

A Taxonomy of Colombia's Informal Labor Market

Cristina Fernández*
Leonardo Villar

Abstract

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 Fernández 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.

Resumen

En este trabajo proponemos una metodología para identificar los cuatro 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, y trabajadores informales con educación media pueden cubrir la informalidad mixta. Por lo tanto, las recomendaciones de política para manejar la informalidad entre cada grupo educativo son diferentes.

Keywords: Informality, Education Level, Policy Recommendations

Palabras clave: Informalidad, Nivel educativo, Recomendaciones de política

Clasificación JEL: J210, J680, I240

Primera versión recibida el 4 de junio de 2016; versión final aceptada el 28 de octubre de 2016

Coyuntura Económica. Volumen XLVI, No. 1, Junio de 2016, pp. 15-50. Fedesarrollo, Bogotá - Colombia

* Researcher and Director of Fedesarrollo, respectively. We want to thank Juan Camilo Medellín and Francisco Fernández for excellent research assistance, and Adriana Arreaza, Fernando Álvarez and Manuel Toledo for their comments. The elaboration of this paper was supported by PAC-DFID ELLA program and CAF.

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. Mixed informality is a combination of subsistence and voluntary informality.

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

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) 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, 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.

The Colombian data used in the majority of the ensuing analysis is from the second and 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.¹ 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).² 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.³ 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-

productivity/Subsistence informality; Section V illustrates the indicators for barriers to formality/Induced informality; Section VI 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 VII 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

¹ 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.

² 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).

³ 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).

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.

We also considered a fourth group: 'Mixed informality', that includes workers that are at the same time voluntary and subsistence informal workers. Those workers have a low chance to find a job in the formal market due to their low productivity, but even if they were able to find a formal job they wouldn't take it. The reasons to prefer informality vary, but they are often related to the existence of social benefits. According to Levy (2009) some workers prefer to be informal in order to do not lose their social benefits. In the case of Colombia, a recent research lead by Stefano Farné (2016), found that without the recent increase in cash transfers and other benefits, informality would have been around 10 percentage points less for those workers that receive this type of help. However, failures in the social policy design might not be the only cause of this new type of informality. There are other reasons that make a worker with low productivity to prefer informal jobs. One of them is the geographical distance to the places where formal jobs are offered as suggested by Hausmann (2014). Another reason might be the lack child care facilities, that restrict women possibility of holding a full type job.

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.

Ulysea (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 only between 10 and 20% of informal workers depending on the co-variables introduced into the model, demonstrated marked signs of segmentation (or not Voluntary informality) 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 cities⁴ 75% of informality is involuntary, in big cities⁵ it accounts for 47% of total informality. Meanwhile, this percentage is 76% for the Caribbean Coast cities.⁶ However, Perry (2007)

⁴ Includes Cúcuta, Montería, Pasto and Villavicencio.

⁵ Includes Bogotá, Medellín, Cali, Bucaramanga, Manizales, Pereira and Ibagué.

⁶ Includes Barranquilla and Cartagena.

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.

III. Choice

We identified three main ways of establishing whether entry into informality is voluntary or involuntary: i) via surveys about worker's preferences, ii) through a counter/pro cyclical 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.

A. 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 infor-

mal workers whether they would like to accept a job in the formal market with the same or higher earnings (including wage plus benefits).⁷ 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),⁸ the preferences of informal work among independent workers 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).⁹ Preferences for informality in the region among the informal salaried 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 Colombia¹⁰ and 59%¹¹ in Argentina. In Mexico¹² only 12% of 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.¹³ When compared with a more international

⁷ 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.

⁸ 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).

⁹ 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*.

¹⁰ 34% of answers, since two choices were available

¹¹ 43% of answers, since multiple choices were available

¹² Based on *Encuesta Anual de Micronegocios, 1994*

¹³ 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.

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).¹⁴ 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).¹⁵ The results suggest that workers with tertiary education, relatively old workers (45 years old or more), and workers living in big¹⁶ and border cities¹⁷ exhibit higher preferences for informal jobs. The later can be explained by the

presence of smuggling activities in border cities. Women reported as second earners, according to their role in the household (41% of the informal working women in 2007¹⁸) also show strong preferences for informality. Women that are head of the household (30% 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 results show higher preferences for formality among low educated workers, males and the urban population; and lower preferences for formality among women, second earners 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.

¹⁴ Based on Blanchflower (2004), Blanchflower, Oswald and Stutzer (2001) and own calculations.

¹⁵ The odds ratio is a 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.

¹⁶ Bogotá, Medellín and Cali.

¹⁷ Pasto and Cúcuta.

¹⁸ Only 3% of the male informal workers describe themselves as second earners of the household, and they do not show significant preferences for informal jobs.

