Wp

A TAXONOMY OF COLOMBIA'S INFORMAL LABOR MARKET

Fernández, C.
Villar, L.

CAF DEVELOPMENT BANK OF LATIN AMERICA

A TAXONOMY OF COLOMBIA'S INFORMAL LABOR MARKET

Fernández, C.
Villar, L.

ABSTRACT

A taxonomy of the informal labor market is extremely important to understand and handle informality, particularly in a country as Colombia where this phenomenon is large and heterogeneous. As we will argue in this paper, it is possible to identify four different types of informality, according to the reasons to be informal: low productivity of the worker (subsistence informality), barriers to formality (induced informality), choice (voluntary informality) and both choice and low productivity (mixed informality). The policy recommendations to handle informality varies according to the target type of informality. While induced informality might be reduced by the removal of formal employment barriers or by the implementation of active policies to reduce segregation in society, structural informality requires other kinds of policies, such as a focus on improving educational outcomes. Similarly, in the case of voluntary informality, providing economic incentives to formalize and controlling informality might be more effective, whereas mixed informality is more related to wrong incentives created by social benefits. In this paper we propose a methodology to identify the four types of informality to the case of Colombia that follows what we did in Fernandez et al. (2016) but with greater emphasis on the education level. Although the correspondence is far from perfect, we show that in general terms, informal workers with primary education or less can be classified in the Subsistence informality group, informal workers with secondary education can be included in the Induced informality group, informal workers with tertiary education or more can be treated as Voluntary informal workers and informal workers with middle school education can cover mixed informality. Hence, the policy recommendations to handle informality among each education group are different.

TAXONOMÍA DEL MERCADO LABORAL INFORMAL EN COLOMBIA

Fernández, C.
Villar, L.

RESUMEN

Una taxonomía del mercado de trabajo informal es sumamente importante para entender y manejar la informalidad, particularmente en un país como Colombia, donde este fenómeno es grande y heterogéneo. Como se argumentará en este trabajo, es posible identificar tres tipos de informalidad, según las razones por las cuales se es informal: baja productividad del trabajador (informalidad de subsistencia), barreras a la formalidad (informalidad inducida) y elección (informalidad voluntaria). Las recomendaciones de políticas para manejar la informalidad varían de acuerdo con el tipo de informalidad a combatir. Mientras que la informalidad inducida podría reducirse mediante la eliminación de las barreras al empleo formal, o mediante la implementación de políticas activas para reducir la discriminación en la sociedad, la informalidad estructural requiere otro tipo de políticas, como mejoras en la educación. Similarmente, en el caso de la informalidad voluntaria, podría ser más eficaz proporcionar incentivos económicos para formalizar y controlar la informalidad. En este trabajo proponemos una metodología para identificar los tres tipos de informalidad en caso colombiano, que sigue lo que hicimos en Fernández et al. (2016), pero con mayor énfasis en el nivel educativo. Aunque la correspondencia está lejos de ser perfecta, mostramos que en términos generales, los trabajadores informales con educación primaria o menor pueden clasificarse en el grupo de informalidad de subsistencia, los trabajadores informales con educación secundaria pueden ser incluidos en el grupo de informalidad inducida, los trabajadores informales con educación terciaria o mayor pueden ser tratados como trabajadores informales voluntarios. Por lo tanto, las recomendaciones de política para manejar la informalidad entre cada grupo educativo son diferentes.

A Taxonomy of Colombia's Informal Labor Market

Cristina Fernandez
Leonardo Villar[1]

A taxonomy of the informal labor market is extremely important to understand and handle informality, particularly in a country as Colombia where this phenomenon is large and heterogeneous. As we will argue in this paper, it is possible to identify three different types of informality, according to the reasons to be informal: low productivity of the worker (subsistence informality), barriers to formality (induced informality) and choice (voluntary informality). The policy recommendations to handle informality varies according to the target type of informality. While induced informality might be reduced by the removal of formal employment barriers or by the implementation of active policies to reduce segregation in society, structural informality requires other kinds of policies, such as a focus on improving educational outcomes. Similarly, in the case of voluntary informality, providing economic incentives to formalize and controlling informality might be more effective.

In this paper we propose a methodology to identify the three types of informality to the case of Colombia that follows what we did in Fernandez et al. (2016) but with greater emphasis on the education level. Although the correspondence is far from perfect, we show that in general terms, informal workers with primary education or less can be classified in the Subsistence informality group, informal workers with secondary education can be included in the Induced informality group and informal workers with tertiary education or more can be treated as Voluntary informal workers. Hence, the policy recommendations to handle informality among each education group are different.

I.   INTRODUCTION:

According to Fernandez, Lilenstein, Oothuzien and Villar (2016) it is possible to implement a taxonomy of informality by analyzing the reasons to be informal: (1) Low productivity: Informal workers do not possess the necessary skills in order to produce at the level required by the formal labour market; (2) Barriers: Informal workers have a level of productivity comparable to workers in the formal labour market but they are prevented from entering due to formality barriers, that can be explicit, as in the case of payroll taxes or implicit, as in the case of gender or race discrimination; and (3) Choice: Informal workers have a level of productivity comparable to workers in the formal labour market but they choose to be informal after a cost-benefit analysis. We call these three types of informal workers: Subsistence, Induced and Voluntary, respectively.

A perfect taxonomy of informality is not feasible even on theoretical terms. In fact, it is possible to have a single informal worker facing the three types of informality at the same time: a worker that is both segregated by low productivity and high formal market barriers, and who at the same time values to be informal. However, it is possible to find indicators to estimate how much the characteristics of informality in the country as a whole resemble the key characteristics of each type of informality. We identified before that those three key characteristics are the following: the level of choice, to identify whether workers are being moved towards informality by self-decisions; the composition of informality by productivity level, to identify whether workers are moved to informality by lack of productivity; and the level of barriers to formality, to identify whether workers are being moved to informality by excessive protection to the formal workers or by discrimination.

We analyzed these three characteristics (choice, productivity and barriers) for the three aggregates but also by level of education. Although the correspondence is far from perfect, we show that in general terms, informal workers with primary education or less can be classified in the Subsistence informality group, informal workers with secondary education can be included in the Induced informality group and informal workers with tertiary education or more can be treated as Voluntary informal workers. Hence, the policy recommendations to handle informality among each education group are different.

The Colombian data used in the majority of the ensuing analysis is from the third quarter 2007 and 2015 data of the *Gran Encuesta Integrada de Hogares (Widescale Integrated Household Survey)* (GEIH, 2007-2016), provided by the Department of Statistics (Dane). When feasible we present the results using three different aggregates of this survey: total, rural and 13 main metropolitan areas, which is the aggregate most commonly used by the Colombian authorities[2]. When needed, we also use the *Encuesta Continua de Hogares (Continuous Household Survey)* (ECH, 2002-2006) provided by Dane and the *Encuesta Longitudinal Colombiana de la Universidad de los Andes (Longitudinal Survey of Colombia by the Andes University)* (ELCA, 2010-2013)[3]. Throughout this section we mostly applied the firm definition of informality, that includes workers employed in firms with no more than five employees; unpaid family helpers or housekeepers; self-employed with the exception of independent professionals and technicians; and business owners or firms with no more than five workers.[4] However, to check robustness we also included the results using the legal definition of informality, which includes workers that have access and contribute to the health and pension systems.

The paper is structured as follows: Section II reviews previous literature; Section III illustrate the indicators for choice/Voluntary informality; Section IV illustrates the indicators for low-

---

[2] The GEIH total aggregate covers 23 cities with rural areas, gathering information on more than 62 thousand households per quarter. The 13 metropolitan areas aggregate represents 60% of the total urban population according to the 2005 census, gathering information on more than 30 thousand households per quarter. The GEIH total sample includes not only the rural and the 13 metropolitan areas survey but also other urban areas.

[3] The ELCA, although not statistically representative, gathered information from around 5,000 urban households per year and was applied in a panel structure for 2010 and 2013. In this chapter when we refer to the ELCA, we classify informal workers as those who do not contribute either to state health or pension systems (legal definition).

[4] This criterion changed from 10 workers or less (ILO10) to 5 workers or less (ILO5) showing a higher correlation with other measures of informality (Bernal, 2009). Since 1999 the Delhi Group established the ILO5 as the standard measurement of informality (Central Statistical Organization, 1999).

productivity/Subsistence informality; Section VI illustrates the indicators for barriers to formality/Induced informality; Section VII gathers the indicators of the previous three sections to provide a taxonomy of informality, attempts to identify the specific shares of each type, and illustrate the close empirical relationship that can be found between our taxonomy and the levels of education in the case of Colombia; and Section V concludes.

II. LITERATURE REVIEW

This taxonomy closely resembles the old discussion between segmented and integrated markets. 'Subsistence informality' resembles the idea of segmented markets pursued by Lewis (1954) in his famous dual sector theoretical model where a "capitalist" sector develops by taking labour from a non-capitalist backward "subsistence" sector. In the "subsistence sector" there is unlimited supply of labour from which the capitalist sector takes advantage to expand without the need of raising wages, and where subsistence workers queue for a job in the "capitalist" sector. It is also related to the Harris & Todaro (1970) tradition with their rural-urban migration model, where the driver of migration from the rural sector to the urban sector is the expected urban real income. This implies that rural-urban migration in a context of high urban unemployment can be economically rational if expected urban income exceeds expected rural income. The model assumes that unemployment is non-existent in the rural agricultural model; therefore, workers could queue for an urban (formal) job in the rural (informal) sector or in urban unemployment.

'Induced Informality' resembles the idea of De Soto (2000) according to whom low productive capital does not transit to high productive capital because of prevailing law, and voracious governments acting as an unnecessary tether for reaching higher productivity. In other words, workers are segregated from the formal market due to extremely high barriers to formality. We also include in this group workers segregated from the formal market by discrimination given that they are well prepared to participate in the formal market and they want to participate on it, but they can't, due to the existence of social implicit barriers.

'Voluntary Informality' resembles the idea of integrated markets in which workers continuously flow from/to formality to/from informality according to a cost benefit analysis (Maloney, 2004 and Levy 2008), where the informal sector is viewed as an unregulated micro-entrepreneurial sector, instead of a disadvantage residual of segmented labour markets.

The heterogeneity of the informal markets has been pursued by some authors, but it is not really popular in the economic literature since it implies the simultaneous use of three structurally different models of labour economics. However, as we will show later, it is a more realistic approach. Some of the authors that follow the informal heterogeneity approach are the following:

The 2008 World Bank Flagship (Perry, 2007) analyses the most important aspects of informality in Latin American countries, understanding informality as a dynamic combination of three types of informality, very closely related to the ones described in this chapter.

