Labor Market Flows of Young Workers in Colombia*

Nicolás Mancera †

June, 2021

Abstract

Identifying the effect of policies focused on transitions towards formality of young adult workers is important to understand the effectiveness of labor policy instruments. This research article examines how payroll tax cuts affect labor market flows of young workers in Colombia through the analysis of the First Job Act (Law 1429 of 2010), which encouraged hiring of workers under 28 years of age. I present a conceptual framework, based on the on-the-job search model of Narita (2020), to explain channels through which a payroll tax cuts affect labor transitions in developing economies. Then, I use data from the Colombian National Labor Survey (Gran Encuesta Integrada de Hogares) to assess the effect of a reduction in payroll taxes, taking advantage of the age specific eligibility cutoff (28 years of age or younger). The main result shows a positive effect of 2,8 percentage points in fostering transitions into formal employment of workers between 20 and 28 years. The effect is larger for men in the same age group and individuals with less school attainment. Finally, I present evidence of positive effects in flows between other employment states as unemployment to self-employment and out of labor force to formal employment. These results suggest that payroll tax reductions have a positive effect on the flow of young workers into formal employment.

Keywords: Labor market flows, Payroll taxes, On-the-job search model

JEL Codes: J08, J32, J64, J68

*I am grateful to Juan Camilo Chaparro and Eduardo Lora for their helpful feedback, support, and encouragement.

†Msc. student in Economics at the University EAFIT. Email: jnmanceraa@eafit.edu.co

1 Introduction

In labor economic research, many studies have been interested in how individuals change between labor states and which elements explain that change (Dube et al., 2016; Elsby et al., 2010). In countries with high informality rates, it is urgent to formulate public policies that increase transitions into labor formality, since formal employment offers better conditions to workers in terms of social security and income stability. Efforts in this direction use policy instruments such as payroll tax cuts creating incentives from the demand side of the labor market to hire more formal workers. However, it is difficult to know if such policies accomplish their desired goals. One example is the First Job Act ("*Ley de Primer Empleo.* Ley 1429 de 2010") in Colombia, which gave benefits to hire young employees and encourage the registration of small firms. The law established ideal conditions to evaluate the effects on labor market flows of payroll tax reductions under sensible identification assumptions.

Building on theoretical labor search and empirical regression discontinuity design literature, I exploit the discontinuity entailed by the First Job Act on workers above and below 28 years of age. In particular, I use the model of on-the-job search proposed by Narita (2020) to describe dynamic labor market flows and how a change in payroll taxes can affect such flows. I extend the model of Narita (2020) by including an additional labor state (out of the labor force). Then, using data from the Colombian National Labor survey, I compare changes in transition rates of individuals just below the age cutoff with changes in transition rates for individuals who are just above the age cutoff. I find that this policy affects positively the transition rates from informality to formality of young individuals. The estimates are robust to specifications allowing for time fixed effects and they are robust to changes in different bandwidth specifications. The effect is stronger for men and for individuals with a high school degree or less.

This research paper contributes to a body of research on formalization policies targeted towards young workers and labor market flows. On one hand, studies in labor market flows have analyzed the dynamics of changes between labor states and the labor matching process.

On the other hand, studies on the effects of payroll taxes have reviewed the incidence of this change on employment and wages. Fundamental for my research paper is the previous work by Narita (2020), which presents an on-the-job search model to show dynamic responses to changes in payroll taxes. My research paper also contributes to this literature using microdata from household surveys, to provide an analysis on how retrospective questions can be used to distinguish labor flows with cross section data and when it is not possible to follow individuals across time.

My research paper is closely related to the study of Becerra-Camargo (2016) and Galiani et al. (2017). The paper by Galiani and coauthors (2017) finds that reductions on formalization cost do not have long term effects on firm's decisions regarding entry into formality, even if labor institutions encourage firms to participate in meetings to know benefits of tax reductions. On the other hand, the paper by Becerra-Camargo (2016) studies labor demand responses instead of flow changes into labor states. Also, his paper uses data from administrative records of formal employment allowing to follow individuals in different periods. However, his study excludes from the analysis the response on unemployment, inactivity and self-employment.

The remainder of this research paper is organized as follows. Section 2 provides a brief the literature review of labor market flows and the study of flows in young adults. Section 3 presents the institutional background related to the First Job Act. In section 4, I describe the data used in the identification strategy and include and analysis of labor market flows for the Colombian population. Section 5, presents the conceptual framework of labor market flows to investigate channels through payroll tax changes affect labor flows. Section 6 describes the identification strategy to estimate the effect of payroll taxes reductions in labor market flows of young adults. Section 7 reports the main findings and robustness checks. Finally, section 8 concludes.

2 Literature Review

The study of labor market flows has a long standing in economics beginning with the seminal paper of Kaitz (1970) with the analysis of unemployment spans. This literature has presented different views of labor flows. On the one hand, Blanchard and Diamond (1992) present the most traditional view of labor market flows, where much of this view explores cyclical behavior of labor market flows through different states as employment, unemployment and out of the labor force. A similar view separates measures of workers flows and measures of job flows to study their relation with employed and unemployed stocks (Davis and Haltiwanger, 1998). On the other hand, there is a view of labor flows in the context of labor search models (Mortensen and Pissarides, 1994; Pissarides, 2000; Burdett and Mortensen, 1998), that consider job creation and job destruction as a result of matching process between unemployed workers and firms with available jobs or vacancies, where wages are a result of a bargaining process between workers and employers. There are also further studies that addresses the study of flows with the second view (Shimer, 2007; Hagedorn and Manovskii, 2008; Shimer, 2012).

In this paper, I study labor flows in the context of labor search models, specifically, following the conceptual groundings of Burdett and Mortensen (1998) and Narita (2020). Under this view, there exists a dispersed wage distribution in the equilibrium of a model of labor search which is explained by market frictions even if workers and jobs are homogeneous. From this starting point, literature has developed equilibrium search models that account for several facts of labor markets as heterogeneous employment status. For instance, the work of Van den Berg and Ridder (1998) shows a dispersed equilibrium wage distribution as a result of endogenous decisions of job seekers and hiring firms and this result explains job and unemployment spells. In addition, Bontemps et al. (2000) also propose a model of dispersed equilibrium wage distribution, but allowing for a continuous distribution of firms' productivity.

More recent papers use the model of Burdett and Mortensen (1998), including extensions

to capture particular facts about labor market flows and the determination of employment stocks (Meghir et al., 2015). These extensions range from expanding labor states to including human capital as a determinant of different types of workers. In this context, I use the conceptual framework presented by Narita (2020) to describe how changes in payroll taxes can affect flows into labor market flows. Following the model of Burdett and Mortensen (1998), Narita (2020) estimates a life-cycle job search model for Brazil with three employment states (formal employment, informal employment and self-employment) and unemployment state, accounting for heterogeneous worker ability and payroll taxes. Therefore, it seems appropriate to use this work for the following reasons: (1) the model focuses on Latin America; (2) there are several labor states including self-employment and (3) This work uses data from labor surveys to validate empirically its model.

In the context of Latin American countries, most studies of labor market flows have been carried on from the empirical view of Blanchard and Diamond (1992) and Davis and Haltiwanger (1998). Two of relevant works are the papers of Bosch and Maloney (2006) and (Bosch et al., 2007), which characterize job flow dynamics in Brazil and Mexico using micro-data from labor market and household surveys. These studies reached three conclusions: First, the unemployment rate is countercyclical, essentially because job separations of informal workers increase dramatically in recessions. Second, the share of formal employment is countercyclical because of the difficulty of finding formal jobs from inactivity, unemployment and other informal jobs during recessions rather than because of increased separation from formal jobs. Third, flows from formality into informality are not countercyclical, but, if anything, pro-cyclical. From the view of labor search models, Meghir et al. (2015) include informality in an equilibrium wage-posting model too and find that markets with search frictions in the informal sector have important detrimental effects by endogenously segmenting the labor market, it reduces the competition for workers and makes it harder for workers to relocate to higher productivity firms.

In Colombia, there are few papers that study labor market flows (Prada, 2012; Lasso-

Valderrama, 2012; López Castaño and Lasso Valderrama, 2015; Flórez et al., 2018). Specially, Lasso-Valderrama (2020) explains the flow dynamics of Colombian workers between salaried, non-salaried, unemployed and inactive statuses. He founds that flows from inactivity to formal jobs are larger than flows from unemployment to formal jobs even if the probability of the former is lesser than the later and possible explanations are higher labor incomes or falls in reservation wages. Since this paper focuses on the transitions of young adult workers, literature concerning them gives little information about the labor flows of younger workers who are entering the labor market and how those transitions respond in the face of labor formalization policies. Instead, there is a substantive literature which aims to explain returns to young adults from experiences during childhood (Hotz et al., 2002; Bai and Wang, 2020; Beam et al., 2020) without the description of the current and future movements into the labor market. In this context, there is identification summary of first job policies made by the International Labor Organization (ILO, 2015) which includes a complete description of these policies in Latin American countries. In Colombia, one of the main policies identified by the ILO is the First Job Act (Law 1429 of 2010).

Although there is little evidence on the labor market flows effects of formalization laws, the effect of changes in payroll taxes on employment and wages has been an important research question in developing countries. However, there is no consensus on the answer to that question and it depends in the level of labor market frictions that affects the wage' pass-through mechanism of the tax change (Kugler and Kugler, 2009). For instance, some studies of this relationship report that payroll taxes reforms increase total employment by between 0.3% to 0.5% and formal employment by between 3.4% to 3.7% (Antón, 2014), meaning that there is a movement between labor states driven by formalization laws. Also, Galiani et al. (2017) study the effect of Colombian First Job Act in the probability that a firm becomes formal, they show that in some scenarios the probability that a firm operates in formality increase by 5.5 percentage points and also this effect does not persist over time. In addition, the study of Becerra-Camargo (2016) analyzes labor demand responses to payroll

taxes in Colombia, through the same law of Galiani et al. (2017), and found that a reduction of payroll taxes increased formal-sector demand for young workers by 3.38%.