Table 1
LOGIT: PREFERENCES FOR INFORMALITY, ODDS RATIO

<i>Depvar: preferences for informality</i>	Firm definition			Legal definition		
	National	13-areas	Rural	National	13-areas	Rural
Elementary or less	0.908 * [-2.3]	0.903 * [-2.0]	0.949 [-0.4]	0.912 * [-2.1]	0.905 [-1.9]	0.945 [-0.5]
High school	1,132 ** [2.7]	1,113 * [2.0]	0.925 [-0.5]	1,086 [1.7]	1,035 [0.6]	0.918 [-0.5]
Tertiary or more	1,706 *** [9.3]	1,617 *** [7.4]	2,385 * [2.5]	1,642 *** [8.0]	1,541 *** [6.2]	2,300 * [2.2]
Women (second earner)	1,378 *** [7.6]	1,383 *** [6.5]	1,524 *** [4.0]	1,426 *** [8.3]	1,417 *** [6.8]	1,565 *** [4.3]
Women (other)	0.932 [-1.8]	0.941 [-1.4]	0.902 [-0.9]	0.934 [-1.7]	0.933 [-1.4]	0.916 [-0.8]
Less than 24 years	0.621 *** [-8.3]	0.734 *** [-4.3]	0.459 *** [-5.7]	0.636 *** [-7.7]	0.748 *** [-3.9]	0.464 *** [-5.6]
45-55 years	1,341 *** [7.7]	1,257 *** [5.0]	1,545 *** [4.7]	1,332 *** [7.3]	1,259 *** [4.7]	1,530 *** [4.6]
56+ years	2,259 *** [19.2]	1,991 *** [13.1]	2,616 *** [10.9]	2,163 *** [17.3]	1,849 *** [10.9]	2,509 *** [10.3]
Big city	1,942 *** [18.4]	1,942 *** [19.8]		1,971 *** [17.9]	1,986 *** [19.3]	
Border city	1,671 *** [13.1]	1,822 *** [15.2]		1,681 *** [12.8]	1,850 *** [15.0]	
Rural	1,110 * [2.4]			1,125 ** [2.7]		
Constant	0.342 *** [-25.7]	0.356 *** [-22.5] [-8.8]	0.351 *** [-25.6]	0.332 *** [-22.2]	0.348 *** [-8.8]	0.347 ***
Number of obs	64,098	29,643	6,916	60,534	27,376	6,856
F	104	84	29	91	75	27
df_m	11	10	8	11	10	8
df_r	64,097	29,642	6,915	60,533	27,375	6,855

t-statistics in parenthesis. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. 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.

Source: GEIH (2nd and 3rd quarter 2007).

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,¹⁹ 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 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 36% of the informal workers in Colombia were voluntary informal workers in 2015. Applying the same procedure yield that 51% of the informal workers in the 13-areas and 25% of the informal workers in the rural areas were voluntary informal workers.^{20,21}

B. Countercyclicality

Counter cyclicity 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 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 cyclicity enable us to discern between voluntary and involuntary informality.

¹⁹ 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 (Annex A).

²⁰ The respective percentages are 31% for the total expanded survey, 46% for the 13 metropolitan areas and 24% for the rural areas, using the health and pension contributions definition.

²¹ Alternatively, it is possible to assume that employees do not have preferences for informality, which yields to lower percentages of voluntary informality. However, according to Arias and Bustelo (2007) the preferences for informal employment among salaried workers in Colombia tend to be similar to those of the independent workers - 40% and 41% of the informal workers have preferences for informality among the salaried and the independent workers respectively.

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.

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 cyclical 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-cyclical in the region (Loayza and Rigolini, 2006 and Tornarolli *et al.*, 2014). However, there is strong evidence of pro-cyclical 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-cyclical of informality by level of education, showed in Figure 1, shows that informality among workers with low levels of education 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 education²⁵ or less and workers with high school education, respectively.²⁶ Meanwhile, there is no evidence of counter-cyclical 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.