Ulyssea (2013), for the case of Brazil, develops a general equilibrium model in which there are three types of informality according to their productivity, approximated by wages. The author finds that each type of informality should be approached in a different way and alerts on the possible negative results on welfare of applying the same recipe to all types of informality. As an example the author claims that enforcing formality in a labour market where voluntary informality exists might have a positive impact on welfare because of the new tax revenue, while enforcing formality in a market where structural informality exists might have a negative impact on welfare.

Alcaraz, Chiquiar and Salcedo (2012) using a model to identify voluntary and involuntary informality (which they call "model of segmentation and self-selection") for the case of Mexico, find that between 10 and 20% of informal workers depending on the co-variables introduced into the model, demonstrated marked signs of segmentation based on their personal characteristics -as education or age- and on their households characteristics –as the households composition-.. While this result provides evidence of the presence of segmentation in the Mexican labour market, it suggests that it is quite low and that an important proportion of workers in the informal sector self-select into it.

Similarly, for the case of Colombia, Galvis (2012) uses a wage gap approach to characterize informality. One important finding is that for the high earning workers, especially those in the highest 90th percentile of income, the wage gap between formal and informal workers is the smallest over all the distribution. This might be evidence of voluntary informality among this group. Garcia (2014) also used wage gaps as a criterion for segmentation inside the Colombian labour market in between regions and cities, and found that in the less developed cities[5] 75% of informality is involuntary, in big cities[6] it accounts for 47% of total informality. Meanwhile this percentage is 76% for the Caribbean Coast cities[7]. However, Perry (2007) argues that wages do not necessarily demonstrate segmentation but can in fact signal differences in unobservable characteristics among workers, as well as preference for informality.

Other authors have focussed on the prevailing type of informality. In the case of Colombia, there is some evidence supporting the idea that integrated markets existed during the period 1991-1996 and became more segmented after the nineties, a decade characterised by increases in payroll taxes and low productivity (Perry, 2007; Maloney, 2004). Peña (2013) goes in the same direction. The author documents how the high levels of payroll taxes in Colombia in addition to a high minimum wage have deepened the labour market segmentation into formal and informal sectors. In particular, the paper finds that during the 1998-2000 recession, informal sector wages lost 20 percentage points as compared to formal sector wages, meanwhile there was a decline in formal employment and an increase in the informal employment, this giving the notion of a segmented labour market. Mondragon, Peña & Wills (2010) find that the increases in payroll taxes and in the minimum wage have driven the formal sector to be less able to adjust to economic cycles. In other words, the aforementioned rigidities make the formal sector adjust to the economic cycle through quantities instead of salaries, increasing the size of the informal sector and lowering their salaries, a marked sign of a segmented labour market.

---

[5] Includes Cúcuta, Montería, Pasto and Villavicencio.
[6] Includes Bogotá, Medellín, Cali, Bucaramanga, Manizales, Pereira and Ibagué.
[7] Includes Barranquilla and Cartagena.

III. CHOICE:

We identified three main ways of establishing whether entry into informality is voluntary or involuntary: (i) via surveys about workers preferences, (ii) through a counter/pro cyclicality analysis, and (iii) by revising the frequency of transitions between informality and formality. As it is shown in this section the evidence of voluntary informality in Colombia is higher among workers with tertiary education or more, as informal workers with those characteristics show higher preferences for informality. However, voluntary informality in Colombia is relatively weak when compared to other countries in Latin America, particularly among workers with low levels of education.

PREFERENCES FOR INFORMALITY REVEALED BY SURVEYS

Surveys are the easiest way to establish whether the choice to work in the informal labour market is voluntary. Asking informal workers why they chose to work in the informal labour market can establish whether they perceive informal work as beneficial compared to formal work, or whether they are simply unable to access the formal labour market.

Unfortunately, surveys and data about preferences are scarce. However, the 2007 Colombian Household Survey (GEIH) includes two useful questions to identify whether workers are informal out of choice or necessity. One question asks informal workers whether they would like to accept a job in the formal market with the same or higher earnings (including wage plus benefits)[8]. The other question asks respondents the reason they are informally employed. If the worker answers negatively to the first question or does not state the impossibility of finding work as the reason to be informal, then the worker is considered a voluntary informal worker. According to the expanded data, 36% of informal workers in Colombia reveal preferences for informal jobs. This proportion increases to 41% using the 13-areas aggregate and decreases to 32% in rural areas, indicating that voluntary informality is more of an urban issue.

These percentages are not low per se, but they look rather low in a regional perspective. According to Perry (2007)[9], the preferences of informal work among self-employment are 41% in Colombia, which compares with 60% in Argentina, 74% in Bolivia, 75% in the Dominican Republic in Brazil (68% of males and 56% of women)[10]. Preferences for informality in the region show less variance across countries: 43% in Argentina, 52% in Bolivia, and 57% in the Dominican Republic, compared with 40% in Colombia. The responses to the question about the reasons to be informal confirm the previous results. The percentage of informal independent workers that report the impossibility of finding a job in the formal market is 25% in Bolivia, 44% in the Dominican Republic, 55% in Colombia[11] and 59%[12] in Argentina. In Mexico[13]only 12% of

---

[8] Another available question is if they would accept a formal job but with lower earnings. We decided to include as voluntary informal workers those that wouldn't accept that job even if the pay was higher, as we considered that it related better to our purposes. It is important to note that this question was only asked to employees and self-employment. In this portion we are assuming that salaried workers show similar preferences than independent workers.

[9] Based on Arias and Bustelo (2007), Arias, Landa and Yañez (2007). Except in Colombia the question was: if you were able to choose would you rather be a salaried or an independent worker? The data for Colombia refers to the 2006 fall Survey. Also reported by Bernal (2009).

[10] This percentage correspond to the independent workers that answered that they will not leave their independent job for a job with a signed job contract. From the pesquisa Nacional por Amostra de Domicilios 1990.

[11] 34% of answers, since two choices were available

the informal male workers and 6% of the informal female workers reported the impossibility of finding a formal job as the main reason to be informal. However, these results are less comparable among countries since the asked questions varied[14]. When compared with a more international group, preferences for informality are also in the middle low range, with Norway, Denmark, and Russia showing lower preferences for informality, Spain and Japan similar preferences and most of the developed countries showing higher preferences for informality (Perry, 2007)[15]. These results show that there is an important variance in preferences for informality, with Colombia showing the highest preference for formal employment and Mexico the lowest, among Latin American countries where data is available.

Table 1 shows the determinants of preferences for informality, that is to say, what are the determinants of voluntary labor informality among informal workers. The estimations were made with a logit regression using for the 2007 GEIH survey (the table displays the odds ratio, using Middle School education as the baseline variable for comparisons)[16]. The results suggest that workers with tertiary education, relatively old workers (45 years old or more), and workers living in big[17] and border cities[18] exhibit higher preferences for informal jobs. The later can be explained by the presence of smuggling activities in border cities. Women reported as spouses, according to their role in the household (37% of the informal working women in 2007[19]) also show strong preferences for informality. Women that are head of the household (31% of the informal working woman in 2007) or that occupied any other position at the household, as daughters or other relatives, do not have significant preferences for informality. There are also preferences for informality in the rural areas. All these preferences are robust to the aggregated and to the informality definition used. Bernal (2009) performed a similar exercise using the 2006 survey, and only including the willingness of independent workers to accept a formal job at a lower wage as a dependent variable. Bernal´s the results show higher preferences for formality among low educated workers, males and the urban population; and lower preferences for formality among women, spouses and heads of the household. The results are not strictly comparable but highly compatible with ours, excepting that we also found higher preferences for informality among older workers.

Unfortunately, the 2015 GEIH Survey does not include questions on preferences for informality (that were only featured in the 2007 GEIH). However, we can still use the 2007 GEIH to make predictions in the 2015 GEIH, using a similar logit regression to the one displayed in Table 1 and described before[20], which enable us to determine the profile of the informal workers that have preferences for informal jobs and, therefore, allows us to estimate the probability of each 2015 informal worker to prefer informality. For the estimation, after computing the estimated

---

[12] 43% of answers, since multiple choices were available

[13] Based on Encuesta Anual de Micronegocios, 1994

[14] The survey asked for the main reason in Bolivia, up to two reasons in Colombia, and the two most important reasons in the Dominican Republic, whereas in Argentina the question permitted multiple responses.

[15] Based on Blanchflower (2004), Blanchflower, Oswald and Stutzer (2001) and own calculations

[16] The odds ratio are transformation from the original coefficients that are expressed as a ratio of the missing category. Therefore, the missing category has an odd ration of 1. The t-statistics presented in the table, correspond to the original logit regression and this explains why they do not have the same sign as the coefficient.

[17] Bogotá, Medellín and Cali.

[18] Pasto and Cúcuta.

[19] Only 3% of the male informal workers describe themselves as spouses, and they do not show significant preferences for informal jobs.

[20] The logit regression that we used for this purpose uses as determinants for preferences of informality not only a more detailed list of observable characteristics of the worker, but also some endogenous variables as work satisfaction, and the economic sector. See Annex A.

probability of being informal given observable characteristics of the worker, we assumed that those informal workers in 2015 that had a predicted probability of preferring informal jobs higher than 50% were voluntary informal workers. According to this procedure, we estimated that 41% of the informal workers in Colombia were voluntary informal workers in 2015. Applying the same procedure yield that 50% of the informal workers in the 13-areas and 27% of the informal workers in the rural areas were voluntary informal workers[21].

Table 1. Logit: Preferences for informality, odds ratio.