The main conclusion of the literature is that, while labor flows have been extensively studied through labor search models, it is not clear whether changes in payroll taxes can affect the labor flows of young adults in a context of labor flows. Instead, literature focuses on the description of flows through their relationship with the business cycle, in particular, using employment and unemployment and informality as labor states. In the context of policy evaluation literature, changes in payroll taxes could affect the stock of formal employment and the hiring behavior of firms but there is not conclusion about the labor force sources that change the formality level. To my knowledge, no study has explored the effects of payroll taxes reduction on labor market flows for young adults, therefore, this is the slot that this research wants to fill.

3 Institutional Background

The encouragement to increase formal jobs is foremost for labor ministries, mainly in Latin American countries where high levels of informality and unemployment of young adults persist. With this background, Latin American countries such as Chile and Colombia have applied focalized reductions in payroll taxes to ensure better conditions to both employers and workers in order to increase formal employment, wages and the quality of jobs. In 2010, the Colombian government proposed the First Job Act (Law 1429 of 2010 known as "Ley del Primer Empleo") which came into effect in 2011 with the aim of generating new formal jobs, formalize current informal jobs and increase incomes of disadvantaged groups (Ministerio de Trabajo, 2012).

In Colombia, monthly payroll taxes finance public vocational training institutions, the healthcare system and pension funds. In particular, payroll taxes represent 11% of the salary

paid to each worker where 2% goes to SENA[1], 3% to ICBF[2], 4% to Family Compensation Funds, 1.5 to Health Funds and 0.5% to Pension and retirement funds, plus variable contributions depending the contribution base income of each worker where 12.5% goes to health and 16% goes to pension payments. Articles 9, 10, and 11 of the First Job Act dictate that formal firms that operate before December 29 of 2010 can discount from firm income tax, the value of payroll taxes of SENA, ICBF, Family Compensation Funds, Health Funds and Pension funds paid for each new hired worked that belongs to one of the following disadvantaged groups:

- Workers younger than 28 years old.
- Workers in forced displacement condition.
- Women above 40 years old and without formal employment in the past twelve months.
- Heads of household eligible for social assistance programs.
- Workers with a 1.5 Colombian minimum wage and without previous experience in the formal sector.

The First Job Act applied for firms that hire workers between January 2011 and December 2014 and the discounts were effective until two years after the worker was hired and if a worker meets more than two criteria, the discount only applies for one of these. Also, firms must comply the following conditions:

- The employer must increase the number of employees in relation to the number of employees that contribute to social security in December to the previous year.
- The employer must increase the total payroll value of the firm in relation to the total payroll value in December to the previous year.

[1]Servicio Nacional de Aprendizaje (SENA): Colombian institute for vocational employment training
[2]Instituto Colombiano de Bienestar Familiar (ICBF): Colombian institute for children protection.

- Payroll taxes that are going to be discounted were fully paid.
- The new worker must not to be hired to replace an old worker. (This means that hiring requires new jobs).
- New jobs do not result from firm's merge.

Since the focus of this dissertation is the group of young workers, I use the condition for new workers younger than 28 years old to estimate the effect of the First Job act on worker flows. Besides benefits for new hires, the First Job Act includes corporate taxes and registration fees discounts to new formal firms established after 2010, however, I do not estimate the effect on labor market flows of discounts to new formal firms.

4 Data

To study the labor market flows of young workers in Colombia, I use monthly data from the Colombian National Labor Survey (Gran Encuesta Integrada de Hogares, GEIH) which is conducted by the Department of National Statistics of Colombia (DANE). This survey has been carried on since January of 2007 and contains information of individual characteristics, household characteristics and labor force components. DANE surveys around 70.000 households very month that represents the entire Colombian population. For the purpose of this dissertation, I keep individuals between 20 and 36 years old that sums a total of 2.505.858 individuals that appear once between 2008 and 2019 Moreover, I keep information of age, gender, schooling level, current and retrospective questions of labor market and social security information.

The GEIH is a cross-section dataset that surveys different individuals each month, therefore, I cannot follow individuals from one month to another which is a limitation to construct flows between states. To overcome this limitation, I use retrospective questions to know individual labor states twelve months in the past. In this sample, people can be unemployed,

formally employed, informally employed, self-employed, or out of the labor force [3]. Here, formal employment differs from the DANE definition of formality due to the lack of retrospective questions on firm size of previous employers. Hence, I define formality according to the type of work. Formal employment groups individuals who work in private and public firms with more than one worker, and informal employment holds individuals that work in firms with more than one worker different from private and public firms.[4] Self-employment considers individuals that make their own contributions to social security and work in a firm with only one worker. Finally, inactivity corresponds to individuals who are out of the labor force. At the time of the interview, I observe one of the five states of each individual and define its status twelve months ago with retrospective questions.[5]

To estimate the empirical strategy, I construct four age subgroups divided into 20 to 24 years, 25 to 28 years, 29 to 32 years, and 33 to 36 years. Also, I categorize the education groups into no schooling, primary, secondary, high school, vocational education, college degree, and graduate degree. Table 1, displays a description of the stock of individuals in every state and education level, also divided by gender. Over the estimation period, there is a large quantity of formal employment compared with the other states. Also, there are more formal male employees than women even with more women in the total sample. The second stock of employment is self-employment with 647.412 individuals that is almost three times larger than informal employment and accounts for the difference in the informality definition in this paper with respect to other studies. In the case of inactivity, the larger number of female workers in this state highlights the distinctive pattern of labor markets in developing economies. Additionally, the sample shows that most individuals completed high school level that corresponds to eleven or twelve grade in the Colombian educational system and the smallest category is the graduate degree group.

In Table 2, there are additional descriptive statistics by work status, including information

[3]Also, there are other states as migration states and dead flows that I do not consider in this analysis

[4]This includes the following type of worker: Domestic employee, unpaid family worker, unpaid worker in companies or businesses of other households, day laborer and other type of employment.

[5]In Appendix 8, I present the questions used to define labor states.

Table 1: Total number of observations between 2008 - 2019

	All	Gender	
		Male	Female
Number of individuals	2.505.858	1.164.251	1.341.607
Employment Formal	923.454	515.728	407.726
Informal employment	130.752	37.946	92.806
Self-employment	647.412	377.678	269.734
Unemployment	303.595	118.724	184.871
Inactivity	500.645	114.175	386.470
Education			
None	127.419	66.181	61.238
Primary	398.543	204.978	193.565
Secondary	147.811	72.642	75.169
High School	1.181.880	563.912	617.968
Technical	353.772	138.584	215.188
University	246.782	97.686	149.096
Post University	49.651	20.268	29.383

Notes: This table shows the total number of individuals that appears once on a month between 2008 and 2009. Labor states correspond to the states analyzed in this dissertation. *Source*: GEIH, own calculations.

of labor income and information related to social security contributions. The final sample includes only individuals between 20 and 36 years old, there are no big differences between states in age. However, individuals out of the labor force (inactivity state) are younger on average than individuals in other states, suggesting that these individuals have not finished schooling. As mentioned before, there is a big proportion of individuals who have completed their high school studies in comparison to other states. Except for informality and inactivity states the second educational level completed in every state is vocational education level, showing that informal workers can be associated with a less productive labor force. In terms of labor incomes, formal employment incomes are bigger than informal employment and self-employment states, as expected, but an interesting result is that there is more dispersion in informal workers' labor income. Despite self-employment is considered in many cases as informal workers that do not contribute to social security in this sample there is a big proportion of workers that contribute to social security compared to the proportion of

Table 2: Summary of individual data

	Formal Employment	Informal Employment	Self-Employment	Unemployment	Inactivity
Age (mean)	28,07	27,52	28,52	26,44	25,75
Schooling					
None	0,86	0,67	1,93	0,39	1,24
Primary	3,76	1,53	5,69	1,56	3,37
Secondary	1,54	0,42	1,90	0,68	1,35
High School	16,82	2,13	10,64	5,82	11,75
Vocational education	7,38	0,35	2,73	2,03	1,63
College degree	5,24	0,12	2,42	1,49	0,58
Graduate degree	1,25	0,01	0,53	0,15	0,05
% Females	16,27	3,70	10,76	7,38	15,42
Labor Income					
Mean	$ 1.695.660	$ 581.253	$ 671.182	-	-
Std.dev	$ 2.602.678	$ 3.587.242	$ 915.771	-	-
% Social Security	34,41	4,39	21,48	9,00	17,63

Notes: This table shows composition and descriptive statistics of each state. Labor states correspond to the states analyzed in this dissertation. *Source*: GEIH, own calculations.

informal, unemployed, and inactive individuals.

4.1 Bi-proportional matrix balancing (RAS method)

Given the number of observations and the data quality presented above, an immediate question is whether all analysis constructed in the empirical setting can be generalized to the entire Colombian population using expansion factors. In this section, I investigate whether aggregated data has the same properties of the individual data in terms of labor market flows of the young workers. Although there is good reason to think that GEIH data represent the entire population, it is not true in the construction of labor market flows due to measurement errors in retrospective questions combined with expansion factors. Instead, to study the behavior of labor market flows for Colombia in aggregated terms, I construct a transition matrix in which each cell represents the number of flows between an origin state in $t-1$ and destination state in t.

Thus, to solve the problem of empty cells in a transition matrix and measurement errors, I use a statistical procedure called bi-proportional matrix balancing or RAS method (Lahr and De Mesnard, 2004) which consist of an algorithm of matrix scaling, in which the method uses an initial matrix and a particular margins targets and through row and column scaling sequentially it finds a final matrix with the target margins. Formally, Consider two rectangular matrices, an 'initial' matrix $\boldsymbol{Z}$ of dimension n by m where $a_{ij} \in \mathbb{R}^+$ and an unknown 'target' matrix $\boldsymbol{Z}^*$ with the same dimensions and other mathematical properties. Only the margins of the final matrix are known. The problem is to find a third matrix $\widetilde{\bar{\boldsymbol{Z}}}$ not only with the same dimension and mathematical properties as $\boldsymbol{Z}^*$ but also with the same known margin.