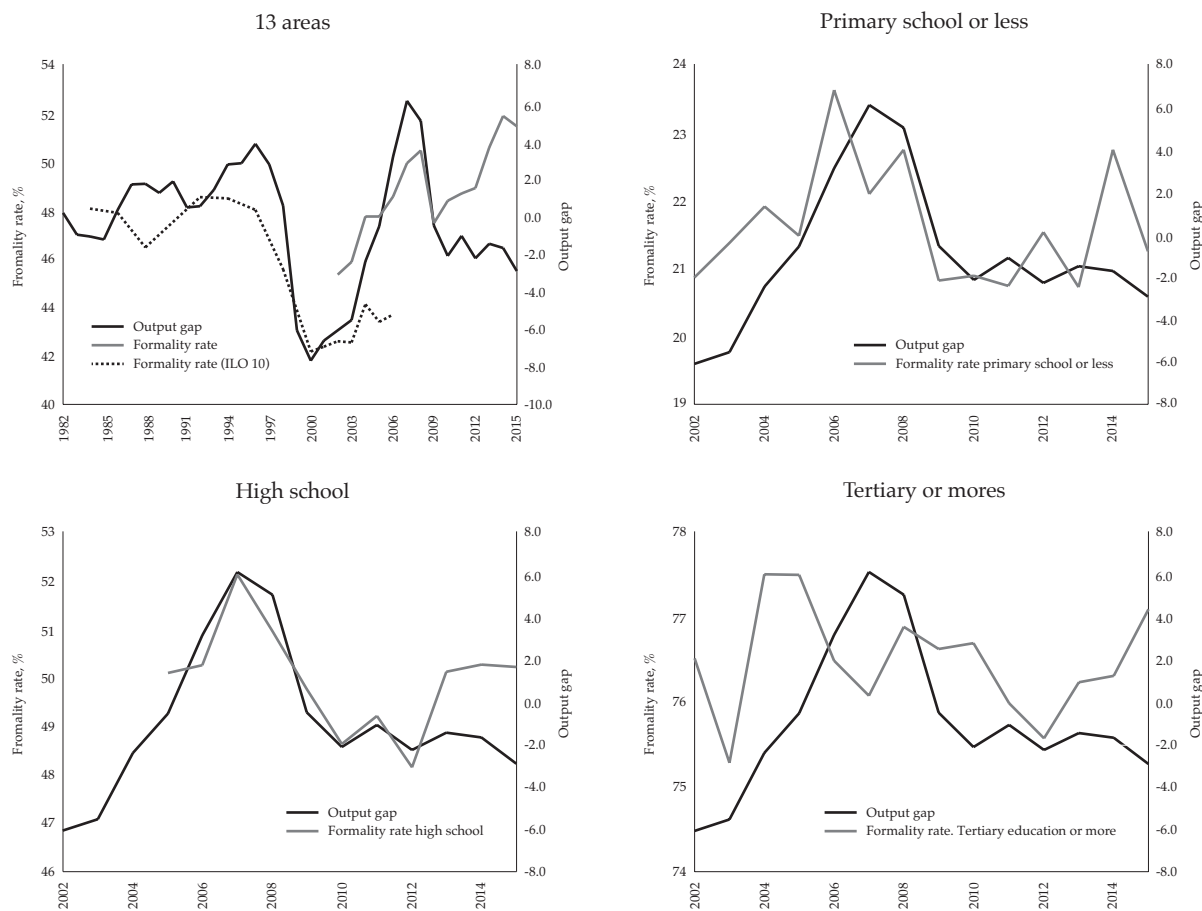
²³ Defining the formality rate as one minus the informality rate. We only performed this exercise using the aggregated 13 metropolitan area data since the wider sample suffered significant changes in the number of interviewed households. Note that in the case of the 13 areas aggregate, the formality 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 áreas)



	Firm definition (2002-2015)	Legal definition (2202-2015)	Mondragon <i>et al.</i> (1984-2006)	Primary school or less	High school	Tertiary educa- tion or more
Output gap (2002-2015)	0.46 **	0.46 *	0.74 ***	0.59 **	0.75 ***	0.13
Output gap (2002-2013)	0.74 ***	0.56 *	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.

C. 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 24-60²⁹ 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 state. 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

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

Note: legal definition of informality. Secondary education includes workers with high school and middle school studies.

Source: Own calculations based on ELCA 2010 and 2013 waves.

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.

A. 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 pro-

ductivity 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.³¹ 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.³²

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.³³ A coefficient of 1.4 for primary education or less means that workers in this group are 1.4 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.

³¹ Alternatively, we performed a multi-logit model for the determinants of informality, where the other alternatives considered were: unemployment, inactivity and formal employment. The results were very similar but their interpretation is more complex, so we prefer to stick in this paper to the simple logit results.

³² The rates are 85%, 56% and 26% respectively if we include rural and other urban areas.

³³ 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.

Table 3
COLOMBIAN INFORMALITY RATES BY WORKER CHARACTERISTICS

	Firm definition			Legal definition		
	Total (%)	13 Cities (%)	Rural (%)	Total 1 (%)	3 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
Second earner	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	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

Note: See the annex B for the ranking of cities and sectors.

Source: Gran Encuesta Integrada de Hogares (GEIH).

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	1,414** [10.0]	1,350*** (8.36)	0.887** [-2.9]	1,355*** [7.2]	0.880* [-2.5]	1,410*** [3.8]	0.949 [-0.4]
High school	0.450*** [-26.9]	0.448*** (-26.02)	1,127** [2.6]	0.443*** [-23.5]	1,104 [1.9]	0.519*** [-7.2]	0.925 [-0.5]
Tertiary or more	0.123*** [-68.4]	0.133*** (-64.30)	1,783*** [10.2]	0.134*** [-56.0]	1,679*** [8.1]	0.123*** [-18.8]	2,385* [2.5]
Less than 24 years	1,423*** [13.1]	1,322*** (10.07)	0.613*** [-8.6]	1,264*** [7.5]	0.719*** [-4.6]	1,395*** [4.1]	0.459*** [-5.7]
45-55 years	1,275*** [9.9]	1,309*** (10.70)	1,351*** [8.0]	1,458*** [12.8]	1,269*** [5.2]	1,201* [2.4]	1,545*** [4.7]
56+ years	2,354*** [27.4]	2,295*** (26.10)	2,262*** [19.4]	2,717*** [26.1]	2,016*** [13.5]	2,738*** [10.1]	2,616*** [10.9]
Rural	1,789*** [17.7]	1,415*** (9.94)	0.978 [-0.5]				
Productive city	0.595*** [-24.2]	0.614*** (-22.39)	1,418*** [9.8]	0.850*** [-7.1]	1,290*** [6.8]		
Less productive city	1,122*** [5.2]	1,140*** (5.82)	0.884** [-2.8]	1,579*** [20.2]	1,595*** [13.6]		
Women (second earner)	1,910*** [24.0]	1,819*** (21.75)	1,392*** [7.9]	1,981*** [21.1]	1,404*** [6.8]	1,941*** [7.2]	1,524*** [4.0]
Women (other)	1,495*** [18.7]	1,485*** (18.00)	0.951 [-1.3]	1,514*** [16.3]	0.966 [-0.8]	1,196* [2.5]	0.902 [-0.9]
Productive sector		0.321*** (-13.53)					
Less productive sector		2,673*** (46.10)					
Constant	2,220*** [26.3]	1,620*** (15.27)	0.394*** [-22.9]	1,494*** [11.5]	0.450*** [-16.7]	3,961*** [16.5]	0.351*** [-8.8]
N	182,636	182,636	64,098	89,119	296,431	7,843	6,916
F	1,159.50	1,040.00	83.18	753.51	51.09	113.73	29.14
df_m	11.00	13.00	11.00	10.00	10.00	8.00	8.00
df_r	182,635	182,635	64,097	89,118	29,642	17,842	6,915

