidden and can be confounded with the place owners live in. This functioning can be due to multiply reasons, the first is that the activity do not need a special place to work properly, another hypothesis is explained by lack of funds, thirdly, we could think of a conscient choice in order to be able to avoid some controls or easily escape the taxes system.

When the owner of the firm does not operate alone different type of workers can take part to the activity: 59% of them are in the category paid workers, 14% are partners and 26% are under the category family and friends, they are unpaid workers.

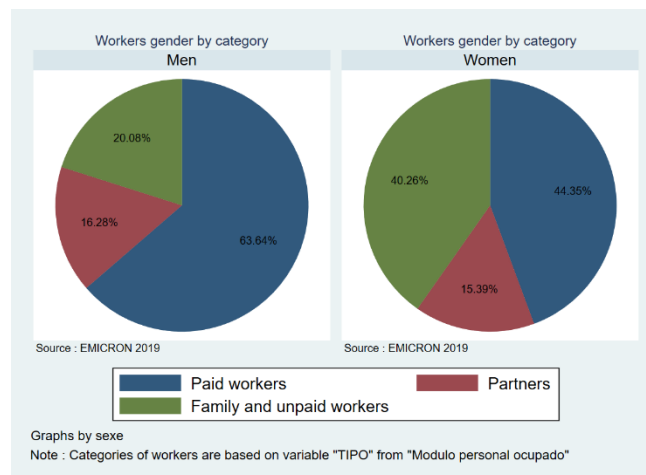


Figure 5: Categories of workers according to gender

In general, men are paid workers and women belongs to the category of unpaid family. With these pie charts it is difficult to conclude whether women tend to work less than men, or if they intervene in the household activity only when it is needed. However, we can suppose that they tend to be more vulnerable than men concerning their jobs.

Colombian informal work is characterized by self-employment. That is consistent with the fact that 74.1% of respondents declare they operate alone in the process of creation of their firm. They usually own only one firm, and, due to difficult access to credit in the informal sector 61.5% of Colombian micro firms are created on the basis of owner's personal funds.

34% of owners declare they were facing a lack of alternatives. It represents a total amount of 2 046 766 people. It is consistent with the survival view of firm informality: these workers are not able to enter the formal sector. 28.8% of them saw it as a business opportunity (1 691 548) and tend more to formalize even though they operate in a small firm.

A.2) Choices and definition of the variables of interest

Defining firm informality

Informal sector (DANE, 2021, from ILO's definition [2013]): Set of units engaged in the production of goods or the provision of services with the primary purpose of creating jobs and generating income for people involved in that activity. These units usually operate on a small scale, with a rudimentary organisation, in which there is little or no distinction between labour and capital as factor of production. Employment relationship - when they exist - are based on casual employment, kinship or personal and social relations, rather than on

In order to define firm informality, we had two potential variables:

RUT (*Registro Unico Tributario*): It corresponds to the registration of the firm at the competent authorities. After this formality, firms receive a number of identification and would declare their activity with this number. It must be done for once and it is free.

CC (*Camara de Comercio*): It corresponds to a kind of operating license. It has to be updated each year. It represents a certain cost. Thus, this is a more restrictive criteria of informality.

If we compare these approaches of firm informality we see (using sampling weights)⁶ that when 24.3% of firms have a RUT, only 12.2% comply to CC. Thus, CC is a more restrictive criteria of informality. If we refer to DANE's definition of informality, we made the choice of entry formality as relevant criteria.

One should keep in mind that informality is far from being a binary concept. Indeed, we could consider: holding of computability (68.9% of firms in EMICRON do not have one), tax declaration (only 21.9% of them actually declare revenues). Another approach could be through the renovation of the CC registration but it appears that usually, when firms register once, they tend to renovate their registration: 75.3% of registered firms updated their registration at CC in 2019. To make estimations easier we have decided to hold the criteria of compliance to *Camara de Comercio* as determinant of firm informality. It is possible to refer to appendix 2 in order to have an idea of the different degrees of informality.

Defining workers informality

Worker informality (DANE, 2021, from ILO's definition [1993]):
Employees and workers in establishments, businesses or enterprises employing up to "five" (are considered up to nine workers in EMICRON) persons in all their agencies and branches; unpaid family workers; domestic servants; own-account workers, except self-employed professionals; and employers in enterprises with ten or fewer workers. Government workers and employees are excluded from this sector.

There are also different approaches of workers informality. In our wish to be as close as possible to Ulyssea (2018) approach we looked for a variable that could be close to his use of *Brazilian carteira de trabalho* as a proxy for a formal labor contract. In Brazil, this document give access to fundamental workers' rights such as contribution to health and retirement, right to paid

⁶ Such as the obtained percentages are computed on the actual national headcount.

holidays, health security and so on. In EMICRON the closest variable to this approach is the one about health and retirement. We learn that, among paid workers (because we consider that unpaid workers are always informal) 80.4% of them do not benefit from health and retirement contribution.

We could also have considered the type of contract. Usually, temporary workers are employed without a contract. There also exist a variable about social prestation. It shows that 88.2% of paid workers do not pay them. Along a discussion with Rosario's academics we learnt that the difference between the chosen variable (health and retirement) and the one about social prestation is the fact that the worker himself choose to pay his social prestation. The decision to contribute to health and retirement belongs to the owner of the firm. Another variable we could have think of, but which is more specific is ARL (*Administracion de Riesgo Laborales*) it is a contribution to protect workers against accidents that could occur at work (only 18.3% of firms' owners contribute to ARL).

Finding information over workers skills

The sample is built on the basis of previous answers to GEIH. If people declare they work in a firm with nine or less people they identify the owner of this firm and he gets interviewed a second time for the purpose of EMICRON.

Common keys for merging process:

- *DIRECTORIO: equivalent of the locality*
- *SECUENCIA_ENCUESTA: equivalent of the surveyed house*
- *SECUENCIA_P: number of individuals who are owner of a micro business in the house*

These keys exist in both database but do not correspond to the same information (above description holds for EMICRON). Variables had to be renamed the same way and must be computed in order to generate a unique identifier for each firm. After this work, correct matching was possible.

In Ulysea (2018) the author exploits information over workers' skills. EMICRON 2019 does not provide such elements. However, as the sample is built over GEIH it makes it possible to merge both datasets. The sample in EMICRON is way smaller than in GEIH, first because it is only applied to small firms, secondly because only owners of firms are asked. In GEIH, all Colombian inhabitants are surveyed.

This process allowed us to have information over employed workers. However, it was found that observations over skills are only available for owners of the firm.

A representative sample

Potential bias in the scope of the sample can't be ignored. EMICRON survey is said representative at the national level. However, remoteness of certain regions can make it difficult to access and some are better represented than other. In order to make possible the analysis of the whole population concerned by the survey, we can use sampling weights⁷. However, it should be borne in mind the fact that all estimations have to be considered with a certain degree of confidence.

B. Results

Some of the figures and tables below are directly inspired from Ulyssea (2018). Initially, we wanted to be able to lead a comparison between Colombian and Brazilian micro-firms. However, we should be conscious about the fact that:

1. Ulyssea didn't consider administration or government activities. We choose to keep the whole EMICRON sample.
2. Ulyssea restricted his observations to firms younger than one year. In our analysis, the majority of firms are aged of more than ten years and we would have lost too many observations if we made his restriction choice.
3. Ulyssea has skills information we do not have and we had to led some variations and only observe owner's education level.