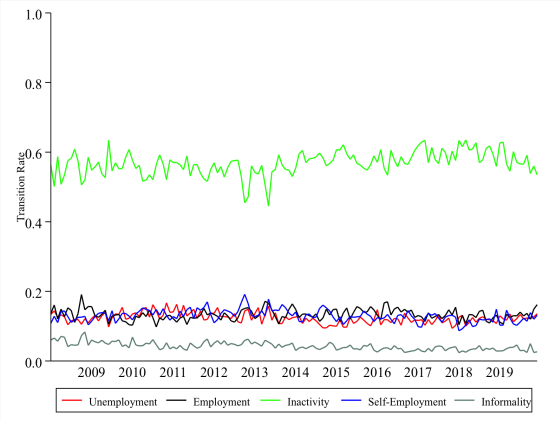
$$\sum_{i=1}^{n} \bar{z}^*_{ij} = \sum_{i=1}^{n} \widetilde{\bar{z}_{ij}} \quad and \quad \sum_{j=1}^{m} \bar{z}^*_{ij} = \sum_{j=1}^{m} \widetilde{\bar{z}_{ij}} \tag{1}$$

Where lowercase letters indicate margins and i and j are rows and columns, respectively. In this way, the objective of this problem is to obtain a $\widetilde{\bar{\boldsymbol{Z}}}$ that is as close as possible to $\boldsymbol{Z}^*$. The Bi-proportional matrix implementation allows me to complete the whole matrix of transitions of the Colombian population and consequently the young adult's population. As an example, consider transitions from unemployment (U_{t-1}) to formal employment (E_t) and define the number of persons that change from U_{t-1} to E_t as $U_{t-1}E_t$ in the year t. With this notation the flow transition rate is,

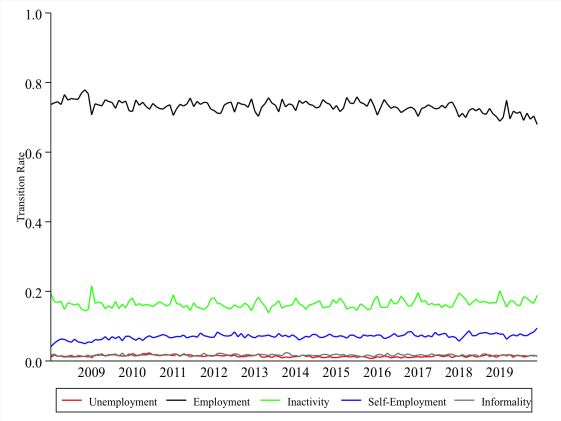
$$\lambda_t^{UE} = \frac{U_{t-1}E_t}{U_{t-1}} \tag{2}$$

Figure 1, shows time series of the transition rates for Colombian young adults aggregated with expansion factors and corrected by the RAS method. This figure includes five panels, each panel shows transition rates to five states. The origin transition rates are Out of labor force (panel a), formal employment (panel b), informal employment (panel c), self-employment (panel d) and unemployment (panel e).

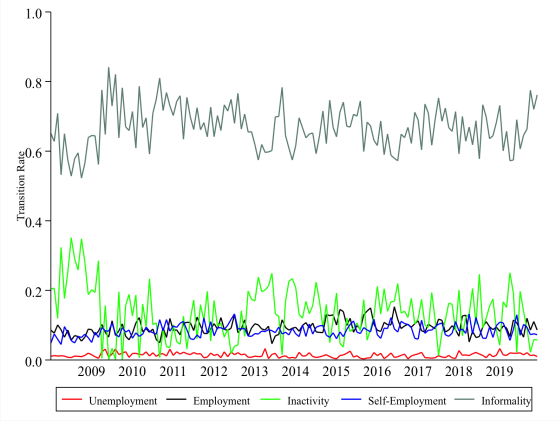
Figure 1: Transition rates for young adults in Colombia (January 2008 – December 20219)



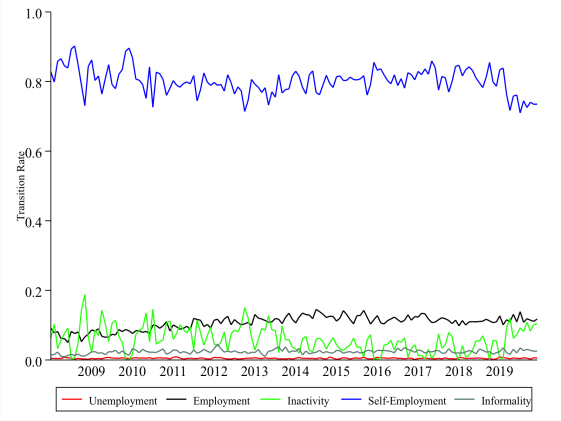
(a) Inactivity



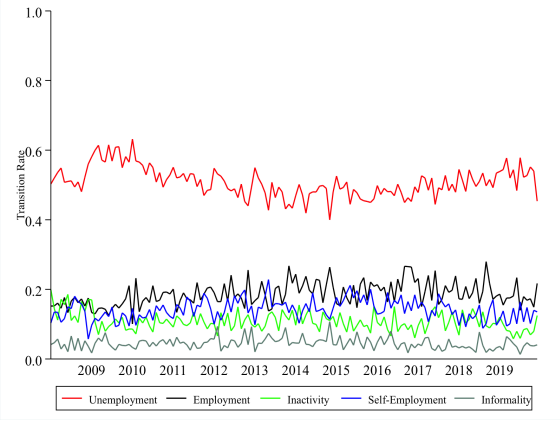
(b) Formal Employment



(c) Informal Employment



(d) Self-Employment



(e) Unemployment

This figure serves to understand the Colombian labor market dynamics and its behavior in the last decade. In Figure 1, the flows to the same state have bigger probabilities that change to another state. Also, in most cases, transition rates are stable across the entire period, however, rates with origin informal employment (panel c) and unemployment (panel e) present more variation relative to other states. Finally, I do not include the transition rates plot of other defined states as migration or deaths due that states are not studied in this dissertation.

5 Conceptual Framework

There is a limited literature that examines the effect of changes of payroll taxes on labor market flows. However, the starting point to study of labor market flows in a context to labor search is the model of Burdett and Mortensen (1998) because its structure allows to include heterogeneity across workers and jobs. This heterogeneity makes it possible to include other states of employment as informality or include self-employment as an extension of informal jobs. The framework of Narita (2020), presented in the literature section, makes it possible to study worker flows in a context of search equilibrium approach and estimate the effect of changes in payroll taxes in these flows. This framework is also important because it includes self-employment and heterogeneity in self-employment ability of workers. Hence, the following conceptual framework follows closely the work of Narita (2020), but I aggregate small changes in the model to include out of labor force state, which is a state that can be defined using data from the Colombian National Survey.

The model of Narita (2020) starts by defining potential work experience ϵ which has a discrete support and is defined in terms of educational achievement. Also, time is defined discrete and ϵ can take integer values in $\{1,\ 2,\ \ldots, E\}$ where E is exogenously determined by the age of retirement. Then, she defines the following labor states:

- Unemployment (state 0)

- Formal employment (state 1)
- Informal employment (state 2)
- Self-employment (state 3)
- Out of labor force state (state 4)

in any point in time t, I include out of labor force (state 4) as a possible state. Each worker has self-employment ability (θ) which is an endogenous parameter and the probability to change from one state to another depends on it. In particular, the probability to change from states $i \in \{1,\ 2,\ 3,\ 4\}$ to unemployment, for a worker with ϵ labor experience, is δ_{ϵ}^{i}. When a worker is unemployed, he receives offers from formal and informal firms, therefore, the probability to receive offers from employment states is λ_{ϵ}^{0j} with $j \in \{1,\ 2\}$. The probability of becoming self-employed is denoted by λ_{ϵ}^{k3} where $k \in \{0,\ 1,\ 2,\ 4\}$. Finally, a worker who is self-employed or out labor force can change to employment j with the probabilities λ_{ϵ}^{3j} and λ_{ϵ}^{4j} respectively.

One particularity of this model is the consideration of worker experience ϵ as the time index instead of taking the time t. According to Narita (2020) individuals have static expectations, consequently, their best guesses about future parameters are the same parameters for older individuals, namely, individuals in $\epsilon+1$. Thus, the time index t is omitted from the analysis. The next step in the model is the definition of incomes of every state. First, when and individual is unemployed, he receives and non-labor income $b_{\epsilon}(\theta)$ which is interpreted as unemployment benefits. Second, employees in formal or informal states receive wage w from a sampling distribution of wages $F^{j}(w)$. Third, an individual in self-employment generates $\pi_{\epsilon}(\theta)$ from the specific self-employment activity. Finally, individuals who are out of labor force creates a home production value $h_{\epsilon}(\theta)$. At first sight, incomes in every state are monetary values, despite the value of home production $h_{\epsilon}(\theta)$, nevertheless, this home production income could be valued in monetary terms with the basis of household sector is essential

to understanding different labor market outcomes (Bridgman et al., 2012; Bridgman, 2016) (Bridgman et al., 2012; Bridgman 2016)

