Effects of Labor Regulation on Informal Labor Markets

by

Oscar Becerra-Camargo

B.Sc., Universidad Nacional de Colombia, 2005
M.Sc., Universidad Nacional de Colombia, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Economics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2016

Abstract

This thesis examines the effects of labor regulation on formal (regulated) labor markets in Latin America. It is divided in three chapters, in which I analyze the effects of pension programs on formal-sector labor supply and the effects of payroll taxes on formal-sector labor demand.

The first two chapters analyze how future pension benefits affect formal-sector labor supply. Since formal-sector jobs comply with labor regulation, including contributions to pension plans, formal-sector workers receive long-run benefits in the form of pensions. If workers account for such benefits when they search for formal-sector jobs, the pension system affects formal-sector labor supply before the retirement age. In Chapter 1, I estimate the causal link between future pension benefits and formal-sector labor supply by using a cohort-based reform undertaken in Colombia. I demonstrate that workers with higher pension gains are more willing to work in formal-sector jobs, rather than working in unregulated businesses or by themselves. The result is consistent with a life-cycle model of formal-sector labor supply presented in Chapter 2, where pension benefits are an amenity of working in the formal sector. The results suggest that pension reforms may have large effects on the labor market that should be taken into account in the design of pension programs.

Chapter 3 analyzes the effect of payroll taxes on formal-sector labor demand in the presence of wage rigidity. In particular, I study the impact of a reduction of payroll taxes on the creation of formal-sector jobs in Colombia, where about 40 percent of formal-sector workers earn the minimum wage. Using a reform that granted tax credits to firms hiring workers younger than 28 years of age, I obtain estimates of the effect of payroll taxes on formal-sector employment and wages. I show that payroll tax incidence is borne by formal-sector employers. The reduction in payroll taxes increased formal-sector employment and had no effects on wages. Using the estimation results, I recover an estimate of the elasticity of the formal-sector labor demand of -0.44. This result implies that a 10 percent increase in the minimum wage reduces formal-sector employment by 4.4 percent.

Preface

This dissertation is original, unpublished, independent work by the author, Oscar Becerra.

Table of Contents

Appendices

List of Tables

List of Figures

Acknowledgements

I would like to express my deepest gratitude to Thomas Lemieux, Joshua Gottlieb, and Florian Hoffmann for their outstanding guidance in the preparation of this document. Their invaluable advice and support made this project possible.

I would also like to thank Vancouver School of Economics' professors Paul Beaudry, Nicole Fortin, Nancy Gallini, David Green, Kevin Milligan, Marit Rehavi, Tomasz Swiecki, and Yaniv Yedid-Levi for their useful comments and the valuable discussion in and out of the classroom. I would like to give a special acknowledgment to my fellow classmates in the UBC Public Finance Reading Group: Alix Duhaime-Ross, Timea Laura Molnar, Daniel Shack, Lori Timmins, and Tzu-Ting Yang for all the hours we spent together discussing this project.

Finally, I thank Carlos Prada and Luz Angela Malagón from the Colombian Statistics office (DANE), and Mauricio Olivera and Adriana Hernandez from Colpensiones for facilitating access to the Census, PILA and Colpensiones datasets. This work represents my sole views and does not necessarily reflect those of DANE or Colpensiones.

To Liceth, the best teammate I could have ever asked for.

Chapter 1

Pension Incentives and Formal-Sector Labor Supply

1.1 Introduction

Does the prospect of future pension benefits determine workers' choices in developing economies? Workers in these economies are able to respond to public policies by changing their search strategies for finding work between the formal and informal sectors. If workers respond to the prospect of pension benefits by changing their search for formal-sector jobs, then the pension system affects formal-sector labor supply.

The informal sector encompasses the set of firms and workers that do not comply with government regulation, including the payment of mandated contributions (e.g., pension) and taxes (Perry, Maloney, Arias, Fajnzylber, Mason, and Saavedra-Chanduvi, 2007).[1] As a result, informal-sector workers are not covered by mandated benefits and other insurance included in the regulation. About 50 percent of Latin American workers work in the informal sector, which is mostly less-educated people working as a self-employed worker or as a salaried-worker in a small firm (Perry et al., 2007). Empirical evidence suggests that many informal-sector workers are part of an integrated labor market in which they move between sectors depending on the benefits in each sector and the cost of finding formal-sector jobs (Maloney, 2004).

[1] As Santa María, García, and Mujica (2009) point out, the decision of firms to operate in the informal sector is the outcome of weak enforcement by tax authorities, high regulation costs for registering, and low valuation of the benefits of operating in the formal sector.

Workers' behavioral responses along the formal-informal margin are an important consideration for the design of retirement policies in Latin American countries. In Latin America, approximately 50 percent of workers do not contribute to the pension system. Since most of these workers have a low income, the lack of pension contributions exacerbates income inequality after retirement (Frölich, Kaplan, Pagés, Rigolini, and Robalino, 2014). Latin American policymakers have implemented major pension reforms in response to concerns about low coverage and high inequality in benefits. These reforms have included new types of funding, changes in qualifying conditions for receiving a pension, and the introduction of pension assistance programs.[2] Yet these reforms have a potential offsetting cost. They may reduce the workers' expected gains from retirement contributions, thereby reducing the incentive to search for formal-sector jobs.

Despite the importance of workers' behavioral response to retirement policies, the empirical evidence establishing a causal link between pension incentives and formal-sector labor supply is scarce. The main empirical challenge is that the observable determinants of a worker's expected pension benefits are likely correlated with unobservable determinants of the worker's current labor choices. In absence of non-linear patterns in the pension benefit formulas, the variation in determinants of future expected benefits does not identify the causal link between pension-related incentives and formal-sector labor supply (Liebman, Luttmer, and Seif, 2009).

In this paper, I estimate the causal link between pension-related incentives and formal sector labor supply. I overcome the identification problem using quasi experimental variation from a cohort-based reform to the Colombian pension system. In 1993, the Colombian government increased the pension contribution rate and changed the minimum qualifying conditions for receiving a pension in the defined-benefit system. However, the reform did not change the qualifying conditions for eligible men born before April

[2] Ten countries in the region implemented major reforms to their pension systems: Chile (1981 and 2008), Peru (1993), Colombia (1994), Argentina (1994, 2008), Uruguay (1996), Mexico (1997), Bolivia (1997), El Salvador (1998), Costa Rica (2000), Nicaragua (2000), and Dominican Republic (2003). A detailed list and discussion of other non-contributory pension programs is in Bosch, Melguizo, and Pagés (2013).

1954 and eligible women born before April 1959. Compared with younger workers, eligible workers could retire contributing for fewer years (20 years instead of up to 25), and at an earlier age (55 years for women and 60 years for men, instead of 57 and 62). In this way, the reform permanently changed the long-run gains from a formal-sector job depending on the worker's birth date.

The difference in qualifying conditions by date of birth provides a source of exogenous variation to estimate the causal link between pension-related incentives and formal-sector labor supply. To estimate this effect, I implement a two-stage procedure. First, I use a regression discontinuity design (RD) on two new confidential datasets from 2005 and 2011. I compute the difference between formal-sector outcomes for workers born just before and just after the eligibility cutoffs. If there is no other economic or institutional factor to explain a discontinuous change in formal-sector labor supply at the cutoff, the difference is an estimate of the causal link between pension-related incentives and formal-sector labor supply. Second, I use additional assumptions to recover the elasticity of the formal-sector labor supply with respect to the net-of-tax share, a measure of the efficiency costs of pension taxes (Feldstein and Liebman, 2002).

To understand the impact of the change in qualifying conditions on formal-sector labor supply, I develop a model that characterizes workers' decisions about retirement and job search in the formal and informal sectors. The model builds on the framework proposed by Chetty (2006) for unemployment insurance and adapted by Gerard and Gonzaga (2014) to include an informal sector. I modify the model to incorporate a defined-benefit pension system, where the worker is entitled to a pension after reaching a minimum retirement age and a minimum number of years of contributions (the vesting period). In the model, workers search for formal-sector jobs because these jobs increase the likelihood of getting pension benefits in the future. Within an age group, the long-run gains from working in the formal-sector are a nonlinear function of the years of contribution. The higher gains concentrate among workers who are just below the vesting period, since they are the ones more likely to see the vesting period as binding.

The comparative statics of the model show that the effect of an increase in the minimum qualifying conditions on labor supply is heterogeneous, and that even its sign is ambiguous. The direction of the response depends on the worker's previous contributions relative to the new vesting period. On the one hand, workers who are a long way from satisfying the new vesting requirement reduce their search effort for formal-sector jobs, given their low likelihood of ever vesting. On the other hand, workers who are close to satisfying the new vesting requirement increase their search effort for formal-sector jobs to secure their pension benefits. The magnitude of the effect depends on the worker's age and opportunities to find a formal-sector job.

I present four main empirical findings. First, there is a sizable and significant response in formal-sector labor supply to changes in pension incentives. The effect is concentrated among men. For men, the average effect of harder qualifying conditions on salaried-formal labor supply is negative 12 percent in 2005, while it is positive 7 percent in 2011. The change is consistent with the insights provided by the model. By 2005, many workers born before April 1954 had not reached the minimum 20 years of contribution for a pension, giving them more incentives to work in the formal-sector than workers born after April 1954, who have to contribute to the system a minimum of 25 years. By 2011 many workers born before April 1954 had already met the required 20 years of contribution, thereby losing their incentive to contribute. In addition, I find little evidence suggesting that the changes in the formal-sector labor supply is offset by changes in wages.

Second, the change in formal employment is related to a shift from self-employment to salaried-formal employment, with no response along the extensive margin. The estimated effect of harder qualifying conditions on self-employment (which is mostly informal) is positive and of similar magnitude to the negative effect on salaried-formal employment. The estimated effect on labor force participation is not significant. These results are similar to those of Almeida and Carneiro (2012), who found that higher mandated benefits with no wage adjustment generate an incentive for self-employed workers to switch to salaried-formal jobs.

Third, the response of formal-sector labor supply to pension incentives is heterogeneous, and depends on the worker's labor market opportunities. The effect is concentrated among workers for whom the minimum qualifying conditions are binding. I analyze the response for groups with different propensities to work in the formal sector (e.g., education and region). In the analysis by educational attainment, I find that workers with primary and post-secondary education are less responsive to pension incentives than workers with secondary education. This result is consistent with the model predictions; workers for whom the minimum qualifying conditions are not binding are less sensitive to changes in pension incentives. Intuitively speaking, increasing the likelihood of securing future pension benefits is not a relevant factor for workers with no prospect of getting a pension, or workers who know with certainty they will get a pension. I obtain consistent results for other subsamples, such as those based on regional variation and marital status.

Fourth, I estimate an elasticity of the formal-sector labor supply with respect to the net-of-tax share of 1.7. Using variation by region and education, I regress the change in formal-sector employment at the discontinuity on the change in the predicted net-of-tax share for workers at the discontinuity. Consistent with the predictions of the model, workers with higher pension incentives along the formal-informal margin also exhibit higher responses in formal-sector labor supply. The estimate is likely a lower bound of the actual elasticity, suggesting large behavioral responses.

The paper is organized as follows: Section 1.2 describes the labor market institutions and the Colombian pension system. Section 1.3 presents the conceptual framework that provides insights about the expected sign and sources of heterogeneity in the results. Section 1.4 discusses the identification strategy and the data sources. Section 1.5 reports the estimation results, and Section 1.6 concludes.

1.2 Institutional background

1.2.1 Labor market institutions

The Colombian government mandates that employers provide benefits to their employees, and that self-employed workers contribute to the pension and contributory health care systems. Workers covered by mandated benefits are considered formal-sector workers.

Formal-sector jobs generate two types of gains for workers. First, formal-sector workers have access to mandated benefits. For salaried workers, formal-sector jobs provide the following: insurance from the pension and contributory health care systems, paid vacations (two weeks per year), severance payments (an additional monthly wage per year of tenure), a maximum number of working hours (48 per week), maternity leave (14 weeks), a 13th month of pay each year, access to subsidies for children's education, and compliance with the minimum wage. For self-employed workers, the gain from paying their contributions is limited to the insurance provided by the pension and the contributory health care systems.[3]

Second, formal-sector workers earn higher wages. As La Porta and Shleifer (2014) show, formal-sector firms tend to pay higher wages than informal-sector firms, and formal self-employed workers tend to have more education. Using data from the household surveys (described in Section 1.4.1), Table 1.1 presents the average wage and distribution of urban workers aged 20 to 65 who work at least 30 hours per week. In the table, I define a formal-sector worker as a worker making a contribution to the pension system and covered by the contributory health care system.[4] The average formal-to-informal wage gap is 75 percent for salaried workers and 100 percent for self-employed workers. The wage gap is positive regardless of the

[3]The minimum contributions for the pension and contributory health care systems are 16 and 12 percent of the minimum wage.

[4]The coverage of pension and contributory health care systems is a widely-used measure of formal employment (Perry et al., 2007). Because workers are not subject to penalties for being without coverage, it is unlikely that they would misreport their coverage status. Moreover, the formal employment indicators are consistent with aggregate statistics obtained from administrative data.

worker's level of education and type of employment. Nonetheless, as Figure 1.1 shows, the wage for a large fraction of informal-sector workers is above the minimum wage.

Despite the gains available from working in the formal sector, other supply and demand factors may prevent workers from working in this sector. On the labor supply side, workers may find it optimal to work in the informal sector (Maloney, 2004). Studies show that low valuation of mandated benefits, social programs that substitute the mandated benefits, and preferences for independent work reduce the incentives to search for formal-sector jobs (Levy, 2008; Perry et al., 2007). On the labor demand side, firms may find it optimal to operate informally in response to high regulation costs and weak enforcement, thereby reducing workers' chances of finding a formal-sector job. The effect of labor market regulation is evident in Figure 1.1, as a large fraction of formal-sector workers earn the minimum wage.

1.2.2 The Colombian pension system

In 1993, the Colombian government introduced the General Pension System (GPS), a new system intended to increase coverage and equality in retirement benefits while improving the financial viability of the system. The GPS integrates two pension systems: a new system to cover all new entrants, men born after March 1954, and women born after March 1959; and a transition system to cover all other workers.

For young workers and new entrants, the GPS allows workers to choose between two pension systems. All public and private sector workers must contribute to one system, and their choice determines their pension eligibility and benefits.[5] The first system, the social insurance system, allows workers to contribute to a defined-benefit pension plan managed by Colpensiones (the public pension fund). In the defined-benefit plan, the pension benefits are the maximum between a fraction of the worker's wage and the minimum wage, while eligibility is based on the worker's age and years of contribution. The second system, the individual account system, allows workers to

[5] Workers can switch systems every five years, up to the last ten years before the minimum retirement age (62 for men and 57 for women).

contribute to a defined-contribution plan managed by private pension funds. In the defined-contribution plan, the pension fund invests the worker's contributions in the capital market, and the principal and returns constitute the worker's savings for retirement. The worker's benefits and eligibility are based on the accrued capital. The defined-contribution plan also includes a guaranteed minimum pension of a monthly minimum wage. Eligibility for the guaranteed minimum pension is based on the worker's age and years of contribution.[6]

For all other workers, the GPS is the transition system where workers contribute to a defined-benefit plan managed by Colpensiones. Unlike the social insurance system, the transition system retains the pre-reform eligibility and benefits for eligible workers. The eligibility criteria were based on age and contributions at the time the reform took effect (April 1, 1994). Originally, three groups of employees were eligible for the transition system: men born before April 1954 who had contributed before April 1994, women born before April 1959 who had contributed before April 1994, and younger workers who had contributed to the system for at least 750 weeks (just over 14 years). A 2005 reform required eligible workers to have contributed more than 750 weeks by July 2005 and meet the qualifying conditions by 2014.[7]

Table 1.2 and Figure 1.2 summarize the main characteristics of the GPS. Workers in the three systems face the same contribution rate (16 percent)[8] but different minimum qualifying conditions. Compared with younger workers and new entrants, an eligible worker could retire having contributed for fewer years (20 years instead of up to 25) and at an earlier age (55 years for women and 60 years for men instead of 57 and 62). Although the transition

[6] When workers do not accumulate enough time (and capital) for being entitled to a pension, their contributions are refunded in a lump-sum payment. In the social insurance system they receive their contributions adjusted by inflation, in the individual account system they receive the accrued capital plus interest.

[7] In the rest of the paper, I focus on the eligibility criteria based on the date of birth of the worker. The criterion based on 750 weeks by 1993 has a limited effect on younger workers, given that the requirement implies that men younger than 40 and women younger than 35 would have worked by at least 14 years in the formal sector. Moreover, after the 2005 reform, men born after 1954 and women born after 1959 became ineligible for the transition system, even though they could have met the original eligibility criteria.

[8] The contribution rate before the 1993 reform was 6 percent.

system has a higher replacement rate than the social insurance system, the minimum pension guarantee implies that low wage workers face the same effective replacement rate in both systems (Figure 1.2). This is a relevant feature of the system given that 90 percent of workers in the GPS report earnings between one and two times the minimum wage.

Relevance of the minimum qualifying conditions. The importance of differences in the minimum qualifying conditions for a pension depends on whether the workers take these conditions into account when making their retirement decisions. Figure 1.3 shows that most workers claim their pension benefits as soon they meet the requirements. Based on statistics from Colpensiones, Figure 1.3 displays the distribution by age and weeks of contribution for non-retired men and women who contributed to Colpensiones up to December 2013. I focus on workers around the minimum retirement age, 60 years for men and 55 years for women. The distribution exhibits a clear discontinuity at the minimum retirement age when the number of weeks is above 1,000 (the minimum for the transition system) and the discontinuity widens as the number of weeks increases.

Interactions with other programs. The introduction of the General Pension System generated cohort differences in the minimum qualifying conditions and pension benefits received by workers. However, the differences in the conditions may have a limited effect on the workers' behavior if there were other cohort-based assistance programs targeted to the same population.

In recent years, Colombia has expanded several non-contributory social assistance programs. Three of the most significant programs are the non-contributory health care system, conditional cash transfers for families with children aged 0 to 17, and a subsidy of approximately 20 percent of the minimum wage for old-age population. The eligibility for these programs is based on a poverty score index computed by the Colombian government, and the programs' eligibility threshold changes by program. An important feature of the poverty score index is that it does not depend on whether the person has a formal-sector job, or whether the person is eligible for the transition system. Consequently, other assistance programs do not offset the cohort differences caused by eligibility for the transition system.

1.3 Pension incentives and formal-sector labor supply

To understand the incentives that workers consider when making their labor supply decisions, I present a model of workers' decisions with respect to retirement and formal-sector participation. In this model, a representative worker chooses between retiring and searching for a job, given a defined-benefit pension plan and a labor market with an informal sector.