⁷ Appendix 4

B.1) Descriptive statistic

Descriptive statistics - Employees									
	Formal			Extensive Margin			Intensive Margin		
	Mean	SD	Total	Mean	SD	Total	Mean	SD	Total
Log (wage)	12.37	4.09	4 638	5.47	6.40	15 386	8.71	6.37	13 721
<i>Sectoral composition (dummy)</i>									
Agriculture	0.03	0.19	4 638	0.29	0.45	15 386	0.06	0.24	13 721
Mining & Manufacturing	0.15	0.36	4 638	0.13	0.33	15 386	0.17	0.37	13 721
Construction	0.10	0.28	4 638	0.09	0.28	15 386	0.08	0.27	13 721
Mechanics	0.35	0.47	4 638	0.22	0.41	15 386	0.31	0.46	13 721
Transports	0.03	0.17	4 638	0.01	0.11	15 386	0.02	0.14	13 721
Food services	0.12	0.32	4 638	0.19	0.39	15 386	0.16	0.37	13 721
Telecom	0.01	0.12	4 638	0.01	0.09	15 386	0.01	0.12	13 721
Education, Admin. & Health	0.17	0.37	4 638	0.02	0.15	15 386	0.09	0.29	13 721
Arts & Others	0.02	0.15	4 638	0.03	0.18	15 386	0.06	0.24	13 721
Age	36	11.04	4 638	36.29	14.61	15 386	35.70	13.01	13 721
Sexe (dummy)	0.59	0.49	4 638	0.60	0.48	15 386	0.58	0.49	13 721

Source: EMICRON 2019

Notes: See Infographics « Database composition »: Formal corresponds to formal firms and their employees with social security, Extensive margin corresponds to workers in informal firms, the Intensive margin of informality refers to informal workers (no health and retirement contribution) in formal firms. The sample includes employees (paid employees, partners, family and unpaid workers) between 10 and 80 years old (N=33 745). In order to assess for wage gap amongst all employees, we attribute value 1 to unpaid workers, all others wages take +1.

Table 1: Descriptive statistics

Like in Ulyssea, we consider as employees the total amount of workers in a firm: the owner is said “self-employed” and other workers are considered as “employed by the owner of the firm”.

Here, we can compare different variables following the different margins of informality. It is possible to confront them to formal ones. As we only consider workers we have a total amount of observations of 33 745 observations. Formal firms represent a small minority of the sample.

Having a look at mean age and gender we can say that there is not difference between formal and informal sector. It means that age and sexes does not seem to be a determinant of informality. However, if we have a look at wage gaps it is in average twice bigger in formal firms than in informal firms. In the intensive margin, average wage is higher than in the extensive margin. It can be due to the fact that the firm avoids important wage gap so that formal and informal workers cannot be distinguished on the basis of their remuneration.

There is no dualism in the type of activity between formal and informal firms. They both operate in all sectors of the economy. However, they are not distributed the same way. Indeed, mechanics and car selling represent 35% of formal firms when it is only 22% of informal firms. Also, 30% of informal firms operate in the agricultural sector but represent less than 1% of formal firms. On the intensive margin of informality, 16% of informal workers in formal firms operate in food services and 17% of them operate in manufacture. These activities occupy the biggest share of informal workers in formal firms.

B.2) Gender approach

By type of occupation⁸

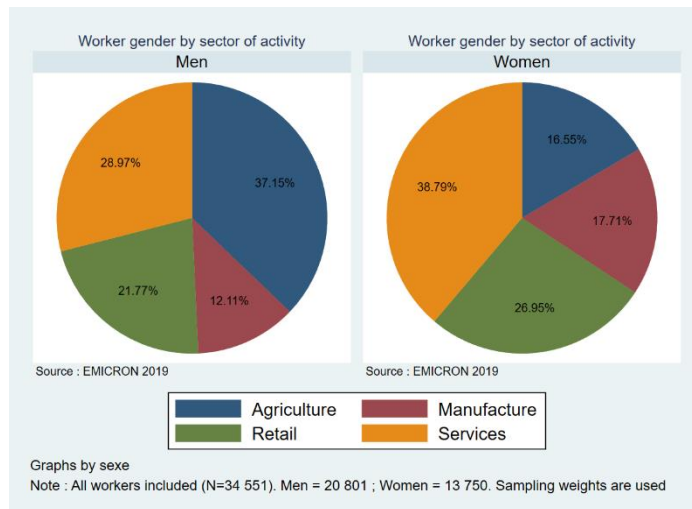


Figure 7: Worker Gender by sector of activity

In formal and informal sectors confounded men and women do not work in the same sectors of activity. While 37% of men are occupied in agricultural activities, there are only 16.5% of women in this sector. The majority of women are occupied in services activities.

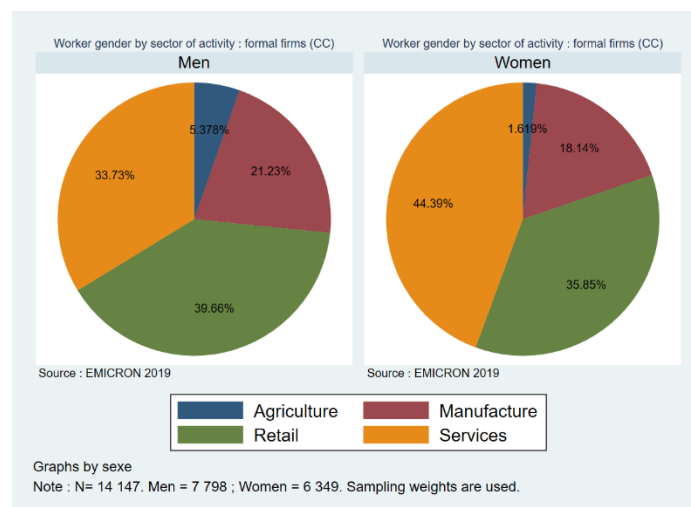


Figure 8: Worker gender by sector of activity in formal firms

⁸ See Appendix 5 for the approach through RUT criteria

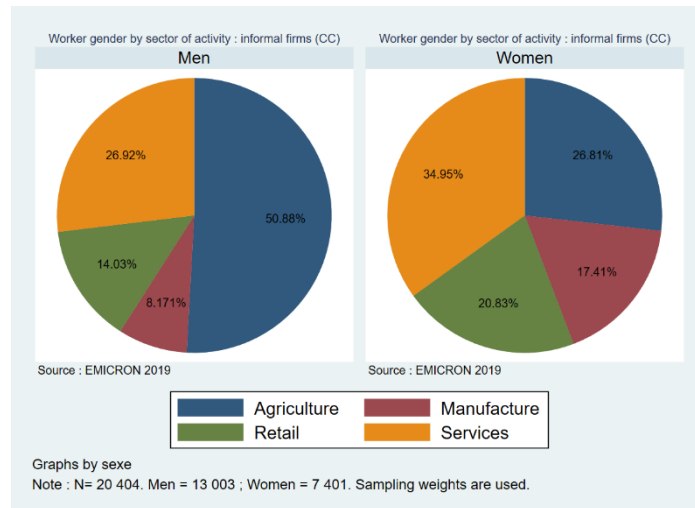


Figure 9: Worker gender by sector of activity in informal firms

If we make a distinction between formal and informal sectors we can see that the share of occupied persons in agriculture shrinks in the formal sector. However, it occupies 50% of the masculine workforce in the informal sector.