The model of Narita (2020) defines the expected discounted lifetime income when a worker is unemployed as $U_\epsilon(\theta)$ with experience ϵ and self-employment ability θ. In the same way, expected discounted lifetime income when the worker is formal employee or informal employee is $W_\epsilon^j(w,\theta)$ with $j \in \{1,\ 2\ \}$ and when he is self-employed as W_ϵ^3. Finally, taking into account the addition of out of the labor force state, the discounted lifetime income is $H_\epsilon(\theta)$. Now, with the terms defined, there is a reservation wage policy (i.e., a sequence of decision rules) which specify the offers at which individuals are indifferent between accepting the offer from one of the employment states j with $j \in \{1,2\}$ and remaining unemployed, inactive or self-employed. To ensure the existence of reservation wages, $W_\epsilon^1(w,\theta)$ and $W_\epsilon^2(w,\theta)$ must be increasing in w and also $U_\epsilon(\theta)$ and $H_\epsilon(\theta)$ must be independent of w. Then, there exist unique reservation wages $R_{\epsilon+1}^{01}(\theta)$ and $R_{\epsilon+1}^{02}(\theta)$ for offers arriving from the formal and informal sectors respectively when the individual is unemployed. In the same manner, there exist unique reservation wages $R_{\epsilon+1}^{41}(\theta)$ and $R_{\epsilon+1}^{42}(\theta)$ for offers arriving from the formal and informal sectors respectively when the individual is inactive. And it is true for offers from the formal and informal sectors when the individual is self-employed, in the case I denote these wages $R_{\epsilon+1}^{31}(\theta)$ for offers arriving from the formal sector and $R_{\epsilon+1}^{32}(\theta)$. The reservation wages associated to the present value of unemployment in $\epsilon+1$ are:

$$U_{\epsilon+1}(\theta) = W_{\epsilon+1}^1\left(R_{\epsilon+1}^{01}(\theta),\theta\right) = W_{\epsilon+1}^2\left(R_{\epsilon+1}^{02}(\theta),\theta\right) \tag{3}$$

For the inactivity state, reservation wages for offers arriving from the employment states j, must satisfy:

$$H_{\epsilon+1}(\theta) = W_{\epsilon+1}^1\left(R_{\epsilon+1}^{41}(\theta),\theta\right) = W_{\epsilon+1}^2\left(R_{\epsilon+1}^{42}(\theta),\theta\right) \tag{4}$$

Finally, when an individual is self-employed the reservation wage from the two employment

states must satisfy:

$$W^3_{\epsilon+1}(\theta) = W^1_{\epsilon+1}\left(R^{31}_{\epsilon+1}(\theta), \theta\right) = W^2_{\epsilon+1}\left(R^{32}_{\epsilon+1}(\theta), \theta\right) \tag{5}$$

Since this is an on-the-job model, individuals can receive offers even when they are employed in formal and informal sectors, so it is necessary to define reservation wages at which individuals are indifferent between formal employment and informal employment when she is employed in formal sector or when she is employed in informal Narita (2020). Therefore, these reservation wages must satisfy:

$$W^1_{\epsilon+1}(w, \theta) = W_{\epsilon+1}2\left(R^{12}_{\epsilon+1}, \theta\right) \tag{6}$$

when he is employed in the formal sector, and

$$W^2_{\epsilon+1}(w, \theta) = W_{\epsilon+1}1\left(R^{21}_{\epsilon+1}, \theta\right) \tag{7}$$

when he is employed in the informal sector. Now, with the reservation wage policies defined I describe the Bellman equation for each state. The Bellman equation for unemployment is:

$$\begin{aligned}(1+r)U_\epsilon(\theta) = b_\epsilon(\theta) &+ \lambda^{01}_\epsilon(\theta)\int_{R^{01}_{\epsilon+1}}^{\bar{w^1}} W^1_{\epsilon+1}(\chi, \theta)\, dF^1(\chi) \\ &+ \lambda^{02}_\epsilon(\theta)\int_{R^{02}_{\epsilon+1}}^{\bar{w^2}} W^2_{\epsilon+1}(\chi, \theta)\, dF^2(\chi) \\ &+ \lambda^{03}_\epsilon(\theta)\, 1\{W^3_{\epsilon+1}(\theta) > U_{\epsilon+1}(\theta)\} W^3_{\epsilon+1}(\theta) \\ &+ \lambda^{04}_\epsilon(\theta)\, 1\{H_{\epsilon+1}(\theta) > U_{\epsilon+1}(\theta)\} H_{\epsilon+1}(\theta) \\ &+ \left[1 - d^0_\epsilon(\theta)\right] U_{\epsilon+1}(\theta)\end{aligned} \tag{8}$$

Where $d^0_\epsilon(\theta) = \lambda^{01}_\epsilon(\theta)\bar{F^1}\left(R^{01}_{\epsilon+1}(\theta)\right) + \lambda^{02}_\epsilon(\theta)\bar{F^2}\left(R^{02}_{\epsilon+1}(\theta)\right) + \lambda^{03}(\theta)\, 1\{W^3_{\epsilon+1}(\theta) > U_{\epsilon+1}\} + \lambda_{04}(\theta)\, 1\{H_{\epsilon+1}(\theta)(\theta) > U_{\epsilon+1}\}$ is the probability to leave unemployment state, $1\{.\}$ is an

indicator function and $\bar{F^j} = 1 - F^1$ with $j \in \{1,\ 2\}$ are survival functions. In this case an unemployed individual receives non-labor income plus the continuation value to receive offers from the formal employment sector plus the continuation value to receive offers from the informal sector plus the continuation value to the opportunity to become self-employed plus the continuation value to the opportunity to go out from the labor market. The second Bellman equation correspond to the out of labor force state which is:

$$\begin{aligned}(1+r)\,H_\epsilon(\theta) = h_\epsilon(\theta) &+ \lambda_\epsilon^{41}(\theta)\int_{R_{\epsilon+1}^{41}}^{\bar{w^1}} W_{\epsilon+1}^1(\chi,\theta)\,dF^1(\chi) \\ &+ \lambda_\epsilon^{42}(\theta)\int_{R_{\epsilon+1}^{42}}^{\bar{w^2}} W_{\epsilon+1}^2(\chi,\theta)\,dF^2(\chi) \\ &+ \lambda_\epsilon^{43}(\theta)\,1\{W_{\epsilon+1}^3(\theta) > H_{\epsilon+1}(\theta)\}W_{\epsilon+1}^3(\theta) \\ &+ \delta_\epsilon^4(\theta)\,1\{U_{\epsilon+1}(\theta) > H_{\epsilon+1}(\theta)\}U_{\epsilon+1}(\theta) \\ &+ \left[1 - d_\epsilon^4(\theta)\right] H_{\epsilon+1}(\theta)\end{aligned} \tag{9}$$

With $d^4(\theta) = \lambda_\epsilon^{41}(\theta)\,\bar{F^1}\left(R_{\epsilon+1}^{41}(\theta)\right) + \lambda_\epsilon^{42}(\theta)\,\bar{F^2}\left(R_{\epsilon+1}^{42}(\theta)\right) + \lambda^{43}(\theta)\,1\{W_{\epsilon+1}^3(\theta) > H_{\epsilon+1}(\theta)\} + \delta^4(\theta)\,1\{U_{\epsilon+1}(\theta) > H_{\epsilon+1}(\theta)\}$ being the probability to leave out of labor state. Unlike the unemployment state, an out of the labor force individual receives home production income $h_\epsilon(\theta)$ plus the continuation value of receive offers from formal or informal sectors plus offers that arrive from self-employment sector including the possibility to enter into unemployment being an out of the labor force person. This addition of out of labor force state is the main change that I perform on the model of Narita (2020). Now, I describe the bellman equation

of a formal employee, in the following equation:

$$\begin{aligned}(1+r)\,W_{\epsilon}^{1}(w,\theta) = w + \delta_{\epsilon}^{1}\left(U_{\epsilon+1}(\theta) + UI + sw\right) \\ +\lambda_{\epsilon}^{11}(\theta)\int_{w}^{\bar{w^1}} W_{\epsilon+1}^{1}(\chi,\theta)\,dF^{1}(\chi) \\ +\lambda_{\epsilon}^{12}(\theta)\int_{R_{\epsilon+1}^{12}(\theta)}^{\bar{w^2}} W_{\epsilon+1}^{2}(\chi,\theta)\,dF^{2}(\chi) \\ +\lambda_{\epsilon}^{13}(\theta)\,1\{w < R_{\epsilon+1}^{13}(\theta)\}W_{\epsilon+1}^{3}(\theta) \\ +\lambda_{\epsilon+1}^{14}(\theta)\,1\{w < H_{\epsilon+1}(\theta)\}H_{\epsilon+1}(\theta) \\ +\left[1-d^{1}\right]W_{\epsilon+1}^{1}(w,\theta)\end{aligned} \tag{10}$$

as before $d^{1}(\theta)$ is the probability to leave the formal sector, defined by $d^{1}(\theta) = \delta^{1} + \lambda_{\epsilon}^{11}(\theta)\,\bar{F^{1}}(w) + \lambda_{\epsilon}^{12}(\theta)\,\bar{F^{2}}\left(R_{\epsilon+1}^{12}(w,\theta)\right) + \lambda_{\epsilon}^{13}(\theta)\,1\{w < R^{13}(\theta)\} + \lambda_{\epsilon}^{14}(\theta)\,1\{w < H_{\epsilon+1}(\theta)\}$. According to Narita (2020) when a formal employee become unemployed, she receives unemployment insurance (UI) plus severance payment (sw) that depends of the wage received in her working period and the continuation value of become unemployed. Additionally, a formal worker receives wage w if she keeps her employment, the value if she changes to another formal employment, the value of the opportunity to receive offers form the informal sector, the opportunity to become self-employed and the opportunity to leave labor market. Next, I describe the Bellman equation of the informal sector:

$$\begin{aligned}(1+r)\,W_{\epsilon}^{2}(w,\theta) = w + \delta_{\epsilon}^{2}U_{\epsilon+1}(\theta) \\ +\lambda_{\epsilon}^{21}(\theta)\int_{R_{\epsilon+1}^{21}(w,\theta)}^{\bar{w^1}} W_{\epsilon+1}^{1}(\chi,\theta)\,dF^{1}(\chi) \\ +\lambda_{\epsilon}^{22}(\theta)\int_{w}^{\bar{w^2}} W_{\epsilon+1}^{2}(\chi,\theta)\,dF^{2}(\chi) \\ +\lambda_{\epsilon}^{23}(\theta)\,1\{w < R_{\epsilon+1}^{32}(\theta)\}W_{\epsilon+1}^{3}(\theta) \\ +\lambda_{\epsilon}^{24}(\theta)\,1\{w < H_{\epsilon+1}(\theta)\}H_{\epsilon+1}(\theta) \\ +\left[1-d_{\epsilon}^{2}(w,\theta)\right]W_{\epsilon+1}^{2}(\theta)\end{aligned} \tag{11}$$