The representative worker lives for $T - a_0 + 1$ periods, indexed by $a = a_0, \ldots, T$. Each period, the worker chooses whether to retire and leave the labor market permanently. If he retires and is eligible for retirement benefits, he receives a fraction b of the wage in the formal sector w^f along with other benefits valued θ^r (e.g., health care). If he retires and is not eligible for retirement benefits, he gets zero income. Thus, the retiree's earnings at age a are $e_a(\tau_{a-1})\, bw^f$. The variable $e_a(\tau_{a-1})$ is an indicator of the worker's pension eligibility, where τ_{a-1} stands for the number of periods the worker has worked in the formal sector. To be entitled to retirement benefits, the worker is required to work for at least τ^* periods in the formal sector and be at least R periods old, and thus $e_a(\tau_{a-1}) = 1\{\tau_{a-1} \geq \tau^*\} \cdot 1\{a \geq R\}$.

If the worker does not retire, he draws a random shock of searching for a formal-sector job $\psi_a \in \mathbb{R}$, which follows an i.i.d distribution with cumulative distribution $G(\cdot)$. ψ_a is measured in utility units and it is used as a catch-all variable summarizing the relative cost of searching and workers' preferences for formal-sector jobs. After drawing ψ_a, the worker decides between working in the formal sector and working in the informal sector. If he chooses to work in the formal sector, he will receive a wage w^f along with mandated benefits (valued θ^f), will face the utility shock ψ_a, and will be liable to pay the pension tax rate t^{nom}. Additionally, his cumulative number of periods with a contribution to the pension system will increase by one period (so $\tau_a = \tau_{a-1} + h_a$, where h_a is an indicator of whether the worker searches for a formal-sector job). If he chooses to work in the informal sector, he will receive a wage w^i (I assume that $w^i \leq w^f (1 - t^{nom})$ and $\theta^r \geq \theta^f$). For simplicity, I assume that the worker loses his job at the end of the period, and that he cannot save. In Chapter 2 I show that the model implications are robust to more general assumptions.

Since the worker does not save, so his consumption per period is equal to his income. Let r_a denote an indicator of whether the worker retires at the beginning of period a. Given τ_{a-1}, the worker's problem at the beginning of period a is

$$v_a (\tau_{a-1}) = \max_{r_a \in \{0,1\}} \{v_a^w (\tau_{a-1}), v_a^r (\tau_{a-1})\} \tag{1.1}$$

where

$$\begin{aligned} v_a^w (\tau_{a-1}) =& \mathbb{E} \max_{h_a \in \{0,1\}} \Big\{ u\left(w^i\right) + \beta v_{a+1} (\tau_{a-1}), \\ & \qquad u\left(w^f \left(1 - t^{nom}\right)\right) + \theta^f - \psi_a + \beta v_{a+1} (\tau_{a-1} + 1) \Big\} \\ v_a^r (\tau_{a-1}) =& u\left(e_a (\tau_{a-1}) b w^f\right) + e_a (\tau_{a-1}) \theta^r + \beta v_{a+1}^r (\tau_{a-1}) \end{aligned}$$

and $\tau_{a_0-1} = 0$. In the definitions above, $u(c)$ is the worker's utility over current consumption, which I assume to be continuous, strictly increasing, concave, and state-independent; $0 < \beta < 1$ is the discount factor, and $v_{T+1}^w (\tau_T) = v_{T+1}^r (\tau_T) = 0$.

The model encompasses the two common views in the literature about the incentives for workers to work in the informal sector (Gerard and Gonzaga, 2014). First, workers may choose to work in the informal sector because the perceived gains from formal-sector jobs are low. In the model, low gains from searching are represented by a low formal-to-informal wage gap and a low valuation of the mandated benefits provided by a formal-sector job (i.e., low θ^f and β). Second, workers may choose to work in the informal sector because finding a formal-sector job is difficult due to labor market rigidities and other structural characteristics (e.g., preferences for independent work). In the model, less favorable labor market opportunities are represented by a distribution of search costs with a heavier right tail. When $G(\psi)$ has a heavy right tail, it is likely that the worker draws a large value of ψ, high enough to offset the gains from working in the formal sector. It is common to have these two forces interact and reinforce each other. For example, workers with narrower wage gaps may also face higher search costs, further reducing the incentive to work in the formal sector.

1.3.1 Retirement and formal-sector participation decisions

The labor supply plan that solves the worker's problem can be obtained by backward induction. Given the value function, the worker's labor supply and retirement decisions can be obtained in a two-stage procedure. In the first stage, the worker finds the optimal plan for searching for a formal-sector job and the value function from working. In the second stage, the worker compares the value function from working with the value function from retiring, and determines the optimal retirement decision policy.

In the first stage, given a realization of the search cost ψ_a, the worker searches for a job in the formal sector as long as the gains from the search are greater than the costs. Thus, the worker searches for a formal-sector job (sets $h_a = 1$) if

$$\bar{u}_a(\tau_{a-1}) = \tilde{u} + \beta \Delta v_{a+1}(\tau_{a-1}+1) \geq \psi_a. \tag{1.2}$$

where $\tilde{u}$ and $\Delta v_{a+1}(\tau_{a-1}+1)$ are defined as

$$\tilde{u} = u\left(w^f\left(1-t^{nom}\right)\right) + \theta^f - u\left(w^i\right)$$
$$\Delta v_{a+1}(\tau_{a-1}+1) = v_{a+1}(\tau_{a-1}+1) - v_{a+1}(\tau_{a-1}).$$

In inequality (1.2), $\bar{u}_a(\tau_{a-1})$ summarizes the gains from working in a formal-sector job. The first term represents the short-run gains, that is, the utility gains determined by the differences in wages in both sectors along with the mandated benefits. The second term represents the long-run gains, that is, the gain that one additional period of working in the formal sector has on the likelihood that the worker will receive pension benefits in the future.

From inequality (1.2), the ex ante probability that a worker works in the formal sector is

$$P\left(h_a = 1 \mid \tau_{a-1}\right) = G\left(\bar{u}_a\left(\tau_{a-1}\right)\right), \tag{1.3}$$

and the value function from working is

$$\begin{aligned} v_a^w\left(\tau_{a-1}\right) = u\left(w^i\right) + v_{a+1}\left(\tau_{a-1}\right) \\ + G\left(\bar{u}_a\left(\tau_{a-1}\right)\right)\mathbb{E}\left(\bar{u}_a\left(\tau_{a-1}\right) - \psi_a \middle| \psi_a \leq \bar{u}_a\left(\tau_{a-1}\right)\right). \end{aligned} \tag{1.4}$$

In the second stage, the worker retires if the value function from retiring is greater than the value function from working. Thus, the worker retires (sets $r_a = 1$) if

$$v_a^r\left(\tau_{a-1}\right) \geq v_a^w\left(\tau_{a-1}\right). \tag{1.5}$$

1.3.2 Model implications

The model provides four useful predictions that contribute to understand the empirical results of the paper. The first three predictions are discussed in detail in the Appendix A.1.

The first prediction of the model is that, when the replacement rate equals one, the worker retires as soon as he meets the qualifying conditions. The retiree receives the wage in the formal sector as pension and does not have to pay the search cost. Since $b = 1$ is the effective rate faced by Colombian low-wage workers, the result is consistent with the patterns reported in Section 1.2.2, where workers retire as soon as they meet the minimum requirements of age and years of contribution.[9] Values of b lower than one may delay the retirement decision, depending on the value function conditional on working.

The second prediction of the model is that the long-run gains from a formal-sector job are heterogeneous and depend on the worker's employment

[9] An alternative explanation is that the worker is myopic or information constrained. If so, he may take the requirement conditions as target values regardless of the incentive to delay his retirement. In Chapter 2 I develop a version of the model in which the minimum retirement age is exogenous, and the predictions hold.

history (τ_{a-1}). Equations (1.2) and (1.3) imply that the worker's search for formal-sector jobs depends on the long-run gains from working in the formal sector. However, not all workers have the same long-run gains. Workers who cannot accumulate enough years to meet the vesting requirement will not receive pension benefits; therefore their long-run gains from a formal-sector job are zero ($\Delta v_{a+1}(\tau_{a-1}+1)=0$). Similarly, workers who already met the vesting requirement do not have long-run gains from working an extra period in the formal sector. For the remaining workers, the long-run gains from working in the formal sector are positive. Because the probability of working in the formal sector is an increasing function of the long-run gains from a formal-sector job, the result implies that workers with positive long-run gains search more actively for formal-sector jobs. Nonetheless, some workers with no long-run gains continue to work in the formal sector, but their decision is motivated by short-run gains only.

The third prediction of the model is that a change in the minimum retirement age R or the vesting period τ^* affects formal-sector labor supply. The effect of an increase in R on formal-sector labor supply is negative, since it reduces the long-run gains from working in the formal sector. In contrast, the effect of an increase in τ^* on formal-sector labor supply is ambiguous. The increase in τ^* shifts the long-run gains from working in the formal sector to the right, generating two opposite effects depending on the worker's employment history. On the one hand, workers who are close to meeting the new vesting requirement increase their search efforts to reach the new threshold. On the other hand, workers with a few existing years of contribution reduce their search efforts because it is unlikely that they will meet the new requirement.

Finally, the fourth prediction of the model is that the magnitude of the response to changes in the qualifying conditions for retirement depends on the worker's labor market opportunities. The response is smaller in labor markets with low formality rates (a low value of $\tilde{u}$ and a search cost distribution with a heavy right tail), and labor markets with high formality rates (a high value of $\tilde{u}$ and a search cost distribution with a light right tail). When $P(h_a=1 \mid \tau_{a-1}) \to 0$, workers cannot reach the vesting requirement and the long-run gains from searching are zero. When $P(h_a = 1 \mid \tau_{a-1}) \to 1$, workers always reach the vesting requirement and the long-run gains from searching for formal-sector jobs are zero. The effect of changes in the qualifying conditions is concentrated among workers who struggle to meet the vesting requirement but for whom reaching this threshold is still possible.

Figure 1.4 shows the expected value function ($v_{a+1}(\tau_a)$, top) and the probability of searching for a formal-sector job ($G(\bar{u}_a(\tau_{a-1}))$, bottom) by years of contribution for two simulated cohorts aged $a = 50$. Both cohorts face the same labor market opportunities, but different defined-benefit pension plans. One plan sets a minimum retirement age of $R = 60$ years, a vesting period of $\tau^* = 20$ years, and a replacement rate of $b = 1$. The other plan sets $R = 62$, $\tau^* = 25$ and $b = 1$.[10] Since $b = 1$ in both plans, workers facing both schemes retire as soon as they meet the requirements. For both cohorts, the long-run gains from working one more period in the formal sector are higher in years just below the vesting period, as working in the formal sector increases more the likelihood of securing pension benefits. As a result, the probability of working in the formal sector is higher in years just below the vesting period. An increase of the minimum qualifying conditions shifts the expected value function to the right (workers have to work more time to reach the vesting period) and reduces its level (workers receive the pension benefits for less time).

The probability of working in a formal-sector job given the two pension plans is presented in the bottom panel of Figure 1.4. The figure shows that an increase in the minimum qualifying conditions has a heterogeneous effect on the formal-sector labor supply, depending on the years of formal-sector experience. Workers with a few years of experience are not sensitive to different qualifying conditions, as they are too far from reaching the vesting requirements. Similarly, workers with many years of formal-sector experience are not sensitive to different qualifying conditions, as they already secured their pension benefits. For the rest of workers, the increase of the minimum

[10]In addition, I assume that workers' work from $a_0 = 20$ up to $T = 75$ years and their utility is linear. I also assume that $w^f(1 - t^{nom}) = 1.2$, $w^i = 1$, $\theta^r = \theta^f = 0$, $\psi \overset{i.i.d}{\sim} \mathcal{U}(0, 0.5)$, and $\beta = \frac{1}{1.05}$.

qualifying conditions has two types of effects on formal-sector employment. On the one hand, harder qualifying conditions discourage workers with a few periods of formal-sector experience, given the difficulty in reaching the new vesting requirement. On the other hand, more difficult qualifying conditions encourage workers approaching the new vesting requirement to search for formal-sector jobs, given that they are required to contribute additional years to reach the new vesting period. The sign of the overall effect of changes in minimum qualifying conditions on formal-sector labor supply is ambiguous and depends on the distribution of the workers' formal-sector experience.

1.3.3 General equilibrium

The model shows that future pension benefits create incentives for workers to work in the formal sector. In general equilibrium, though, changes in the long-run gains from working in the formal sector should be offset by changes in wages. Because future pension benefits are attractive to workers, workers would be willing to give up part of their wage in order to get the long-run gains from a formal-sector job (Summers, 1989).

The previous observation implies that, in equilibrium, the wage gap should exhibit an inverse pattern to that observed in Figure 1.4. Changes in qualifying conditions would be reflected in wages, leaving formal employment unchanged. However, the result requires that the extra benefits can be passed on to workers by way of lower wages, which are determined by the wage-setting process (Saez, Matsaganis, and Tsakloglou, 2012). Institutional factors such as minimum wage laws, search based on posted earnings, unobservable employment history, and pay fairness norms may prevent firms from setting differential wages among workers. If firms are not able to set a different wage scale for workers who do similar work, the response of the formal-sector labor supply changes the wage gap for all workers. As a result, the comparative statics of changes in the minimum retirement age and the vesting period exhibit similar patterns to the presented in Figure 1.4.

1.4 Data and empirical approach

1.4.1 Data

To measure the effects of changes in qualifying conditions on formal-sector labor supply, I combine two new sources of confidential data. The first source is the microdata from the long-form questionnaire of the Colombian Census of 2005. This is a cross-sectional dataset including information about labor market outcomes, pension and health care coverage, and demographic and household characteristics. The second source is the PILA dataset of 2011, an administrative dataset that collects information on all workers and earnings in the formal sector.

These two datasets have limitations but are highly complementary: I am able to analyze factors with the Census dataset that I am not able to analyze with the PILA dataset, and vice versa. In particular, the Census dataset does not include information about workers' earnings, while the PILA dataset does not include information about informal employment or demographic characteristics. Moreover, using both datasets, I am able to study the response of the formal-sector labor supply to changes in the pension incentives as the workers age.

Neither the Census nor PILA datasets includes the worker's employment history. I complement the information from the Census and PILA datasets with the distribution of years of contributions from the Colombian household surveys.

Colombian Census (2005)

The long-form questionnaire of the Colombian Census for 2005 collates information from 2 million households and 9.7 million people, approximately 20 percent of Colombian households. The dataset includes date of birth (in months), demographic information, type of employment, contributions to the pension system, and health care system coverage. The information about date of birth is reliable since the interviews were carried out in person and the interviewer was able to verify the date of birth from the respondent's identification card. In the absence of an identification card, the date of birth was either provided by the respondent or inferred on the basis of the reported age. The birth date for 92 percent of the urban population was established on the basis of their identification card.

The sample used in this paper is based on people living in urban areas, with a known date of birth, and born up to four years before or after the date of eligibility for the transition system (April 1950 to April 1958 for men, and April 1955 to April 1963 for women). The final samples sizes are 129,061 for men and 178,990 for women.

PILA dataset (2011)

The PILA dataset is a new dataset designed to collect information from the system used by firms and independent workers to pay for mandated benefits. Since formal-sector workers must be covered by mandated benefits, the dataset collects information for all formal-sector workers, and includes identifiers for employer and employee, basic wage, job location, gender and date of birth (in days) of the employee. The worker's date of birth and gender are added by the Ministry of Health based on the employee's identification card number. This dataset also includes information about firm ownership (public or private) and type of worker (independent or employee). It includes approximately eight million employer-employee pairs per month.

I select information from the entire dataset about all private-sector employees between February and December 2011 (66 percent of total formal employment). Although the dataset incorporates all formal workers, there are some problems with the identification numbers for employers and employees. To avoid false transitions in and out of the dataset, I fill in job spells in cases where an employer-employee match is missing and where the dataset records the same match up to three months before and after. In addition, I drop employees who appear only once in the dataset.

The sample used in this paper is based on all workers who were born up to two years before and after the eligibility threshold (April 1952 to March 1956 for men, and April 1957 to March 1961 for women). The final sample sizes are 964,558 for men (about 88,000 observations per month), and 927,961 for women (about 84,000 observations per month).

Household Surveys (2006-2011)

The Colombian Household survey is the official source of employment statistics in Colombia. After a large methodological change in 2006, the dataset includes information about an individual's date of birth (in months), and coverage in terms of pension and contributory health care systems. The surveys also contain information about a worker's earnings and the number of years of contributions made (conditional on contributing). The main limitation of the household surveys for this study is the small sample size for the cohort of interest. The number of observations by birth month in a year is approximately 200 people, and only about 40 people report information regarding years of contributions.

The sample used in this paper is based on all urban workers born up to three years before and after the eligibility threshold. The final sample sizes are 12,222 for men and 19,139 for women.

1.4.2 Identification strategy

To identify the effect of pension incentives on the formal labor market outcomes, I use a two-stage approach. In the first stage, I use a regression discontinuity design (RD) to estimate the effect of harder qualifying conditions on labor market outcomes. These estimations provide evidence on the response of formal-sector labor supply to pension incentives, without making further assumptions about workers' earnings and expectations. In the second stage, I use additional assumptions to recover an estimate of the elasticity of the formal-sector labor supply with respect to the net-of-tax share.

Effect of harder qualifying conditions on labor market outcomes

To identify the causal link between future pension benefits and formal-sector labor supply, I use a sharp regression discontinuity design. Given a cross-sectional sample of population, I run regressions of the form

$$Y_i = \alpha_0 + \rho 1_{\{DOB_i \geq 0\}} + \sum_{k=1}^{K} \left(\alpha_k + \beta_k 1_{\{DOB_i \geq 0\}}\right) \cdot DOB_i^k + \varepsilon_i \tag{1.6}$$

where Y_i is an indicator of the formal-sector labor supply; DOB_i is the normalized date of birth of the individual ($DOB_i = 0$ corresponds to the cutoff for harder qualifying conditions); $1_{\{DOB_i \geq 0\}}$ is a treatment indicator equal to one for people born after the cutoff, who are the ones facing harder qualifying conditions; and $\sum_{k=1}^{K} \left(\alpha_k + \beta_k 1_{\{DOB_i \geq 0\}}\right) \cdot DOB_i^k$ is a control function. Based on the reported date of birth, the relevant cutoff for eligibility for the transition system is April 1954 for men and April 1959 for women. Workers born before those dates were eligible for retirement benefits with 1,000 weeks of contributions and at an age of 55 (women) and 60 (men), while workers born after those dates are required to retire two years later (older), and after contributing up to 300 additional weeks.