Depvar: Preferences for informality.	Firm definition			Legal definition		
		13-areas	Rural	National	13-areas	Rural
Elementary or less	0.886*	0.94	0.809	0.901	0.934	0.817
	(-2.05)	(-0.87)	(-1.42)	(-1.72)	(-0.93)	(-1.34)
High school	1.124	1.169*	0.785	1.089	1.093	0.782
	(1.84)	(2.10)	(-1.18)	(1.30)	(1.14)	(-1.17)
Tertiary or more	1.772***	1.750***	3.125**	1.713***	1.729***	2.937*
	(7.10)	(6.22)	(2.59)	(6.24)	(5.62)	(2.30)
Women (spouse)	1.322***	1.366***	1.333*	1.344***	1.364***	1.358*
	(4.70)	(4.45)	(2.07)	(4.87)	(4.29)	(2.21)
Women (other)	0.931	0.937	0.812	0.925	0.914	0.806
	(-1.39)	(-1.05)	(-1.53)	(-1.45)	(-1.37)	(-1.57)
Less than 24 years	0.632***	0.811*	0.488***	0.645***	0.826	0.496***
	(-5.86)	(-2.09)	(-4.14)	(-5.51)	(-1.87)	(-4.03)
45-55 years	1.325***	1.333***	1.450**	1.295***	1.320***	1.463**
	(5.44)	(4.52)	(3.19)	(4.86)	(4.12)	(3.26)
56+ years	2.125***	2.057***	2.344***	1.987***	1.895***	2.217***
	(12.90)	(9.83)	(7.38)	(11.23)	(8.09)	(6.88)
Big city	1.821***	1.834***		1.856***	1.879***	
	(11.91)	(12.99)		(11.71)	(12.74)	
Border city	1.823***	1.815***		1.831***	1.840***	
	(10.43)	(10.94)		(10.16)	(10.77)	
Rural	1.156*			1.177**		
	(2.53)			(2.79)		
Constant	0.359***	0.339***	0.452***	0.349***	0.334***	0.444***
	(-17.92)	(-16.81)	(-5.39)	(-17.86)	(-16.38)	(-5.47)
Number of obs	33104	15219	3860	31226	14028	3826
F	50	41	14	43	36	13
df_m	11	10	8	11	10	8
df_r	33103	15218	3859	31225	14027	3825

(t-statistics in parenthesis). Source: GEIH (3rd quarter 2007). * $p<0.05$, ** $p<0.01$, *** $p<0.001$. The odds ratio are transformation from the original coefficients that are expressed as a ratio of the missing category. Therefore, the missing category has an odd ration of 1. The t-statistics presented in the table, correspond to the original logit regression and this explains why they do not have the same sign as the coefficient. The base categories are middle school, male, 25-44 years, medium and small cities that are not in the border.

COUNTERCYCLICALITY

Counter cyclicality is another indicator for involuntary informality. According to Tornarolli et al (2014), in the presence of labour market rigidities and involuntary informal sector employment, when the economy enters into recession and a minimum wage exists, some of the formal enterprises retrench workers who subsequently find refuge in the informal sector. Therefore, the ratio of informal to formal workers tends to increase during downturns. Similarly, when the

---

[21] The respective percentages are 32% for the total expanded survey, 46% for the 13 metropolitan areas and 24% for the rural areas, using the health and pension contributions definition

economy grows, the cost of hiring formally becomes relatively lower and the ratio of informal to formal employment decreases (Loayza & Rigolini, 2006). However, in the presence of voluntary informal workers, during upturns, an increase in the informal wage should attract informal workers and increase the size of the sector, parallel or even pro-cyclically to the increase in formal employment[22]. Therefore, indicators of pro/counter cyclicality enable us to discern between voluntary and involuntary informality.

In the case of Colombia, Figure 1 shows that there is a positive relationship between the formality rate and the business cycle, measured as the relative difference between observed and potential GDP[23]. The correlation coefficient of the formality rate and the output gap is 0.46, between 2002 and 2015.[24] Therefore, we argue that the formality rate in Colombia is broadly pro-cyclical and therefore, the informality rate is broadly countercyclical. These results support the counter cyclicality hypothesis of informal employment in Colombia (Loayza & Rigolini, 2006; Tornarolli et al., 2014; and Fiess et al., 2008) which is evidence of a significant portion of involuntary informal workers. Similarly, most of the business cycles literature in Latin America supports the idea of a prevailing counter-cyclicality in the region (Loayza and Rigolini, 2006 and Tornarolli et al., 2014). However, there is strong evidence of pro-cyclicality of informality in Mexico (Fiess et al 2008 and Bosh and Maloney 2006) and results for Brazil are mixed. These results are congruent with the preferences for informality revealed by surveys.

The pro-cyclicality of informality by level of education, showed in Figure 1, shows that informality among workers with low levels of education informality tends to be counter cyclical. The correlation coefficients of the formality rate and the output gap are 0.59 and 0.79, for workers with primary education[25] or less and workers with high school education, respectively[26]. Meanwhile, there is no evidence of counter-cyclicality among workers with tertiary education or more. These results confirm the findings of de la Torre and Ize (2016), according to whom informality among workers with low levels of education tends to be countercyclical whereas informality among workers with very high levels of education tends to be pro-cyclical in Latin America.

---

[22]According to Fiess, Fugazza and Maloney (2008), during booms generated by commodity exports, which promote the informal-intensive service sector, one might expect informal employment to be even more procyclical than the formal employment.

[23] Defining the formality rate as one minus the informality rate. We only performed this exercise using the aggregated 13 metropolitan area data since the wider sample suffered significant changes in the number of interviewed households. Note that in the case of the 13 areas aggregate, the formality rate is calculated using two different ILO methodologies/series since one includes firms with less than 10 worker (ILO10, Mondragon et al, 2010), and the other includes firms with less than 5 workers (ILO5). It should be noted that in the last two years, the behavior of informality in Colombia has been less countercyclical. In fact, the correlation coefficient between formality and the output gap increases for all the aggregates after dropping the last two years of the series. This can probably be explained by the 2014 reduction in payroll tax, implemented on 2013 and 2014, as shown in Fernandez and Villar (2016).

[24] Significant at a 95% level of confidence. It should be noted that in the last two years, the behaviour of informality in Colombia has been less countercyclical. In fact the correlation coefficient between formality and the output gap was 0.74, with a 99 significance level, between 2002 and 2013. This can probably be explained by the 2014 reduction in payroll tax, implemented on 2013 and 2014.

[25] Includes middle school.

[26] Significant at 95% and 99% level, respectively.

Figure 1: Formality rate and output gap (13 Metropolitan Areas)

	Firm definition 2002-2015)	Legal definition (2002-2015)	Mondragón et al (1984-2006)	Primary school or less	High school	Tertiary education or more
Output gap (2002-20015)	0.46**	0.26	0.74***	0.59**	0.75***	0.13
Output gap (2002-20013)	0.74***	0.54*	0.74***	0.65**	0.88***	0.18

***p<0.01, ** p<0.05, * p<0.1

Source: Own calculations based on GEIH (2008-2015) and ECH (2002-2008), Mondragon et al. 2010 and Fedesarrollo (Output pag). Primary education includes Middle School. In the case of the 13 areas aggregate, the formality rate is calculated using two different ILO methodologies/series since one includes firms with less than 10 workers (ILO10, Mondragon et al, 2010), and the other includes firms with less than 5 workers (ILO5). It should be noted that in the last two years, the behavior of informality in Colombia has been less countercyclical.

TRANSITIONS FROM INFORMALITY TO FORMALITY

Another method which allows us to assess whether informality is voluntary or a default option is by observing how frequently individuals transition from informality to formality. Frequent transitions between informality and formality may indicate a degree of voluntariness. On the other hand, if there is little or no movement from informality to formality, this transition is obstructed by either a lack of productivity or formal market barriers[27]. In estimating those transitions, we use panel data from the ELCA urban survey for Colombia (for the years 2010 to 2013) [28]. Table 2 shows that transitions from informality to formality in Colombia are not frequent, with only 14% of informal workers transiting to formality between 2010 and 2013, compared with 20% of formal workers transiting to informality and 19% of workers entering formality from unemployment during the same period. By level of education it is possible to observe that the transitions from informality to formality are much less frequent among workers with low levels of education than among workers with tertiary education or more (26%).

Pages and Stampini (2006), computed yearly transition matrices for males 24-60[29] for six developing countries including Mexico, Argentina and Venezuela. According to these matrices, 26% of the unskilled workers (less than a high school degree) and 42% of skilled workers transited from informality to formality, including both informal salaried and self-employed as informal workers. Among the Latin American countries, Mexico shows the highest transitions between informality and formality (30% for unskilled and 54% for skilled) and Argentina, the lowest (18% for unskilled and 36% for skilled). A similar estimation for the case of Colombia, indicates that 17.5% of unskilled and 21.3% of skilled informal men between 24 and 60 transited to formality between 2010 and 2013, suggesting that transitions in Colombia are rather low, taking also into account that our transition matrix is estimated along a 3 years-period[30].

---

[27] Pages and Stampini (2006) used the distance to the transition matrix of a steady state transition matrix as a measure of segmentation.

[28] The rural survey does not include the same question that we used in this estimation.

[29] As a way to control by other unobservable and observable characteristics

[30] Defining skilled workers as those that have approved grade 11 or more. The results also suggest that transitions are higher for skilled than for unskilled works, however Carmen and Stampini (2006) claim that there are not significant differences between education levels in the distance of these matrices to their steady stead. Note that the definition of skilled does not correspond to our aggregate for tertiary education.

Table 2. Transition Matrices, between 2010 and 2013. Legal definition of informality

Total		2013				Primary or less		2013			
		Unemployed	Formal	Informal	Inactive			Unemployed	Formal	Informal	Inactive
2010	Unemployed	18	35	19	27	**2010**	Unemployed	20	40	9	30
	Informal	4	72	**14**	10		Informal	5	73	**10**	12
	Informal	3	20	72	6		Informal	2	31	60	7
	Inactive	Total	26	3	67		Inactive	4	21	1	74
High school		2013				Tertiary		2013			
		Unemployed	Formal	Informal	Inactive			Unemployed	Formal	Informal	Inactive
2010	Unemployed	11	40	29	20	**2010**	Unemployed	11	30	14	45
	Informal	4	75	**12**	10		Informal	2	66	**26**	6
	Informal	3	25	63	9		Informal	1	12	83	4
	Inactive	4	28	3	65		Inactive	2	30	11	57

Source: Own calculations based on ELCA 2010 and 2013 waves. Note: legal Definition of informality. Secondary education includes workers with high school and middle school studies.

IV. DIFFERENCES IN PRODUCTIVITY

The previous section focused on identifying workers who were voluntarily informal. In this section, we shift our attention to involuntary informality and the distinction between subsistence and induced informality. To make this distinction, we need to determine whether there are substantial differences in productivity among involuntary workers. The two indicators used to this purpose are: the incidence of informality in lower productivity groups and the percentage of workers earning a wage significantly lower than the minimum hiring cost of the formal sector. If wages are productivity-linked, informal workers earning well below such minimum hiring cost of the formal sector are likely to be informal because of their low productivity and not because of segregation or of barriers to formality. In sum, as we will show in this section, there is evidence for a strong presence of subsistence informality in Colombia: high rates of informality amongst low productivity groups of the workforce, and a significant group earning a wage substantially lower (less than half) than the hiring cost of the formal sector.