In this case, $d_{\epsilon}^{2}(w,\theta)=\delta^{2}+\lambda_{\epsilon}^{21}(\theta)\bar{F^{1}}\left(R_{\epsilon+1}^{21}(w,\theta)\right)+\lambda_{\epsilon}^{22}(\theta)\bar{F^{2}}(w)+\lambda_{\epsilon}^{23}(\theta)\mathbf{1}\{w<R^{23}(\theta)\}+\lambda_{\epsilon}^{24}(\theta)\mathbf{1}\{w<H_{\epsilon+1}(\theta)\}$ is the probability to leave informal employment. Now, I describe the composition over the lifecycle of every state, which describes the inflows and outflows of any particular state in terms of change probabilities and the size or measure of individuals in each state. Assume the measure of individuals in self-employment with ability θ and work experience $\epsilon+1$ is the quantity of individuals remaining as self-employees plus the number of individuals that arrives to self-employment from other states. In particular, this measure is:

$$\begin{aligned}
m_{\epsilon+1}^{3}(\theta) &= \left[1-d_{\epsilon}^{3}(\theta)\right]m_{\epsilon}^{3}(\theta)\\
&+\lambda_{\epsilon}^{03}(\theta)u_{\epsilon}(\theta)1\{W_{\epsilon+1}^{3}(\theta)>U_{\epsilon+1}(\theta)\}\\
&+\lambda_{\epsilon}^{13}(\theta)\int_{\underline{w}^{1}}^{R_{\epsilon+1}^{31}(\theta)}dM_{\epsilon}^{1}(\chi,\theta)\\
&+\lambda_{\epsilon}^{23}(\theta)\int_{\underline{w}^{2}}^{R_{\epsilon+1}^{32}(\theta)}dM_{\epsilon}^{2}(\chi,\theta)\\
&+\lambda_{\epsilon}^{43}(\theta)m_{\epsilon}^{4}(\theta)1\{W_{\epsilon+1}^{3}(\theta)>H_{\epsilon+1}(\theta)\}
\end{aligned} \tag{12}$$

In the last equation, $M_{\epsilon}^{j}=m_{\epsilon}^{j}(\theta)G_{\epsilon}^{j}(w,\theta)$ for $j\in\{1,2\}$ is the proportion of workers in each employment state j, where $G_{\epsilon}^{j}(w,\theta)$ is the distribution of accepted contracts in the formal and informal sectors (Meghir et al., 2015). The measure of workers in formal sector with experience $\epsilon+1$ and self-employment ability θ is:

$$\begin{aligned}
M_{\epsilon+1}^{1}(w,\theta) &= \left[1-d_{\epsilon}^{1}(w,\theta)\right]M_{\epsilon}^{1}(w,\theta)\\
&+\lambda_{\epsilon}^{01}(\theta)u_{\epsilon}(\theta)max\{F^{1}(w)-F^{1}\left(R_{\epsilon+1}^{01}(\theta)\right),0\}\\
&+\lambda_{\epsilon}^{21}(\theta)\int_{\underline{w}^{2}}^{\bar{w^{2}}}max\{F^{1}(w)-F^{1}\left(R_{\epsilon+1}^{21}(\chi,\theta)\right),0\}dM_{\epsilon}^{2}(\chi,\theta)\\
&+\lambda_{\epsilon}^{31}(\theta)m_{\epsilon}^{3}(\theta)max\{F^{1}(w)-F^{1}\left(R_{\epsilon+1}^{31}(\theta)\right),0\}\\
&+\lambda_{\epsilon}^{41}(\theta)m_{\epsilon}^{4}(\theta)max\{F^{1}(w)-F^{1}\left(R_{\epsilon+1}^{41}(\theta)\right),0\}
\end{aligned} \tag{13}$$

And the measure of workers in informal sector with experience $\epsilon+1$ and self-employment ability θ is:

$$\begin{aligned} M_{\epsilon+1}^{2}(w,\theta) &= \left[1-d_{\epsilon}^{2}(w,\theta)\right] M_{\epsilon}^{2}(w,\theta) \\ &+\lambda_{\epsilon}^{02}(\theta)\, u_{\epsilon}(\theta)\, max\{F^{2}(w)-F^{2}\left(R_{\epsilon+1}^{02}(\theta)\right),0\} \\ +\lambda_{\epsilon}^{12}(\theta)\int_{\underline{w}^{1}}^{\bar{w}^{1}} & max\{F^{2}(w)-F^{2}\left(R_{\epsilon+1}^{12}(\chi,\theta)\right),0\} dM_{\epsilon}^{1}(\chi,\theta) \\ &+\lambda_{\epsilon}^{32}(\theta)\, m_{\epsilon}^{3}(\theta)\, max\{F^{2}(w)-F^{2}\left(R_{\epsilon+1}^{32}(\theta)\right),0\} \\ &+\lambda_{\epsilon}^{42}(\theta)\, m_{\epsilon}^{4}(\theta)\, max\{F^{2}(w)-F^{2}\left(R_{\epsilon+1}^{42}(\theta)\right),0\} \end{aligned} \tag{14}$$

As in the self-employment case, the measures of formal and informal states depend on the workers that do not change of state plus the workers that come from other states along life cycle. In addition, I define the measure of out of labor force $m_{\epsilon+1}^{4}(\theta)$ with experience $\epsilon+1$ and self-employment ability θ, in the following way:

$$\begin{aligned} m_{\epsilon+1}^{4}(\theta) &= \left[1-d_{\epsilon}^{4}(\theta)\right] m_{\epsilon}^{4}(\theta) \\ &+\lambda_{\epsilon}^{04}(\theta)\, u_{\epsilon}(\theta)\, 1\{H_{\epsilon+1}(\theta)>U_{\epsilon+1}(\theta)\} \\ &+\lambda_{\epsilon}^{14}(\theta)\int_{\underline{w}^{1}}^{R_{\epsilon+1}^{41}(\theta)} dM_{\epsilon}^{1}(\chi,\theta) \\ &+\lambda_{\epsilon}^{24}(\theta)\int_{\underline{w}^{2}}^{R_{\epsilon+1}^{32}(\theta)} dM_{\epsilon}^{2}(\chi,\theta) \\ &+\lambda_{\epsilon}^{34}(\theta)\, m_{\epsilon}^{3}(\theta)\, 1\{H_{\epsilon+1}(\theta)>W_{\epsilon+1}^{3}(\theta)\} \end{aligned} \tag{15}$$

As in the previous equations, the measure of out of labor force state are the individuals who are out of labor force in the period ϵ and keep in that state until $\epsilon+1$ plus the entries from other states in life cycle. The measures m_{ϵ}^{i} with $i \in \{1,2,3,4\}$ are fractions of working age population, then, $u_{\epsilon}(\theta)=1-\sum_{i=1}^{4} m_{\epsilon}^{i}$. Narita (2020) follows the assumption of heterogeneity in firm productivity of Bontemps et al. (2000) to describe firms' behavior. Then, each firm has a productivity p with a distribution function $\Gamma^{1}(p)$ if the firm is

formal and $\Gamma^2(p)$ if the firm is informal with support $\left[\underline{p^j}, \bar{p^j}\right]$. Also, the measure of firms is normalized to one, so n_1 and n_2 are formal and informal fractions of total firms. With these assumptions, Narita (2020) describes that each firm makes initial offers of wages w, and informal firms enters the market as long as there are wages greater than one reservation wage which is $w \geq min\left(R_1^{02}, R_2^{02}, \ldots, R_E^{02}\right)$ an also $w \geq min\left(R_1^{42}, R_2^{42}, \ldots, R_E^{42}\right)$ which correspond to reservation wages to entry in an informal employment from out of labor force state, that is to say that these two conditions represent the constraints of wages offered in informal firms. Due to formal firms are regulated by the government, they an additional constraint which is minimum wage policies, then the formal the wages from these firms must be greater than the minimum wage and the set of reservation wages both of unemployment and out of labor force state which is $w \geq max\left\{w_{min}, min\left(R_1^{01}, R_2^{01}, \ldots, R_E^{01}\right)\right\}$ and $w \geq max\left\{w_{min}, min\left(R_1^{41}, R_2^{41}, \ldots, R_E^{41}\right)\right\}$.

In this framework, firms have to pay corporate taxes σ (fraction of the profits), payroll taxes τ and severance payments to workers who leave firm sxw. Therefore, the steady state profit of formal firms is defined in the following way:

$$\pi_1(p, w) = (1-\sigma) \sum_{\epsilon} \sum_{\theta} \left[p - \left(1 + \tau + \delta_{\epsilon}^1 s\right) w\right] \ell_{\epsilon}^1(w, \theta) \gamma_{\epsilon} q_{\theta} \tag{16}$$

Where γ_ϵ is the probability of potential experience ϵ, q_θ is the probability of potential self-employment ability θ, and $\ell_\epsilon^1(w, \theta)$ is the normalized labor force size of formal firms defined by:

$$\ell_{\epsilon}^1(w, \theta) = \frac{1}{n_1} \frac{dM_{\epsilon}^1(w, \theta)}{dF^1(w)} \tag{17}$$

For the case of informal firms, they do not have to pay either corporate taxes, payroll taxes or severance payments, but they have a cost when the authorities catch them, so this cost is denoted by C. Thus, the steady state profit of informal firms is:

$$\pi_2(p, w) = [p - C - w] \sum_{\epsilon} \sum_{\theta} \ell^2{}_{\epsilon}(w, \theta) \gamma_{\epsilon} q_{\theta} \tag{18}$$

With ℓ_{ϵ}^{2} as the normalized labor size of informal firms equal to:

$$\ell^{2}{}_{\epsilon}\left(w,\theta\right)=\frac{1}{n_{2}}\frac{dM_{\epsilon}^{2}\left(w,\theta\right)}{dF^{2}\left(w\right)} \tag{19}$$

To study the effect of a change in payroll taxes, I start by focusing on the equation 17. This equation describes the equilibrium in formal labor market where $n_1 dF^1\left(w\right)\ell_{\epsilon}^{1}\left(w,\theta\right)$ is the demand for labor and the term $dM_{\epsilon}^{2}\left(w,\theta\right)$ is the labor supply. This equilibrium shows the channel which employer's decisions according to labor demand changes when there is a payroll change. More precisely, formal employers as profit maximizing individuals choose how much employment demand, so when there is a change in one of his parameters in the profit function, they have two options to adjust this change in his behavior. The first one is to change wages in order to keep the same level of employment. The second one is to change the number of individuals employed in his firm and the decision between changes in wages or changes in employment will depend of friction in the labor market. For instance, in a high-regulated country, with wage rigidity as minimum wages, firms cannot change the level of wage to adjust the change in payroll taxes and this will turn in changes of employment.