The identifying assumption in this setup is that unobserved determinants of the formal-sector labor supply evolve smoothly around the eligibility threshold. Under this assumption, ρ can be interpreted as the average effect of harder qualifying conditions on the formal-sector labor supply, defined as

$$\rho = \lim_{c \downarrow 0} \mathbb{E}\left(Y_i | \, DOB_i = c\right) - \lim_{c \uparrow 0} \mathbb{E}\left(Y_i | \, DOB_i = c\right). \tag{1.7}$$

(Imbens and Lemieux, 2008). However, as discussed in Section 1.3, the response of the formal-sector labor supply to changes in minimum qualifying conditions depends on the worker's age (a) and years of contribution (τ_{a-1}). Therefore, ρ corresponds to the weighted average of the effect by previous contributions, i.e.,

$$\rho = \int_{\tau_{a-1}} \left(\lim_{c \downarrow 0} \mathbb{E}\left(Y_i | \, DOB_i = c, \tau'\right) - \lim_{c \uparrow 0} \mathbb{E}\left(Y_i | \, DOB_i = c, \tau'\right) \right) dF_a\left(\tau'\right), \tag{1.8}$$

where $F_a\left(\tau\right)$ represents the distribution of years of contribution at age a.

Since the expected change in qualifying conditions has an ambiguous effect on the formal-sector labor supply, the sign of ρ is ambiguous. As discussed in Section 1.3, the expected effect is positive for workers who are a long way from reaching the new vesting threshold, while it is negative for workers near the new vesting requirement. The sign of the average effect depends on the specific distribution of the number of years of contributions in the population.

Although the distribution of the years of contribution is not observed in the data, the analysis in Section 1.3 provides useful insights about the expected sign and magnitude of ρ. First, ρ should increase with the worker's age, as the distribution of τ shifts toward higher values of τ as workers age, putting more weight on the positive effects. Second, ρ should be smaller (in absolute value) for groups of workers with a low probability of finding a formal-sector job. For them, the estimated average effect should be small since they have low long-run gains from searching and a right-skewed distribution of previous contributions. A similar explanation would apply to the result for workers with a high probability of finding a formal-sector job. Workers with a middle-range probability of finding formal-sector jobs are the most responsive to changes in the minimum qualifying conditions.

An additional assumption is required for the estimation of the effects of harder qualifying conditions on formal-sector labor supply for 2011. In 2011, the sample is restricted to the universe of formal-sector workers. Because of that, regression discontinuity estimates are based on counts of formal-sector workers instead of the size of the formal-sector employment relative to the entire population. The identification strategy assumes that the density of the population by birth date evolves smoothly around the eligibility threshold. If so, the estimates based on counts of formal-sector employees identify a change in formal-sector employment in response to harder qualifying conditions and not to a change in the population by birth date.

To estimate equation (1.6), I run regressions separately by gender, as the cohorts affected by the reform are different. I cluster the standard errors by date of birth in months to account for potential misspecification in the control function (Lee and Card, 2008). I also follow the standard practice of testing the sensitivity of the results to the choice of control functions and bandwidth.

Labor Supply Elasticity with Respect to the Net-of-Tax Share

I next measure the incentive effects of pension benefits on the labor supply by calculating the elasticity of labor supply with respect to the net-of-tax share of income (Liebman et al., 2009). This elasticity is a common measure of the efficiency costs of pension policies, as the deadweight loss of the pension tax is proportional to it (Feldstein and Liebman, 2002).

I estimate the elasticity with respect to the net-of-tax share, one minus the effective pension tax rate, along the formal-informal margin, defined as

$$\sigma = \frac{d \ln L_a^f}{d \ln \left(1 - t_a^{eff}\right)}, \tag{1.9}$$

where L_a^f is the formal-sector labor supply for workers of age a, and t_a^{eff} is the effective pension tax rate for workers of age a,

$$\begin{aligned} t_a^{eff} &= t^{nom} - \beta \frac{\mathbb{E}PW_{a+1}(\tau+1) - \mathbb{E}PW_{a+1}(\tau)}{w} \\ &= t^{nom} - \beta \frac{\Delta \mathbb{E}PW_{a+1}(\tau+1)}{w}. \end{aligned}$$

In the definition of t_a^{eff}, t^{nom} is the pension tax rate, $\mathbb{E}PW_a(\tau)$ stands for the expected pension wealth at age a and τ years of contribution, and w is the worker's wage. In Section 1.5.3, I present a detailed discussion of the procedure used to compute the net-of-tax share.

The net-of-tax share measures the net gains from working in the formal sector in the current period. The share takes into account the pension tax rate paid for a worker and the change in the expected pension wealth derived from working an additional period in the formal sector. Based on the results from Section 1.3, formal-sector labor supply is increasing in the long-run gains from working an additional period in the formal sector. Thus, the expected sign of σ is positive.

To estimate σ, I split the sample into groups characterized by different propensities to work in the formal sector (e.g., by education and region). For each group (denoted by X), I estimate the average change at the discontinuity of the formal-sector employment $\left(\Delta \ln L_{aX}^{f}\right)$ and compute the average change at the discontinuity of the net-of-tax share $\left(\Delta \ln \left(1 - t_{aX}^{eff}\right)\right)$. Then, I estimate σ by running the regression

$$\Delta \ln L_{aX}^{f} = \alpha_0 + \sigma \Delta \ln \left(1 - t_{aX}^{eff}\right) + \varepsilon_X. \tag{1.10}$$

In equation (1.10), two sources of variation identify σ: the variation induced by the change in the minimum qualifying conditions and the variation across groups with different labor market opportunities.

1.5 Estimation results

1.5.1 Identification checks

The identification strategy relies on the assumption that the unobserved determinants of formal-sector labor supply evolve smoothly around the eligibility threshold. This assumption could be undermined in at least two ways: First, workers who are likely to work in the formal sector could manipulate their date of birth in order to appear eligible for the program when they are, in fact, not eligible (McCrary, 2008). Second, the estimated effect of the policy could be confounded by changes in other covariates that might influence the outcome (Imbens and Lemieux, 2008). In this section, I assess these two potential ways in which the identification could be compromised.

I test the manipulation hypothesis by estimating the density of the total population by date of birth above and below the eligibility thresholds, and implementing the test statistic proposed by McCrary (2008).[11] The results are presented in the top panel of Table 1.3. The manipulation hypothesis implies that the estimated difference should be negative, as younger workers

[11]Because the census data are reported by birth month, I run the regressions grouped by date of birth in months and use a bandwidth of 48 months. In all specifications, I use a triangular kernel.

might change their documentation to appear eligible for the transition system. The estimated effect for men is positive and not significant, supporting the idea that men did not manipulate their date of birth to appear eligible for the transition system. In contrast, the estimated effect for women is positive and significant. Although the sign of the estimated effect for women is the opposite of the expected sign under the manipulation hypothesis, the results raise concerns that population characteristics may change sharply around the discontinuity.

To further test the potential for manipulation by women, I run the McCrary density test with a placebo discontinuity ranging from March 1949 to February 1960. The t-statistics for each month are presented in Figure 1.5. The t-statistics exhibit two-year cyclical patterns over time, where the largest (absolute) values occur around March and September. The cyclical pattern occurs for both men and women, and the significant effect in March 1959 also occurs in the density for men. Using the regression estimates of the change in the density for men and women born around March 1959, I test whether the changes at the boundary are equal for the density for men and women, and I fail to reject the null hypothesis (p-value 0.394). The results suggest that the sharp change observed in the density of population by date of birth for women is the result of demographic trends and it is not explained by the eligibility for the transition system. Nevertheless, the discontinuity for women suggests to be cautious when interpreting the results for women.

In addition to the manipulation tests, I look for discontinuities at the eligibility thresholds in other observable variables that might explain the worker's labor supply choice. The variables are indicators for whether individuals have a high school diploma or less, whether they report any disability, and whether they identify as members of an ethnic group (black or indigenous). These variables are predetermined by the time the policy change took place and are correlated with the likelihood that an individual has a formal-sector job. Thus, significant differences in these variables would suggest that there are other unobservable factors that may be driving the labor supply decisions around the discontinuity. The bottom panel of Table 1.3 indicates that there are no significant differences in any of these indicators for either men or women.

Taken together, the results in Table 1.3 and Figure 1.5 provide evidence supporting the assumption that other determinants of the formal-sector labor supply evolve smoothly around the eligibility threshold. Although the distribution by date of birth for women is not continuous around the eligibility threshold, the placebo test suggests that the change is caused by time trends other than changes in the pension eligibility. Nonetheless, the interpretation of the results for women must take into account this caveat.

1.5.2 Results

The results of estimating equation (1.6) are presented in Table 1.4, and the graphical analysis is presented in Figure 1.6.

The top panel of Table 1.4 presents regression discontinuity estimates of the effect of harder qualifying conditions on salaried-formal employment for 2005. I use an indicator of whether the person works as salaried-formal worker as the dependent variable.[12] Thus, the estimated effect is the average effect of harder qualifying conditions on salaried-formal employment rate. This specification is my preferred specification because it is more robust to changes in the population unrelated to workers' self-selection, a particular concern given the results from the identification checks for women. The middle and bottom panels of Table 1.4 show the regression discontinuity estimates for the log of the number of salaried-formal workers for 2005 and 2011. The two panels have the advantage of being comparable over time. In all regressions, I use a quadratic polynomial in date of birth as a control function to account for potential non linearities in the formal-sector employment rate, and I use a bandwidth of 48 months for 2005, and of 730 days for 2011.[13]

The results in Table 1.4 show that Colombian workers actively responded to changes in the pension incentives. For men, the estimated effect is signif-

[12]In 2005, I define a person as a salaried-formal worker when the person worked as a salaried employee, contributed to the pension system, and was covered by the contributory health care system.

[13]The Imbens and Kalyanaraman (2012) optimal bandwidth for the 2005 regression is 55 months.

icant and changes over time. In 2005, the average effect of harder qualifying conditions decreased the salaried-formal employment rate by 2.6 percentage points (on a base of 18 percent). The effect is confirmed by the specification that uses the number of salaried-formal workers for 2005 as dependent variable (panel B of Table 1.4). The regression discontinuity estimates show that the increase in the number of salaried-formal workers at the discontinuity is negative 12 percent. In 2011, the estimated average effect on salaried-formal employment for men is positive and significant, implying an increase of 6.8 percent. The results are robust to the definition of formal worker, and to the choice of control functions, bandwidth, estimators, and controls (Tables A.1 to A.3 in the Appendix A).

The results for men are consistent with the framework presented above, in which the average effect of harder qualifying conditions depends on the worker's age. Table 1.4 presents information about the distribution of the years of contribution for workers born around the eligibility threshold, based on the household surveys for 2006 and 2011.[14] Since workers accumulate more years of experience in the formal sector as they age, the distribution of years of contribution is more concentrated on values above 20 years in 2011 than in 2005. As a result, the average effect for 2011 should be greater than the average effect for 2005, given that fewer eligible workers have long-run incentives to search for formal-sector jobs, as they already met the vesting requirement.

The results for women are intriguing. The 2005 estimates do not show any sizable or significant response. For 2011, however, Table A.2 shows significant results, depending on the specification. Since the 2011 results are not normalized by the population by date of birth, it is not possible to disentangle the potential effect of changes in the policy from the documented changes in the total population around the discontinuity. One explanation for the lack of response by women is that the transition system required workers to have already contributed to the pension system by 1994. This condition limited the applicability of the reform for women because of their relatively low labor force participation prior to this time (62 percent from

[14]The distribution is conditional on making contributions.

1984 to 1993).[15]

General equilibrium. To address the general equilibrium response to changes in pension incentives, I estimate the effect of harder qualifying conditions on the wages of formal-sector workers in 2011. If firms are able to set different wages between workers, wages offset part of the long-run gains from a formal-sector job. Therefore, the expected sign of the average effect of harder qualifying conditions on wages is the opposite to the sign of the average effect on employment. The estimates for the formal-sector wages are presented in the top panel of Table 1.5. The average effect of harder qualifying conditions on formal-sector wages is about negative 3 percent for men and is not significant for women (columns (1) and (5)). Since the effect on employment is positive in 2011, the result on average wages suggests that part of the workers additional search effort is offset by a change in wages.

To understand the sources of the aggregate results, I estimate the average effect of harder qualifying conditions on formal-sector wages and employment by wage range. The results are presented in the bottom panel of Table 1.5. Consistent with the analytical framework used here, low-wage men are the most responsive to changes in pension incentives. This occurs for two reasons. First, low-wage workers are more likely to find the minimum qualifying conditions binding. Second, the replacement rate for low-wage workers is close to one. As a result, they do not have additional long-run gains from working in the formal sector once they meet the requirements. The response from women is not significant.

The top panel of Table 1.5 presents the average effect of harder qualifying conditions on formal-sector wages by wage range. For men, the estimated effects are small and not significant. The difference relative to the aggregate results is driven by a composition effect, as the number of workers earning the minimum wage is larger for younger workers (panel B of Table 1.5). Because the change of the minimum wage around the discontinuity is zero, the average effect on wages goes down. Thus, the results indicate that the

[15] Between 1984 and 1993, the labor force participation rate for men around the discontinuity threshold was 97 percent.

impact of the policy change on wages was limited.

Nonetheless, the results presented in Table 1.5 do not rule out the possibility that changes in pension benefits are offset by changes in wages. The regression discontinuity estimates are intended to identify differential changes in the wages around the eligibility threshold. If the response in wages is associated with spillover effects, the estimates presented above are a lower bound of the actual response of the formal-sector labor supply to pension incentives.

Composition effects. I complement the analysis by testing the effect of pension incentives on the composition of the labor force. Based on information from 2005, I run versions of equation (1.6) for indicators of whether the worker is self-employed, whether the worker works as a salaried-informal worker, and whether the worker is in the labor force.

The estimation results are presented in Table 1.6, while the graphical evidence for men is presented in Figure 1.6. The reduction in the salaried-formal employment for men is associated with increases in informal-sector employment, in particular self-employment. The regression discontinuity estimate for the self-employment indicator is of the same magnitude but opposite sign as that of the estimate for the salaried-formal employment indicator. In contrast, the estimates using salaried-informal and labor force participation indicators as dependent variables are not significant. For women, there is no significant response in labor force participation or type of employment. Similar results have been noted in the literature concerned with the effect of mandated pension benefits on formal-sector labor supply. For instance, Almeida and Carneiro (2012) found that higher mandated benefits with no wage adjustment generate an incentive for Brazilian self-employed workers to switch to salaried-formal jobs.

Heterogeneity analysis. In this section, I analyze the differential effect of harder qualifying conditions on formal-sector labor supply for different groups. Because not all groups exhibit the same propensity to work in the formal sector, the group analysis provides evidence on the mechanisms driving the aggregate results.

I estimate the response of the formal-sector labor supply to changes in pension incentives for subsamples. To estimate this response, I group workers according to three demographic characteristics: educational attainment, household composition (e.g., presence of a spouse in the household), and region. In what follows, I present the results for men. The results for women are not significant and may be affected by changes in the distribution of population by date of birth. Due to data availability, I present the results for educational attainment and household characteristics for 2005, and the regional results for 2005 and 2011.

The first set of results is for workers grouped according to educational attainment. Less-educated workers are more likely to react to pension incentives for two reasons. First, these workers face higher replacement rates with no incentives to contribute beyond the vesting period. Second, they face lower formal-sector employment rates, which makes the condition for minimum years of contribution binding.

The estimation results show that the effect of harder qualifying conditions is concentrated among workers with secondary education (Table 1.7).[16] For workers with secondary education, harder qualifying conditions reduced the salaried-formal employment rate by 9 percentage points (on a 21 percent basis). In contrast, the estimated effects for workers with primary or post-secondary education are smaller and not significant. The third column of Table 1.7 shows the average salaried-formal employment rate by educational attainment. Consistent with the theoretical framework set out in this paper, workers with low or high informality rates are less responsive to changes in pension benefits.

The second set of results is for workers grouped according to the composition of the household. I analyze the response of workers in households with different incentives to search for formal-sector jobs. The samples are defined

[16] My implicit assumption is that workers do not change their schooling as response to the change in the pension qualifying conditions. Since men at the eligibility cutoff were 40 years old when the reform took place, this assumption seems reasonable and is consistent with the evidence presented in the identification checks (Table 1.3).

according to the person's marital status and whether the person is living in a household with only one member in the labor force. Married men and men living in a household with only one member in the labor force should respond more actively to harder qualifying conditions. First, men tend to get married to younger women (the median difference is 5 years). Given that the survivor pension rate is 100 percent, the long-run benefits of getting a pension are higher for households with married couples. Second, men living in households with only one member in the labor force may have limited family support after retirement. A concern with this part of the analysis is that the variables used to select the samples are endogenous to the eligibility for the transition system. However, I find no evidence that household structure changes as result of harder qualifying conditions (Table 1.8).

Table 1.9 reports the results for the different subsamples. The effect of pension incentives varies systematically depending on household characteristics. The response is concentrated among married men, and among households where there is only one member in the labor force.

The third set of results is for workers grouped according to region. Institutional factors and economic development generate differential formal-sector patterns by region (La Porta and Shleifer, 2014). The regional differences provide additional evidence on the relationship between the labor supply response to pension incentives and the labor market opportunities.

Table 1.10 reports the regression discontinuity estimates by region for 2005 and 2011. I group workers based on their departments' (provinces') GDP per capita excluding oil. The developed departments are Bogota-Cundinamarca, Antioquia, and Valle, and the developing departments comprise the rest of the country. The developed regions represent about 60 percent of the total GDP and 45 percent of total population in 2005. The average response to changes in the pension benefits is large and significant for developed regions, which offer most of the formal-sector employment.

In summary, the results presented in this section support the view that the formal-sector labor supply responds to pension incentives. The estimated average responses of formal-sector labor supply to harder qualifying conditions are heterogeneous and depend on labor market opportunities for the worker. The effect is concentrated among workers for whom the minimum qualifying conditions for retirement are binding, workers having higher expected pension wealth, and workers in households with only one member in the labor force.

1.5.3 Elasticity of formal-sector labor supply with respect to the net-of-tax share

To compute the elasticity of the formal-sector labor supply with respect to the net-of-tax share (σ), first I compute the average change in the net-of-tax share at the discontinuity for selected samples. Next, I recover the elasticity by regressing the estimates of average changes in formal-sector employment on average changes in the net-of-tax share.

To estimate σ, first I compute the net-of-tax share for subsamples of workers with different propensities to work in the formal sector. These subsamples are defined according to region and educational attainment for 2005, and region and wage range for 2011 (12 groups).[17] For each subsample (denoted by X), I compute the average change in the net-of-tax share at the discontinuity. To do this, I construct a grid for the expected pension wealth for every combination of age a and years of contribution τ, $\mathbb{E}_X PW_a(\tau)$. I assume that the worker will retire as soon as he meets the conditions for retirement, and that he will enjoy the pension benefits until age 80. The conditions and benefits that the worker receives after retirement are defined by the pension system. If the worker does not meet the retirement conditions by age 65, he will ask for a refund of his contributions to date. For the refund of contributions, I assume that the average contribution rate of a worker over his lifetime is 10 percent, as the pension contribution rate before 1994 was 6.5 percent of the worker's wage. If the worker does not retire, the worker will work an additional period in the formal sector with probability

[17] For 2005, I grouped workers according to their place of residence (developed and developing regions) and their educational attainment (primary, secondary and postsecondary education) for 6 groups in total. For 2011, I grouped workers according to their place of work (developed and developing regions) and their wage range (1, 1-2, and 2+ times the minimum wage) for another 6 groups.