Men and women in the formal sector are between 75 and 80% in retail and services activities. This ratio is different in informal firms but these activities still concerns more than 50% of women.

With these pie charts we can see that even if formal and informal firms are operating in the same sectors of activity there exist ratio differences. Also, men and women are not occupied in the same activities, there is a gender dualism.

According to wages and skills

The sample only includes the category of paid workers. We have a number of observations equal to 20 597. In this sample, only 342 workers are not paid. Their presence has no significant effect on average wage gap. We choose to include them in the analysis.

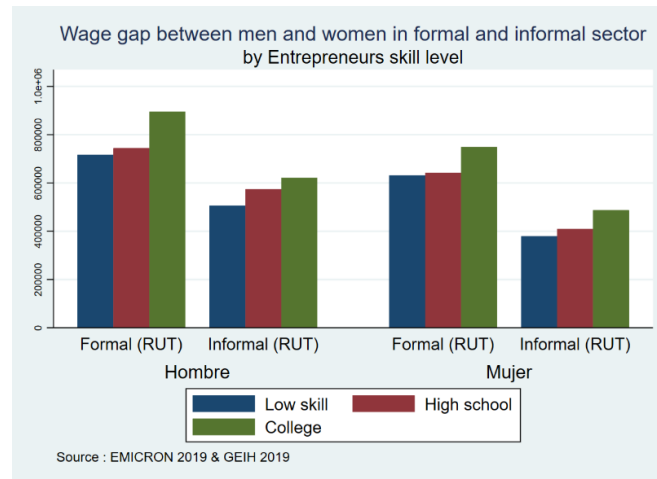


Figure 10: Analysis without sampling weights and no confidence intervals

In the first figure criteria of informality is only based on RUT⁹. Also, we do not apply sampling weights. Here, we can conclude that independently of individual's gender people tend to have a better wage in formal sector than in informal sector. Also, men tend to be paid more than women in formal and informal sectors respectively. However, we cannot draw more conclusion about the gender gap because we have no information concerning the number of hours worked by individuals each week (or month). Though, we can't say if men's wage rate is higher than women's ones.

This graph gives us information on wage level following entrepreneur's skills. We can say that in average, when the owner of the firm is high skilled his workers have a better remuneration. This is less true in the informal sector where wage levels tend to be similar for each gender regardless of the owner's skill level. It could lead us to conclude that returns to school are lower in the informal sector. However, we can't draw such a conclusion in the absence of confidence interval.

⁹ RUT is free and gives the firm a tax identification. On the other hand, *Camara de Comercio* has to be repaid each year and represent a fixed cost for the firm.

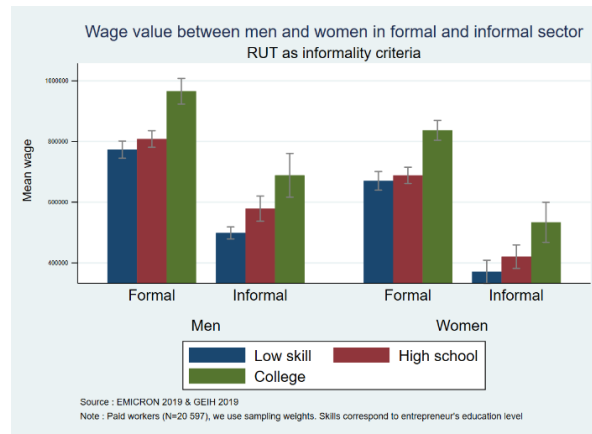


Figure 11: Analysis with sampling weights and confidence interval: wage gap between formal and informal sector following RUT criteria

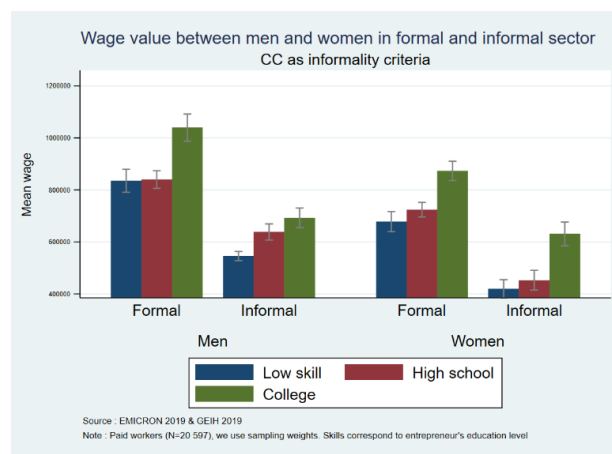


Figure 12: Analysis with sampling weights and confidence interval: wage gap between formal and informal sector following CC criteria

When we use the sampling weights, we can see that the gap increases in the formal and the informal sector for both criteria. Also, confidence intervals allow us to conclude that these gaps are significant.

If we look at the wage gap between men and women in the informal sector we can see the difference is more important than in the formal sector. In the formal sector following *Camara de Comercio* criteria, wages are more important than when RUT criteria is used. It is due to the fact that CC is more restrictive, so, firms who tend to be registered following CC should be more productive or bigger. We can say that they should benefit from registration.

Entrepreneur skill influence worker's wage value. The gap is more important as the owner is more qualified. As this influence is bigger between High school and college in the formal sector, it seems that returns on education are more gradual in the informal sector. How can we interpret this? Confidence intervals are small, this gap is significant. Even if we can say that returns on

education are smaller in the informal sector (for a same level of education the wage gap remains important between formal and informal sector) they exist in spite of everything. So, we can conclude that without considering entrepreneur's skills it is better to be in the formal sector if you expect a more important wage. But, if you have to be hired in the informal sector, you'd better to be hired by a skilled entrepreneur if you expect a wage that wouldn't be too low.

B.2) Regression (See appendix 6)

1. Variable social protection has a high level of significance even when we add some fixed effects. It explains 31% of the wage gap between formal and informal workers on the **extensive margin of informality**. Within **formal firms**, the wage gap between formal and informal workers remains highly significant and explains 20% of wage gap. We can explain the decrease of coefficient's size. Formal firms tend to reduce the gap between formal and informal workers in order to limit the risk of being caught if remuneration differences are too important.
2. Sexe difference can explain 14% of wage gap on the extensive margin of informality when we control for sector and geography. It means that, even we men and women work in the same sector they are still not equally paid. The coefficient size tend to decrease in formal firms
3. Time spent in the firm by the worker is also significant but does not have an important impact on wage gap.
4. Age and Age square variable can tell us experience have not important impact on wages
5. The size of the coefficient for the time spent by the owner in his firm each week is significantly different from zero when we look at columns 1 and 2 but, when we restrict the sample to formal firms it is not the case anymore.
6. Owner's gender also has an impact on wage gap. It means that whether the owner is a man or a woman, worker's wage in firms run by men are in average 14% higher in the extensive margin of informality. In the sample restricted to formal firms this difference is reduced to 8% difference.
7. Owner's education has a high level of significance and explains 10% or 8% of wage gap. It means that in average more educated owners tend to pay workers with better wages.
8. Size of the firm has a significant impact on wages. Coefficients are not very important (4% on the extensive margin and 2% on the intensive margin). It can be interpreted as the fact that as firms get bigger they tend to formalize because the opportunity cost of informality increases.