The main conclusion of the theoretical framework developed by Narita (2020), with the extension of out of labor force state included here, is that decisions that motivate workers to change from one labor state to another are driven by economic forces that makes some states more attractive than others and they are not the result of random events in the labor market. This framework contrast heavily with the study of labor flows using Markov chains where there are no microeconomic foundations of individual decisions. According to the economic forces that drive flows in the labor market, this framework accounts for payroll taxation, which has consequences on labor market equilibria.

6 Identification Strategy

Article 9 of the First Job Act enables me to use a regression discontinuity design to determine the effect of payroll reductions on labor market flows of young adults. Formally, the specific age cutoff of 28 years creates an exogenous variation of individuals to compare labor market flows measured as transition rates below and above this age. In addition, I measure transition rates from one state to another as the flow of individuals with g years in months in month t who change from the origin state, to the destination state and divide the flow to the number of individuals with g years in months in month t in the origin state. Then, transition rates are grouped in individuals who are g years old in months in month t. Though, I have monthly information from 2008 to 2019, to identify the effect I keep information from January 2011 to December 2014, which correspond to the years in which the law was in force and I use remain information to run robustness checks and placebo test.

With the definition of transition rates, the identification assumption is that, for the specific subset of young adults, whether treated (below 28 years) or not treated (above 28 years) is "as good as" random. First, I estimate a model with all the age groups pooled across the for years of the policy. This model has the following structure:

$$y_a = \tau D + f(a) + e_a \tag{20}$$

where y_a is the transition rate from one state to another for age group individuals a, D is a dummy variable that indicates 1 if the age group is treated and 0 otherwise. a is normalized age (in months), $f(a)$ is a polynomial function of age and e_a is the error term. In this specification errors are clustered at age group. In this specification, τ measures the local average treatment effect of the First Job Act. In the robustness checks section, I extend the regression discontinuity design, allowing the effect of the policy to vary over time, then the local average treatment effect τ measures the effect over four years of the First Job Act. This extension allows me to include the effect of a particular year (year fixed effects) and

also increase the number of observations in the specification. The model that includes the time extension has the following specification:

$$y_{at} = \tau D + f(a) + \alpha_t + e_{at} \tag{21}$$

In this case, the depend variable y_{at} is the transition rate from one state to another for age group individuals a in time t (in months), α_t are time fixed effects and e_{at} is the error term in age group a in time t. As before D is a dummy variable that indicates 1 if the age group is treated and 0 otherwise. The identification assumption implies that potential outcomes are continuous around the cutoff.

7 Results

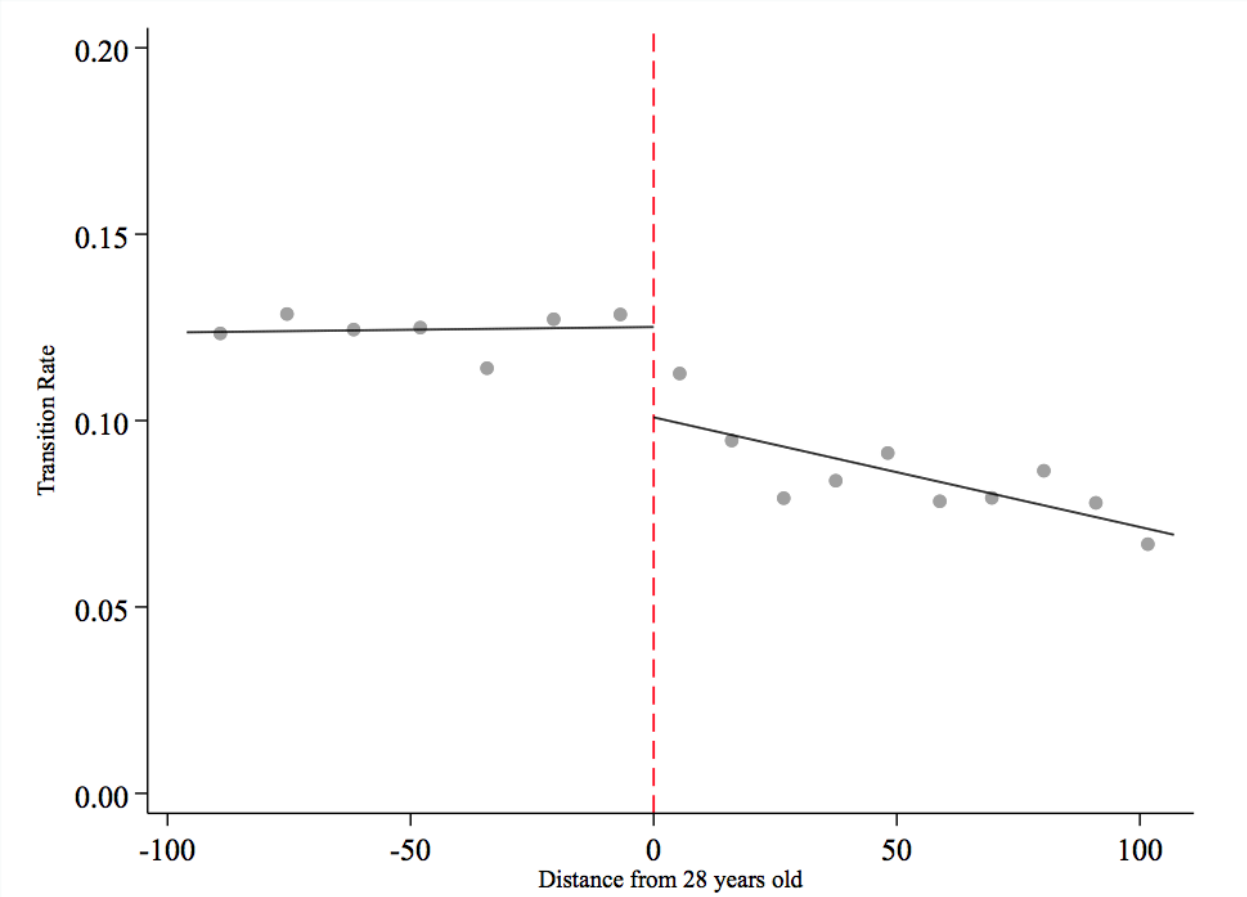
I start by reporting the estimation results of equation 20 in Table 3 for different transition rates and differentiating between the entire sample, men and women including clustered standard errors in parenthesis. In this estimation, I allow manual bandwidth selection to take all observations in the estimation process. Due to treatment means individuals who are below the cutoff, the coefficient interpretation should be from "right to left" graphically or with the opposite sign in the table. According to this, a negative coefficient means a positive effect in the transition rate for young adults. The results, show a significant positive effect of the First Job of 2.8 percentage points in transition rates from informality to formality, which is consistent with the aim of the policy. When I allow the disaggregation between men and women, the results suggest that the entire positive effect is driven by men with and positive effect of 6.2 percentage points in transitions rates from informal employment to formal employment. In the case of women this effect is not significant. Estimation results show that specifications when the dependent variable is different to transition rate from informality to formality, there is not significance. So, the policy had no effects no fostering other labor flows.

Table 3: RD Estimates of First Job Act

Dependent Variable	All	Men	Women
Informality to Formality	-0.0287***	-0.0621***	-0.0110
	(0.00760)	(0.0130)	(0.00868)
Unemployment to Formality	0.00506	-0.00557	0.00685
	(0.0107)	(0.0273)	(0.0124)
Inactivity to Formality	-0.000568	-0.00301	0.00313
	(0.00519)	(0.0125)	(0.00589)
Informality to Self-employment	0.00300	0.0117	-0.00166
	(0.00572)	(0.0128)	(0.00770)
Unemployment to Self-employment	-0.00670	0.00826	-0.0117
	(0.0119)	(0.0188)	(0.0127)
Inactivity to Self-employment	-0.00593	0.00757	-0.00786
	(0.00479)	(0.0132)	(0.00516)
N	204	204	204

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table presents the estimated coefficient from equation 20 and the depend variable is the transition rate between informality, unemployment or inactivity to formal employment or self-employment. I assume polynomial linear order in the relationship between the dependent and the running variable (age). Column (1) is the estimation form the entire sample and columns (2) and (3) is the estimates for men and women respectively. Robust standard errors are shown in parentheses and clustered by age group a.

Figure 2: RD estimate of First Job Act effect on transition rates from informality to formality



Notes: The figure shows the estimated effect of First Job Act on the transition rate from informal employment to formal employment. The x-axis is age in months normalized with respect to the age of 28 years. The estimation corresponds to the entire sample of individuals between 20 and 36 years. Standard errors are clustered at age level

Graphical representation of an RD design is helpful and informative (Lee and Lemieux, 2010) in this case, for policy estimates. In Figure 2, I show the estimates of the transition rate from informality to formality for all individuals using a linear relationship between transition rate and age. The average treatment effect is the difference between the intercept of the right regression line with the cutoff and the intercept of the left regression line with the cutoff. In this figure, there is a positive effect of being to the right to the cutoff which means, the policy increases the probability to make transitions from informal jobs to formal jobs for individuals who are above 28 years.