INCIDENCE OF INFORMALITY IN LOW-PRODUCTIVITY-GROUPS

Survey data help to establish whether there is high incidence of informality amongst low-productivity groups. This involves looking at dimensions of productivity such as education, experience and overall productivity of the city and sector of work. To illustrate this, and in order to ascertain if indicators of low worker productivity are significant correlates of informality, in this section we provide summary statistics of informality according to worker characteristics (Table 3) as well as a regression on the probability of being informal. Tables 4a and 4b, show the coefficients and the T-statistics for each aggregate using the firm and legal definition of

informality, respectively. Next to each column we included the results of a regression for preferences for informality, similar to the one presented in Table 1. This allows us to conjecture if a high incidence of informality is caused by preferences, barriers to informality (as segregation) or by a mismatch with the required characteristics to work in the formal sector.

Education: As is evident in Table 3, using education as a measure of productivity results in the largest differences in informality rates. In Colombia as a whole, the informality rate for workers with primary education is 84%, compared with 56% among workers with a high school education and 27% among workers with a tertiary education.[31] In addition, education is a significant determinant of the probability of informality. Estimates of the probability of a worker being in the informal labour market are given in Table 4. The coefficients shown in the table are the odds ratio of each variable with respect to the base (or missing) category on each classification (in this case, Middle School education), with their respective t-statistic in parenthesis[32] . A coefficient of 2.6 for primary education or less means that workers in this group are 2.6 times more likely to be informal than workers with a middle school education. Similarly, workers with high school education have about one half (and workers with tertiary education about one-sixth) of a middle school worker probability to be informal. As shown in Table 4, the differences in the incidence of informality by educational levels cannot be explained by preferences for informality. In fact, workers with tertiary education or more, show higher preferences for informal jobs and lower probability for being informal. The opposite is true for primary education or less, that there is an important portion of the population with levels of education so low that they are unlikely to find a formal job, which is also consistent with previous results in this section.

Experience: Age is commonly used as a proxy for experience. While the majority of youth aged 15 to 24 are not economically active and likely to be enrolled in education, of those youth who are working, 59% are employed informally (Table 3). As Table 4 shows, these relatively high levels of informality among young workers cannot be explained by education, since the difference continues to hold after controlling for other observable characteristics, and neither by preferences, since young workers report low preferences for informal jobs[33]; but perhaps by experience. Therefore, we can argue that some young workers might not find a job in the formal market due to their low experience, and therefore low productivity.

Geography: The productivity of a worker, and therefore the informality rate, also depends on the location of the worker. This probably explains why a low qualified worker in a developing country shows a higher productivity and a lower probability for being informal in a developing country. In Colombia, the informality rate in the four most productive cities is 45%, in the four least productive cities is 60%[34] (see Table 3), and 84% in the rural areas, differences that are statistically significant after controlling for other observable characteristics (Table 4). In most cases the rates of informality cannot be explained by preferences, suggesting that the overall

---

[31] The rates are 85%, 56% and 26% respectively if we include rural and other urban areas.

[32] Note that the sign of the odds ratio may not coincide with the t-statistic, given that the former is estimated for the coefficient in the logistic regression, and the odds ratio is just a transformation of this coefficients for a more approachable interpretation

[33] These results contrast with the older population that shows high rates of informality, even after controlling for other observable characteristics, but also high preferences for informal jobs.

[34] The 4 most productive cities are Bogota, Medellín Tunja and Bucaramanga. The four least productive cities are Cúcuta, Barranquilla Sincelejo and Quibdó. Relative productivity is measured as average wage over minimum wage.

productivity of the city is playing a role in the productivity of the workers, and perhaps affecting their participation in the formal labor market. The only exception is the case of the less productive cities in the 13-metropolitan areas, that include Cucuta and Pasto. These two cities are also border cities, and therefore impacted by smuggling, showing higher than predicted preferences and levels of informality.

Economic sector
Another determinant of productivity and therefore to informality is the economic sector. We didn't include this variable in the logit model, since we envisage some endogeneity in this variable. However, Table 3, shows that productive sectors as mines and public services tend to be more formal than others as agriculture and retail, restaurants and hotels tend to show higher rates of informality even after correcting for other observable characteristics[35].

---

[35] According to the labour productivity index of Isaza and Rojas (2015).

Table 3. Colombian Informality rates by worker characteristics.

	Firm definition			Legal definition		
	Total* %	13 Cities** %	Rural %	Total* %	13 Cities** %	Rural %
Total	59.9	48.0	83.9	63.3	50.3	87.6
Gender						
Male	59.3	45.1	83.4	63.0	47.9	86.9
Female	60.6	51.3	85	63.8	53.2	89.2
Head of the household	58.2	54.0	80.1	60.9	56.4	66.8
Spouse	62.2	58.9	88.7	65.2	60.0	91.6
Education Level						
Less than primary	92.3	90.2	93.6	94.9	89.7	97.6
Primary	83.7	76.7	90.2	86.2	75.9	93.7
Middle School	75.9	69.8	85.2	80.2	72.6	89.2
Completed Secondary School	56.4	48.4	73.7	59.3	50.0	77.9
Tertiary	26.5	23.9	45.4	31.0	28.2	49.0
Certificate/Diploma	30.3	21.2	41.1	25.2	22.7	42.6
Age						
15-24 years	59.0	42.6	83.5	72.4	58.6	91.8
25-34 years	46.5	34.4	77.7	51.5	38.3	82.3
35-44 years	57.7	46.2	81.7	60.7	47.9	85.2
45-54 years	56.2	56.6	84.4	64.2	53.7	86.8
55+ years[36]	80.0	72.7	92.3	75.3	64.4	91.8
Location						
Rural Area	84.3	-	-	80.3	-	-
Productive Cities	44.5	44.9	-	65.7	45.9	-
Non Productive Cities	59.8	59.2	-	56.9	72.7	-
Economic Sector						
Productive	32.0	3.1	69.4	41.5	12.0	80.9
Non Productive	77.2	66.5	89.4	80.3	66.3	92.3

.Source GEIH. See the annex B for the ranking of cities and sectors

Table 4a. Logit: Probability of being informal and preferences for informality. Logit model (firm definition)

	National			13-areas		Rural	
	Odds ratio	Odds ratio. Including sectors	Preferences for informality 2007	Odds ratio	Preferences for informality 2007	Odds ratio	Preferences for informality 2007
Elementary or less	1.413***	1.336***	0.867*	1.289***	0.912	1.563***	0.809
	(7.12)	(5.74)	(-2.46)	(4.37)	(-1.31)	(3.54)	(-1.42)
High school	0.444***	0.442***	1.13	0.427***	1.174*	0.521***	0.785
	(-19.43)	(-18.62)	(1.95)	(-17.65)	(2.17)	(-5.13)	(-1.18)
Tertiary or more	0.126***	0.136***	1.831***	0.134***	1.806***	0.145***	3.125**
	(-47.75)	(-44.72)	(7.61)	(-40.10)	(6.63)	(-12.03)	(2.59)
Women (spouse)	1.900***	1.810***	1.338***	2.012***	1.382***	2.051***	1.333*
	(17.08)	(15.50)	(4.96)	(15.64)	(4.67)	(5.55)	(2.07)
Women (other)	1.496***	1.492***	0.948	1.506***	0.959	1.087	0.812
	(13.15)	(12.77)	(-1.05)	(11.54)	(-0.68)	(0.77)	(-1.53)
Less than 24 years	1.455***	1.353***	0.622***	1.292***	0.794*	1.551***	0.488***
	(9.91)	(7.77)	(-6.07)	(5.97)	(-2.30)	(3.87)	(-4.14)
45-55 years	1.269***	1.321***	1.335***	1.444***	1.344***	1.145	1.450**
	(6.83)	(7.77)	(5.65)	(8.91)	(4.70)	(1.23)	(3.19)
56+ years	2.387***	2.339***	2.128***	2.824***	2.078***	2.313***	2.344***
	(19.99)	(19.34)	(13.09)	(19.70)	(10.04)	(6.14)	(7.38)
Rural	1.969***	1.498***	1.016				
	(14.39)	(8.08)	(0.29)				
Productive city	0.613***	0.634***	1.334***	0.848***	1.236***		
	(-16.20)	(-14.91)	(5.75)	(-5.20)	(4.09)		
Less productive city	1.151***	1.175***	0.894	1.575***	1.632***		
	(4.51)	(5.05)	(-1.85)	(14.46)	(10.30)		
Productive sector		0.389***					
		(-8.67)					
Less productive sector		2.820***					
		(34.67)					
Constant	2.120***	1.507***	0.413***	1.503***	0.424***	3.940***	0.452***
	(17.52)	(9.08)	(-15.88)	(8.35)	(-12.78)	(11.90)	(-5.39)
N	91500	91500	33104	45011	15219	8530	3860
F	582.50	530.00	38.11	376.20	27.69	52.22	14.03
df_m	11.00	13.00	11.00	10.00	10.00	8.00	8.00
df_r	91,499	91,499	33,103	45,010	15,218	8,529	3,859

t-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1). The odds ratio were computed using Middle School Education as the baseline. The results show the odds ratio, which don't necessarily have the same sign as the t-statistic in the logistic regression, were the signs would match.
Source: Authors calculations based on GEIH 3rd Quarter 2015. Note: See Annex 2 for cities and sectors ranking according to their productivity. The base categories are middle school, male, 25-44 years, medium and small cities that are not in the border and sectors with medium productivity

Table 4b. Logit: Probability of being informal and preferences for informality. Logit model (legal definition)