t-statistics in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1). The odds ratios 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.

Note: See Annex B 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.

Source: Authors calculations based on GEIH 3rd Quarter 2015.

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	1,211 *** [5.6]	1,159 *** (4.28)	0.89 ** [-2.6]	1,349 *** [7.1]	0.884 * [-2.4]	0.713 *** [-4.5]	0.945 [-0.5]
High school	0.443 *** [-27.8]	0.433 *** (-27.30)	1,083 [1.7]	0.396 *** [-26.3]	1,028 [0.5]	0.743 ** [-3.2]	0.918 [-0.5]
Tertiary or more	0.140 *** [-65.7]	0.150 *** (-61.35)	1,715 *** [8.8]	0.145 *** [-53.9]	1,599 *** [6.8]	0.152 *** [-14.9]	2,300 * [2.2]
Less than 24 years	1,386 *** [11.8]	1,363 *** (11.31)	0.629 *** [-8.0]	2,274 *** [26.0]	0.737 *** [-4.2]	0.342 *** [-17.2]	0.464 *** [-5.6]
45-55 years	1,065 ** [2.7]	1,081 ** (3.23)	1,341 *** [7.5]	1,120 *** [3.9]	1,271 *** [5.0]	1,336 *** [3.7]	1,530 *** [4.6]
56+ years	1,532 *** [14.9]	1,440 *** (12.14)	2,159 *** [17.4]	1,547 *** [11.7]	1,870 *** [11.2]	2,488 *** [9.8]	2,509 *** [10.3]
Rural	1,074 * [2.5]	0.819 *** (-6.37)	0.99 [-0.2]				
Productive city	0.487 *** [-34.6]	0.501 *** (-32.09)	1,424 *** [9.4]	0.711 *** [-15.2]	1,302 *** [6.7]		
Less productive city	1,186 *** [7.8]	1,220 *** (8.78)	0.868 ** [-3.0]	1,698 *** [23.5]	1,620 *** [13.5]		
Women (second earner)	1,865 *** [24.9]	1,812 *** (22.90)	1,445 *** [8.7]	1,921 *** [20.5]	1,446 *** [7.2]	2,822 *** [10.7]	1,565 *** [4.3]
Women (other)	1,241 *** [10.2]	1,259 *** (10.67)	0.954 [-1.2]	1,348 *** [11.9]	0.961 [-0.9]	0.698 *** [-6.5]	0.916 [-0.8]
Productive sector		0.541 *** (-7.91)					
Less productive sector		3,170 *** (54.18)					
Constant	3,212 *** [38.6]	2,225 *** (25.58)	0.382 *** [-22.9]	1,986 *** [19.4]	0.440 *** [-16.5]	6,697 *** [22.7]	0.347 *** [-8.8]
N	184,247	184,247	60,534	89,119	27,376	19,454	6,856
F	1,031.66	1,033.80	71.46	705.66	43.84	128.31	27.38
df_m	11.00	13.00	11.00	10.00	10.00	8.00	8.00
df_r	184,246	184,246	60,533	89,118	27,375	19,453	6,855

T-statistics in parentheses. *** p<0.01, ** p<0.05, * p<0.1. The odds ratios 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.

Note: See Annex 2 for cities and sectors ranking according to their productivity.

Source: Authors calculations based on GEIH 3rd Quarter 2015.

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;³⁴ 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%³⁵ (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 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.

B. 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.³⁶

³⁴ 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.

³⁵ 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.

³⁶ According to the labour productivity index of Isaza and Rojas (2015).

C. Percentage of Workers with productivity levels well below the minimum cost of hiring them

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 wage³⁷. 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, 49% 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 35% if we restrict the survey to the main 13 metropolitan areas and 65% 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, 50% for workers with middle school or less and 41% 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 than 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.