In addition, I estimate equation 20 for two groups separated by education. In this manner,

Table 4 reports estimates for individuals who have a high school degree or lesser education attainment, allowing to include individuals between 20 and 36 years old. These estimates show the same pattern of the estimation for the entire sample. However, it is interesting that this policy had a little effect in the transition rates from out of labor force to formal employment in women. This effect corresponds to a negative effect of 1 percentage point in the transition for women which is the opposite effect expected by the policy. Additionally, in Table 5, I show the estimates for individuals with more than a high school degree. In this case, the most relevant result is that the First Job Act reduces the transitions from unemployment to formality of individuals with less than 28 years of age. This is also the case for women, but these estimates are significant at 1%. Table 5 shows that the policy increases transitions from unemployment to self-employment for all individuals together and separately for women, but it had no effect for men.

7.1 Robustness checks

As mentioned before, to assess the effect of the First Job Act under different specifications, I change the pooled specification and account for time, namely, now individuals are grouped by age a and months t. This specification allows me to include monthly fixed effects as expressed in equation 21. The result of the estimation taking into account time is presented in Table 6. The coefficients show the same pattern of the estimates in equation 20, that is, the First Job Act increases the probability of change to formal employment from informal employment and the effect is larger for men. Also, this specification shows no effect on other transition rates.

Finally, to explore the sensitivity of the results to a range of bandwidths, I run equation 20 with different bandwidths. In this case, estimates are consistent to changes from 30 to 100 months, but the estimates losses significance with bandwidths below 45 months. Figure 3 shows the results of this exercise. I also perform a placebo regression using individuals who change form one state to another, but in years after and before the validity of the First Job

Table 4: RD Estimates of First Job Act for individuals with high school degree or less

Dependent Variable	All	Men	Women
Informality to Formality	-0.0299***	-0.0642***	-0.0115
	(0.00765)	(0.0137)	(0.00904)
Unemployment to Formality	-0.00595	-0.0450	0.00465
	(0.0135)	(0.0352)	(0.0142)
Inactivity to Formality	0.00999	0.00353	0.0130*
	(0.00537)	(0.0167)	(0.00629)
Informality to Self-employment	0.00446	0.0118	-0.000284
	(0.00644)	(0.0137)	(0.00873)
Unemployment to Self-employment	-0.00191	0.0282	-0.00984
	(0.0162)	(0.0260)	(0.0179)
Inactivity to Self-employment	-0.00591	-0.0151	-0.00468
	(0.00571)	(0.0148)	(0.00607)
N	204	204	204

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table presents the estimated coefficient from equation 20 and the depend variable is the transition rate between informality, unemployment or inactivity to formal employment or self-employment. I assume polynomial linear order in the relationship between the dependent and the running variable (age). Column (1) is the estimation form the entire sample and columns (2) and (3) is the estimates for men and women respectively. Robust standard errors are shown in parentheses and clustered by age group a.

Table 5: RD Estimates of First Job Act for individuals with more than high school degree

Dependent Variable	All	Men	Women
Informality to Formality	-0.00877	0.0441	0.00756
	(0.0330)	(0.0932)	(0.0337)
Unemployment to Formality	0.0319**	0.0785*	0.0172*
	(0.0197)	(0.0390)	(0.0257)
Inactivity to Formality	-0.00441	-0.00177	-0.00632
	(0.0136)	(0.0280)	(0.0148)
Informality to Self-employment	-0.0142	-0.00176	-0.0133
	(0.0251)	(0.0444)	(0.0272)
Unemployment to Self-employment	-0.0168**	-0.0170	-0.0151*
	(0.0144)	(0.0274)	(0.0189)
Inactivity to Self-employment	-0.00549	0.0434	-0.0167
	(0.0105)	(0.0290)	(0.0125)
N	204	204	204

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table presents the estimated coefficient from equation 20 and the depend variable is the transition rate between informality, unemployment or inactivity to formal employment or self-employment. I assume polynomial linear order in the relationship between the dependent and the running variable (age). Column (1) is the estimation form the entire sample and columns (2) and (3) is the estimates for men and women respectively. Robust standard errors are shown in parentheses and clustered by age group a.

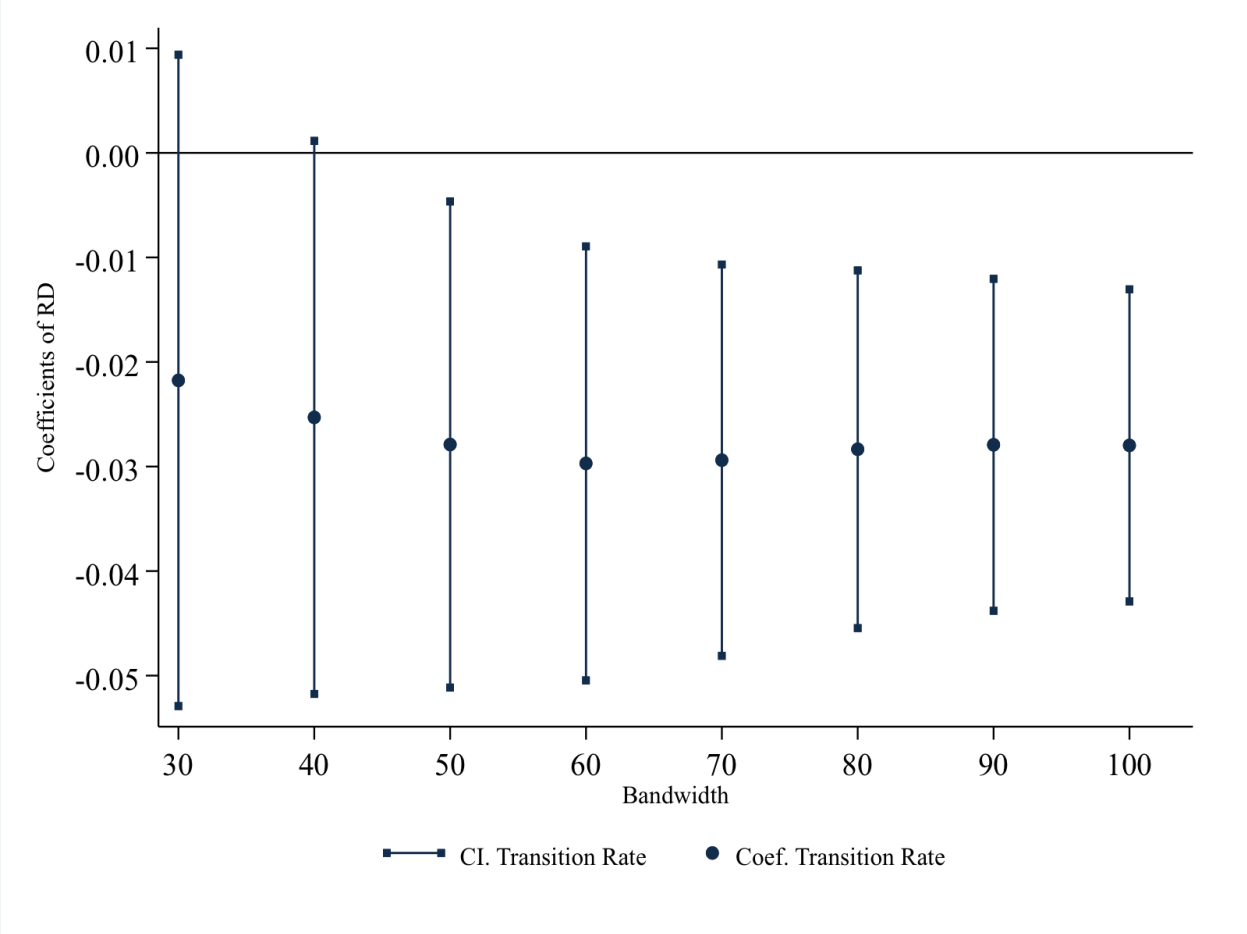
Table 6: RD Estimates of First Job Act including time effects

Dependent Variable	All	Men	Women
Informality to Formality	-0.0270**	-0.0679***	-0.00658
	(0.00994)	(0.0154)	(0.00969)
Unemployment to Formality	-0.00416	-0.00816	0.00279
	(0.0131)	(0.0292)	(0.0148)
Inactivity to Formality	0.00141	-0.00576	0.00331
	(0.00507)	(0.0139)	(0.00585)
Informality to Self-employment	0.00118	0.0106	-0.00329
	(0.00695)	(0.0143)	(0.00871)
Unemployment to Self-employment	0.00607	0.00336	-0.00333
	(0.0166)	(0.0200)	(0.0164)
Inactivity to Self-employment	-0.00442	0.00629	-0.00854
	(0.00469)	(0.0161)	(0.00532)
N	9454	6448	8797

Notes: Standard errors in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. The table presents the estimated coefficient from equation 21 which have into account the effect of time, the depend variable is the transition rate between informality, unemployment or inactivity to formal employment or self-employment. I assume polynomial linear order in the relationship between the dependent and the running variable (age). Column (1) is the estimation form the entire sample and columns (2) and (3) is the estimates for men and women respectively. Robust standard errors are shown in parentheses and clustered by age group a.

Act. This result is presented in Table B1 and shows that there no effects on labor transitions.

Figure 3: RD estimates of First Job Act effect with varying bandwidth



Notes: This figure show the estimates of different bandwidths in the regression discontinuity design specification.

8 Conclusion

This research article has explored the effect of the First Job Act, a policy to encourage flows into formality of young adults in Colombia. I proposed an extension of the on-the-job labor model of Narita (2020) to understand how labor decisions made by firms and workers can affect the flows between states, even when there exist changes in payroll taxes. This theoretical extension is based on the idea that all individuals make their choices on a basis

of rational decisions of employment about their expected value of every state.