$p_X(a)$.

Next, I compute the change in the log net-of-tax share at the discontinuity as

$$\begin{aligned}\Delta \ln\left(1-t^{eff}_{a\tau' X}\right) &= \ln\left(1-t^{nom}+\frac{\beta\Delta\mathbb{E}_X PW^{SI}_{a+1}(\tau'+1)}{w_X}\right)\\ &\quad -\ln\left(1-t^{nom}+\frac{\beta\Delta\mathbb{E}_X PW^{T}_{a+1}(\tau'+1)}{w_X}\right).\end{aligned}$$

where the superscripts SI and T denote that the expected pension wealth is computed using the conditions of the transition and the social insurance systems. I assume a pension tax rate of 4 percent, the contribution paid by salaried-formal workers.[18] Given that the estimates of the changes in employment are observed for men in 2005 and 2011, I compute the change in the log net-of-tax share for workers at age 51 and 57 (the age of the eligible men at the cutoff in 2005 and 2011). Finally, using information about the distribution of the number of years of contribution for the group X at age a, $F_{aX}(\tau)$, I compute the average change of the log net-of-tax share along the formal-informal margin as

$$\Delta \ln\left(1-t^{eff}_{aX}\right)=\sum_{\tau'}\Delta \ln\left(1-t^{eff}_{a\tau' X}\right)dF_{aX}\left(\tau'\right).$$

In the calculation of $\Delta \ln\left(1-t^{eff}_{aX}\right)$, I estimate $p_X(a)$ from the 2005 census and $F_{aX}(\tau)$ from the household surveys of 2006 and 2011. Moreover, I assume that w_X is constant over time and I set it to twice the minimum wage for skilled workers and to the minimum wage for the other groups.[19]

For the second stage of the estimation of σ, I regress the changes in the formal-sector employment on the change of the net-of-tax share. Figure 1.8

[18] In Colombia, the pension tax rate for all workers is 16 percent of the monthly wage. Since for salaried workers the employer pays 12 percentage points, I am assuming that the employers cannot pass through the additional contribution to lower wages. This is likely the case for minimum wage workers. The results are not sensitive to changes in the pension tax rate.

[19] The skilled workers are workers with post-secondary education for 2005 and workers with wages above twice the minimum wage for 2011.

displays a scatterplot with the average changes in log employment (vertical axis) and in net-of-tax share (horizontal axis) at the discontinuity. The groups from 2005 and 2011 are represented by triangles and circles, respectively. Consistent with the predictions of the model, workers with stronger pension incentives along the formal-informal margin also exhibit stronger responses in their formal-sector labor supply. A linear regression on these points yields an estimated elasticity of $\sigma = 1.66$. The estimated elasticity is slightly larger than the values of the same regression when restricted to cross through the origin ($\sigma = 1.60$) and the median value of the elasticity by group ($\sigma = 1.47$). Regardless of the estimator used, the implied values of σ are estimated with low precision.

The implied value of σ is likely a lower bound of the actual elasticity for at least three reasons. First, the estimates of changes in the formal-sector labor supply do not account for spillover effects, for instance, offsetting effects of wages affecting workers born before and after the eligibility threshold. Second, because of the definition of the transition system, a fraction of the population could not take up the benefits (Section 1.2.2). Third, $\Delta \ln \left(1 - t_{aX}^{eff}\right)$ may over-estimate the actual change in the net-of-tax share along the formal-informal margin. In particular, $\Delta \ln \left(1 - t_{aX}^{eff}\right)$ would be smaller (and σ larger) if workers have a lower discount rate β or worker's utility function is concave (Stock and Wise, 1990).

1.6 Final remarks

In this paper, I show that workers take into account their future pension benefits when it comes to making their labor supply decisions. Using the Colombian pension system, I show that a change in future pension benefits generates a large shift between the formal-sector and informal-sector labor supply. In contrast, there is no effect on labor force participation. The response is heterogeneous and depends on the worker's age, employment history, and opportunities to find formal-sector jobs. Using additional assumptions, I obtain an elasticity of formal-sector labor supply with respect to the net-of-tax share of 1.7.

Although the estimation results cannot be generalized to other cohorts or to other countries, the results suggest that the behavioral response to pension incentives may be large. Workers' behavioral responses should be taken into account in the design of pension programs, as such responses may create large efficiency costs. In particular, pension programs that reduce the value of the expected pension benefits have a negative effect on formal-sector labor supply. From a fiscal perspective, the effect of such programs is twofold. On the revenue side, these programs reduce the revenue achieved by way of contributions to the pension system, since fewer workers contribute. On the expenditure side, these programs increase the future expenditure in assistance programs, since more retirees would claim non-contributory pension benefits.

Nevertheless, a comprehensive evaluation of pension programs must take into account other factors that may mitigate their efficiency costs. For example, the welfare gains from the insurance against consumption losses after retirement may be significant. Additionally, the overall effect of pension programs depends on which sector of the population is affected. For instance, non-contributory pension programs for workers with low opportunities of finding formal-sector jobs could be welfare enhancing. For these workers, the behavioral response is small and the extra gains from insurance may be large.

Table 1.1: Labor market composition and average wages, Colombia, 2011

	Composition (percent)			**Average wage to min. wage ratio**		
	High School or less	**Post Secondary**	**Total**	**High School or less**	**Post Secondary**	**Total**
Salaried-employed						
- *Formal*	37.7	65.2	47.7	1.4	2.8	2.1
- *Informal*	17.5	7.3	13.8	1.1	1.4	1.2
Self-employed						
- *Formal*	5.1	11.2	7.3	1.9	3.8	3.0
- *Informal*	39.4	16.2	31	1.3	2.3	1.5
Observations	76,920	38,786	115,706	76,920	38,786	115,706

Notes: The table reports the composition and average wages of urban workers aged 20 to 65 working at least 30 hours per week. To avoid the effect of outliers and misreported information in the wage distribution, I trim the top 1 percent of workers of the wage distribution, and workers with wages below 40 percent of the minimum wage. A formal worker is defined as a worker who is making contributions to the pension system and is covered by the contributory health care system. Source: Colombian Household Surveys, 2011

Table 1.2: General Pension System characteristics

	Transition	Social Insurance	Individual Account
Managed by	Colpensiones	Colpensiones	Private pension funds
Type of system	Defined benefit	Defined benefit	Defined contribution
Eligibility	Workers born before April 1959 (women) or April 1954 (men) with 750 weeks of contributions by July 2005.[†]	All public and private sector workers (including self-employed[††]) not eligible for the transition system.	
Qualifying conditions	Private sector workers: 55 years (women), 60 years (men) AND 1,000 weeks of contributions in any time. Public workers: 50 years (women), 55 years (men), AND 20 years of service.	All workers: 55 years (women), 60 years (men) AND 1,050 to 1,300[†††] weeks of contributions in any time. Starting in 2014, minimum age increased by two years to 57 for women and 62 for men.	All workers: Enough capital to buy an annuity of 1.1 minimum wages, OR 57 years (women), 62 years (men) and 1,150 weeks of contributions in any time for an annuity of a minimum wage.
Total contribution	16% of wage – 11.5% contribution, 4.5% for administrative fees and insurance. (Salaried workers: 12% paid for the employer - 4% paid for the employee. Self-employed workers pay 16%)		

Continues in next page.

General Pension System characteristics (continued)

	Transition	Social Insurance	Individual Account
Replacement rate	Function of length of contributions. From 65% to 85% (See Figure 1.2)	Function of length of contributions and wage. From 65% to 85% (See Figure 1.2).	It depends only on the accrued capital
Pension range	At least 1 Minimum wage	1-25 Minimum wages	At least 1 Minimum wage
Survivor benefits	100 percent	100 percent	100 percent
Contributions refund	Contributions adjusted by inflation (lump-sum payment)		Accrued capital + interest (lump-sum payment)
Coverage Statistics (2005) - Millions			
Total	5.67		5.95
1-2 Min. wage	5.22		5.08
Aged 45+	2.35		0.67
Retirees	0.82		0.02

Notes: † The limit of 750 weeks of contributions by July 2005 was introduced in 2005. †† Contributions for Self-employed workers become compulsory since January 2003. ††† Starting in 2003, the length of contributions needed to qualify for a pension increased gradually from 1,000 weeks in 2004 up to 1,300 weeks in 2015. Coverage statistics taken from the Superintendencia Financiera website. Source: Santa María, Steiner, Botero, Martinez, Millán, Arias, and Schutt (2010), Llano, Cardona, Guevara, Casas, Arias, and Cardozo (2013) and texts of the reforms.

Table 1.3: Identification checks, 2005

A: McCrary's density test

	Men	**Women**
Test Statistic	0.024	0.078
(Bandwidth 48 months)	[0.033]	[0.025]***
Observations	126,095	175,047

B: Balance tests (estimates scaled up by 100)

High School or less indicator	0.68	0.25
(Bandwidth 48 months)	[1.21]	[1.12]
Disability indicator	-0.21	0.48
(Bandwidth 48 months)	[1.03]	[0.71]
Ethnical minority indicator	-0.13	0.99
(Bandwidth 48 months)	[0.74]	[0.64]
Observations	78,655	110,626

Notes: The Table presents estimates for testing factors that affect the validity of the identification assumptions required for the regression discontinuity design described in Section 1.4.2. The top panel presents the estimation results by gender for the test proposed by McCrary (2008), to test potential discontinuities in the density of the running variable (population by date of birth). The bottom panel presents RD estimates for observable determinants of formal-employment and other predetermined variables, to gather evidence about other potential changes that may confound the estimated effect of the policy. Each cell reports an RD estimate based on a separate regression of a variable predetermined by the time of the introduction of the policy as dependent variable versus a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). The selected variables are indicator variables for whether the person's has a high school diploma or less, whether the person reports any disability, and whether the person identifies himself as a member of a ethnic group (black or indigenous). Regressions were computed using the IPUMS Colombian Census dataset. Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 1.4: RD estimation results, 2005 and 2011

	RD estimates		Dist. of years of contribution (%)					
	Men	Women	Men			Women		
			0-10	11-20	21+	0-10	11-20	21+
A: 2005 Results – Dependent variable: Salaried-formal indicator								
Harder qualifying conditions	-2.62	-0.18	20.2	39.8	40.0	27.6	36.7	35.7
(Bandwidth 48 months)	[1.28]**	[1.03]						
Observations	129,061	178,990						
Mean Dep. Variable (%)	18.1	15.7						
B: 2005 Results – Dependent variable: Log salaried-formal workers by date of birth in months								
Harder qualifying conditions	-11.9	8.53	20.2	39.8	40.0	27.6	36.7	35.7
(Bandwidth 48 months)	[8.65]	[7.60]						
Observations	15,349	20,616						
C: 2011 Results – Dependent variable: Log salaried-formal workers by date of birth in days								
Harder qualifying conditions	6.79	2.03	13.0	31.6	55.5	20.8	37.6	41.5
(Bandwidth 48 months)	[2.39]***	[2.17]						
Observations	964,558	927,691						

Notes: All estimates scaled up by 100. Each cell reports an RD estimate based on a separate regression of a labor market indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). Panel A includes the total population and uses as dependent variable an indicator variable of whether the person is a salaried worker making contributions to the pension system and being covered by the contributory health care system – so the RD estimate is an effect on the salaried-formal employment rate. Panels B and C report the RD estimates of regressions in which the dependent variable is the log number of salaried-formal workers for 2005 and 2011. Regressions were estimated using the Colombian Census long-form questionnaire dataset (2005) and the PILA dataset (2011). Standard errors clustered by date of birth (in months) in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. The distribution of years of contribution is conditional on making contributions, and it is based on the Household Surveys data of 2006 and 2011.

Table 1.5: RD estimation results for wages in the formal sector, 2011

	Men				Women			
	All	At W_m	1-2 W_m	2+ W_m	All	At W_m	1-2 W_m	2+ W_m
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: RD Estimates for log wages (estimates scaled up by 100)								
Harder qualifying conditions	-3.11	–	-0.25	-1.51	1.14	–	0.23	1.2
(Bandwidth 730 days)	[0.82]***		[0.56]	[1.46]	[1.51]		[0.83]	[1.85]
Observations	964,558	416,927	287,659	259,972	927,691	365,667	321,405	240,619
B: RD Estimates for log number of workers (estimates scaled up by 100)								
Harder qualifying conditions	6.79	10.51	7.22	0.39	2.03	0.98	1.95	4.99
(Bandwidth 730 days)	[2.39]***	[3.36]***	[2.03]***	[3.03]	[2.17]	[3.41]	[3.47]	[3.84]
Observations	964,558	416,927	287,659	259,972	927,691	365,667	321,405	240,619

Notes: Each cell reports an RD estimate based on a separate regression of a labor market indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). Panel A includes salaried-formal workers for 2011 and reports the RD estimates using as dependent variable the log monthly wage of formal workers. Columns (1) and (5) presents the results for the full sample, while columns (2) to (4) and (6) to (8) show the results for subsamples defined by wage range. By construction, the difference at the discontinuity for workers at the minimum wage is zero. Panel B reports the RD estimates of regressions in which the dependent variable is the log number of salaried-formal workers for 2011 following the same sample selections than panel A. Regressions were estimated using the PILA dataset. Standard errors clustered by date of birth (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 1.6: Estimation results for other labor market outcomes, 2005

A: RD estimates for other labor market outcomes – Men, 2005
Estimates scaled up by 100

	Lab. Force participation	Salaried formal	Salaried informal	Self-employed
Harder qualifying conditions	1.20	-2.62	0.21	2.49
(Bandwidth 48 months)	[1.39]	[1.28]**	[1.47]	[1.33]*
Observations	129,061	129,061	129,061	129,061
Mean dep. variable (%)	78.3	18.1	27.3	25.7

B: RD estimates for other labor market outcomes – Women, 2005
Estimates scaled up by 100

	Lab. Force participation	Salaried formal	Salaried informal	Self-employed
Harder qualifying conditions	-0.69	-0.18	-0.86	0.68
(Bandwidth 48 months)	[1.76]	[1.03]	[1.32]	[0.93]
Observations	178,990	178,990	178,990	178,990
Mean dep. variable (%)	49.7	15.7	18.9	10.9

Notes: Each cell reports an RD estimate based on a separate regression of a different labor market indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). The columns labeled salaried-formal present the baseline RD estimates presented in Table 1.4. The additional columns reports results of RD estimates for labor force participation, salaried-informal employment, and self-employment rate. Regressions were estimated using the Colombian Census long-form questionnaire dataset. Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table 1.7: Estimation results by educational attainment – Men, 2005

A: 2005 Results - Dependent variable: Salaried-formal employment indicator
(Estimates scaled up by 100)

	RD Estimate	Sal-formal emp. rate	Dist. of years of contribution (%) 0-10	11-20	21+
Primary	-0.16	10.6	29.3	35.6	35.1
(Bandwidth 48 months)	[1.58]				
Secondary	-9.32	21.2	26.3	40.8	33.0
(Bandwidth 48 months)	[3.38]***				
Post-Secondary	-0.82	38.4	8.0	41.5	50.5
(Bandwidth 48 months)	[2.91]				
Observations	129,061				
Mean dep. variable (%)	18.1				

Notes: The first column reports RD estimates based on a separate regression of the salaried-formal employment indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables and educational attainment. The cells report the average effect of harder qualifying conditions by educational attainment (primary or less, secondary or at least some secondary, and post-secondary). Regressions were estimated using the Colombian Census long-form questionnaire dataset. Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The distribution of years of contribution is conditional on making contributions, and it is based on the Household Surveys data of 2006 and 2011.

Table 1.8: RD results for indicators of household characteristics, 2005

(Estimates scaled up by 100)		
Dependent variable	Married	Only worker in household
Harder qualifying conditions	0.42	-2.08
(Bandwidth 48 months)	[1.44]	[1.49]
Observations	110,174	106,811
Mean dep. variable (%)	83.1	38.7

Notes: Each cell reports an RD estimate based on a separate regression of the household composition indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). The first column presents RD estimates using as dependent variable an indicator variable for marital status (1 if married, 0 otherwise). The second column restricts the sample to households with at least one person in the labor force, and estimates the model using as dependent variable an indicator for being the only member of the household in the labor force. Regressions were estimated using the Colombian Census long-form questionnaire dataset. Standard errors clustered by date of birth (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 1.9: RD results by household characteristics – Men, 2005

(Estimates scaled up by 100)

	RD Estimate	Sal-formal emp. rate	Dist. of years of contribution (%) 0-10	10-20	21+
A: Estimates for men with spouse in the household					
No Spouse	1.40	21.9	11.5	62.5	26.0
(Bandwidth 48 months)	[3.68]				
Spouse	-3.96	18.3	21.6	36.6	41.8
(Bandwidth 48 months)	[1.60]**				
Observations	110,174				
Mean dep. variable (%)	19.6				
B: Estimates for men in households with one or more members in the labor force					
More than one member in LF	-0.89	16.4	21.7	36.1	42.2
(Bandwidth 48 months)	[1.83]				
Only member in labor force	-6.23	20.3	15.9	50.7	33.4
(Bandwidth 48 months)	[1.91]***				
Observations	110,174				
Mean dep. variable (%)	20.5				

Notes: The first column of the Table reports an RD estimate based on a separate regression of the salaried-formal employment indicator on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). The top panel reports the results for the samples of married and unmarried men, while the bottom panel reports the results for the sample of workers who are the only member of the family in the labor force. Regressions were estimated using the Colombian Census long-form questionnaire dataset. Standard errors clustered by date of birth (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01. The distribution of years of contribution is conditional on making contributions, and it is based on the Household Surveys data of 2006 and 2011.