9. Formal firms pay a share of their workers at minimum wage level (formal workers). Thus, usually, wages are bigger in formal sector. As firms get bigger, they tend to formalize and formal firms pay better wages. Wages are positively correlated with firm size.

10. Urban dummy is not significant when we control for sector and geography. The sign changes in formal firms following *Camara de Comercio* criteria with fixed effects.

B.3) Informality, wages, education

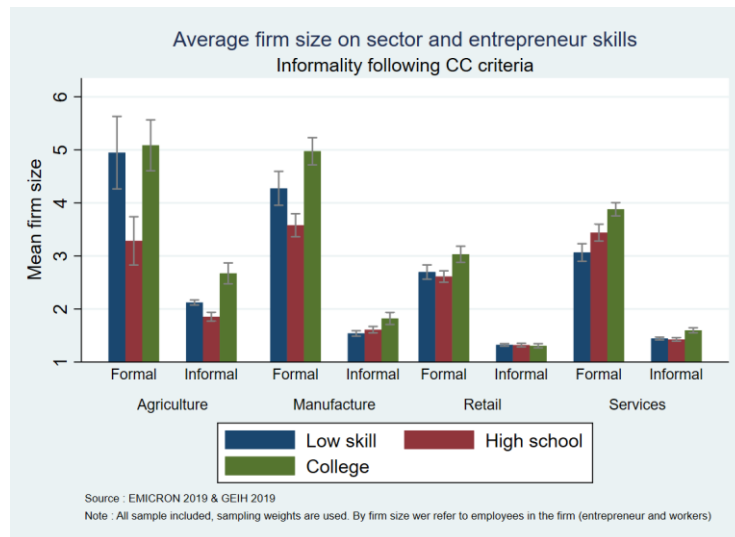


Figure 13: Firm size, sector, and entrepreneur skills

- Firms run by college skilled entrepreneurs are bigger in average and formal firms are bigger than informal firms
- In Informal firm entrepreneur skills does not have a significant impact on firm size, except in informal agriculture where college educated entrepreneurs tend to run bigger firms than less skilled owners
- Formal agricultural firms and manufacture firms are bigger than firms who exercise in retail or services

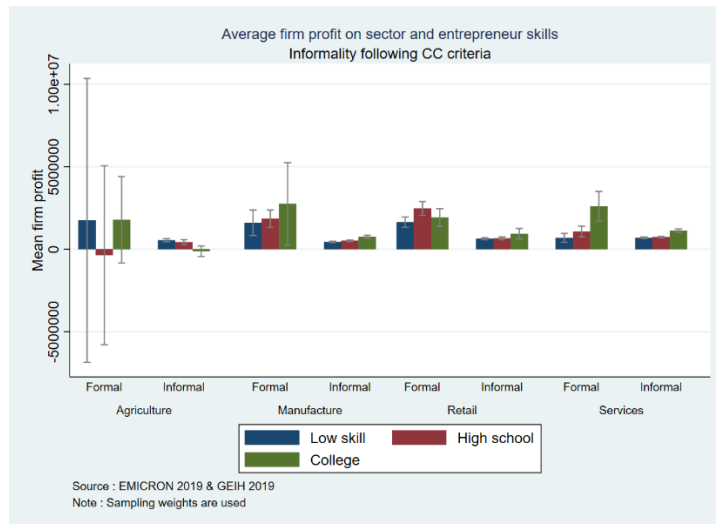


Figure 14: Firm profit, sector and entrepreneur skills

- Firms run by college skilled entrepreneurs are bigger in average & Formal firms are bigger than informal firms
- In Informal firms entrepreneur skills does not have a significant impact on firm size, except in informal agriculture where college educated entrepreneurs tend to run bigger firms than less skilled owners
- Formal agricultural firms and manufacture firms are bigger than firms who exercise in retail or services

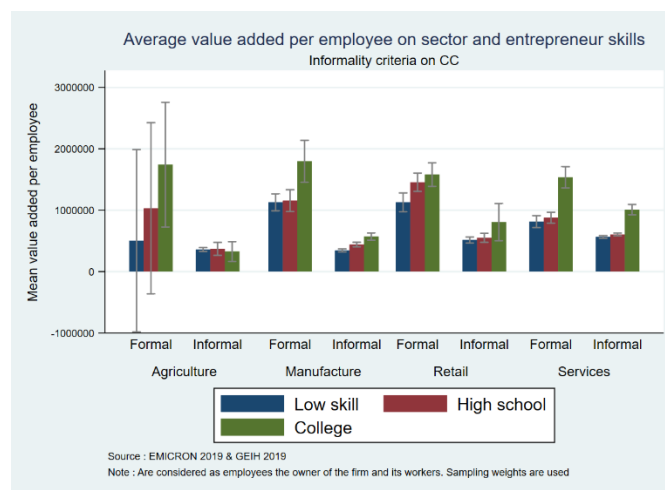


Figure 15: Value added per employee, sector and entrepreneur skills

- Whatever entrepreneurs' skill level, in manufacture, retail and services value added per worker is more important in formal sector. It is not necessarily true in agricultural sector.

- Value added is not always significantly higher in firms run by high skilled entrepreneurs except for those who reached college in formal manufacture and formal services

C. Has the challenge been met?

Results put forward during the internship were consistent with the literature. We found that on average informal firms are less productive and smaller than formal firms. We also found out that formal and informal firms operate in the same sectors of activity. Also, we found evidence that a lot of variables influence wage level, gender but also informality criteria. Access to social security for workers is an important determinant of wage gap between formal and informal workers. Thus, it confirms the importance of reaching objectives of decent work through the generalization of access to social security.

In the limited time we had it was difficult to draw more conclusions. Further investigations on motivations and access to financial intermediation would have enriched our findings. Also, some question remained unanswered: Why Colombian informality is characterized by self-employment? How can we explain that owner's skills influence wage even in informal firms? With absence of information on employed skills it was hard to define the average profile of informal workers, we were only able to conclude that, independent on skills, there is no specific characteristic of informal workers.

IV Conclusion and lessons learned

Working in research field was the opportunity to measure how difficult it is to study such an important subject in a few months. Indeed, my feeling about informality was quite different between the begging and the end of my traineeship. It reminds me of a quote from Benjamin Disraeli saying that "Being aware that one is ignorant is a big step towards knowledge". That is what I realized studying informality. I have barely approached the issues underlying this notion, the mechanisms governing individual decisions have remained obscure and the determination of what is informal and what is not has proved to be more of a "partipris" than an incontrovertible definition. I have realized that research is a long run process and a work of every moment. That it is hard to find answer but asking the right question is even more difficult.