I estimated a regression discontinuity design to evaluate the policy effect on labor flows. While there is not significant effect on flows into formal states from unemployment or inactivity, the reduction in payroll taxes of around 11 percentage points of formal firms in Colombia can raise flows from informality to formality of workers by around 2,8 percentage points. The effect is larger for men whose transition rate increase in 6,2 percentage points. The policy is also effective on individuals with lesser education attainment. These results are robust to changes of bandwidths and the inclusion of time effects.

The result of the effect of the First Job Act into transitions from informality to formality support the hypothesis that payroll tax cuts could be useful to increase the transition of young adults into formality. However, the results show that these policies do not change flows from other states as unemployment or out of the labor force. This raises the question of which policies encourage flows from other states into formality, even in the context of developing countries such as Colombia. This research article is limited by the nature of data and the constraints of quasi-experimental exercises. I cannot include welfare analysis or have a better definition of labor states, but these research constraints should encourage policy makers to formulate policies to improve labor opportunities of young adults such as the First Job Act.

References

Antón, A. (2014). The effect of payroll taxes on employment and wages under high labor informality. *IZA Journal of Labor & Development*, 3(1):1–23.

Bai, J. and Wang, Y. (2020). Returns to work, child labor and schooling: The income vs. price effects. *Journal of Development Economics*, 145:102466.

Beam, E. A., Hyman, J., and Theoharides, C. (2020). The relative returns to education, experience, and attractiveness for young workers. *Economic Development and Cultural Change*, 68(2):391–428.

Becerra-Camargo, O. (2016). *Effects of labor regulation on informal labor markets*. PhD thesis, University of British Columbia.

Blanchard, O. J. and Diamond, P. (1992). The flow approach to labor markets. *The American Economic Review*, 82(2):354–359.

Bontemps, C., Robin, J.-M., and Van den Berg, G. J. (2000). Equilibrium search with

continuous productivity dispersion: Theory and nonparametric estimation. *International Economic Review*, 41(2):305–358.

Bosch, M., Goni, E., and Maloney, W. (2007). The Determinants of Rising Informality in Brazil: Evidence from Gross Worker Flows. *Policy Research Working Paper*.

Bosch, M. and Maloney, W. (2006). *Gross worker flows in the presence of informal labor markets: the Mexican experience 1987-2002*. The World Bank.

Bridgman, B. (2016). Home productivity. *Journal of Economic Dynamics and control*, 71:60–76.

Bridgman, B., Dugan, A., Lal, M., Osborne, M., and Villones, S. (2012). Accounting for household production in the national accounts, 1965–2010. *Survey of current Business*, 92(5):23–36.

Burdett, K. and Mortensen, D. T. (1998). Wage differentials, employer size, and unemployment. *International Economic Review*, pages 257–273.

Davis, S. and Haltiwanger, J. (1998). Measuring gross worker and job flows, in 'labor statistics measurement issues'.

Dube, A., Lester, T. W., and Reich, M. (2016). Minimum wage shocks, employment flows, and labor market frictions. *Journal of Labor Economics*, 34(3):663–704.

Elsby, M. W., Smith, J. C., and Wadsworth, J. (2010). Cep discussion paper no 1058 july 2010 the role of worker flows in the dynamics and distribution of uk unemployment.

Flórez, L. A., Morales-Zurita, L. F., Medina, D., Lobo, J., Florez, L. A., and Morales, L. F. (2018). Labour flows across firm´ s size, economic sectors and wages in colombia: evidence from employer-employee linked panel. *Borradores de Economía; No. 1013*.

Galiani, S., Meléndez, M., and Ahumada, C. N. (2017). On the effect of the costs of operating formally: New experimental evidence. *Labour Economics*, 45:143–157.

Hagedorn, M. and Manovskii, I. (2008). The cyclical behavior of equilibrium unemployment and vacancies revisited. *American Economic Review*, 98(4):1692–1706.

Hotz, V. J., Xu, L. C., Tienda, M., and Ahituv, A. (2002). Are there returns to the wages of young men from working while in school? *Review of Economics and statistics*, 84(2):221–236.

Kaitz, H. B. (1970). Analyzing the length of spells of unemployment. *Monthly Lab. Rev.*, 93:11.

Kugler, A. and Kugler, M. (2009). Labor market effects of payroll taxes in developing countries: Evidence from colombia. *Economic development and cultural change*, 57(2):335–358.

Lahr, M. and De Mesnard, L. (2004). Biproportional techniques in input-output analysis: table updating and structural analysis. *Economic Systems Research*, 16(2):115–134.

Lasso-Valderrama, F. J. (2012). La dinámica del desempleo urbano en colombia. *Capítulo 3. La dinámica del desempleo urbano en Colombia. Pág.: 131-166*.

Lasso-Valderrama, F. J. (2020). La dinámica del mercado laboral colombiano: Cuál es el rol de los flujos de trabajadores? *Borradores de Economía; No. 1120*.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

López Castaño, H. and Lasso Valderrama, F. (2015). Flujos de trabajadores en el mercado laboral colombiano, determinantes de las diferencias entre hombres y mujeres, y futuro laboral esperado. Technical report, IDB Working Paper Series.

Meghir, C., Narita, R., and Robin, J.-M. (2015). Wages and informality in developing countries. *American Economic Review*, 105(4):1509–46.

Mortensen, D. T. and Pissarides, C. A. (1994). Job creation and job destruction in the theory of unemployment. *The review of economic studies*, 61(3):397–415.

Narita, R. (2020). Self-employment in developing countries: A search-equilibrium approach. *Review of Economic Dynamics*, 35:1–34.

Pissarides, C. A. (2000). *Equilibrium unemployment theory*. MIT press.

Prada, C. F. (2012). Análisis de los flujos de trabajadores y la segmentación laboral en colombia. *Revista Ensayos Sobre Política Económica; Vol. 30. No. 68. Junio, 2012. Pág.: 74-113.*

Shimer, R. (2007). Mismatch. *American Economic Review*, 97(4):1074–1101.

Shimer, R. (2012). Reassessing the ins and outs of unemployment. *Review of Economic Dynamics*, 15(2):127–148.

Van den Berg, G. J. and Ridder, G. (1998). An empirical equilibrium search model of the labor market. *Econometrica*, pages 1183–1221.

Appendix A

Table A1: Questions to define labor states

State	Question	Question Code
Formal Employment	¿Have you ever worked for at least two consecutive weeks?	P7430
	¿How long have you been working for the last time? At least for two weeks	P7440
	¿How long have you last looked for a job?	P7456
	¿Have you looked for work for the first time or have you worked before for at least two consecutive weeks?	P7310
	¿How many weeks ago have you stopped working for the last time?	P7320
	¿How many weeks have you been or were you looking for work?	P7250
	¿How long have you been working in this company, business, industry, office, firm or farm continuously?	P6426
	Before the current job, ¿did you have another job?	P7020
	¿How many months were you without a job between your current job and your previous one?	P760
Unemployment	¿Have you ever worked for at least two consecutive weeks?	P7430
	¿How long has it been since you last worked?	P7440
	¿How many weeks ago have you stopped working for the last time?	P7320
	In this last job you were:	P7350
	¿How long have you been working in this company, business, industry, office, firm or farm continuously?	P6426
	In your previous job you were:	P7028
	¿How many months were you without a job or job between your current job and your previous one?	P760
	¿How long have you been working in this company, business, industry, office, firm or farm continuously?	P6426
	In this job you are:	P6430
Informal Employment	¿How many weeks ago have you stopped working for the last time?	P7320
	In this last work it was:	P7350
	¿How long have you been working in this company, business, industry, office, firm or farm? keep going?	P6426
	In your previous job you were:	P7028
	¿How many months were you without a job or job between your current job and your previous one?	P760
	In this job you are:	P6430
Self-Employment	¿How many weeks ago have you stopped working for the last time?	P7320
	In this last work it was:	P7350
	¿How long have you been working in this company, business, industry, office, firm or farm? keep going?	P6426
	In your previous job you were:	P7028
	¿How many months were you without a job or job between your current job and your previous one?	P760
	In this job you are:	P6430
Out of Labor Force	¿Have you looked for work for the first time or had you worked before for at least two consecutive weeks?	P7310
	¿How many weeks ago have you stopped working for the last time?	P7320
	¿How many weeks have you been or were you looking for work?	P7250
	¿How long have you been working in this company, business, industry, office, firm or farm continuously?	P6426
	Before the current job, ¿did you have another job?	P7020
	¿How many months were you without a job or job between your current job and your previous one?	P760
	¿How long did have your last work?	P7440

Source: GEIH

Appendix B

Table B1: RD estimates of placebo sample

Dependent Variable	All	Men	Women
Informality to Formality	0.00140	0.00321	-0.000367
	(0.00532)	(0.0104)	(0.00543)
Unemployment to Formality	-0.0172	-0.0259	-0.0152
	(0.00730)	(0.0159)	(0.00851)
Inactivity to Formality	-0.00643	0.00209	-0.00500
	(0.00348)	(0.0107)	(0.00373)
Informality to Self-employment	-0.00911	-0.00897	-0.00902
	(0.00701)	(0.0128)	(0.00699)
Unemployment to Self-employment	0.00278	-0.0136	0.00845
	(0.0162)	(0.0119)	0.00845
Inactivity to Self-employment	-0.00553	0.00852	-0.00746
	(0.00332)	(0.00907)	(0.00372)
N	204	204	204

Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The table presents the estimated coefficient from equation 20 and the depend variable is the transition rate between informality, unemployment or inactivity to formal employment or self-employment. I assume polynomial linear order in the relationship between the dependent and the running variable (age). Column (1) is the estimation form the entire sample and columns (2) and (3) is the estimates for men and women respectively. Robust standard errors are shown in parentheses and clustered by age group a.