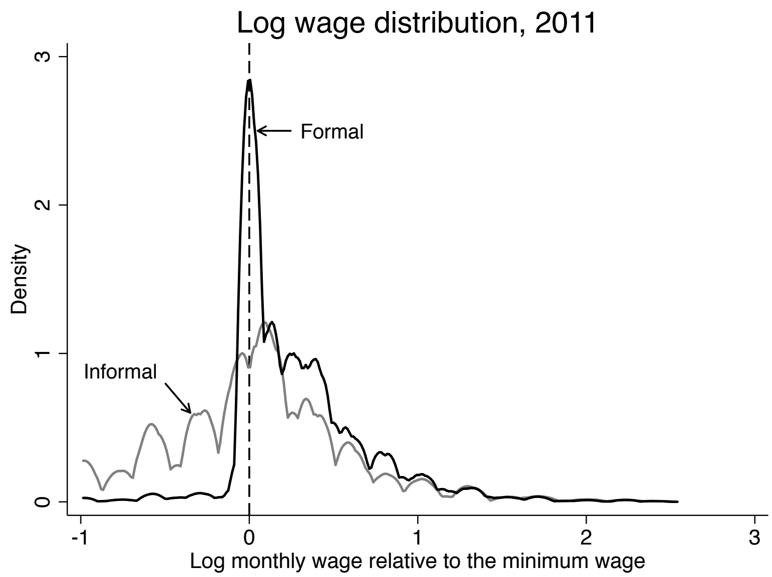
Table 1.10: Estimation results by region – Men, 2005 and 2011

(Estimates scaled up by 100)

	RD Estimate	Sal-formal emp. rate	Dist. of years of contribution (%) 0-10	10-20	21+
A: 2005 Results - Dependent variable: Salaried-formal employment indicator					
Developing regions	-0.80	13.3	16.7	38.9	44.4
(Bandwidth 48 months)	[1.66]				
Developed regions	-3.99	21.3	22.4	40.3	37.3
(Bandwidth 48 months)	[1.66]**				
Observations	129,061				
Mean dep. variable (%)	18.1				
B: 2011 Results - Dependent variable: log number of workers					
Developing regions	1.58	18.1	14.0	38.4	47.7
(Bandwidth 730 days)	[5.68]				
Developed regions	8.21	23.3	12.7	29.6	57.7
(Bandwidth 730 days)	[2.38]***				
Observations	964,558				

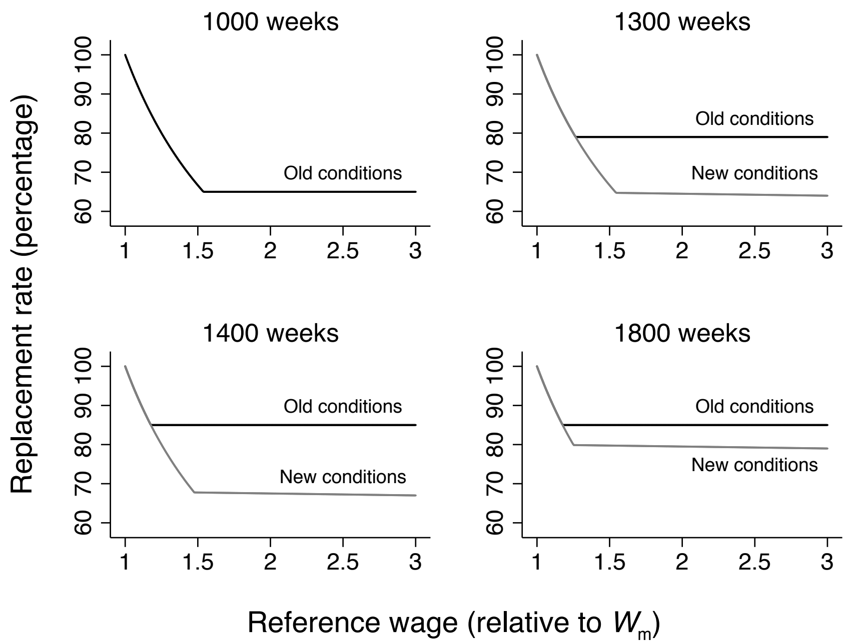
Notes: The first column reports RD estimates based on a separate regression of a salaried-formal employment variable on a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables by region (See equation (1.6)). The top panel presents results by region for 2005, while the bottom panel reports the results for 2011. I defined developed regions as the departments (provinces) with the highest GDP per capita excluding oil, namely, Bogota and Cundinamarca, Antioquia, and Valle, and less developed regions are the other provinces. Regressions were estimated using the Colombian Census long-form questionnaire (2005) and the PILA (2011) dataset. Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$. The salaried-formal employment rate for 2011 is based on Household Surveys data. The distribution of years of contribution is conditional on making contributions, and it is based on the Household Surveys data of 2006 and 2011.

Figure 1.1: Distribution of wages for workers with High School diploma or less, 2011



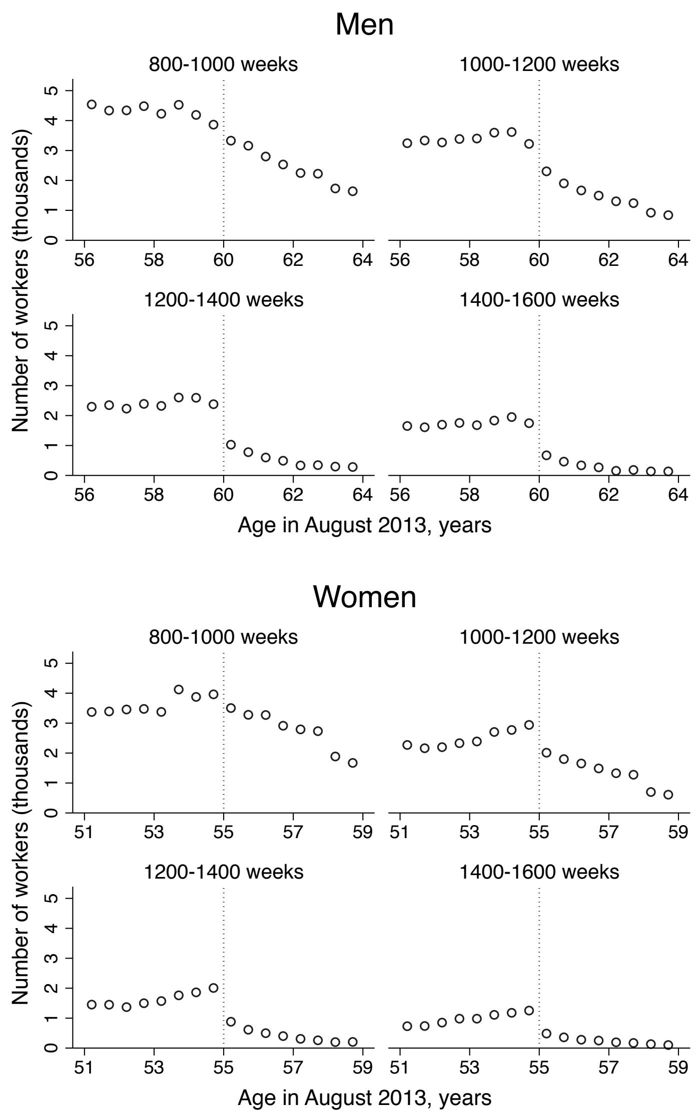
Notes: The figure displays kernel estimates of the density of the log monthly wage relative to the minimum wage for the formal (black line) and informal (grey line) sector. The selected sample includes all urban men and women aged 20 to 65, with High School diploma or less, working at least 30 hours per week. To minimize misreporting errors, I drop the top 1 percent wages and wages below 40 percent the minimum wage. Formal-sector workers are defined as workers who contributed to the pension and are covered by the contributory health care system.

Figure 1.2: Replacement rate for the defined-benefit systems by weeks of contributions



Notes: The figure displays the replacement rates as percentage of the reference wage for the social insurance (gray line) and the transition (black line) systems. Each panel represents the particular value of the formula defining the replacement rate by weeks of contributions. For the insurance system the vesting period is 1,300 weeks.

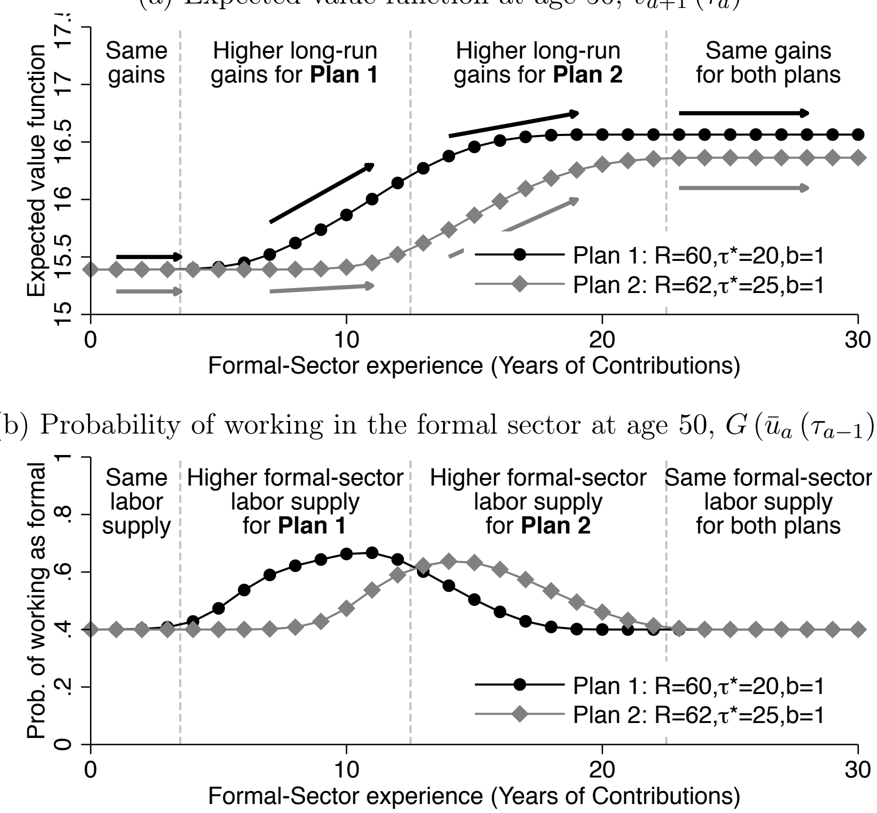
Figure 1.3: Distribution of workers by age and number of weeks



Notes: The figure presents the distribution by age, weeks of contribution and gender for non retired workers who have made contributions to the public pension system throughout their lifetime up to December 2013, based on Colpensiones administrative data. Once the workers claim the pension benefits they are excluded from the dataset. The reference date of birth is calculated relative to August 2013, as the expected processing time for awarding retirement benefits is four months.

Figure 1.4: Probability of working in the formal-sector at age $a = 50$

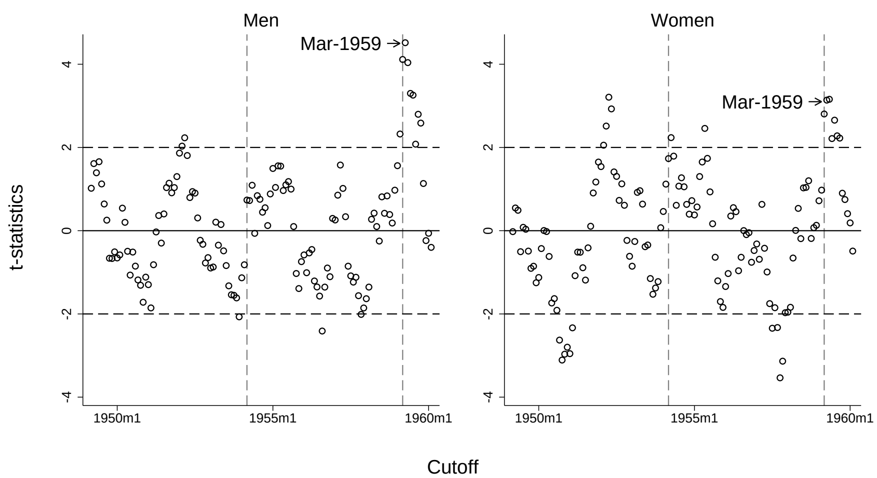
(a) Expected value function at age 50, $v_{a+1}(\tau_a)$



(b) Probability of working in the formal sector at age 50, $G(\bar{u}_a(\tau_{a-1}))$

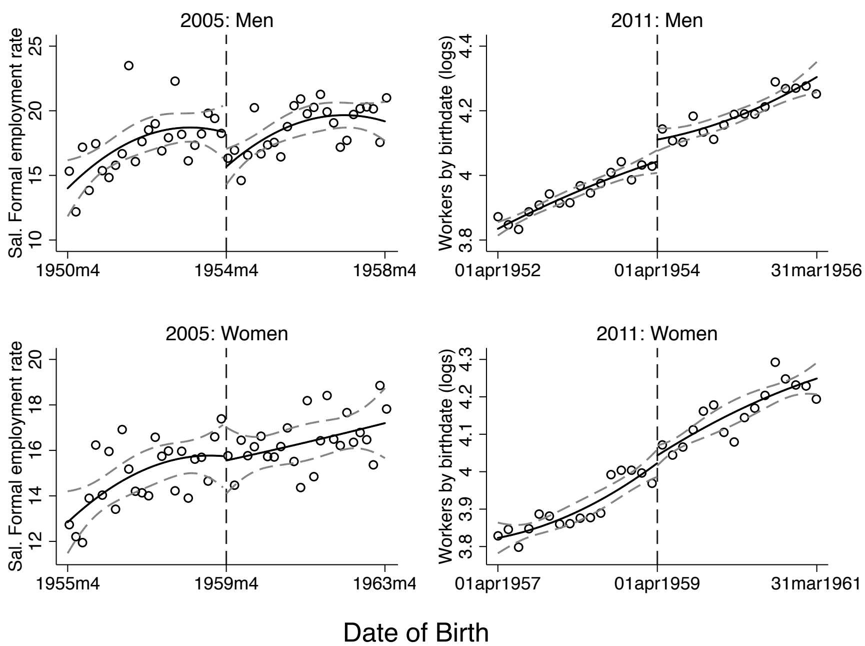
Notes: The figure shows an example of the expected value function ($v_{a+1}(\tau_a)$, top) and the probability of searching for a formal-sector job ($G(\bar{u}_a(\tau_{a-1}))$, bottom) by years of contribution for two simulated cohorts aged $a = 50$, given two defined-benefit pension plans. For the example, I assume that workers live from $a_0 = 20$ until $T = 75$ and their utility function is linear. I also assume that $w^f(1 - t^{nom}) = 1.2$, $w^i = 1$, $\theta^r = \theta^f = 0$, $\psi \overset{i.i.d}{\sim} \mathcal{U}(0, 0.5)$, and $\beta = \frac{1}{1.05}$. The pension plans are given by $R = 60$, $\tau^* = 20$ and $b = 1$ (labeled Plan 1) and $R = 62$, $\tau^* = 25$ and $b = 1$ (Plan 2). Panel (a) shows the workers' expected value function by pension plan. Increasing the minimum qualifying conditions reduces the workers' expected utility and changes the long-run gains from working in the formal sector (the slope of the expected value function). Panel (b) shows the probability of working in the formal sector by pension plan. Harder qualifying conditions shift to the right the long-run gains from working one more period in the formal sector, as it takes longer to reach the vesting period. The response of formal-sector labor supply to harder qualifying conditions is heterogeneous, and depends on the workers' formal sector experience.

Figure 1.5: Rolling t-statistics for testing the manipulation in date of birth, 2005



Notes: The figure displays the t-statistics using the test proposed by McCrary (2008) for testing discontinuities in the density of the running variable in the regression discontinuity setup. Each panel represents the value of the t-statistics changing the cutoff point, where the vertical dashed lines show the relevant cutoff dates for harder qualifying conditions in Colombia (April 1954 for men and April 1959 for women).

Figure 1.6: RD estimation results, 2005 and 2011

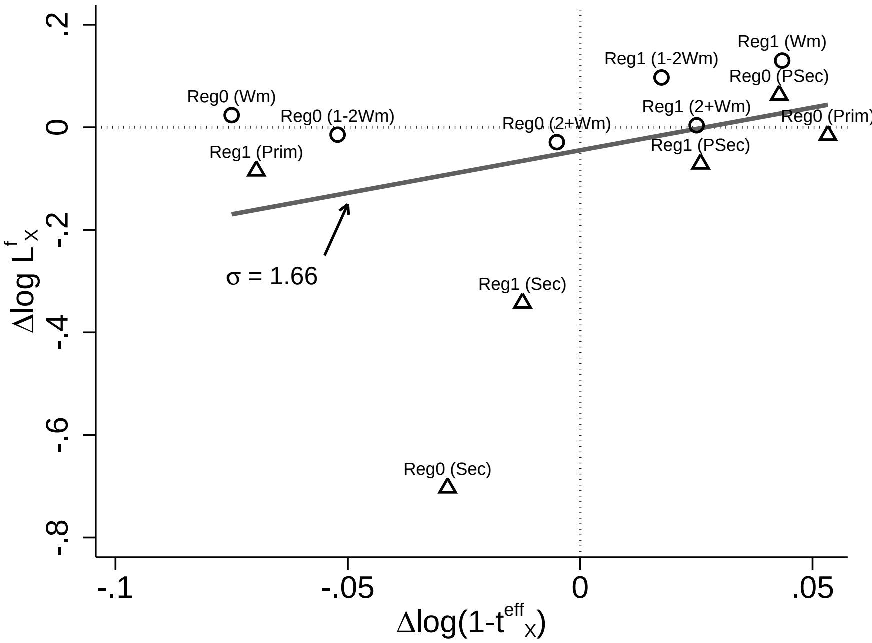


Notes: The figure presents the salaried-formal employment indicators by gender and date of birth. Each point represents the 2-month average of the salaried-formal employment rate by month in 2005 and the 47-days average of the log number of workers by age in 2011. The regression estimates on the graphs are based on the estimates reported in panels A and C of Table 1.4. Confidence bands are computed based on standard errors clustered by date of birth (in months).

Figure 1.7: Labor force participation, salaried-informal employment and self-employment rates for Men, 2005

Notes: The figure presents the labor force participation rate, salaried-informal employment and self-employment rate for men by date of birth. Each point represents the 2-month average of the specific labor market outcome. The regression estimates on the graphs are based on the quadratic fit of the microdata (Table 1.6) of the Colombian Census long-form questionnaire dataset. Confidence bands are computed based on standard errors clustered by date of birth (in months).

Figure 1.8: Elasticity of the formal-sector labor supply to changes in the net-of-tax share



Notes: The Figure displays the average change in the log net-of-tax share (horizontal axis), computed in Section 1.5.3, and the average change in the salaried-formal labor supply (vertical axis), derived from the results obtained in Section 1.5.2. Each point represents a combination of regions (developing and developed, denoted Reg0 and Reg1) and educational attainment (primary, secondary and post-secondary) or wage range (1, 1-2, 2+ Minimum wages). The regression slope corresponds to an estimate of the elasticity of the formal-sector labor supply with respect to the net-of-tax share along the formal-informal margin.

Chapter 2

A Life-Cycle Model for Formal-Sector Labor Supply

2.1 Introduction

In Chapter 1, I provide empirical evidence of the importance of pension programs in a worker's decision to participate in the formal (regulated) sector.[20] In this chapter, I propose a general life-cycle analysis to study the interaction between pension incentives and formal-sector labor supply before retirement. Using this analysis, I identify the gains from working in the formal sector and characterize their relationship with financial and pension wealth.