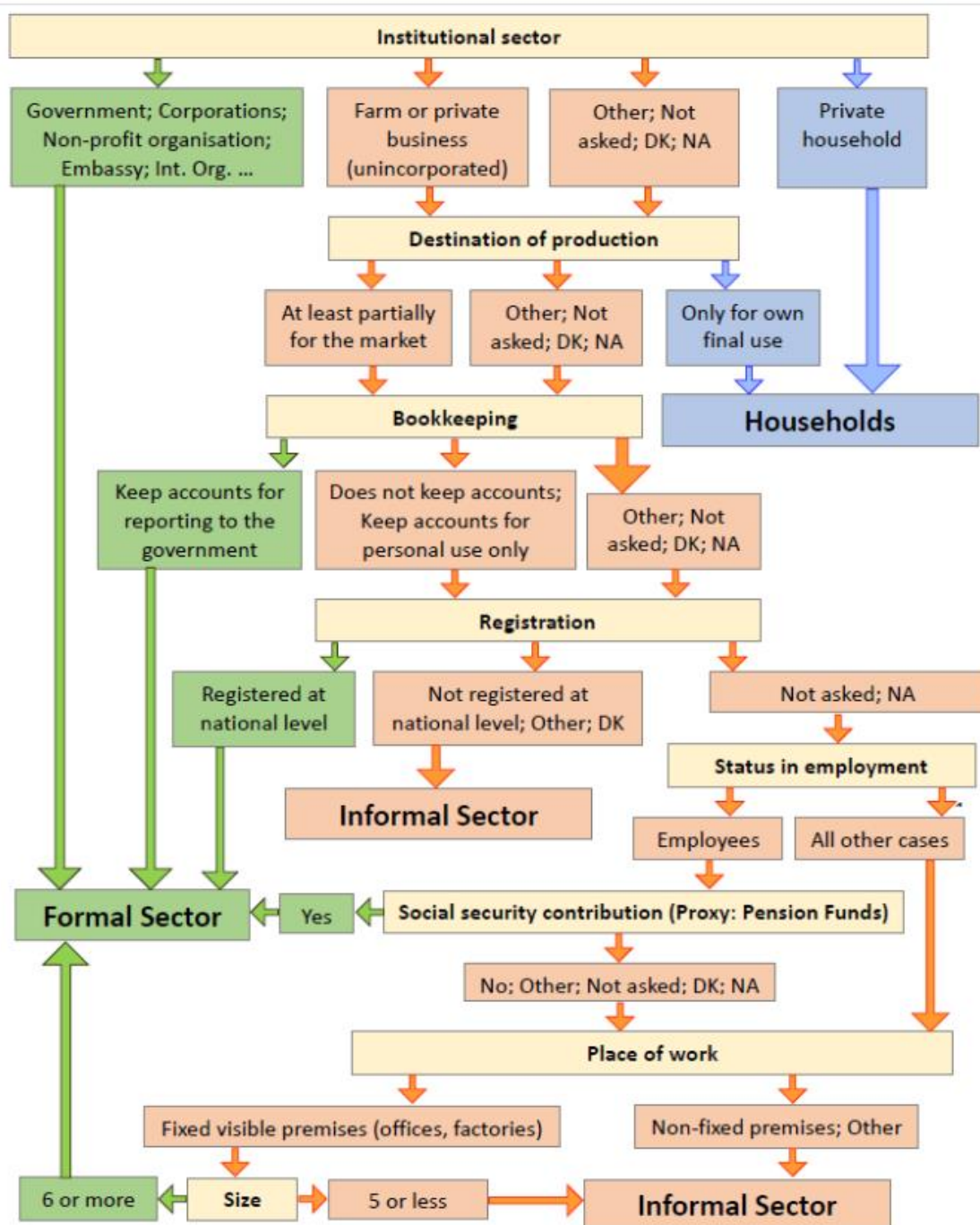
I feel like I have learnt a lot about the necessity to have precise daily objectives and to ask very precise questions in order to keep in being concise and relevant. My obsession with completeness has been hardly mistreated. I had to realize that in the limited time we had,

complete understanding and total exploitation of the database was anything but possible. Accepting the fact that I couldn't allow myself to work with multiple analysis angles was part of the job. Important lessons were learned, Mr Desdoigts repeatedly told me to be humble with the information I gathered and to avoid drawing too much conclusions from a single observation.

Difficulties to overcome were important, as this type of work was new I struggled to find an equilibrium between data analysis and literature research. At the beginning, I also had difficulties to talk about unachieved part of my work and present results I was not confident about. It is part of research work I had to get used to. As we had weekly meetings with academics of Bogota I learnt to put forward and present personal work even if it was unachieved, I also learnt to ask about some help and guidance. This interaction was important in the process of setting research priorities. Thanks to these exchanges I felt more confident about the quality of the work I can make, also about the ideas I can have concerning ways to introduce data (see the Infographics – Appendix 2).

Appendix

Appendix 1: ILO's definition of informality: Histogramms



Formal and Informal Sectors
Source: ILOSTAT

Step 1: Institutional sector = juridical organisation and property. Government, NGOs are in formal sector, owners who are physic persons or households are considered households

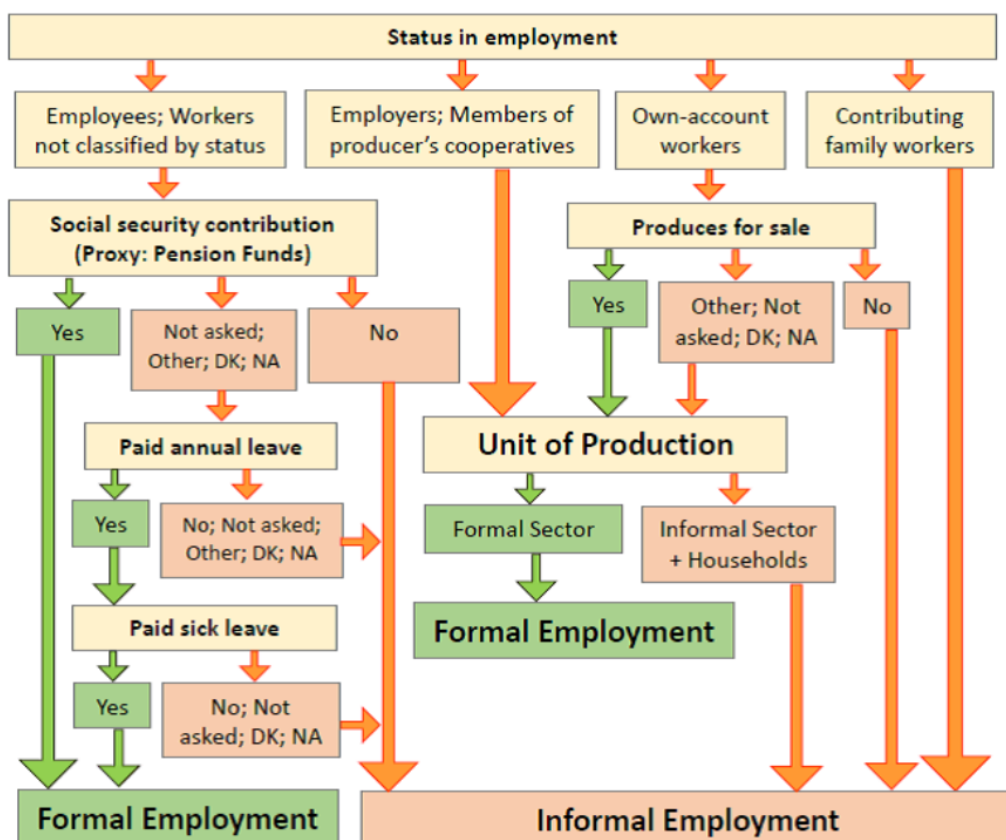
Step 2: Production destination identify if some goods are sold

Step 3: Bookkeeping in order to know if the economic unit has some balance sheets or assessment required by law. **If there's some official accounts, it's enough to be considered formal.**

Step 4: Registration in front of any authority: social insurance, tax revenue, chamber of commerce. Use of proxy for employees concerning social security

Step 5: Place of work: fixed, visible...