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

³⁷ Including vacations, transport subsidy, severance 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 not reported, about 10% of the informal workers. The percentage is 44% if we restrict the survey to the main 13 metropolitan areas.

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	49	35	65	48	35	64
Primary or less	59	43	67	58	44	66
Middle school	50	39	64	50	40	63
High school	41	31	61	42	32	60
Tertiary or more	28	23	44	27	22	42

** Labor income imputed when not available.

Source: GEIH.

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.

A. 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 second earners 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 second

earner, 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.

Ethnicity: In the 2005 Colombian census, around 14% of the population declared itself an

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	Percentage difference	Women Head or other)	Male	Percentage difference	Women	Male Head or other)	Percentage difference
Firm definition									
Total	48	51	-6	39	37	5	74	78	-5
Tertiary	22	21	5	19	17	12	33	40	-18
High school	58	48	21	52	38	37	68	66	3
Low	86	77	12	80	64	25	90	84	7
Legal definitio									
Total	51	55	-7	40	40	0	80	81	-1
Tertiary	24	27	-11	20	22	-9	35	43	-19
High school	61	49	24	53	39	36	76	68	12
Low	89	82	9	79	68	16	95	89	7

Source: GEIH.

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.

B. 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.

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). For analyze the impact by level of education, we compare the results among male workers 25 to 50. The total impact among this group was 8 p.p.; among workers with primary education or less was 10 p.p., 13 p.p. among workers with high school studies and not significant 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,³⁹

Table 7
BARRIERS TO FORMALITY

	Labor tax and contributions (% of profit)	Other barriers	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		

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 and Mandatory severance pay. 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.

Source: Doing Business Indicators, 2016. World Bank, OECD and Heritage Foundation, 2016.

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 wages⁴⁰ 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 in-

formality. 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).⁴¹

Other barriers to formality. Other barriers to formality include 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.⁴² When asked about the main constraints, 39% of employees mention firing cost, a percentage that is almost as high as the 44%

³⁹ In particular, it assumes a cashier, age 19, with one year of work experience.

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

⁴¹ 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 a role in reducing inequality.

⁴² 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.

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-cyclical behavior 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 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 49% 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%).

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). In this section we attempt to go a step further in this analysis and to 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.

In fact, the motivations to be informal reported by surveys that we mentioned in Section III help

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 36%	Low 51%	Low 33%	Low 34%	Medium 51%
	Cyclical: coefficient of correlation	-0.42%**	Countercyclical -0.59**	Countercyclical -0.71***	Countercyclical -0.13	No sig.
Subsistence	Percentage of workers earning below half the minimum hiring cost	High 49%	High 35%	High 59%	High 41%	Medium 28%
	Relative probability of being informal, low productivity groups	High	High			
	Informality rate gap (woman that prefer to be formal vs male, 25 to 55 years old)	Medium	Medium	Medium 12%	High 21%	Low 5%
	Impact of reducing payroll taxes (p.p.) Male 25 to 55		Medium -8.0	High -10	High -13	Low (n.s)

Source: Author's estimations.

us to distinguish voluntary from involuntary informality. According to our estimations, about 36% of the informal workers are Voluntary by this indicator. In turn, workers with earnings well below the minimum wage (49% 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, in contrast with those in induced informality.

However, the direct use of the numbers mentioned above may lead to an identification problem

as far as about 13% 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. According to Levy (2009) some workers prefer to be informal in order to do not lose their social benefits. 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). However, failures in the social policy design might not be the only cause of this new type of informality. There are other reasons that make a worker with low productivity to prefer informal jobs like geographical distance to the places where formal jobs are offered (Hausmann, 2014) and lack child care facilities, that restrict women possibility of holding a full type job. We classified the workers in a separated group called mixed informal workers. Table 9 shows the final distribution of our three types of informality.

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 second earners, 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. The fact that males are more likely than females to be in this groups demonstrates that discrimination is not the main factor driving Induced Informality. This result is coherent with the results of Fernandez and Villar (2016) in terms of the tax reform having a higher impact on workers with high school level of education and males.

Table 9
SHARES OF EACH TYPE OF INFORMALITY

	Firm definition			Legal definition		
	Total (%)	13-areas (%)	Rural (%)	Total (%)	13-areas (%)	Rural (%)
Subsistency	36	20	52	38	23	54
Induced	29	29	23	29	30	22
Voluntary	22	36	12	19	31	11
Mixed	13	15	13	14	16	13

Source: GEIH. Authors Calculations.