This chapter contributes to the literature on formal-sector labor supply by introducing the life-cycle component into the analysis. In the model, a representative agent's decision to participate in the formal sector relies on the gains from working in the formal-sector, his preferences for formal-sector jobs, and the labor market conditions that may prevent him from getting a formal-sector job (Gerard and Gonzaga, 2014; Meghir, Narita, and Robin, 2015; Rauch, 1991). This approach is consistent with available evidence for Latin American economies, which suggests that workers are mobile across sectors and that many informal-sector workers exhibit high levels of satisfaction with their job (Maloney, 1999, 2004; Perry et al., 2007). Compared with an informal-sector job, a formal-sector job generally pays higher wages

[20]In what follows, I define informal sector as the set of firms and workers that do not comply with government regulation, such as the payment of mandated contributions and taxes. A firm operating in the informal does not pay taxes (including payroll taxes), but it is subject to fines if it is inspected. Regarding workers, informal-sector workers do not pay contributions to the mandated benefit and insurance systems, but they are not covered by mandated benefits and other insurance included in the regulation.

and provides benefits in the form of insurance and future pension benefits. Nonetheless, the representative agent's decision depends on his own preferences for formal-sector jobs, his net valuation of the benefits from working in the formal sector, and the availability of formal sector jobs.

In the model setup, I explicitly account for the relative gains and costs from working in the formal sector. I specifically account for the gains from working in the formal sector by incorporating sector-specific wages and a defined-benefit pension plan that entitles the representative worker to pension benefits after retirement. With respect to the costs from working in the formal sector, I build on Gerard and Gonzaga (2014) and Eissa, Kleven, and Kreiner (2008) and introduce formal-sector participation shocks. These shocks are a catch-all variable that includes the worker's preferences for formal-sector jobs, the worker's net valuation of the benefits from working in the formal sector, and the availability of formal sector jobs.

The gains from working in the formal-sector are summarized by the threshold $\bar{u}$. This threshold represents the lifetime utility gains from working one more period in the formal sector. When the representative worker's utility gains are larger than the utility shock, he chooses to work in the formal sector. In this setup, $\bar{u}$ is a function of the worker's age, his potential earnings in both sectors, his financial wealth (assets), and his future pension benefits.

The model results are qualitatively similar in all cases, but they depend on the basic assumptions on the relationship between pension and financial wealth and the type of pension plan. When I assume that financial and pension wealth are perfect substitutes, I show that the short and long-run gains from working in the formal sector cannot be separated. Working in the formal sector increases the workers' lifetime income, as it increases present earnings and the future pension wealth. Then, consumption smoothing leads workers to increase their present and future consumption, having a positive impact on their utility in both the short and the long-run. In contrast, when I assume that workers cannot save, the short and long-run gains are separated. The short-run utility gains come from an increase in present earnings while the long-run utility gains come from an increase in the pension wealth. Finally, when I change the pension plan from a defined-benefit to a defined-contribution pension plan, I show that workers internalize the long-run gains from contributing to the pension system and reduce their financial wealth by the amount they contribute to the plan. As a result, in a defined-contribution system the gains from working in the formal sector come from the increase in present earnings only.

A direct implication of the model is that, when the pension system is a defined-benefit pension system, the change in the pension wealth is the main driver of the long-run gains from working in the formal sector. Nonetheless, the level of wealth determines the sensitivity of the workers to the utility gains from working in the formal sector. In particular, I show that when the per-period utility function is concave, the threshold $\bar{u}$ is a decreasing function of the level of financial wealth. Thus, when workers have more assets, an increase in lifetime income from working in the formal-sector has a smaller impact on lifetime utility.

The rest of the chapter is organized as follows. Section 2.2 describes workers' incentives to participate in the formal-sector. Sections 2.3 introduces the model setup, and characterizes formal-sector labor supply. Sections 2.4 and 2.5 present a version of the model where workers cannot save and a version of the model where workers lose their job with certain probability. Section 2.6 presents the results of a model with savings and pension wealth assuming a defined-contribution pension plan. Finally, Section 2.7 sets out the conclusions.

2.2 The environment

The model assumes a representative agent lives for three periods (young worker, adult worker, and retiree), indexed by $a \in \{1, 2, 3\}$. Every period, the agent chooses how much of his income to consume, how much to save, and whether to work in the formal or in the informal sector. The agent has a set of exogenous characteristics (e.g., education and ability), denoted by X, that will determine his wage in each sector and formal-sector participation shocks while working. I assume that the agent always finds a job in the sector he chooses, and that he loses his job by the end of the period.

Because the formal sector is the only one complying with labor regulation, working in either the formal or the informal sector entitles the agent to a different set of benefits and costs, summarized in Table 2.1. If he works in the formal sector, he receives a wage of $w_a^f(X)$ and pays mandatory contributions to the pension system and other insurance programs at rates t_p and t_c. Mandatory contributions entitle the worker to receive non-monetary benefits in the short-run (e.g., health care) and increase the working experience used to compute pension benefits in the long-run. If the agent works in the informal sector, he receives a wage of $w_a^i(X)$, does not pay payroll taxes, yet he may be eligible for social assistance programs in the short and long-run (e.g. public health care and a social pension). In what follows, I assume that the formal-sector wage is greater or equal than the informal-sector wage (i.e., $w_a^f(1-t_c-t_p) \geq w_a^i$), which is a common feature of informal labor markets in Latin America (Albrecht, Navarro, and Vroman, 2009; Meghir et al., 2015).

At the beginning of the first two periods, the agent draws an i.i.d. random utility shock $\psi_a \in \mathbb{R}$ with cumulative distribution $G(\cdot|X)$. ψ_a encompasses three unobservable determinants of the worker's participation in the formal sector frequently found in the literature. First, there is evidence suggesting that workers have a low valuation of the benefits provided by formal-sector jobs because they have access to social programs that substitute those benefits (e.g., public health care) (Camacho, Conover, and Hoyos, 2013; Galiani and Weinschelbaum, 2011; Levy, 2008). Second, workers may prefer to work in the informal-sector because informal-sector jobs are in line with their needs for job independence and time flexibility (Maloney, 1999, 2004). Finally, labor market frictions and regulations may prevent workers from finding formal-sector jobs (Joubert, 2015; Meghir et al., 2015; Ulyssea, 2010). I introduce ψ_a as a utility cost. In this setup, a higher value of ψ_a means that the worker has less incentives to participate in the formal sector. The model assumes that the distribution of ψ_a is conditional on the agent's characteristics X. For example, if college graduates find jobs in the formal-sector easier than high-school graduates, then $G(\psi_a|\,College) \geq G(\psi_a|\,HighSchool)$ for all ψ_a.

The final element of the model is the pension plan. The agent retires at the mandatory retirement age $a = 3$, leaving the labor force permanently. After retirement, he receives a per-period benefit of $B(\tau_2, X)$, which is a non-decreasing function of the number of periods contributed to the pension system while he worked, denoted by τ_2. If the worker never contributed to the pension system, he receives a social pension equal to $B(0, X)$. Finally, the model assumes that the worker does not have any bequest motive, so he exhausts his income and savings by age $a = 3$. The definition of the pension plan intends to capture the basics from a defined contribution system, which is the most common system used in Latin American countries (Bosch et al., 2013).

The interaction of sector-specific benefits from working, formal-sector participation shocks, and the pension plan, provides the basis to characterize the agent's formal- sector labor supply.

2.3 The model

Each period, a representative agent chooses consumption (denoted by $c_a \in \mathbb{R}^+$), assets ($A_a \in \mathbb{R}$), and formal-sector labor supply ($h_a \in \{0, 1\}$) to maximize his expected lifetime utility. In what follows, I assume that lifetime utility is time-separable, with a per-period utility function $u(c)$ strictly continuous and concave. Although the representative worker's decisions are conditional on his characteristics X, I omit them to simplify notation.

The agent's decision variables depend on his age. When he is a worker (ages 1 and 2), he chooses a consumption and formal-sector labor supply plan that solves

$$v_a(\tau_{a-1}, A_{a-1}) = \max_{c_a, A_a, h_a} u(c_a) - \psi_a h_a + \beta \mathbb{E} v_{a+1}(\tau_a, A_a) \qquad (2.1)$$

subject to

$$A_a = (1+r) A_{a-1} + w_a^i + h_a \left(w_a^f (1 - t_c - t_p) - w_a^i \right) - c_a \tag{2.2}$$

$$\tau_a = \tau_{a-1} + h_a \tag{2.3}$$

with A_{a-1} and τ_{a-1} given. In equations (2.1) and (2.2), r and β represent the interest rate and the discount factor, respectively. Additionally, $\mathbb{E}v_{a+1}(\tau_a, A_a)$ is the expected value function, where the expectation is also a function of τ_a and A_a. For tractability, I assume $\beta = \frac{1}{1+r}$.

Equations (2.1) to (2.3) formalize the setup discussed in Section 2.2. If the agent works in the formal sector ($h_a = 1$), he faces the utility shock ψ_a, receives a formal-sector wage net of contributions $w_a^f (1 - t_c - t_p)$ and increases his formal-sector experience by one more period. If he works in the informal-sector ($h_a = 0$), he receives a informal-sector wage w_a^i, does not incur in the utility shock, and does not increase his formal-sector experience.

When the agent retires (age 3), he chooses a consumption plan that solves

$$v_3(\tau_2, A_2) = \max_{c_3, A_3} u(c_3) \tag{2.4}$$

subject to

$$A_3 = (1+r) A_2 + B(\tau_2) - c_3 \tag{2.5}$$

where $B(\tau_2)$ represents the retiree's pension wealth (i.e., his income from the pension plan). Because the retiree does not have bequest motives, the value function after age 3 is equal to zero.

The optimal consumption and formal-sector labor supply plan can be obtained by backward induction. To begin with, the optimal consumption plan for the retiree is to consume all his wealth, setting $c_3(A_2, \tau_2) = (1+r) A_2 + B(\tau_2)$. As a result, the value function for period 3 is equal to

$$v_3(\tau_2, A_2) = u(c_3(\tau_2, A_2)). \tag{2.6}$$

From the properties of the utility function and the pension plan, the value function defined in (2.6) is an increasing function of financial wealth and formal-sector experience.

The definition of the value function $v_3(\tau_2, A_2)$ allows to characterize the optimal consumption and formal-sector labor supply plans before retirement. Conditional on formal-sector choice, the first order conditions of the worker's problem implies that $u'(c_2) = u'(c_3)$ and so $c_2 = c_3$.

Let c_2^f, A_2^f, c_2^i, and A_2^i denote the optimal consumption and savings plan conditional on working in the formal and informal sector. Using the condition $c_2 = c_3$ and the budget constraint (2.2), the consumption and savings plan conditional on sector choice is

$$c_2^i = \frac{1}{1+\beta}\left((1+r)A_1 + w_2^i + \beta B(\tau_1)\right) \tag{2.7}$$

$$c_2^f = \frac{1}{1+\beta}\left((1+r)A_1 + w_2^f(1-t_c-t_p) + \beta B(\tau_1+1)\right) \tag{2.8}$$

$$A_2^i = \frac{\beta}{1+\beta}\left((1+r)A_1 + w_2^i - B(\tau_1)\right) \tag{2.9}$$

$$A_2^f = \frac{\beta}{1+\beta}\left((1+r)A_1 + w_2^f(1-t_c-t_p) - B(\tau_1+1)\right), \tag{2.10}$$

while the value function conditional on sector choice is

$$v_2^i(\tau_1, A_1) = (1+\beta)\,u\left(c_2^i\right) = \tilde{v}_2^i(\tau_1, A_1) \tag{2.11}$$

$$v_2^f(\tau_1, A_1) = (1+\beta)\,u\left(c_2^f\right) - \psi_2 = \tilde{v}_2^f(\tau_1, A_1) - \psi_2. \tag{2.12}$$

Finally, using equations (2.11) and (2.12), it is possible to characterize the worker's formal-sector labor supply for age 2. Because the worker faces the utility shock only when he works in the formal sector, he chooses to work in the formal sector as long the utility gains are larger than the participation costs. Thus, the worker sets $h_1 = 1$ if

$$\bar{u}_2(\tau_1, A_1) = \tilde{v}_2^f(\tau_1, A_1) - \tilde{v}_2^i(\tau_1, A_1) \geq \psi_1. \tag{2.13}$$

The threshold $\bar{u}_2(\tau_1, A_1)$ encompasses the utility gains from working in the formal sector. Working in the formal sector increases the worker's lifetime income, as it increases the current income by the formal-to-informal wage gap $\left(w_a^f\left(1-t_c-t_p\right)-w_a^i\right)$ and increases his future pension benefits. Consumption smoothing implies that the additional income is allocated between present and future consumption. Thus, in absence of utility shocks, the worker's lifetime utility is higher when he works in the formal sector (i.e., $\bar{u}_2\left(\tau_1, A_1\right) \geq 0$). The worker's final choice depends on his preferences for formal-sector employment, his net valuation of the mandated benefits, and the availability of formal-sector jobs, all of them summarized by ψ_a.

Equation (2.13) is also informative about the effect of a worker's financial wealth and his level of formal-sector experience on the utility gains from working in the formal sector. This equation indicates that the long-run gains from working one more period in the formal sector depend on the change of the pension wealth. However, equation (2.13) also shows that the level of wealth makes the worker less sensitive to long-run gains from working in the formal sector. Intuitively, working in the formal sector increases lifetime income, yet this increment is relatively less important when the worker has more financial or pension wealth. To see this, note that the partial derivative of $\bar{u}_2\left(\tau_1, A_1\right)$ with respect to A_1, the marginal change of the worker's utility gains to a change in financial wealth is

$$\frac{\partial \bar{u}_2\left(\tau_1, A_1\right)}{\partial A_1}=(1+r)\left(u'\left(c_2^f\right)-u'\left(c_2^i\right)\right), \tag{2.14}$$

which is negative as long as $u''(c)<0$. Thus, when the worker has more assets, the curvature of the utility function implies that the worker becomes less sensitive to the gains from working one more period in the formal sector.

A second implication of the model is that the gains from working in the formal sector are a non-monotonic function of the formal-sector experience, as they depend on the curvature of the pension plan $B(\cdot)$ around τ_1. The partial derivative of equation (2.13) with respect to τ_1 is

$$\frac{\partial \bar{u}_2\left(A_1, \tau_1\right)}{\partial \tau_1}=\beta\left(u'\left(c_2^f\right) B'\left(\tau_1+1\right)-u'\left(c_2^i\right) B'\left(\tau_1\right)\right). \tag{2.15}$$

Thus, when $B(\cdot)$ is concave around τ_1, then $B'\left(\tau_1+1\right) \leq B'\left(\tau_1\right)$ and the derivative in (2.15) is negative. When $B(\cdot)$ is convex around τ_1, then the sign for (2.15) is ambiguous and depends on the relative size of the product between the marginal utility of consumption and the marginal change in $B(\cdot)$.

After characterizing the optimal consumption and formal-sector labor supply plan of the worker, the ex-ante probability of working in the formal sector is

$$P\left(h_2 = 1 | \tau_1, A_1\right) = P\left(\bar{u}_2\left(\tau_1, A_1\right) \geq \psi_2\right) = G\left(\bar{u}_2\left(\tau_1, A_1\right)\right) \tag{2.16}$$

while the expected value function for age 2 is

$$\begin{aligned} \mathbb{E}v_2\left(A_1, \tau_1\right) = & \left(1+\beta\right) u\left(c_2^i\right) + \\ & G\left(\bar{u}_2\left(A_1, \tau_1\right)\right) \mathbb{E}\left(\bar{u}_2\left(A_1, \tau_1\right) - \psi_2 \mid \psi_2 \leq \bar{u}_2\left(A_1, \tau_1\right)\right). \end{aligned} \tag{2.17}$$

Using (2.17) it is possible to characterize the optimal consumption and labor supply plan for the young worker ($a = 1$). However, the plan $\{c_1, A_1, h_1\}$ has no analytical solution. I show in Appendix B.1 that the properties of the solution described in this section also hold for the young worker.

2.3.1 Numerical example

To illustrate the characteristics of the utility gains from working in the formal sector, $\bar{u}_2\left(\tau_1, A_1\right)$, Figure 2.1 presents a numerical example for a hypothetical adult worker (age 2). In this example, I assume that the worker's utility function is logarithmic, that the wages in the formal and informal sector are $w^f = 1.1$ and $w^i = 1$, and that the worker does not have to contribute to the mandated benefits and pension systems ($t_c = t_p = 0$). Regarding the pension plan, I assume a replacement rate (the fraction of the formal-sector wage the retiree receives as a pension) that is an increasing function of formal-sector experience: $0.5 \times \left(1 + \left(1 + e^{-2(\tau_1 - 1)}\right)^{-1}\right)$. I choose a logistic function because it allows me to show the results under a pension plan that is convex or concave depending on the formal-sector experience (see Figure 2.2). In the simulation, I use small changes in the formal-sector experience to characterize in detail the shape of $\bar{u}_2\left(\tau_1, A_1\right)$.

Figure 2.1 shows the main points presented in the model. Gains from working in the formal sector are non negative and they are determined by the agent's financial wealth (A_1) and formal-sector experience (τ_1). When workers have low financial wealth, for instance $A_1 = -0.5$, the gains are higher, as working one more period in the formal sector increases substantially their lifetime income. In contrast, workers with high financial wealth are less sensitive to the utility gains from working in the formal sector. For example, assuming $\tau_1 = 0$, the increase in the lifetime income by working one more period in the formal-sector is the wage gap only (the change in the replacement rate is close to zero). The additional wage received from working in the formal sector $\left(w^f - w^i = 0.1\right)$ increases consumption by 10 percent for workers with no financial wealth ($A_1 = 0$), while it increases consumption by 2.3 percent for workers with higher levels of financial wealth ($A_1 = 5$).[21] The additional utility gains are associated with the increase of future consumption via an increase in savings.

Moreover, the utility gains from working in the formal sector are a function of the agent's formal-sector experience, and follows closely the slope of the pension benefits received in age $a = 3$. Despite the long-run gains from working in the formal sector are close to zero for low and high values of formal-sector experience, the utility gains for workers with a few periods of formal-sector experience are larger. This is the result of the workers' ability to save: analogous to the example with financial wealth, workers with higher pension wealth are less sensitive to the increase in the lifetime income associated with working one more period in the formal sector. Because workers with a few periods of formal-sector experience have lower pension wealth, the increase in lifetime income from working in the formal sector is relatively more important for them.

[21] The agents' consumption plan conditional on sector choice and financial wealth are $c_2^f\left(\tau_1 = 0, A_1 = 0\right) = 0.8585$, $c_2^i\left(0,0\right) = 0.7805$, $c_2^f\left(0,5\right) = 3.5474$, and $c_2^i\left(0,5\right) = 3.4695$.

2.4 No savings

In the previous section, I characterize the optimal consumption and labor supply plan given the gains from working in the formal sector. The previous setup assumed that the worker can save, and therefore he uses the increase in lifetime income from working in the formal sector to increase present and future consumption. Nonetheless, this may not be an appropriate assumption for workers in Latin American economies, especially the ones more likely to work in the informal sector. Empirical evidence from Latin America shows that low income population exhibit low or negative saving rates, and their access to adequate financial instruments to save is rather limited (Cavallo and Serebrisky, 2016).

Due to these limitations to saving behavior, I examine a version of the model in which the workers cannot save, and so $A_a = 0$ for all a. Except for this assumption, the model setup is the same as that presented in Section 2.3. Again, the consumption and formal-sector labor supply plan can be obtained by backward induction. In the last period, the retiree consumes all his income, that in this case is equivalent to his pension wealth. Therefore, the value function for the retiree is $v_3(\tau_2) = u(B(\tau_2))$.