Step 6: Firm size (threshold = 6)



Formal and Informal Employment
 Source: ILOSTAT

Step 1: Employees: Access to social security, annual leave, sick leaves

Step 2: Self-employed workers: in informal sector, subsistence production, help from family, domestic workers

Appendix 2: Informality as a continuum rather than a binary outcome (Shown by Christina Fernandez)

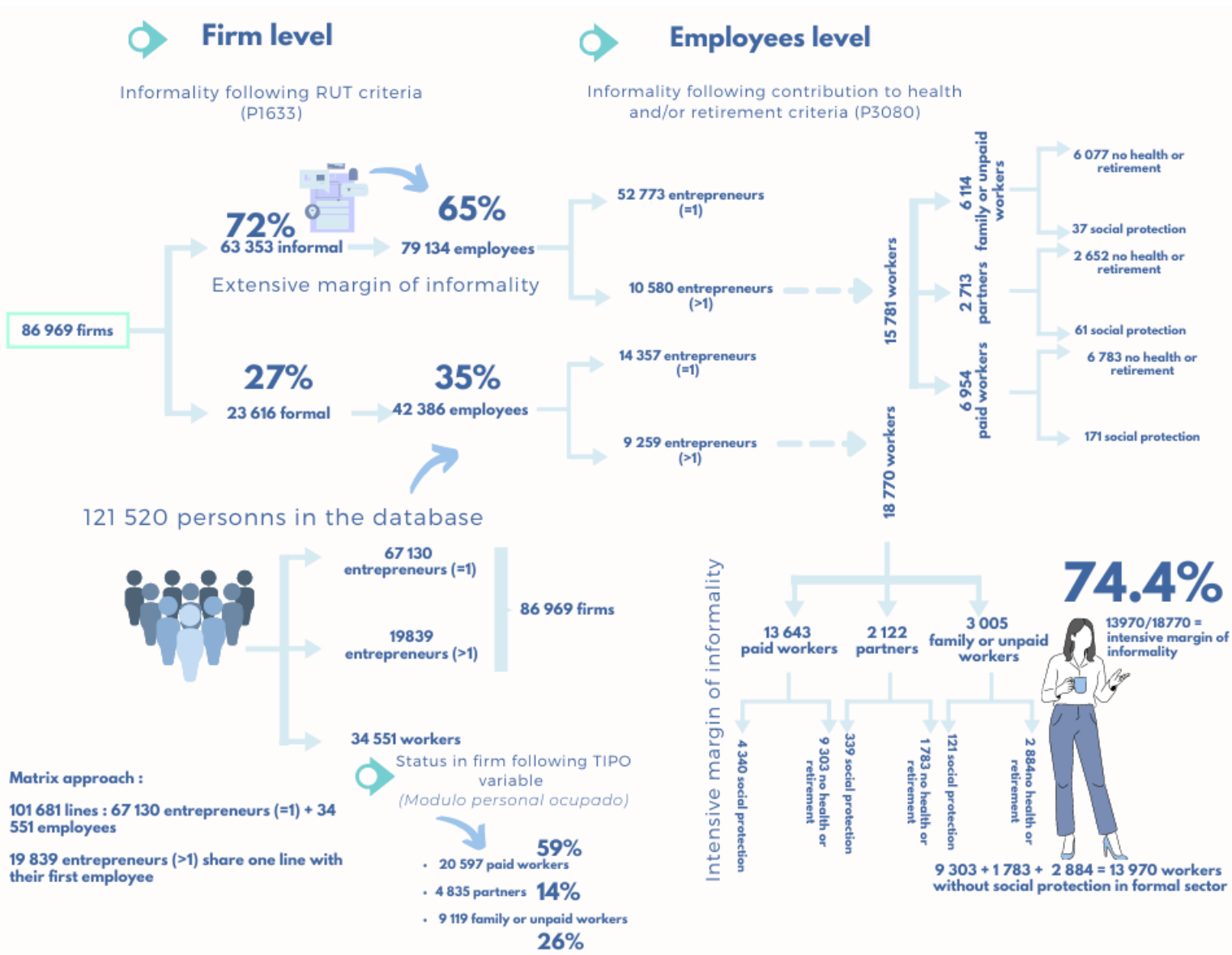


Appendix 3: Infographics: database composition

Following infographics represent the database composition according to the two margins of informality defined by Ulysea. We can distinguish firm informality with criteria of registration in front of the competent authorities, that is to say whether they comply to *Registro Unico Tributario* or pay *Camara de Comercio*. We can also have a view on workers informality with the criteria of contribution to health and retirement.