	National			13-areas		Rural	
	Odds ratio	Odds ratio. Including sectors	*Preferences for informality 2007*	Odds ratio	*Preferences for informality 2007*	Odds ratio	*Preferences for informality 2007*
Elementary or less	1.165**	1.114*	*0.881*	1.291***	*0.906*	0.650***	*0.817*
	(3.15)	(2.20)	*(-2.11)*	(4.39)	*(-1.34)*	(-3.86)	*(-1.34)*
High school	0.415***	0.406***	*1.096*	0.367***	*1.101*	0.703*	*0.782*
	(-21.10)	(-20.49)	*(1.41)*	(-20.44)	*(1.24)*	(-2.55)	*(-1.17)*
Tertiary or more	0.130***	0.139***	*1.771***	0.137***	*1.787***	0.137***	*2.937*†
	(-47.58)	(-44.45)	*(6.70)*	(-39.60)	*(6.01)*	(-10.20)	*(2.30)*
Women (spouse)	1.891***	1.845***	*1.363***	2.002***	*1.382***	2.939***	*1.358*†
	(18.21)	(16.77)	*(5.17)*	(15.77)	*(4.54)*	(7.78)	*(2.21)*
Women (other)	1.239***	1.263***	*0.943*	1.382***	*0.935*	0.609***	*0.806*
	(7.04)	(7.55)	*(-1.12)*	(9.23)	*(-1.04)*	(-6.18)	*(-1.57)*
Less than 24 years	1.342***	1.332***	*0.638***	2.369***	*0.813*†	0.300***	*0.496***
	(7.41)	(7.36)	*(-5.68)*	(19.57)	*(-2.02)*	(-13.47)	*(-4.03)*
45-55 years	1.033	1.058	*1.306***	1.071	*1.334***	1.329*	*1.463***
	(0.99)	(1.63)	*(5.08)*	(1.69)	*(4.33)*	(2.50)	*(3.26)*
56+ years	1.467***	1.379***	*1.990***	1.478***	*1.925***	2.446***	*2.217***
	(9.74)	(7.79)	*(11.41)*	(7.64)	*(8.35)*	(7.01)	*(6.88)*
Rural	1.104*	0.818***	*1.03*				
	(2.40)	(-4.50)	*(0.52)*				
Productive city	0.507***	0.522***	*1.329***	0.729***	*1.217***		
	(-23.34)	(-21.43)	*(5.40)*	(-10.10)	*(3.58)*		
Less productive city	1.203***	1.244***	*0.880*†	1.690***	*1.628***		
	(5.99)	(6.77)	*(-1.98)*	(16.69)	*(9.83)*		
Productive sector		0.653***					
		(-4.29)					
Less productive sector		3.263***					
		(39.11)					
Constant	3.334***	2.266***	*0.402***	2.045***	*0.425***	7.863***	*0.444***
	(27.94)	(18.23)	*(-15.83)*	(14.44)	*(-12.26)*	(16.69)	*(-5.47)*
N	92306	92306	*31226*	45011	*14028*	9336	*3826*
F	525.80	539.10	*31.07*	362.40	*23.11*	69.29	*12.94*
df_m	11.00	13.00	*11.00*	10.00	*10.00*	8.00	*8.00*
df_r	92,305	92,305	*31,225*	45,010	*14,027*	9,335	*3,825*

T-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Source: Authors calculations based on GEIH 3rd Quarter 2015.). * p<0.05, ** p<0.01, *** p<0.001.The odds ratio were computed using Middle School Education as the baseline. The results show the odds ratio, which don't have the same sign as the t-statistic in the logistic regression, were the signs would match. The base categories are middle school, male, 25-44 years, medium and small cities that are not in the border, and sectors with medium productivity. See Annex 2 for cities and sectors ranking according to their productivity.

It is very likely that a worker that has a productivity level significantly below the minimum wage would not be able to find a formal job, even if barriers and excessive costs were removed. Therefore, we can argue that this worker is being segregated from the formal labour market because of low productivity rather than because of barriers to the formal market.

In Colombia, the cost of hiring a worker is approximately 1.5 times the minimum wage[37]. Using the average wage as a proxy for labor productivity of an informal worker, we tried to approximate the number of workers that are left out of the formal job because of their low productivity. According to Table 5, 48% of informal workers in Colombia earn less than half of the total hourly cost of hiring[38], i.e. less than 75% of the minimum wage. The percentage is 34% if we restrict the survey to the main 13 metropolitan areas and 64% in the rural area. Results are robust to the definition of informality used in the estimations. We will refer to these groups of workers in the following sections, as subsistence informality workers.

According to Table 5, the percentage of subsistence informal workers among informal workers is 59% for workers with primary education or less, 56% for workers with middle school or less and 38% for workers with high school and 28% for workers with tertiary education of more. Similar but lower results are observed for the 13 areas aggregate. Therefore, subsistence informality is prevalent (higher that 50%) among workers with middle school education or less. In the 13 metropolitan areas aggregate, percentages are a little lower, but we argue that the same conclusion can be achieved assuming reasonable percentages for induced and voluntary informality.

Table 5: Percentage of informal workers that earn less than half the hiring cost (Subsistence informal workers)

	Firm definition			Legal definition		
	Total	13-areas	Rural	Total	13-areas	Rural
Total	48%	35%	65%	48%	35%	64%
Primary or less	59%	44%	67%	58%	44%	66%
Middle school	50%	40%	64%	51%	42%	64%
Less than middle school	56%	42%	66%	55%	43%	65%
High school	38%	29%	57%	40%	30%	58%
Tertiary or more	28%	23%	49%	27%	22%	45%

Source: GEIH. ** Labor income imputed when not available

---

[37] Including vacations, transport subsidy, severage and their interest, yearly bonus, pension contributions, risk insurance and the *Caja de Compensación Familiar*.
[38] We imputed the salary, in those cases when it was no reported, about 10% of the informal workers. The percentage is 44% if we restrict the survey to the main 13 metropolitan areas.

V. BARRIERS TO FORMALITY

Barriers to formality can be separated into implicit or explicit obstacles to entry into formal employment. Implicit barriers may come from discrimination or from custom. Therefore, looking at the incidence of informality by ethnicity and gender, controlled by observable characteristics as education, can provide an indication of whether implicit barriers to formality exist in a country. In turn, explicit barriers to formality may arise from legislation and regulation. International comparisons of the barriers to formality, for example in the form of payroll taxes or excessive minimum wages, help to establish the extent of explicit formal market barriers within a country. Both types of barriers are preventing productive workers from being employed in the formal sector, where they could be even more productive in the long run. Therefore, the removal of these barriers to the formal labour market would stimulate important productivity and distributional gains for these economies.

In this chapter, we found that Colombia shows some degree of induced informality either in the form of discrimination or in the form of legal barriers to formality. In fact, Colombia shows some signs of discrimination against women and race, even after controlling for education, preferences and other observable characteristics. With respect to the legal barriers to informality, Colombia has a relatively high level of payroll taxes, despite the recent reduction implemented by the government, and some rigidities in the minimum wage setting mechanism. However, other types of worker protection are rather flexible in the country.

IMPLICIT BARRIERS TO FORMALITY

An important number of productive and prepared workers are being segregated from the formal labour market because of their gender or their race.

Gender: According to Table 3 the informality rate in Colombia is 61% among women and 59% among men at the national level, and 51% and 45%, respectively in the main 13 metropolitan areas (Table 3). Similarly, the unemployment rate is 7% among women and 5% among men (GEIH, 2015). These figures show some bias against women in the labour market, in spite of the fact that the levels of education are similar between both genders. In fact, as shown in Table 4, these results are robust after controlling for other observable characteristics. This higher female participation in informality can only be partially explained by preferences. As Table 4 shows, women reported as spouses show high preferences for informality, whereas those that are household heads or that occupy any other place at the household do not show any significant preferences for informality. However, both groups show higher rates of informality than men, signaling some kind of discrimination in the labour market.

Supporting this, Table 6, shows the informality rates of women that are head of the household or occupy a place in the household other than spouse, between 25 and 55 years old; by age, geography and education. The percentage difference of the informality rates between this group and men is a good indicator for gender discrimination, since both groups show similar preferences for informality. Table 6 indicators demonstrate that gender discrimination is low or inexistent for workers with tertiary education, medium among workers with primary education or

less, and high for workers with high school studies; particularly in the 13 main Metropolitan Areas. In the rural areas, the results do not show signs of discrimination in terms of informality.

Table 6. Informality rates of males and females with similar preferences to informality (25 to 55 years old.

		Total			13 areas			Rural		
		Women (Head or other)	Male	% difference	Women (Head or other)	Male	% difference	Women (Head or other)	Male	% difference
Firm definition	Total	47%	51%	-8%	38%	37%	2%	75%	79%	-5%
	Tertiary	22%	21%	6%	19%	18%	2%	43%	42%	3%
	High school	55%	47%	17%	49%	38%	29%	65%	68%	-4%
	Low	85%	78%	10%	78%	63%	23%	89%	85%	4%
Legal definition	Total	51%	55%	-8%	40%	40%	1%	78%	83%	-5%
	Tertiary	24%	26%	-6%	21%	23%	-10%	42%	41%	3%
	High school	60%	49%	23%	52%	38%	37%	69%	70%	-2%
	Low	89%	83%	7%	81%	69%	18%	92%	90%	2%

Source: GEIH

Ethnicity: In the 2005 Colombian census, around 14% of the population declared itself an ethnic minority: 11% Afro-Colombians and 3% indigenous. The GEIH does not ask questions related to ethnicity, however using the 2006 special report on informality, Bernal (2009) found that the probability of working in the informal sector is 5.4 percentage points higher for indigenous people and 2.2 percentage points higher for afro-Colombians, after controlling for other observable characteristics. Similarly, the ELCA survey indicates that estimated informality rates vary significantly, more than 15 percentage points, with the color of the worker's skin (Fernandez & Villar, 2005). This large gap in informality rates cannot be explained by preferences for informality, as in the case of gender. In fact, Bernal (2009) found that ethnic minorities are 8% more likely to prefer a formal job than the rest of the self-employed population.

EXPLICIT BARRIERS TO FORMALITY

The existence of explicit formal market barriers is an unambiguous symptom of induced informality. In analyzing these barriers, we look at payroll taxes, hiring and firing costs and high minimum wages. In order to compare Colombian statistics internationally, we constructed a rank of the main barriers indicators in Table 7[39],

---

[39] Source: Doing Business Indicators, 2016. World Bank and Heritage Foundation, 2016. Notes: Data from Brazil, Mexico and Nigeria corresponds to the average of the two biggest cities. The labour rigidity index is estimated as 1 - the Labour Freedom index, from the Heritage Foundation. The levels are assigned according to the following rule: Greater than the average plus a standard deviation, very high; between the average plus a half a standard deviation and the average plus one standard deviation, high; between the average and the average plus one standard deviation, medium high; between the average minus one half standard deviation and the average, medium low; between the average minus one standard deviation and the average minus half a standard deviation, low; and lower than the average minus one standard deviation, low.