When the agent is still working, he chooses a consumption and formal-sector labor supply plan that solves

$$v_a(\tau_{a-1}) = \max_{c_a, h_a} u(c_a) - \psi_a h_a + \beta \mathbb{E} v_{a+1}(\tau_a) \tag{2.18}$$

subject to

$$c_a = w_a^i + h_a\left(w_a^f(1 - t_c - t_p) - w_a^i\right) \tag{2.19}$$

$$\tau_a = \tau_{a-1} + h_a \tag{2.20}$$

with τ_{a-1} given. Without savings, the representative worker's optimal consumption plan is to consume all his income per period and, given a realization of the utility shock ψ_a, the worker works in the formal sector as long as the gains from the search are greater than the costs. Thus, the worker works in the formal-sector if

$$\bar{u}_a(\tau_{a-1}) = u\left(w_a^f(1-t_c-t_p)\right) - u\left(w_a^i\right) + \beta\left(\mathbb{E}v_{a+1}(\tau_a+1) - \mathbb{E}v_{a+1}(\tau_a)\right) \geq \psi_a \tag{2.21}$$

and the ex-ante probability that a worker works in the formal sector is given by

$$P(d_a = 1 \mid \tau_{a-1}) = G(\bar{u}_a(\tau_{a-1})). \tag{2.22}$$

As in the model in which the worker can save, $\bar{u}_a(\tau_{a-1})$ summarizes the gains from working in a formal-sector job. The first term represents short-run gains, that is, utility gains that come from the wage gap. The second term represents long-run gains, that is, the utility gains that come from increasing the pension benefits due to an increase in formal-sector experience.

Due to the assumptions on the wage gap and the pension plan, $\bar{u}_a(\tau_{a-1})$ is a non-negative, non-monotonic function of τ_{a-1}. For example, when $a = 2$, the derivative of $\bar{u}_2(\tau_1)$ with respect to τ_1 is

$$\frac{\partial \bar{u}_2(\tau_1)}{\partial \tau_1} = \beta\left(u'(B(\tau_1+1))\,B'(\tau_1+1) - u'(B(\tau_1))\,B'(\tau_1)\right) \tag{2.23}$$

which depends on the concavity or convexity of $B(\cdot)$ around τ_1.

In summary, the main characteristics of the model holds when I assume that the worker cannot save. The main difference with respect to the model with savings is that in this case the worker cannot smooth consumption, and therefore the short and long-run utility gains from working in the formal sector are clearly separated.

2.4.1 Numerical example

Figure 2.3 presents a numerical example of the utility gains for the model with no savings. The example uses the assumptions listed in Section 2.3.1, except that in this case I assume that the representative agent cannot save. Relative to the utility gains presented in Figure 2.1, the dependence of the utility gains with respect to formal-sector experience is associated with the changes of pension wealth, regardless of the level of pension wealth. The increase in current consumption by working in the formal sector is 10 percent (the wage gap), while the gains associated with future consumption depend on the change of pension wealth only. For example, for workers with a few periods of formal- sector experience, the changes of the replacement rate are close to zero, implying that the utility gains they perceive are only the increase in current consumption.

2.5 Separation rate less than one

The versions of the model presented above assume that the worker loses his job by the end of the period. In this section, I use the framework from Section 2.4 to analyze a scenario in which the worker loses his job with probability $q < 1$. To simplify the analysis, I assume that the separation probability is equal in both the formal and the informal sector.

Because the worker loses his job with probability q, the worker's decision also depends on the sector he worked in $a-1$. Let x_a denote an i.i.d. Bernoulli random variable that indicates whether the worker lost his job at the end of the previous period (i.e., $q = P(x_a = 1)$).

Since the retiree is not affected by the separation results, his optimal consumption plan is the same as that discussed in Section 2.4. When the agent is a worker, though, his consumption and labor supply plan depends on whether he loses his job. If the worker does not lose his job, then he chooses a consumption and formal-sector labor supply plan that solves

$$v_a(\tau_{a-1}, h_{a-1}, x_a = 0) = \max_{c_a} u(c_a) + \beta \mathbb{E} v_{a+1}(\tau_a, h_{a-1}, x_{a+1}) \tag{2.24}$$

subject to

$$\begin{aligned} c_a &= w_a^i + h_{a-1}\left(w_a^f(1 - t_c - t_p) - w_a^i\right) \\ \tau_a &= \tau_{a-1} + h_{a-1}. \end{aligned}$$

Similarly, if the worker loses his job at the end of $a-1$, his problem becomes

$$v_a\left(\tau_{a-1},h_{a-1},x_a=1\right)=\max_{c_a,h_a}u\left(c_a\right)-h_a\psi_a+\beta\mathbb{E}v_{a+1}\left(\tau_a,h_a,x_{a+1}\right) \tag{2.25}$$

subject to

$$\begin{aligned}c_a&=w_a^i+h_a\left(w_a^f\left(1-t_c-t_p\right)-w_a^i\right)\\ \tau_a&=\tau_{a-1}+h_a.\end{aligned}$$

Thus, when the representative worker does not lose his job in the previous period, he does not take any new decision about his formal-sector labor supply; when he loses his job, the previous formal-sector labor supply does not affect his decision.

Because the worker cannot save, his optimal consumption plan in both cases is equal to his income. Conditional on $x_a=0$, the value function is

$$\begin{aligned}v_a\left(\tau_{a-1},h_{a-1},0\right)=u\left(w_a^i+h_{a-1}\left(w_a^f\left(1-t_c-t_p\right)-w_a^i\right)\right)+\\ \beta\mathbb{E}v_{a+1}\left(\tau_{a-1}+h_{a-1},h_{a-1},x_{a+1}\right).\end{aligned} \tag{2.26}$$

Additionally, conditional on $x_a=1$, the worker consumes his entire income and chooses to work in the formal sector if

$$\begin{aligned}\bar{u}_a\left(\tau_{a-1}\right)=u\left(w_a^f\left(1-t_c-t_p\right)\right)-u\left(w_a^i\right)+\\ \beta\left(\mathbb{E}v_{a+1}\left(\tau_{a-1}+1,1,x_{a+1}\right)-\mathbb{E}v_{a+1}\left(\tau_{a-1},0,x_{a+1}\right)\right)\geq\psi_a,\end{aligned} \tag{2.27}$$

and therefore the ex-ante probability that a worker works in the formal sector is

$$P\left(h_a=1|\,\tau_{a-1},x_a=1\right)=G\left(\bar{u}_a\left(\tau_{a-1}\right)\right). \tag{2.28}$$

Using equations (2.26) to (2.28), the expected value function conditional on τ_{a-1} and $h_{a-1} = h \in \{0, 1\}$ is

$$\begin{aligned}
\mathbb{E}v_a\left(\tau_{a-1}, h, x_a\right) = {} & (1-q)\, u\left(w_a^i + h\left(w_a^f\left(1 - t_p - t_c\right) - w_a^i\right)\right) + \\
& (1-q)\, \beta \mathbb{E}v_{a+1}\left(\tau_{a-1} + h, h, x_{a+1}\right) + \\
& q\left(u\left(w_a^i\right) + \beta \mathbb{E}v_{a+1}\left(\tau_{a-1}, 0, x_{a+1}\right)\right) \\
& qG\left(\bar{u}_a\left(\tau_{a-1}\right)\right) \mathbb{E}\left(\bar{u}_a\left(\tau_{a-1}\right) - \psi_a \middle| \bar{u}_a\left(\tau_{a-1}\right) \geq \psi_a\right)
\end{aligned} \tag{2.29}$$

Since the worker consumes all his pension wealth when retired, the difference $\mathbb{E}v_3\left(\tau_2 + 1, 1, x_3\right) - \mathbb{E}v_3\left(\tau_2, 0, x_3\right) = u\left(B\left(\tau_2 + 1\right)\right) - u\left(B\left(\tau_2\right)\right)$ is non-negative. Using this result and equation (2.29),

$$\mathbb{E}v_a\left(\tau_{a-1} + 1, 1, x_a\right) - \mathbb{E}v_a\left(\tau_{a-1}, 0, x_a\right) = (1-q)\, \bar{u}_a\left(\tau_{a-1}\right) \geq 0. \tag{2.30}$$

Therefore, the long-run utility gains from working in the formal sector are always non-negative. As a result, $\bar{u}_a\left(\tau_{a-1}\right)$ possesses the same properties as the threshold defined in equation (2.21). The main difference in this case is that the aggregate formal-sector labor supply will be determined by the combination of past and present decisions.

2.6 Defined-contribution pension plan

The previous sections analyze the effect of a defined-benefit pension plan on formal-sector labor supply. Although defined-benefit plans are common across Latin America, some countries have changed from defined-benefit pension plans to combinations of defined-benefit and defined-contribution (individual account) plans. For instance, Chile has had an exclusive individual account plan since early 1980s, and Colombia and Peru implemented dual systems in which both defined-contribution and defined-benefit plans coexist (Bosch et al., 2013).

In this section, I use the framework presented in the previous section assuming an individual account pension plan, in which a pension fund invests the worker's contributions in the capital market, and the principal and returns constitute the worker's income after retirement.

The setup of the model is the same as that presented in Section 2.3, except for the definition of the pension plan. Let B_a denote the pension wealth of the representative agent at age a. When the agent retires ($a = 3$), his optimal consumption and saving plan is the solution of

$$v_3 \left(B_2, A_2\right) = \max_{c_3, A_3} u\left(c_3\right) \tag{2.31}$$

subject to

$$A_3 = \left(1 + r\right)\left(A_2 + B_2\right) - c_3. \tag{2.32}$$

Because the agent does not have bequest motives, the optimal solution is to consume all his income, setting $A_3 = 0$ and $c_3 = \left(1 + r\right)\left(A_2 + B_2\right)$. In what follows, I assume that financial assets and pension wealth have the same return in the financial market.

When the agent is working ($a \in \{0, 1\}$), his optimal consumption, saving, and formal-sector labor supply plan is the solution of

$$v_a \left(B_{a-1}, A_{a-1}\right) = \max_{c_a, A_a, h_a} u\left(c_a\right) - \psi_a h_a + \beta \mathbb{E} v_{a+1}\left(B_a, A_a\right) \tag{2.33}$$

subject to

$$A_a = \left(1 + r\right) A_{a-1} + w_a^i + h_a \left(w_a^f \left(1 - t_c - t_p\right) - w_a^i\right) - c_a \tag{2.34}$$

$$B_a = \left(1 + r\right) B_{a-1} + h_a t_p w_a^f \tag{2.35}$$

with B_{a-1} and A_{a-1} given. The agent's pension wealth is the agent's accumulated pension wealth up to age $a - 1$ plus the additional per-period contribution that he makes when working in the formal sector (equation (2.35)).

The solution of the agent's problem follows similar arguments as those presented in Section 2.3. Using backward induction, the agent's optimal consumption and saving plan conditional on h_2 is

$$c_2 = \frac{1}{1+\beta}\left(\left(1+r\right)A_1 + w_2^i + h_2\left(w_2^f\left(1-t_c\right) - w_2^i\right)\right) + \frac{1+r}{1+\beta}B_1 \quad (2.36)$$

$$A_2 = \frac{\beta}{1+\beta}\left(\left(1+r\right)A_1 + w_2^i + h_2\left(w_2^f\left(1-t_c\right) - w_2^i\right)\right) - \frac{1+r}{1+\beta}B_1 - h_2 t_p w_2^f \quad (2.37)$$

The optimal consumption and saving plan differs from the plan presented in equations (2.7) to (2.10). Under the individual account plan the agent internalizes the relationship between his present contribution to the pension system and its impact on future pension benefits. As a result, the agent's consumption does not depend on the pension contribution rate t_p (equation (2.36)), and an increase of pension wealth of $t_p w_2^f$ is offset by a reduction in financial wealth (equation (2.37)).

The value function conditional on sector choice is

$$v_2^i\left(B_1, A_1\right) = \left(1+\beta\right)u\left(c_2^i\right) = \tilde{v}_2^i\left(B_1, A_1\right) \quad (2.38)$$

$$v_2^f\left(B_1, A_1\right) = \left(1+\beta\right)u\left(c_2^f\right) - \psi_2 = \tilde{v}_2^f\left(B_1, A_1\right) - \psi_2, \quad (2.39)$$

which implies that the agent chooses to work in the formal sector if

$$\bar{u}_2\left(B_1, A_1\right) = \tilde{v}_2^f\left(B_1, A_1\right) - \tilde{v}_2^i\left(B_1, A_1\right) \geq \psi_2. \quad (2.40)$$

Equation (2.40) summarizes the gains from working in the formal sector. In contrast to the environment with a defined-benefit pension plan, the gains in the model with an individual account plan are associated with the increase of the agent's lifetime income by the wage gap $w_2^f\left(1-t_c\right) - w_i$. Because financial assets and pension wealth are perfect substitutes, other potential long run gains associated with the increase of the pension wealth are offset by a reduction in financial assets. Following a similar analysis that the presented in the Appendix B.1, it is possible to show that this behavior holds when the worker is young ($a = 1$).

Although the long run gains from working in the formal sector are offset by changes in financial wealth, the utility gains from working in the formal sector are positive as long as $w_2^f(1-t_c) > w_i$. Moreover, using equations (2.38) and (2.39), the utility gains are decreasing in both, the level of financial assets and pension wealth.

The results presented in this section show that an individual account pension plan has a limited effect on formal-sector labor supply when financial and pension wealth are perfect substitutes. However, this one-to-one relationship between financial and pension wealth depends on frictions affecting the economy. For example, credit constraints, minimum pensions guarantees, and differential returns of assets and pension wealth, may affect the degree of substitution between financial and pension wealth and may generate effects on formal-sector labor supply.

2.7 Final remarks

This chapter describes the analysis of formal-sector labor supply for workers under a life-cycle setting. The analysis extends the discussion presented in Chapter 1, in which I provide empirical evidence of the effects of changes in pension wealth on pre-retirement formal-sector labor supply.

The central piece of the model is the threshold $\bar{u}$, which is the valuation the worker places to the gains from working one more period in the formal-sector. The gains from working in the formal sector are divided in two: short-run gains, represented by the wage gap the worker receives, and long-run gains, represented by the increase in future pension benefits. As I show under different specifications, as long as either the wage gap or the change in the (defined-benefit) pension plan is positive, working one period in the formal sector represents a gain in utility. Everything else constant, the gains from working in the formal sector are decreasing in the level of financial wealth (assets), while its relationship with respect to the formal-sector experience depends on the specific pension plan. In contrast, when the pension benefits are related to an defined-contribution pension plan, the long-run gain from working in the formal sector is offset by a one-to-one reduction of financial assets. As a result, the utility gains in an individual account system comes from the increase in the lifetime income of the wage gap.

Although the model presents the formal-sector labor supply in a stylized framework, its structure and implications can be extended to other contexts. For example, the model allows to study the interaction of pension programs, such as individual account pensions with minimum pension guarantees, and social pensions. The features of the model also provide a basic setup for the study of welfare consequences of pension programs in economies with a large informal sector.

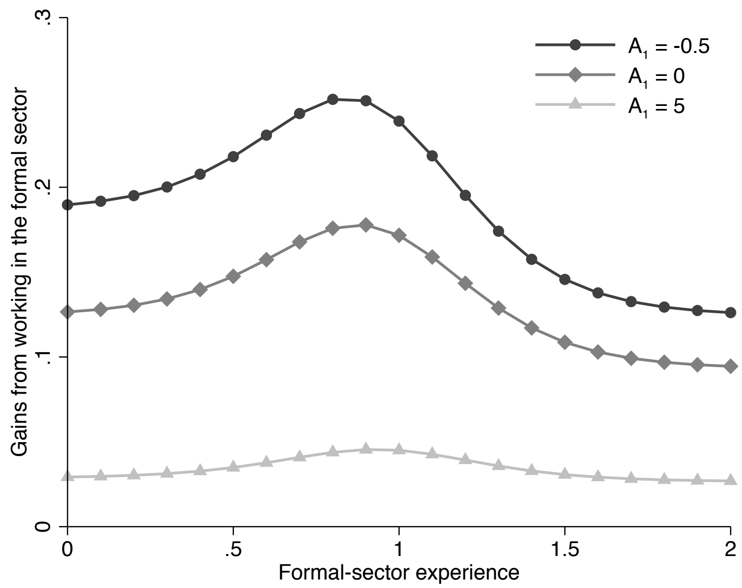
Table 2.1: Pros and cons of working in the formal sector

Pros	Cons
– Higher wages	– Pay contribution
– Mandated benefits	– Preference for informal jobs*
– Pension benefits	– Low valuation of benefits*
	– Formal-sector jobs are scarce*

**Participation shock in the model*

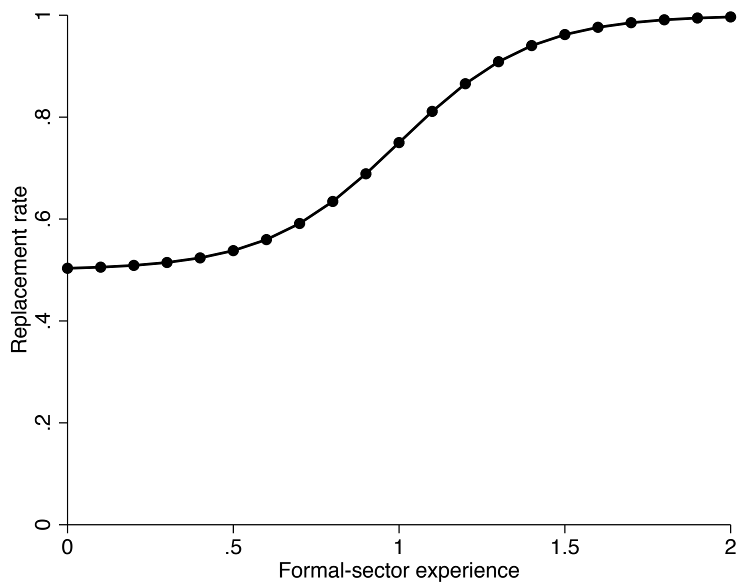
Notes: This table presents the factors in favor and against working in the formal sector. For workers, a formal-sector job provides a wage greater than the wage in the informal sector and mandated benefits, such as health care and severance payments. In addition, formal-sector workers may be entitled to pension benefits in the future, depending on their time of contribution. However, there are factors that prevent workers from getting a formal-sector job. First, working in the formal sector implies that workers have to pay contributions to the mandated benefits system. Second, some workers may prefer to work in the informal sector, due to time flexibility and desires of being independent. Third, because of the existence of substitutes for the mandated benefits, workers may have low valuation for the benefits provided by the formal sector. Finally, labor market frictions and regulations may prevent workers from finding formal-sector jobs.