Source : Author's own creation

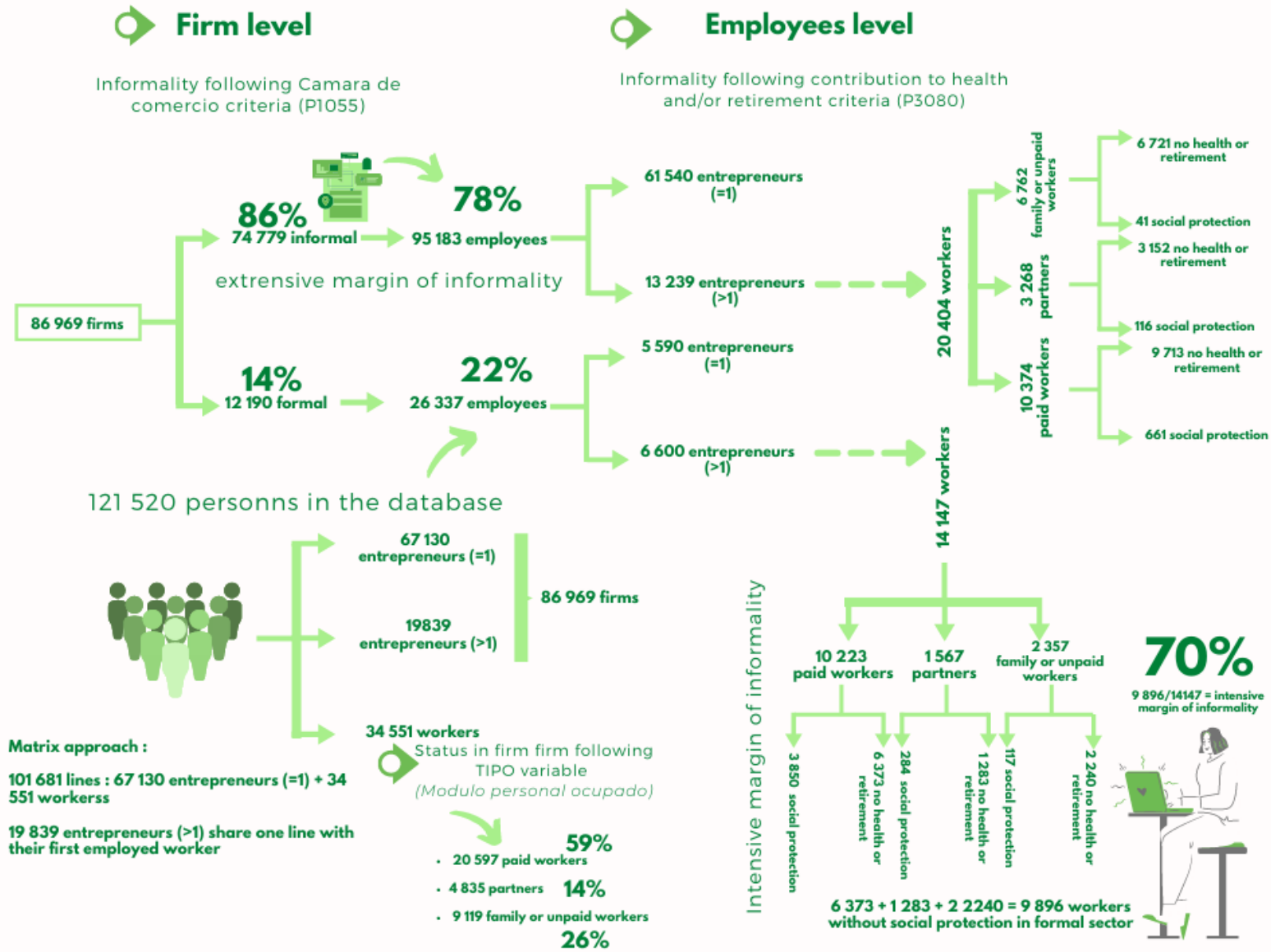
EMICRON 2019 - DATABASE COMPOSITION



Definitions : Ulyssea (p. 2030-2031) : "Firm's minimum size is one employee".

We consider two type of employees : Entrepreneurs who are employed in their own firm + Workers employed by entrepreneurs

EMICRON 2019 - DATABASE COMPOSITION



Definitions : Ulyseea (p. 2030-2031) : "Firm's minimum size is one employee".

We consider two type of employees : Entrepreneurs who are employed in their own firm + Workers employed by entrepreneurs

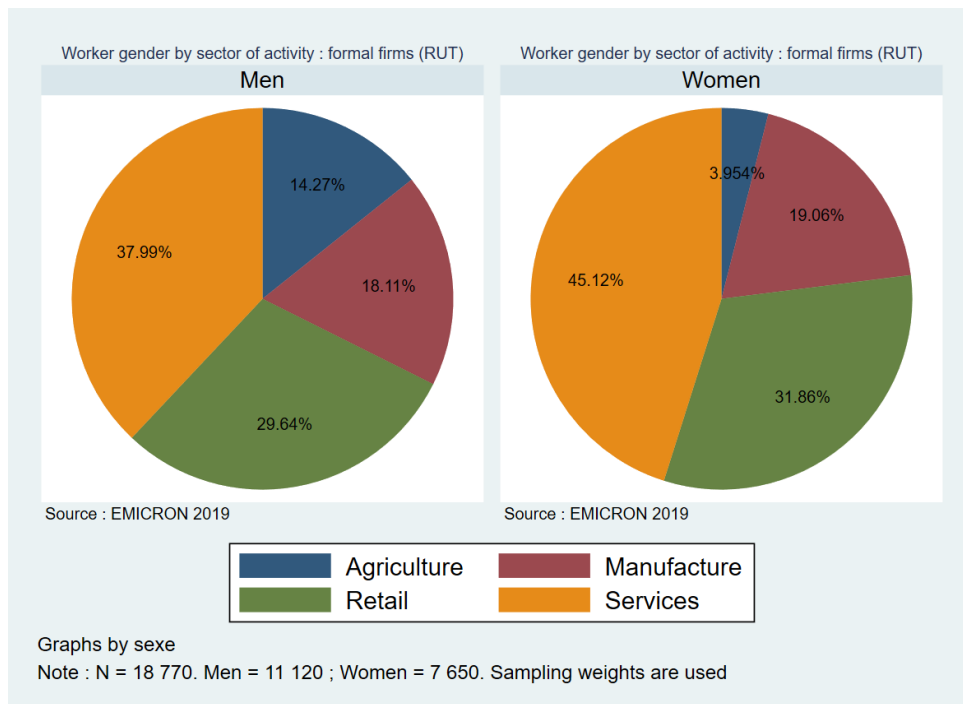
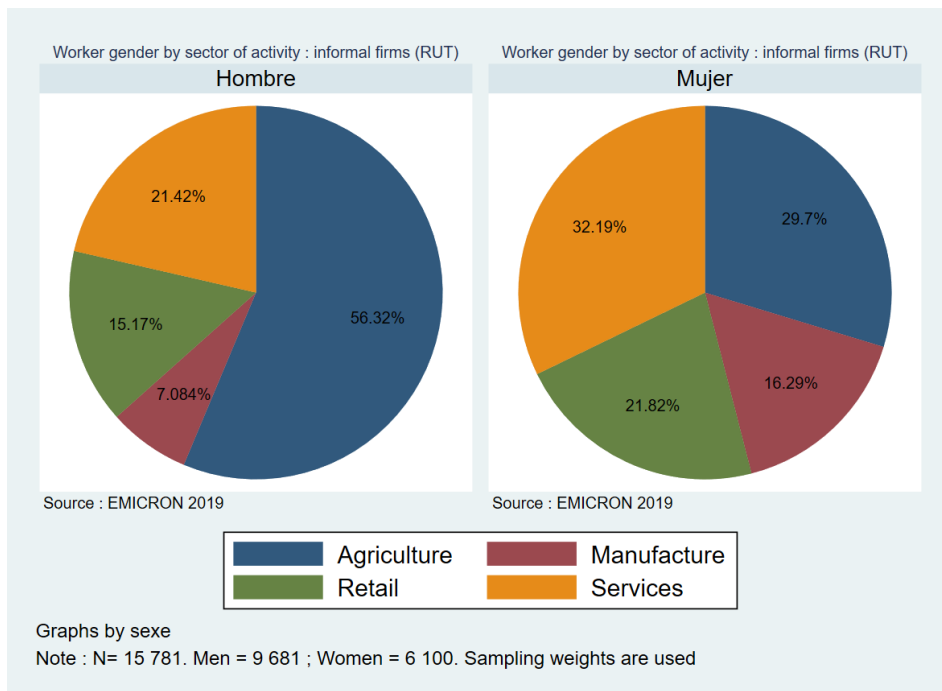
Appendix 4: Survey coverage (in number of persons)