Table 7. Barriers to formality

	Labor tax and contributions (% of profit)		Other barriers		Minimum wage					
					Ratio of minimum wage to value added per worker		Minimum/median		Minimum/mean	
Argentina	29.3	Very high	68	High	63%	High	115%	Very high	69%	Very high
Bolivia	18.8	Medium high	79	Very high	51%	Medium high			69%	Very high
Brazil	40.3	Very high	65	High	31%	Medium low	66%	Medium high	36%	Medium low
Chile	4	Very low	58	Medium low	23%	Medium low	68%	Medium high	45%	Medium high
Colombia	18.6	Medium high	43	Very low	34%	Medium low	77%	High	66%	Very high
Ecuador	13.7	Medium low	62	Medium high	51%	Medium high	83%	High	57%	High
Mexico	25.4	High	65	High	14%	low	38%	Low	29%	low
Panama	20	Medium high	72	Very high	34%	Medium low	73%	High	45%	Medium high
Paraguay	18.6	Medium high	78	Very high	72%	Very high				
Peru	11	Medium low	56	Medium low	33%	Medium low			53%	High
Uruguay	15.6	Medium low	57	Medium low	29%	Medium low	54%	Medium low		

Source: Doing Business Indicators, 2016. World Bank, OECD and Heritage Foundation, 2016.
Notes: Doing business data from Brazil and Mexico corresponds to the average of the two biggest cities. The minimum wage to value added per worker assumes a cashier, age 19, with one year of work experience, when there is not minimum wage. The other labour market rigidities is the arithmetic inverse of the Freedom of labour from the Heritage Foundation, adjusted to exclude the minimum wage to value per worker index. It includes the following factors Hindrance to hiring additional workers, Rigidity of hours, Difficulty of firing redundant employees, Legally mandated notice period, Mandatory severance pay, and the Labour force participation rate. The minimum wage/median or mean wage are taken from OECD, and ILO if the country does not belong to the OECD. The levels are assigned according to the following rule. Greater than the average plus a standard deviation, very high; between the average plus a half a standard deviation and the average plus one standard deviation, high; Between the average and the average plus one standard deviation, medium high; between the average minus one half standard deviation and the average, medium low; between the average minus one standard deviation and the average minus half a standard deviation, low; and lower than the average minus one standard deviation, low. This ranking was performed over

Payroll taxes. Despite of a recent tax reform in Colombia that reduced payroll taxes by 14 percentage points, Colombia continues to be in the upper half of the distribution when looking at labour taxes as percentage of commercial profits. According to Table 6, the level of labour taxes and contributions as a proportion of commercial profits is 18.5% in Colombia, compared with an overall world average of 16%, a Latin America average of 14% and a OECD average of 24% (World Development Indicators, 2016). The relative high impact of the tax reform is another symptom of the high incidence of induced informality in Colombia. In fact, according to Fernandez & Villar (2016), the reduction in the payroll taxes lowered the informality rate of those affected by the reform by 7.4 percentage points, which translates into a reduction of the informality rates of about 4%. This result is similar to previous estimates of the impact of payroll taxes over the informality rate in Colombia (Anton, 2014; Kugler & Kugler, 2009; Mondragon et al., 2010) and at a global level (Albrecht et al., 2009; Hazans, 2011; Slonimczyk, 2011; Lora & Fajardo, 2012). When analyzed by level of education, the impact over workers with primary education or less was 10.1 p.p. among workers with primary education or less, 11.9 p.p. among workers with high school studies and 4.6 p.p. among workers with tertiary education or more.

Minimum wage: Very low minimum wages can be negative in terms of income distribution. However, high minimum wages in relation to value added might encourage informality and particularly, induced informality. The Doing Business Indicators of the World Bank estimates

the ratio of the minimum wage to the value added per worker[40], taking into account not only the countries with a national minimum wage, but also those countries that set wages in collective bargaining on at least 50% of the private sector. According to this index Colombian minimum wage is not particularly high. However, this index is affected by income distribution and by the distribution of human capital, so it does not capture how binding is the minimum wage for poorly educated workers. We also included in Table 6 indicators of the minimum wage over the mean/median wages. These indexes, rank high and very high, respectively. The minimum wage is 66% of average wages[41] and 77% of the median wage.

Another important feature is the way the minimum wage is set. The minimum wage in Colombia is established at a national level and, according to the country's Constitution, it is increased annually based on past inflation plus increases in productivity. According to Hazans (2011) the European countries that set the minimum wage at a national level tend to have higher rates of informality. However, the convenience of a national or federal minimum wage is not settled due to the impact on income distribution (Centro Nacional de Consultoria, 2014)[42].

Other barriers to formality. Other barriers to formality include the hiring cost other obstacles to hiring the rigidity of working hours, the difficulty of firing redundant workers and mandatory severance pay. As in the case of minimum wages, it is not necessarily optimal to have an extremely flexible labour legislation because it can be a symptom of unprotected labour. However, legislation which is too rigid is also related to high induced informality, particularly if the payroll taxes and minimum wages are also high. Colombia ranks very low in the Rigidity of Labour Index (Heritage Foundation, 2015) that includes other aspects of labour legislation, such as working hours and hiring, firing and severance costs[43]. When asked about the main constraints, 39% of employees mention firing cost, a percentage that is almost as high as the 44% that mention payroll taxes and contributions, and higher than the 18% that mention minimum wages.

VI.   A TAXONOMY OF INFORMALITY IN COLOMBIA

Table 7 summarizes the main findings of the previous three sections (III, IV and V).

According to section III, the motivations for informality reported in the surveys, the low frequency of transition between informality and formality, and the evidence in favor of counter-cyclicality of informality, show weak support for the hypothesis of voluntary informality prevailing in Colombia. However, there is stronger evidence in favor of voluntary informality among workers with tertiary education or more: they demonstrate high preferences for informal

---

[40] In particular, it assumes a cashier, age 19, with one year of work experience.

[41] Only including workers that work more than 40 hours of work per week.

[42] The impact of the minimum wage as a reference not only for formal workers but also for informal workers, or the lighthouse impact, imply that the minimum wage may have an impact on those workers that are not legally affected by it. Therefore, the simplicity of the minimum wage at a national level may have may have a role in reducing inequality.

[43] One possible explanation for the high barriers related to monetary hiring cost and low barriers is given by the OECD (2015), according to whom, the unionization process in Colombia is weak, forcing them to concentrate their power on minimum wage negotiations, leaving aside other aspects of work protection. In fact, the percentage of employees that are union members as proportion of total employment is 4.5% in Colombia compared with a OECD average of 17% (OECD database). As a result, the minimum wage has increased above the productivity level of the economy, increasing informality despite the feeble power of the unions.

jobs, and high probability of transitions from informality to formality. There is also no evidence for countercyclical behavior on this group.

According to Section IV, there is substantial evidence of subsistence informality in this country: about 48% of the workers in Colombia earn less than 50% of the cost of hiring them in the formal sector. We argue that these workers are not really segregated from the labour market because of formal labour market barriers or because of their preferences, but because their low levels of productivity. Our findings on the incidence of informality on groups with low levels of education and experience and in non-urban and non-productive areas confirm the relevance of subsistence informality in the country. At a disaggregate level, the workers with lower levels of education, show the highest the percentage of workers earning less than half of the cost of hiring a worker (59%).

Table 8. Indicators for Each Type of Informality.

	Marker	Total	13-areas	Primary or less	High School	Tertiary
Voluntary	Transitions informality to formality		Low 14%	Low 10%	Low 12%	Medium 26%
	Preferences for informality	Low 35.9%	Low 41.5%	Low 40%	Low 41%	Medium 52%
	Cyclicality: coefficient of correlation		Counter-cyclical -0.42%**	Counter cyclical -0.59**	Counter -cyclical -0.71***	No sig. -0.13
Subsistence	% of workers earning below half the minimum hiring cost	High 48%	High 35%	High 59%	Medium 38%	Medium 28%
	Relative probability of being informal, low productivity groups	High	High			
Induced	Informality rate gap (woman that prefer to be formal vs male, 25 to 55 years old)	Medium	Medium	Medium 10%	High 17%	Medium 10%
	Impact of reducing payroll taxes (p.p.)		Medium -8.0	High -10.2	High -13	Low -(n.s)

Finally, regarding induced informality on section V, we find good evidence for the relevance of this type of informality in Colombia, mainly related to explicit barriers to formality, such as payroll taxes, but also with some signs of discrimination. By level of education, we found some evidence for discrimination among the workers with high school studies, in the sense that those women who would like to have a formal job show higher levels of informality, correcting by education. This group was also the most affected by the tax reform that recently reduced the payroll taxes (Fernandez and Villar, 2016), signaling that this group was the most affected by the regulatory barriers.

As a result, we argue that in Colombia, there is a heterogeneous distribution of informality, consistent with Perry (2007) and Garcia (2014). However, in this section we attempt to go a step

further in this analysis and identify the shares and characteristics of the workers that face each type of informality. As we claimed before, this step implies a risk of oversimplification, given that a precise identification is not even possible at a theoretical level.

The motivations to be informal reported by surveys that we mentioned in Section III help us to distinguish voluntary from involuntary informality. According to our estimations, about 39% of the informal workers are Voluntary by this indicator. In turn, workers with earnings well below the minimum wage (48% in the total survey), as explained in section IV, may be classified in the group of Subsistence informality. All the remaining workers can be classify as Induced informal workers.

However, the direct use of the numbers mentioned above may lead to an identification problem as far as about 15% of the informal workers are at the same time classified as both voluntary and subsistence informal workers. Those workers earn less than half the cost of hiring and at the same time prefer to have an informal job. A plausible explanation for this is the perverse incentives caused by an incorrect design of the social policy. This is the case of workers remaining or becoming informal to get access to government benefits programs as cash transfers, health protection, Colombia Mayor (oriented to the older population), Jovenes en Acción (oriented to young workers). We classified the workers in this group as Subsistence informal workers, if they have less than high school studies; and as Voluntary informal workers if they have more than high school studies. Table 9 shows the final distribution of our three types of informality.

Table 9. Shares of each type of informality.

	Firm definiton			Legal definition		
	Total	13-areas	Rural	Total	13-areas	Rural
Subsistency	48%	34%	65%	47%	33%	63%
Induced	27%	29%	20%	31%	32%	24%
Voluntary	24%	36%	15%	22%	35%	13%

Source: GEIH. Authors Calculations

The relevance of obtaining these shares is not the numbers per se, since they are the result of assumptions and there is not benchmark to be compared. The advantage to obtain these shares is that they allow us to find the determinants and characterize each type of informality. Running a logit model of each type of informality against the determinants of informality (Tables 10a and 10b), we can make the following characterization of each type of informality. Results are robust to the type of definition of informality implemented.

- **Voluntary informality:** The workers with tertiary education or more, show a higher probability to belong to this group than any other education group; the older population and people living in productive areas and the group of women that are reported as spouses, also show high incidence of voluntary informality.