Table 10A
DETERMINANTS OF EACH TYPE OF INFORMALITY - FIRM DEFINITION

	Subsistence			Induced			Voluntary			Mixed		
	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural
Elementary or less	1,527 *** [12.60]	1,427 *** [7.41]	1,431 *** [5.52]	0,900 ** [-3.03]	1,150 ** [3.08]	0,767 *** [-3.40]	0,671 *** [-10.08]	0,709 *** [-7.34]	0,643 *** [-4.24]	0,984 [-0.37]	0,978 [-0.42]	1,084 [0.80]
High school	0,838 *** [-4.99]	0,774 *** [-5.42]	0,973 [-0.35]	1,258 *** [6.69]	1,247 *** [5.19]	1,280 ** [2.76]	1,229 *** [5.16]	1,221 *** [4.44]	0,897 [-0.83]	0,653 *** [-8.49]	0,674 *** [-6.89]	0,727 * [-2.29]
Tertiary or more	0,431 *** [-17.94]	0,403 *** [-14.75]	0,435 *** [-5.48]	1,075 [1.77]	0,876 ** [-2.69]	1,435 * [2.35]	2,566 *** [21.65]	2,469 *** [18.31]	2,621 *** [5.38]	0,665 *** [-6.62]	0,622 *** [-7.10]	0,893 [-0.52]
Women [second earner]	1,430 *** [13.13]	1,525 *** [11.21]	1,450 *** [6.10]	0,369 *** [-27.54]	0,395 *** [-20.48]	0,404 *** [-10.41]	1,108 ** [2.97]	1,121 ** [2.73]	0,847 [-1.76]	3,649 *** [32.00]	2,566 *** [18.17]	4,696 *** [20.21]
Women [other]	1,065 [1.91]	1,220 *** [4.22]	0,848 ** [-2.72]	0,597 *** [-18.87]	0,672 *** [-11.54]	0,529 *** [-8.25]	0,755 *** [-8.93]	0,808 *** [-5.92]	0,462 *** [-6.62]	2,133 *** [20.31]	1,736 *** [12.26]	2,286 *** [9.62]
Less than 24 years	1,938 *** [19.32]	2,148 *** [16.31]	2,278 *** [12.31]	0,688 *** [-10.51]	0,829 *** [-4.13]	0,734 *** [-3.95]	0,453 *** [-16.12]	0,489 *** [-13.40]	0,187 *** [-9.38]	1,433 *** [6.81]	1,515 *** [6.41]	0,684 *** [-3.36]
45-55 years	0,739 *** [-9.77]	0,717 *** [-7.36]	0,835 ** [-3.10]	0,880 ** [-4.29]	0,815 *** [-5.37]	1,007 [0.10]	1,494 *** [11.71]	1,337 *** [7.28]	1,353 *** [3.44]	1,293 *** [5.75]	1,316 *** [5.06]	1,077 [0.80]
56+ years	0,477 *** [-22.15]	0,446 *** [-15.58]	0,620 *** [-8.36]	0,382 *** [-27.35]	0,299 *** [-25.92]	0,623 *** [-6.35]	2,225 *** [22.05]	1,855 *** [13.91]	1,648 *** [5.82]	3,989 *** [33.61]	3,478 *** [24.03]	2,804 *** [12.33]
Productive city	0,339 *** [-30.58]	0,567 *** [-14.59]		0,728 *** [-10.38]	0,725 *** [-9.82]		2,978 *** [36.15]	1,836 *** [19.20]		1,227 *** [5.25]	1,062 [1.46]	
Less productive city	1,225 *** [7.47]	1,788 *** [18.61]		1,655 *** [18.77]	1,364 *** [10.53]		0,485 *** [-20.15]	0,446 *** [-25.38]		0,444 *** [-17.75]	0,653 *** [-10.76]	
Rural	1,753 *** [19.24]			0,605 *** [-15.13]			0,723 *** [-8.23]			0,978 [-0.54]		
Constant	0,502 *** [-19.88]	0,290 *** [-25.62]	0,834 ** [-2.82]	0,788 *** [-6.84]	0,728 *** [-7.32]	0,466 *** [-9.90]	0,207 *** [-40.05]	0,359 *** [-21.95]	0,213 *** [-15.51]	0,0664 *** [-53.42]	0,0965 *** [-38.37]	0,0708 *** [-25.19]
N	107,889	47,090	15,068	107,889	47,090	15,068	107,889	47,090	15,068	107,889	47,090	15,068
F	349.80	185.30	49.09	265.40	178.50	34.21	406.60	291.00	33.58	253.50	131.30	74.05
df_m	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00
df_r	107,888	47,089	15,067	107,888	47,089	15,067	107,888	47,089	15,067	107,888	47,089	15,067

Source: GEIH and author's calculations.