**Encuesta de Micronegocios -EMICRON-
Cobertura en personas 2019**

Total		Urbano		Rural	
Pereira	98,45%	Pereira	98,12%	Santa Marta	99,70%
Santa Marta	98,16%	San Andres	98,10%	Neiva	99,16%
Tunja	98,15%	Tunja	97,95%	Ibagué	99,11%
San Andres	98,10%	Armenia	97,64%	Montería	98,95%
Sincelejo	97,96%	Cúcuta	97,21%	Pereira	98,79%
Armenia	97,81%	Sincelejo	97,19%	Sincelejo	98,73%
Neiva	97,75%	Barranquilla	96,77%	Barranquilla	98,54%
Cúcuta	97,70%	Santa Marta	96,62%	Tunja	98,36%
Barranquilla	97,65%	Florencia	96,55%	Cúcuta	98,19%
Ibagué	97,61%	Villavicencio	96,46%	Riohacha	98,13%
Manizales	97,15%	Neiva	96,34%	Manizales	98,01%
Montería	97,00%	Manizales	96,30%	Armenia	97,98%
Florencia	96,88%	Ibagué	96,11%	Total Rural	97,37%
Riohacha	96,80%	Quibdó	96,06%	Medellín	97,33%
Villavicencio	96,75%	Total Urbano	96,06%	Florencia	97,21%
Total Nacional	96,71%	Pasto	95,83%	Pasto	97,20%
Pasto	96,51%	Bucaramanga	95,76%	Villavicencio	97,04%
Medellín	95,83%	Popayán	95,58%	Bogotá	96,62%
Bucaramanga	95,78%	Riohacha	95,47%	Cartagena	96,30%
Bogotá	95,74%	Montería	95,04%	Bucaramanga	95,80%
Quibdó	95,71%	Bogotá	94,86%	Quibdó	95,36%
Cartagena	95,03%	Cali	94,53%	Cali	94,99%
Cali	94,76%	Medellín	94,33%	Valledupar	94,52%
Popayán	94,54%	Cartagena	93,76%	Popayán	93,51%
Valledupar	93,65%	Valledupar	92,79%		

Fuente: DANE. EMICRON-2019

Appendix 5: Type of occupation according to gender (Informality with RUT criteria)



Appendix 6: *Log(wage) regression over independent variables in the extensive and the intensive margin of informality*

VARIABLES	Log(wage)			
	(1) All firms	(2) All firms (with FE)	(3) Formal firms (CC)	(4) Formal firms (CC with FE)
<i>Worker's variables</i>				
Social protection	0.4216*** (0.027)	0.3101*** (0.026)	0.2576*** (0.028)	0.1975*** (0.027)
Sexe (<i>dummy</i>)	0.1415*** (0.024)	0.1430*** (0.024)	0.1179*** (0.021)	0.1040*** (0.019)
Time spent in the firm	0.0008*** (0.000)	0.0011*** (0.000)	0.0009*** (0.000)	0.0011*** (0.000)
Age	0.0197*** (0.005)	0.0186*** (0.005)	0.0275*** (0.006)	0.0251*** (0.005)
Age square	-0.0002*** (0.000)	-0.0002*** (0.000)	-0.0003*** (0.000)	-0.0003*** (0.000)
<i>Entrepreneur's variables</i>				
Hours worked (per week)	0.0077*** (0.001)	0.0062*** (0.001)	0.0004 (0.001)	0.0009 (0.001)
Sexe (<i>dummy</i>)	0.1628*** (0.032)	0.1438*** (0.030)	0.1027*** (0.034)	0.0805*** (0.029)
Education	0.1529*** (0.017)	0.1237*** (0.015)	0.1021*** (0.018)	0.0833*** (0.016)
Firm size	0.0371*** (0.009)	0.0432*** (0.008)	0.0251*** (0.007)	0.0234*** (0.006)
Urban (<i>dummy</i>)	0.3157*** (0.043)	0.0569 (0.057)	0.0237 (0.055)	-0.0234 (0.061)
Observations	19,438	19,438	9,661	9,661
R-squared	0.220	0.311	0.166	0.249
Firm FE	No	No	No	No
Sector and geography FE	No	Yes	No	Yes
Sample	All	All	All	All

Robust standard errors in parentheses
 *** p<0.01, ** p<0.05, * p<0.1

Source: EMICRON 2019, GEIH 2019. Author's own calculation.
 Note: We have two different samples. Columns 1 and 2 include workers within formal and informal firms, columns 3 and 4 include workers within formal firms following the *Cámara de Comercio* criteria. We exclude partners (socios) and family (familiares) categories and only consider workers with positive wage (i.e., n=19 438 & n=9 661). Residuals are clustered at the firm level (Moulton, 2006).

References

- Bhowmik. (2005). Street Vendors in Asia: A Review. *Economic and Political Weekly* 40(22):2256-2264. DOI: [10.2307/4416705](https://doi.org/10.2307/4416705)
- CONPES 3956 (2019). Politica de formalizacion empresarial
- DANE (2021). Direccion de metodologia y produccion estadistica – DIMPE. Sistema de informacion para actividades economicas informales – SIECI. *Plan general del sistema de informacion para actividades informales*.
- DANE. (2019). Encuesta de Micronegocios.
- Feige. (1990). Defining and Estimating Underground and Informal Economies: The New Institutional Economics Approach. *World Development* 18(7):989-1002. DOI: [10.1016/0305-750X\(90\)90081-8](https://doi.org/10.1016/0305-750X(90)90081-8)
- FERNÁNDEZ, CRISTINA, et LEONARDO VILLAR. « The Impact of Lowering the Payroll Tax on Informality in Colombia ». *Economía* 18, no 1 (2017): 125-55. Fields, Gary S. « Labour Market Modelling and the Urban Informal Sector: Theory and Evidence », 1 janvier 1990. <https://ecommons.cornell.edu/handle/1813/75798>.
- ILO. (2018). 20th International Conference of Labour Statisticians: Revision of the 15th ICLS resolution concerning statistics of employment in the informal sector and the 17th ICLS guidelines regarding the statistical definition of informal employment.
- La Porta, Rafael, et Andrei Shleifer. « Informality and Development ». *Journal of Economic Perspectives* 28, no 3 (septembre 2014): 109-26. <https://doi.org/10.1257/jep.28.3.109>.
- Levenson, Alec, et William Maloney. « The Informal Sector, Firm Dynamics, and Institutional Participation », 30 octobre 1998. Levine, Ross. « Financial Development A
- Loayza. (2016). “[Informality in the Process of Development and Growth](#),” *The World Economy* 39 (12): 1856-1916 ([Policy Research Working Paper WPS7858](#), World Bank)
- Mondragon-Velez, Camilo, Pena Ximena et Wills Daniel. Labor Market Rigidities and Informality in Colombia. Documentos CEDE, Universidad de los Andes (2010). OECD/ILO (2019), Tackling Vulnerability in the Informal Economy, Development Centre Studies, OECD Publishing, Paris, <https://doi.org/10.1787/939b7bcd-en>.
- Perry, Guillermo E., Omar Arias, Pablo Fajnzylber, William F. Maloney, Andrew Mason, et Jaime Saavedra-Chanduvi. Informality: Exit and Exclusion. The World Bank, 2007. <https://doi.org/10.1596/978-0-8213-7092-6>. Quintin, Erwan. « Contract Enforcement and the Size of the Informal Economy ». *Economic Theory* 37, no 3 (2008): 395-416
- Ulyssea. (2018). Firms, Informality and Development: Theory and Evidence from Brazil. *American Economic Review* 2018, 108(8): 2015-2047

Websites:

- infobae (site internet). « Informalidad laboral en Colombia alcanzó 49,2 %: DANE », de Abril 2021. /america/colombia/2021/04/13/informalidad-laboral-en-colombia-alcanzo-492- dane/.
- IRD’s website : <https://www.ird.fr/lird-presentation-et-historique>

Discussions:

Numerous discussions with the Colombian team which contributed a lot to my work.