- **Induced informality:** Workers with secondary education, males, workers in their most productive years 25-45 (the base category) and people living in less productive cities, show higher incidence of induced informality. This result is coherent with the results of Fernandez and Villar (2016) in terms of

Table 8a. Determinants of each type of informality. Firm definition

	Subsistence informality			Induced informality			Voluntary informality		
	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural
Elementary or less	1.607***	1.428***	1.783***	0.878**	1.072	0.713**	0.604***	0.662***	0.583***
	[10.3]	[6.1]	[6.0]	[-2.6]	[1.1]	[-3.1]	[-9.0]	[-6.3]	[-3.8]
High school	0.450***	0.361***	0.612***	1.183***	1.203**	1.167	2.164***	2.062***	1.889***
	[-16.1]	[-16.9]	[-4.1]	[3.4]	[3.1]	[1.2]	[13.9]	[11.7]	[3.9]
Tertiary or more	0.217***	0.194***	0.237***	0.786***	0.681***	0.927	5.293***	4.879***	6.518***
	[-22.9]	[-19.8]	[-6.6]	[-4.1]	[-5.5]	[-0.3]	[26.6]	[22.7]	[7.8]
Women (spouse)	1.845***	1.664***	1.797***	0.446***	0.471***	0.509***	1.146**	1.297***	0.816
	[12.9]	[8.5]	[5.8]	[-15.5]	[-12.1]	[-5.3]	[2.9]	[4.6]	[-1.6]
Women (other)	1.900***	1.759***	2.428***	0.594***	0.691***	0.489***	0.847***	0.879**	0.513***
	[16.7]	[11.4]	[9.4]	[-13.2]	[-7.6]	[-6.1]	[-3.9]	[-2.6]	[-5.4]
Less than 24 years	2.238***	2.566***	2.358***	0.748***	0.871*	0.792*	0.502***	0.480***	0.243***
	[15.8]	[14.2]	[8.2]	[-5.8]	[-2.2]	[-2.1]	[-10.9]	[-10.5]	[-7.6]
45-55 years	0.827***	0.839**	0.831*	0.911*	0.793***	1.154	1.431***	1.487***	1.164
	[-4.3]	[-3.0]	[-2.1]	[-2.2]	[-4.4]	[1.4]	[7.5]	[7.1]	[1.2]
56+ years	1.229***	1.398***	1.211*	0.279***	0.331***	0.192***	2.322***	1.732***	2.465***
	[4.4]	[5.3]	[2.0]	[-23.3]	[-15.9]	[-10.5]	[16.2]	[8.6]	[7.4]
Rural	1.809***			0.514***			0.872**		
	[13.8]			[-13.6]			[-2.6]		
Productive city	0.462***	0.677***		0.805***	0.771***		2.560***	1.747***	
	[-17.6]	[-8.2]		[-5.0]	[-5.7]		[22.3]	[12.7]	
Less productive city	1.091*	1.553***		1.686***	1.318***		0.406***	0.440***	
	[2.2]	[10.2]		[13.3]	[6.6]		[-18.6]	[-18.9]	
Constant	0.615***	0.456***	0.953	0.809***	0.758***	0.444***	0.203***	0.314***	0.245***
	[-10.0]	[-12.6]	[-0.5]	[-4.2]	[-4.5]	[-7.4]	[-27.8]	[-17.6]	[-10.3]
N	53473	23629	7193	53473	23629	7193	53473	23629	7193
F	238.03	136.01	34.88	165.41	88.28	26.70	268.57	192.45	41.64
df_m	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00
df_r	53,472	23,628	7,192	53,472	23,628	7,192	53,472	23,628	7,192

Source: GEIH and author's calculations

Table 8b. Determinants of each type of informality. Legal definition

	Voluntary informality			Induced informality			Subsistence informality		
	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural
Elementary or less	1.533***	1.363***	1.617***	0.889*	1.082	0.748**	0.612***	0.679***	0.641**
	[9.5]	[5.3]	[5.1]	[-2.4]	[1.3]	[-2.8]	[-8.6]	[-5.8]	[-3.0]
High school	0.491***	0.400***	0.601***	1.249***	1.321***	1.237	1.938***	1.790***	1.905***
	[-14.9]	[-15.5]	[-4.4]	[4.7]	[4.8]	[1.7]	[12.0]	[9.4]	[3.9]
Tertiary or more	0.205***	0.171***	0.225***	0.943	0.830**	1.106	4.935***	4.644***	6.763***
	[-25.5]	[-22.7]	[-7.2]	[-1.1]	[-3.0]	[0.5]	[26.7]	[23.1]	[7.9]
Women (spouse)	1.884***	1.744***	1.871***	0.421***	0.474***	0.450***	1.267***	1.271***	0.951
	[13.8]	[9.5]	[6.4]	[-17.8]	[-12.6]	[-6.6]	[5.1]	[4.3]	[-0.4]
Women (other)	1.979***	1.920***	2.352***	0.596***	0.728***	0.534***	0.830***	0.777***	0.494***
	[18.2]	[13.4]	[9.4]	[-14.0]	[-6.9]	[-6.1]	[-4.4]	[-5.2]	[-5.4]
Less than 24 years	1.999***	2.317***	2.109***	0.746***	0.909	0.789*	0.595***	0.519***	0.285***
	[14.7]	[13.8]	[7.5]	[-6.5]	[-1.7]	[-2.3]	[-9.3]	[-10.8]	[-7.1]
45-55 years	0.871**	0.92	0.852	0.912*	0.828***	1.132	1.387***	1.339***	1.139
	[-3.2]	[-1.4]	[-1.8]	[-2.2]	[-3.6]	[1.3]	[6.8]	[5.1]	[1.1]
56+ years	1.225***	1.336***	1.240*	0.388***	0.358***	0.502***	2.116***	1.795***	1.661***
	[4.2]	[4.4]	[2.2]	[-17.5]	[-14.7]	[-5.8]	[14.0]	[8.7]	[4.0]
Rural	1.840***			0.580***			0.756***		
	[14.6]			[-11.8]			[-5.3]		
Productive city	0.502***	0.734***		0.810***	0.835***		2.460***	1.542***	
	[-16.0]	[-6.6]		[-5.1]	[-4.1]		[21.8]	[10.0]	
Less productive city	1.134***	1.616***		1.605***	1.373***		0.373***	0.388***	
	[3.3]	[11.5]		[13.0]	[8.0]		[-20.3]	[-22.0]	
Constant	0.589***	0.424***	0.986	0.864**	0.719***	0.495***	0.185***	0.342***	0.193***
	[-11.2]	[-14.2]	[-0.2]	[-3.0]	[-5.6]	[-6.7]	[-30.3]	[-16.7]	[-11.6]
N	56732	24748	7501	56732	24748	7501	56732	24748	7501
F	235.79	130.28	33.70	136.31	83.88	17.01	260.55	190.40	31.88
df_m	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00
df_r	56,731	24,747	7,500	56,731	24,747	7,500	56,731	24,747	7,500

Source: GEIH and author's calculations

VII. Conclusions

One of the main conclusions in this paper is that informality in Colombia is large and heterogeneous. Informal workers in the country range from poorly educated individuals, which may all be classified as structurally informal; to highly educated young adults living in productive cities, who can be classified as voluntarily informal.

This paper applies a methodology to identify the three types of informality to the case of Colombia making emphasis on the educational level. Although the correspondence is far from perfect, we show that in general terms, informal workers with primary education or less can be treated as part of the Subsistence informality, informal workers with secondary education are associated with the Induced informality and informal workers with tertiary education or more belong as a general rule to the Voluntary informality group. This correspondence indicates that it is useful to analyze and understand informality differentiating by education level.

This analysis also brings important policy implications. Although there is an important bulk of the informal population that is affected by the formal employment barriers, there is also a component of informality that is structural in nature and that needs to be tackled with other kind of policies, such as education. In the case of voluntary informality, imposing constraints on remaining informal and economic incentives to formalise might be effective, whereas the same policies applied to structural informality are likely to compromise the inclusiveness of growth.

VIII. Bibliography

Albrecht, J., Navarro, L. & Vroman, S. 2009. The effects of labour market policies in an economy with an informal sector. The Economic Journal 119 (539), 1105–1129.

Anton, A. 2014. The effect of payroll taxes on employment and wages under high labor informality. IZA Journal of labor and Development.

Arias, O. and Bustelo, M. 2007. *Profiles and Dynamics of Informal Employment in Latin America..* World Bank, Washington D.C.

Arias, O., Landa, F., and Yáñez, P. 2007. Movilidad Laboral e Ingresos en el Sector Formal e Informal de Bolivia." Documento de Trabajo, UDAPE, La Paz, Bolivia.

Alcaraz, C., Chiquiar, D. and Salcedo, A. 2015. Informality and Segmentation in the Mexican Labor Market. Working Papers. No 2015-25. Banco de México, Mexico City.

Bernal, R. 2009. The informal labor market in Colombia: identification and characterization. Desarrollo y sociedad, (63), 145-208.

Bosch, M., & Maloney, W. F. 2008. Cyclical movements in unemployment and informality in developing countries. World Bank Policy Research Working Paper Series.

Centro Nacional de Consultoria, 2014. Encuesta de Opinión Empresarial Reforma Tributaria.

Departamento Nacional de Estadística (DANE). 2002-2015. Gran Encuesta Integrada de hogares (GEIH) 2002-2015 [dataset]. DANE.

Departamento Nacional de Estadística (DANE). 2002-2008. Encuesta Continua de Hogares (ECH) 2002-2015 [dataset]. DANE.

De Soto, H. 2000. The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else. Basic Books, New York.

De la Torre, A. Ize, A. Empleos, salarios, sistribución y la desaceleración latinoamericana. Montevideo, Third Ridge Summer Schoool.

Fernández, C. y Villar, L. 2016. Rethinking the Effect of Informality on Inclusive Growth: Lessons from Colombia and South Africa for their Regions.. Fedesarrollo & DPRU.

Fernández, C. y Villar, L. 2015. Informality and Inclusive Growth in Latin America, with emphasis in the case of Colombia. Fedesarrollo.

Fiess, N., Fugazza, M. and Maloney, W. F. 2008. Informality and Macroeconomic Fluctuations. IZA Discussion Paper 3519.IZA, Bonn.

Galvis, L., 2012. *Informalidad laboral en las áreas urbanas de Colombia.* Documentos de trabajo sobre economía regional. No. 164. Banco de la República. Bogotá.