Table 10B
DETERMINANTS OF EACH TYPE OF INFORMALITY - LEGAL DEFINITION

	Subsistence			Induced			Voluntary			Mixed		
	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural	National	13-areas	Rural
Elementary or less	1,412 *** [10.97]	1,225 *** [4.94]	1,484 *** [5.81]	0,943 [-1.72]	1,128 ** [2.70]	0,787 ** [-3.14]	0,652 *** [-12.63]	0,722 *** [-7.55]	0,811 ** [-2.64]	0,879 ** [-2.94]	0,981 [-0.35]	1,095 [0.88]
High school	0,715 *** [-10.46]	0,665 *** [-10.10]	0,864 [-1.83]	1,317 *** [8.30]	1,364 *** [7.55]	1,230 * [2.39]	0,884 *** [-3.51]	0,906 * [-2.39]	0,791 * [-2.33]	0,640 *** [-8.96]	0,662 *** [-7.22]	0,713 * [-2.49]
Tertiary or more	0,381 *** [-24.47]	0,387 *** [-19.86]	0,364 *** [-7.37]	1,344 *** [7.91]	1,087 [1.85]	1,798 *** [4.19]	1,641 *** [12.91]	1,779 *** [12.86]	1,780 *** [3.65]	0,671 *** [-6.66]	0,656 *** [-6.53]	0,915 [-0.37]
Women (second earner)	2,006 *** [27.41]	1,925 *** [20.24]	2,174 *** [11.74]	0,338 *** [-31.42]	0,347 *** [-24.33]	0,383 *** [-11.38]	1,058 * [2.10]	1,046 [1.37]	0,787 ** [-3.04]	4,171 *** [35.44]	2,991 *** [21.22]	4,947 *** [21.24]
Women (other)	2,163 *** [25.20]	2,070 *** [18.19]	2,120 *** [11.09]	0,597 *** [-19.88]	0,672 *** [-12.20]	0,575 *** [-7.64]	2,619 *** [31.85]	2,301 *** [21.46]	2,705 *** [15.73]	1,850 *** [15.86]	1,807 *** [13.05]	1,431 *** [3.64]
Less than 24 years	1,964 *** [21.37]	1,935 *** [16.52]	2,040 *** [10.13]	0,699 *** [-11.13]	0,748 *** [-7.20]	0,810 ** [-2.84]	0,786 *** [-6.85]	0,836 *** [-4.29]	0,380 *** [-10.80]	1,437 *** [7.33]	1,458 *** [6.35]	0,805 * [-2.00]
45-55 years	0,882 *** [-4.35]	0,925 * [-1.99]	0,901 [-1.73]	0,894 *** [-3.81]	0,774 *** [-6.84]	1,071 [0.99]	1,442 *** [11.80]	1,567 *** [11.88]	1,064 [0.90]	1,226 *** [4.51]	1,325 *** [5.11]	0,998 [-0.02]
56+ years	1,176 *** [5.20]	1,318 *** [6.47]	1,104 [1.61]	0,482 *** [-21.51]	0,364 *** [-22.10]	0,706 *** [-4.84]	2,958 *** [33.14]	3,627 *** [29.91]	1,705 *** [8.06]	2,803 *** [23.59]	3,111 *** [21.53]	1,812 *** [6.83]
Productive city	0,531 *** [-21.88]	0,753 *** [-9.06]		0,813 *** [-7.13]	0,871 *** [-4.39]		2,627 *** [34.50]	1,465 *** [12.68]		1,338 *** [7.31]	0,996 [-0.10]	
Less productive city	0,914 *** [-3.58]	1,311 *** [9.87]		1,595 *** [18.39]	1,424 *** [12.60]		0,441 *** [-26.65]	0,410 *** [-31.31]		0,517 *** [-14.55]	0,645 *** [-11.19]	
Rural	1,775 ** [19.96]			0,594 *** [-16.29]			0,881 *** [-4.03]			1,105 * [2.37]		
Constant	0,618 *** [-15.04]	0,456 *** [-18.94]	1,024 [0.36]	0,813 *** [-6.20]	0,736 *** [-7.33]	0,472 *** [-9.94]	0,290 *** [-35.65]	0,504 *** [-16.17]	0,314 *** [-14.83]	0,0650 *** [-54.74]	0,0915 *** [-41.02]	0,0771 *** [-24.65]
N	114,240	49,388	15,670	114,240	49,388	15,670	114,240	49,388	15,670	114,240	49,388	15,670
F	348.10	149.10	48.69	259.40	160.10	32.47	467.90	355.10	75.33	200.70	124.80	67.66
df_m	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00	11.00	10.00	8.00
df_r	114,239	49,387	15,669	114,239	49,387	15,669	114,239	49,387	15,669	114,239	49,387	15,669

Source: GEIH and author's calculations.

- ***Subsistence informality***: Workers with primary education, women, young fellows (less than 24 years old) and workers living in rural areas are more likely to be part of these groups. This also is coherent with our previous results.
- ***Mixed informality***: Among this group we found a prevalence of women (particularly those registered as second earners), young and relatively older workers. Given that those are precisely the groups that receive more benefits from the state, it is highly probable that the reason behind this type of informality lies in failures in the design of the social benefits policy.

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. Workers with mixed informality (both subsistence and voluntary informality) tend to have middle school studies. 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. Mixed informality seem to be more related to failures in the benefits policies.

Bibliografía

- 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 Consultoría (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 School.
- Farné S., Rodríguez D. y P. Rios (2015). Impacto de los subsidios estatales sobre el mercado laboral en Colombia. Universidad del Rosario, Bogotá.
- 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.
- Hausmann R. (2014). "The Logic of the Informal Economy". Article published by Project Syndicate.
- 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.
- Hausman (2015)
- Isaza and Rojas (2015) Diagnóstico de la heterogeneidad estructural y el desempeño macroeconómico en Colombia on Macroeconomía 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.
- Levy, S. (2008). *Good Intentions, Bad Outcomes: Social Policy, Informality, and Economic Growth in Mexico*. Brookings Institution Press, Washington D.C.
- 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.