García, G.A. 2014. Informalidad regional en Colombia: evidencia y determinantes. CIDSE

García, G. 2014. *Labor Informality: Choice or Sign of Segmentation? A Quintile Regression Approach at the Regional Level for Colombia.* Archivos de Economía. Documento 412. National Planning Department, Government of Colombia, Bogotá.

Haanwinckel, D y Soares, R. 2013. A Compensating Differentials Theory of Informal Labor Markets: Quantitative Model and Implications for a Developing Country. LACEA conference paper.

Harris, J.R., Todaro, M.P. 1970. *Migration, Unemployment, and Development: A Two-Sector Analysis*. The American Economic Review, 60(1), 126-42.

Hazans, M. (2011). Informal workers across Europe: Evidence from 30 countries.

Heritage Foundation, Labour Freedom, accessed on 14 April 2016.

Isaza and Rojas (2015) Diagnóstico de la heterogeneidad structural y el desempeño macroeconómico en Colombia on Maroeconomía y empleo en Colombia. OIT.

Izasa, J., Rojas, N., Cubillos, R., Farné, S. 2015. Macroeconomía y empleo en Colombia. OIT.

Kugler, A., Kugler, M. 2009. Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia. University of Chicago.

Lewis, W. A. 1954. *Economic Development with Unlimited Supplies of Labour*. The Manchester school, 22(2), 139-191.

Loayza, N., & Rigolini, J. 2006. Informality trends and cycles. World Bank Policy Research Working Paper, (4078).

Lora, E. A., & Fajardo, D. J. 2012. Employment and taxes in Latin America: An empirical study of the effects of payroll, corporate income and value-added taxes on labor outcomes.

Maloney, W. F. 2004. *Informality revisited*. World development, 32(7), 1159-1178.

Mondragon-Velez, C., Peña, X., Wills, D., Kugler, A. 2010. Labor Market Rigidities and Informality in Colombia. Economía, 65-101.

OECD. 2015. OECD Employment Outlook 2015, OECD Publishing, Paris.

Pagés, C., & Stampini, M. (2009). No education, no good jobs? Evidence on the relationship between education and labor market segmentation. Journal of Comparative Economics, 37(3), 387-401.

Perry, G. (Ed.). 2007. *Informality: Exit and exclusion*. World Bank Publications.

Peña, X. 2013. The formal and informal sectors in Colombia. ILO.

Pratap, S., Quintin, E. 2006. The informal sector in developing countries. United Nations University.

Slonimczyk, F. 2011. The effect on taxation on informal unemployment. Evidence from the Russian flat tax reform. London School of Economics.

Tornarolli, L., Battistón, D., Gasparini, L. and Gluzmann, P. 2014.Exploring Trends in Labor Informality in Latin America, 1990-2010. CEDLAS Universidad Nacional de la Plata, La Plata.

Universidad de los Andes. 2010-2013. Encuesta Longitudinal Colombiana de la Universidad de los Andes (ELCA). Universidad de los Andes.

Ulyssea, G. 2013. Firms, Informality and Welfare. University of Chicago.

Uribe, J., Ortiz, C., García. (2007). La Segmentación del Mercado Laboral Colombiano en la Década de los noventa. Revista de economía institucional, vol. 9, No. 16. Universidad Externado de Colombia. Bogotá.

World Bank's Enterprises Survey, accessed 13 April 2016.

World Bank's Doing Business Indicators, accessed 13 April 2016.

World Bank's Social Protection and Labor Indicators, accessed 14 April 2016.

World Bank's World Development Indicators, accessed 14 April 2016.

IX. ANNEX A

This table displays the results for a logistic regression made to predict preferences for informal work in 2015, given the results of the 2007 GEIH. The regressions is quite similar to that shown in Table 1, but, due to the nature of this logit regression (whose principal objective was to predict rather than to find and analyze relations), endogenous variables where included, and the variables included in the regression were established to obtain the best possible prediction for 2015 voluntary informal workers.

Equation for prediction of preferences for informal jobs					
(t-statistics in parenthesis)	Raw	Clear		Raw	Clear
Elementary or less	0.860*	0.821***	Barranquilla	0.537***	0.540***
	(-2.38)	(-3.78)		(-7.46)	(-8.15)
High school	1.069		Cartagena	0.234***	0.233***
	(0.99)			(-14.57)	(-15.56)
Tertiary or more	1.525***	1.489***	Monteria	1.335***	1.343***
	(4.79)	(5.05)		(3.51)	(3.94)
Women (spouse)	0.938		Villavicencio	2.287***	2.306***
	(-1.10)			(9.43)	(10.39)
Women (other)	1.185*	1.221**	Pasto	1.749***	1.773***
	(2.51)	(3.18)		(6.15)	(6.78)
Less than 24 years	0.743***	0.739***	Cucuta	1.691***	1.707***
	(-3.66)	(-3.74)		(6.21)	(6.92)
45-55 years	1.221***	1.218***	Pereira	1.829***	1.847***
	(3.59)	(3.56)		(6.36)	(6.94)
56+ years	1.666***	1.666***	Bucaramanga	1.181	
	(8.39)	(8.42)		(1.91)	
Rural	0.938		Ibague	1.552***	1.558***
	(-0.78)			(4.95)	(5.42)
Job satisfaction	2.796***	2.805***	Cali	1.719***	1.735***
	(21.18)	(21.53)		(6.49)	(7.23)
Self-employment	0.488***	0.490***	Tunja	1.119	
	(-9.71)	(-9.72)		(1.09)	
Agriculture	1.263*	1.296***	Florencia	1.586***	1.599***
	(2.28)	(3.44)		(5.10)	(5.62)
Mines	0.388		Popayan	0.432***	0.435***
	(-1.91)			(-8.69)	(-9.22)
Manufacturing	1.513***	1.584***	Valledupar	0.262***	0.264***
	(5.00)	(6.42)		(-13.25)	(-14.06)
Public services	1.407		Quibdo	0.362***	0.357***
	(0.41)			(-6.15)	(-6.46)
Retail, restaurants and hotels	1.420***	1.498***	Neiva	0.836*	0.840*
	(5.13)	(7.58)		(-2.06)	(-2.20)
Constructions	0.815		Riohacha	0.694***	0.700***
	(-1.69)			(-3.71)	(-3.89)
Transportation and communications	0.956		Santa Marta	0.688***	0.691***
	(-0.50)			(-4.33)	(-4.67)
Medellin	1.6989***	1.700***	Armenia	1.006	
	(6.57)	(7.40)		(0.07)	
Bogota	1.405***	1.414***	Sincelejo	0.351***	0.355***
	(3.89)	(4.30)		(-10.20)	(-10.75)
Constant	0.373***	0.357***	Sincelejo	0.351***	0.355***
	(-6.61)	(-10.78)		(-10.20)	(-10.75)
Number of obs	33104	33104	Design df	41	30
Population size	60	82	Prob > F	33103	33103

X. ANNEX B

A proxy for the cities' productivity was constructed by dividing the average minimum wage earned in each of Colombia's 23 cities by the legal minimum wage. Then we ranked the cities from the most productive to the least productive city. As figure 1 exhibit, in general, the most productive cities also present higher formality rates. When considering all the 23 Colombian cities, Bogotá, Medellín, Tunja and Bucaramanga are the most productive ones; and Barranquilla, Sincelejo, Cucutá and Quibdó, the least productive ones. When we restrict the analysis to the 13 most important areas, Bogotá, Medellin, Bucaramanga and Villavicencio are the most productive cities, while Pasto, Monteria, Barranquilla and Cucutá are the least productive areas. In 2007 the most productive cities: Medellin Barranquilla y Bogota; the non-productive cities: Tunja, Quibdo and Armenia and in the 13 areas: the most productive cities: Medellin Barranquilla y Bogota, and the non- productive cities: Villavicencio, Pasto, Cucuta and Pereira

Figure 1. Cities' productivity vs Formality rates
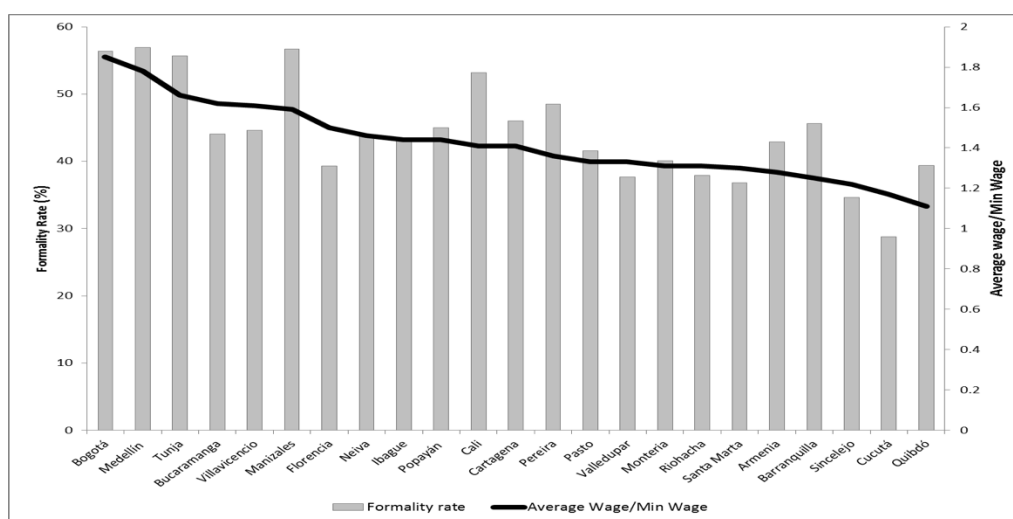
Source: Authors calculations based on GEIH 3rd Quarter 2015.

For the productive vs non-productive sectors' selection, we used a historic ranking constructed by Isaza et al. (2015), were they exploit data coming from Groningen Growth and Development Centre (GGDC). The authors estimate the economic sectors' average productivity, by dividing the aggregate value of the sector over the number of occupied workers per sector. Table 1 exhibits the index's results. Following the results of 2013, we chose as productive sectors Mining and Domiciliary public services, and as non-productive sectors Agriculture and Commerce.

Table 1. Labour productivity Index by economic sector (1980-2013)

	1950	1968	2000	2013
High productivity				
Mining	493	607	702	394
Domiciliary public services	595	1376	1388	1241
Medium productivity				
Transportation and communication	165	178	167	156
Manufacturing	147	150	146	138
Construction	231	202	107	108
Finance and insurance business services	150	159	143	129
Services	121	61	63	66
Low productivity				
Agriculture	35	31	40	53
Commerce	406	82	57	50

Nota: Data taken from Izasa, Rojas, Cubillos, Farné (2015)