Annex A

EQUATION FOR PREDICTION OF PREFERENCES FOR INFORMAL JOBS

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 were included, and the variables included in the regression were established to obtain the best possible prediction for 2015 voluntary informal workers.

(t-statistics in parenthesis)	Raw	Clear		Raw	Clear
Elementary or less	0.904 *	0.854 ***	Cartagena	0.251 ***	0.252 ***
	[-2.2]	[-4.2]		[-19.4]	[-20.6]
High school	1.083		Monteria	1,377 ***	1,387 ***
	[1.6]			[5.4]	[6.0]
Tertiary or more	1,427 ***	1,373 ***	Villavicencio	2,364 ***	2,384 ***
	[5.6]	[5.6]		[14.0]	[15.4]
Women (second earner)	1,291 ***	1,300 ***	Pasto	2,032 ***	2,057 ***
	[5.3]	[5.9]		[10.8]	[11.7]
Women (other)	0.972		Cúcuta	1,685 ***	1,699 ***
	[-0.7]			[8.5]	[9.4]
Less than 24 years	0.740 ***	0.733 ***	Pereira	1,898 ***	1,906 ***
	[-5.0]	[-5.2]		[9.7]	[10.4]
45-55 years	1,241 ***	1,237 ***	Bucaramanga	1,231 ***	
	[5.3]	[5.2]		[3.3]	
56+ years	1,799 ***	1,789 ***	Ibague	1,458 ***	1,465 ***
	[13.2]	[13.1]		[5.9]	[6.4]
Rural	0.935		Cali	1,872 ***	1,884 ***
	[-1.1]			[10.4]	[11.4]
Job satisfaction	2,672 ***	2,681 ***	Tunja	1.073	
	[27.8]	[28.0]		[0.9]	
Self-employment	0.445 ***	0.445 ***	Florencia	1,967 ***	1,984 ***
	[-14.7]	[-14.8]		[10.4]	[11.4]
Agriculture	1,275 **	1,242 ***	Popayan	0.453 ***	0.456 ***
	[3.3]	[3.9]		[-11.5]	[-12.1]
Mines	0.422 **		Valledupar	0.305 ***	0.307 ***
	[-3.2]			[-16.4]	[-17.3]
Manufacturing	1,491 ***	1,503 ***	Quibdo	0.462 ***	0.456 ***
	[6.7]	[7.8]		[-6.6]	[-6.9]
Public services	2,975		Neiva	0.789 ***	0.794 ***
	[1.6]			[-3.8]	[-4.0]
Retail, restaurants and hotels	1,469 ***	1,491 ***	Riohacha	0.749 ***	0.757 ***
	[7.9]	[10.4]		[-4.0]	[-4.1]
Constructions	0.957		Santa Marta	0.770 ***	0.774 ***
	[-0.5]			[-4.4]	[-4.7]
Transportation and	1.007		Armenia	1,055	
	[0.1]			[0.8]	
Medellín	1,913 ***	1,924 **	Sincelejo	0.381 ***	0.384 ***
	[11.3]	[12.5]		[-13.1]	[-13.8]
Barranquilla	0.578 ***	0.583 ***	Constant	0.360 ***	0.368 ***
	[-9.3]	[-10.1]		[-11.6]	[-14.1]
Bogotá	1.542 ***	1.553 ***			
	[7.0]	[7.7]			
Number of obs	64,098	64,098	Design df	40	30
	118	156		64,097	64,097

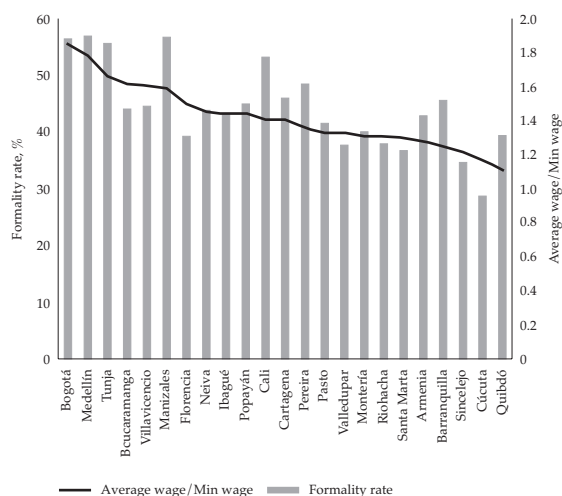
Annex B

CITIES' PRODUCTIVITY VS FORMALITY RATES

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, Medellín, Bucaramanga and Villavicencio are the most productive cities, while Pasto, Montería, Barranquilla and Cucutá are the least productive areas. In 2007 the most productive cities: Medellín Barranquilla y Bogota; the non-productive cities: Tunja, Quibdo and Armenia and in the 13 areas: the most productive cities: Medellín Barranquilla y Bogota, and the non-productive cities: Villavicencio, Pasto, Cucuta and Pereira.

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.



Source: Authors calculations based on GEIH 3rd Quarter 2015.

Table 1
Labour productivity Index by economic sector

	1950	1968	2000	2013
High productivity				
Mining	493	607	702	394
Domiciliary public services	595	1,376	1,388	1,241
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).