Figure 2.1: Simulation results, model with savings



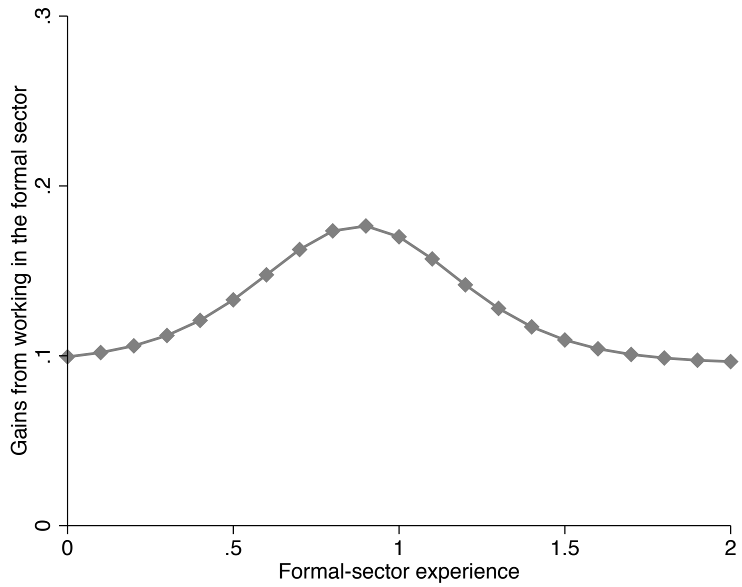
Notes: This Figure shows the utility gains from working in the formal sector ($\bar{u}_2(\tau_1, A_1)$) in a simulated scenario where the representative agent can save. The utility gains are computed for age 2, following equations (2.7) to (2.13) from Section 2.3 for three different values of A_1. I assume a logarithmic utility function, $w^f = 1.2$, $w^i = 1$, $t_c = t_p = 0$, and $B(\tau_1) = \left(1 + \left(1 + e^{-2(\tau_1 - 1)}\right)^{-1}\right) \frac{w^f}{2}$. The utility gains from working in the formal sector are a function of the agent's formal-sector experience, and follows closely the slope of the pension benefits received in age $a = 3$. Moreover, it is a decreasing function of the level of financial wealth. Keeping everything else constant, the marginal gains from the wage gap and the increase in pension wealth become less important when the agent has more financial wealth.

Figure 2.2: Replacement rate



Notes: The Figure shows the replacement rate (the fraction of the formal-sector wage the retiree receives as pension benefits) used in the examples presented in Sections 2.3 and 2.4.

Figure 2.3: Simulation results, model with no savings



Notes: This Figure shows the utility gains from working in the formal sector ($\bar{u}_2(\tau_1)$) in a scenario where the representative agent cannot save. The utility gains are computed for age 2, following equation (2.21) from Section 2.4. I assume a logarithmic utility function, $w^f = 1.2$, $w^i = 1$, $t_c = t_p = 0$, and $B(\tau_1) = \left(1 + \left(1 + \exp^{-2(\tau_1 - 1)}\right)^{-1}\right) \frac{w^f}{2}$. In this case, the utility gains from working in the formal sector are a function of the wage gap and the change of the pension wealth, regardless of the level of the agent's pension wealth.

Chapter 3

Labor Demand Responses to Payroll Taxes in an Economy with Wage Rigidity

3.1 Introduction

A major challenge faced by developing economies is how to create a strong social insurance system while minimizing its distortionary effects on the economy (Levy, 2008). In this challenge, payroll taxation plays a prominent role as a policy instrument. On one hand, a payroll tax provides benefits to workers in the form of insurance and may be used to finance the provision of public goods. On the other hand, if the incidence of payroll taxes is borne by registered employers complying with regulation (formal employers),[22] a payroll tax may discourage the creation of formal-sector jobs. A payroll tax may increase the cost of labor in the formal sector, reducing formal-sector labor demand and reallocating labor towards less-productive, low-quality jobs in the informal (unregulated) sector. Empirical evidence from Brazil and Colombia shows that the increase of payroll taxes has been a determinant in the rise of informal-sector employment in both countries (Santa María et al., 2009; Ulyssea, 2010).

The literature has identified three main determinants of the incidence

[22] In what follows, I define informal sector as the set of firms and workers that do not comply with government regulation, such as the payment of mandated contributions and taxes. A firm operating in the informal does not pay taxes (including payroll taxes), but it is subject to fines if it is inspected. Regarding workers, informal-sector workers do not pay contributions to the mandated benefit and insurance systems, but they are not covered by mandated benefits and other insurance included in the regulation.

of payroll taxes: the tax-benefit link of payroll taxes, the elasticity of labor supply, and the existence of factors that prevent wages from adjusting, such as the minimum wage (Gruber, 2000). When wages are flexible, a one-to-one valuation of the benefits funded from payroll tax revenues or an inelastic labor supply will allow payroll taxes to pass-through fully to wages with no employment effects. Ultimately, the incidence of payroll taxes depends on the interaction of those three factors, which is an empirical question.

In this chapter, I analyze the incidence of payroll taxes in Colombia, an economy in which the labor market institutions prevent wages from adjusting to changes in payroll taxes. Colombia's economy is characterized by major distortions in the wage adjustment process. To begin with, the Colombian minimum wage is binding for a large fraction of the population (Bell, 1997). About 40 percent of workers in the formal sector work for the minimum wage. In addition, about 50 percent of the labor force works in the informal sector. The large informal sector mitigates the pass-through from payroll taxes to wages, as a reduction in the formal-sector wage reduces the gains from working in the formal sector. The wage rigidity suggests that payroll tax incidence is borne mostly by Colombian formal-sector employers.

In this paper, I estimate incidence of payroll taxes on the Colombian labor market by using an exogenous reduction of payroll taxes. In 2011, the Colombian government introduced the First Job Act, which reduced payroll taxes for new workers younger than 28 by 11 percentage points (on a basis of 42 percent). Since the reduction in payroll taxes had no effect on the benefits that workers received (the deducted taxes were used to finance public goods), I interpret this reduction as a shock to the formal-sector labor demand. Using the exogenous variation caused by the First Job Act, I implement regression discontinuity and differences-in-differences identification strategies on a new source of administrative data for formal-sector employment. In both strategies, I compare the number of new workers and hiring wages for workers younger and older than 28 years of age. Consistent with the idea that the incidence of payroll taxes is borne by employers, I find that the reduction of payroll taxes brought about by the First Job Act increased formal-sector labor demand for young workers by 3.4 percent while having no significant effect on wages. The estimated impacts are similar across firms of all sizes, and are concentrated more in workers with no previous experience in the formal sector, men, and workers living in less developed regions.

This chapter contributes to the literature on the incidence of payroll taxes by examining a context where labor market institutions lead to wage rigidity. As a result, the incidence of payroll taxes is borne by employers. The literature on the incidence of payroll taxes includes a number of papers that look at the incidence of payroll taxes in developing economies. Two of the most relevant papers are those by Gruber (1997), who analyzes the impact of the reduction of payroll taxes in Chile in the early 1980s, and Kugler and Kugler (2009), who analyze the impact of the increase of payroll taxes in Colombia in the early 1990s. While Gruber (1997) finds full pass-through from taxes to wages in Chile, Kugler and Kugler (2009) find partial pass-through and employment effects in Colombia. Kugler and Kugler (2009) highlight the importance of wage rigidity as a potential driver of their results. By using data at the individual level, I am able to analyze those effects directly. The empirical approach of this Chapter is similar to the used in Cruces, Galiani, and Kidyba (2010), as I identify the incidence of payroll taxes by comparing the response of similar workers facing different payroll tax rates in the same time period, an identification strategy frequently unavailable in previous studies.

As an additional contribution to the literature on the incidence of payroll taxes, I recover the elasticity of formal-sector labor demand. Based on the estimation results and a standard economic model, I obtain an estimate of the elasticity of the formal-sector labor demand of -0.44. Because the wage distribution of new workers is concentrated around the minimum wage, the elasticity is informative of the labor demand response around the minimum wage. Thus, my results imply that an increase of 10 percent of the minimum wage reduces formal-sector employment by 4.4 percent. This implication is specific for the Colombian context, as the minimum wage is binding for a large fraction of formal-sector workers (about 40%), which makes likely that the results are driven by changes in the formal-sector labor demand only.

The rest of the chapter is organized as follows: Section 3.2 discusses the basic framework used in the analysis of the incidence of payroll taxes. Section 3.3 presents the institutional setting of the Colombian labor market, payroll taxes, and the First Job Act. Section 3.4 describes the data used in the estimation and the identification strategy. Section 3.5 shows the estimation results, and Section 3.6 sets out the conclusions.

3.2 Conceptual framework

The literature that examines the incidence of payroll taxes has a long history. The framework is set out by Summers (1989) and Gruber and Krueger (1991). It emphasizes that the incidence of payroll taxes depends on the extent to which these can be passed-through to wages. In particular, if a change in payroll taxes is offset by a change in wages, payroll taxes do not generate distortions in the labor market. Further, the extent of the pass-through from payroll taxes to wages depends on the elasticity of the labor supply and the tax-benefit link, i.e., the workers' valuation of the benefits they perceive from payroll taxes.

In a competitive labor market with homogeneous agents and employer payroll taxes, the market equilibrium is given by the relationship

$$D\left(w\left(1+t\right)\right)=S\left(w\left(1+\alpha t\right)\right) \tag{3.1}$$

where $D\left(w\left(1+t\right)\right)$ and $S\left(w\left(1+\alpha t\right)\right)$ represent the aggregate labor demand and supply, w is the equilibrium wage, t is the employer payroll tax rate, and α represents the valuation that workers have for the benefits financed with payroll taxes. The inclusion of payroll taxes implies that the worker's wage and the cost that the firm pays for this worker are different, as the firm has to pay the payroll tax rate t. Similarly, the benefits perceived for the worker are higher than the worker's wage, as he receives the wage plus the benefits financed with payroll taxes, that the worker values at rate α. As a result, labor demand is a function of the total labor cost $w\left(1+t\right)$, while labor supply is a function of the total worker's compensation $w\left(1+\alpha t\right)$.

Gruber (1997) shows that, under this setup, a change in the employer payroll tax has effects on the equilibrium level of employment and wages. Using total differentiation on equation (3.1), the response of the equilibrium wages and employment to a change in the employer payroll tax is given by

$$\frac{\frac{dw}{w}}{dt} = \frac{\alpha\,(1+t)\,\varphi - (1+\alpha t)\,\eta}{(\eta-\varphi)\,(1+\alpha t)\,(1+t)} \tag{3.2}$$

$$\frac{\frac{dD}{D}}{dt} = \eta\left(\frac{\frac{dw}{w}}{dt} + \frac{1}{1+t}\right) = \frac{\eta}{1+t}\left(\frac{\varphi\,(\alpha-1)}{(\eta-\varphi)\,(1+\alpha t)}\right). \tag{3.3}$$

In equations (3.2) and (3.3), $\eta = D'\frac{w(1+t)}{D}$ and $\varphi = S'\frac{w(1+\alpha t)}{S}$ stand for the elasticity of labor demand and supply respectively.[23]

From the previous analysis, employer payroll taxes do not have effects on employment (i.e., $\frac{\frac{dD}{D}}{dt} = 0$) as long as the pass-through from taxes to wages is equal to $-\frac{1}{1+t}$. This result holds in two cases: First, if labor supply is perfectly inelastic ($\varphi = 0$), then all the incidence of payroll taxes is on workers. Second, if the worker's valuation from the payroll tax equals the cost paid by the employer ($\alpha = 1$) then the increase in taxes is offset by a proportional reduction in wages, leaving employment unchanged (Summers, 1989).

Along with an inelastic labor supply and a one-to-one tax-benefit link, a full pass-through from payroll taxes to wages requires that wages can adjust to changes in payroll taxes. However, this is not always the case in developing economies. In particular, the existence of a binding minimum wage and a large informal (unregulated) sector prevents wages from adjusting to changes in taxes, which in turn generates employment effects even when the labor supply is inelastic.

A binding minimum wage is a common characteristic of Latin American economies, particularly Colombia (Bell, 1997; Maloney and Nuñez, 2004). If the minimum wage is binding, payroll taxes cannot pass-through to wages, and the incidence of payroll taxes is borne by employers (Gruber, 2000). The extent of the effect of the minimum wage on tax incidence depends on

[23] Equation (3.2) differs from the equation presented by Gruber (1997) because I assume a positive employer payroll tax and a zero employee payroll tax. I show in Section 3.3.1 that these assumptions are a good approximation for the Colombian case.

how binding the minimum wage is, an effect which is country-specific. For example, in Gruber's (1997) examination of payroll taxes in Chile, he argues that the minimum wage is not a relevant factor given that it is relatively low and affects only a small fraction of workers. In contrast, Kugler and Kugler (2009) find a limited pass-through in Colombia, which is consistent with the fact that the minimum wage is binding for a large fraction of Colombian workers (Bell, 1997).

The model presented above predicts that, when the minimum wage is binding, the employment effect of an increase of payroll taxes is negative. Using equation (3.3), the employment effect of a change in payroll taxes is

$$\frac{\frac{dD}{D}}{dt} = \frac{\eta}{1+t}. \tag{3.4}$$

A second characteristic frequently found in developing economies is a formal (regulated) sector co-existing with an informal sector. Typically, the informal sector is composed of small firms and self-employed workers that survive in the market by evading taxes and other regulations (La Porta and Shleifer, 2014; Meghir et al., 2015). Most remain unregistered because they are not productive enough to afford the cost of regulation, they are small enough to avoid detection by tax authorities, or they do not see the benefit of registering (Maloney, 2004; Perry et al., 2007).

The existence of the informal sector may mitigate the pass-through from payroll taxes to wages. To illustrate the effect of the informal sector on the pass-through from payroll taxes to wages, I follow Levy (2008) and analyze a two-sector labor market where one sector (the informal) does not comply with labor regulation. I assume that workers do not have a preference for working in either of the two sectors, and that they do not value the benefits from payroll taxes. As a result, the equilibrium wage is the same for both sectors, and the equilibrium in the labor market is given by

$$D^f\left(w\left(1+t\right)\right) + D^i\left(w\right) = S\left(w\right), \tag{3.5}$$

where $D^i\left(\cdot\right)$ and $D^f\left(\cdot\right)$ represent the labor demand in the formal and informal sector, and $S(\cdot)$ is the aggregate labor supply.

Figure 3.1 presents an example of the equilibrium effect of a reduction of payroll taxes in an economy with an informal sector. I assume that the aggregate labor supply is inelastic and equal to L_m, and that a worker always gets a job in either the formal or the informal sector. The formal-sector labor demand is represented by the curve D_0^f (drawn from left to right), while the informal-sector labor demand is represented by the curve D_0^i (drawn from right to left, starting at L_m). The initial equilibrium is denoted by the point A, where the wage received by workers in both sectors is the same $\left(w_0^{*i} = w_0^{*f}\right)$.

A reduction of payroll taxes shifts formal-sector labor demand curve to the right by $\frac{\eta^f}{1+t}dt$ to D_1^f, where η^f stands for the elasticity of formal-sector labor demand and t is the employer payroll tax. In the new equilibrium, the reduction of payroll taxes increases wages in both sectors and reallocates employment from the informal to the formal sector (point B). Overall, the effect on formal-sector employment caused by the reduction of the payroll taxes $\left(L_1^{f*} - L_0^{f*}\right)$ is smaller than that observed in a case of a binding minimum wage $\left(\frac{\eta^f}{1+t}dt\right)$, but larger than that observed in a case with full pass-through from taxes to wages and no informal sector (0). The magnitude of the effect depends on the relative elasticity of the formal and informal sector labor demand curves.

In general, using total differentiation on equation (3.5), the wage and formal-sector employment effects of a change in payroll taxes are

$$\frac{\frac{dw}{w}}{dt} = \frac{-\delta\eta^f}{\left(\delta\eta^f + (1-\delta)\,\eta^i - \varphi\right)(1+t)} \tag{3.6}$$

$$\frac{\frac{dD^f}{D^f}}{dt} = \eta^f\left(\frac{\frac{dw^f}{w^f}}{dt} + \frac{1}{1+t}\right) = \frac{\eta^f}{1+t}\left(\frac{(1-\delta)\,\eta^i - \varphi}{\delta\eta^f + (1-\delta)\,\eta^i - \varphi}\right), \tag{3.7}$$

where $\delta = \frac{D^f(w(1+t))}{D^f(w(1+t)) + D^i(w)}$ is the fraction of workers employed in the formal sector, and η^f and η^i are the elasticity of labor demand in the formal and informal sectors. Equation (3.7) implies that payroll taxes have employment effects even with an inelastic aggregate labor supply ($\varphi = 0$). Assuming

$\varphi = 0$, the magnitude of the effect depends on the relative elasticity of labor demand and the size of the formal sector. Payroll taxes do not have an effect on the size of the formal sector if the demand for labor in the informal sector is relatively inelastic $\left(\frac{\eta^i}{\eta^f} \to 0\right)$ or the size of the informal sector is relatively small $(\delta \to 1)$.

3.3 Institutional background

3.3.1 Colombian payroll taxes and labor market

The labor market institutions in Colombia exhibit characteristics that suggest that the incidence of payroll taxes is borne by formal-sector employers. A weak tax-benefit link, a binding minimum wage, and a large informal sector lead to a potentially large effect of payroll taxes on the generation of formal-sector employment.

Table 3.1 presents a summary of payroll taxes levied in 2010. It shows the employer and employee tax rate on the basis of contribution, and notes whether the contribution rate is applied to the provision of benefits for workers. The total payroll tax rate represents between 46 to 54 percent of a worker's monthly wage, and is divided into three components: insurance, family benefits, and public goods. The insurance component forms the largest part of the payroll tax rate (37 to 45 percentage points), and provides insurance for workers in the event of negative health shocks, old age, disability, and unemployment. Of the 12.5 percent deducted for health care insurance, 2 percentage points go to finance the public health care system. The family benefits component (4 percentage points) goes to Family Benefits funds, which are non-profit organizations responsible for providing benefits to workers, such as child allowances, access to recreation facilities, and subsidies for housing. The public goods component of the contribution (5 percentage points) funds a public education institution with a focus on technical programs (SENA), and the government agency responsible for providing child protection and family services (ICBF). Most of the payroll tax rate is paid by the employer (38 to 46 percentage points).

Although the level of the Colombian payroll tax rate is similar to other developed and Latin American economies (Corbacho, Fretes Cibils, and Lora, 2013; OECD., 2016), the structure of the payroll tax system may result in a distortionary effect on the labor market. To begin with, the tax-benefit link is weak, given that the benefits deriving from payroll taxes depend on being able to work in the formal sector and personal characteristics, and are not proportional to the worker's contribution. For example, because the health care insurance covers the worker and his family, benefits from health care insurance depend on the worker's family size and their health rather than his actual contribution. Moreover, unless workers place enough value on the social benefit provided by the public good component of the payroll taxes, they will not give up part of their wage to fund them (Summers, 1989).

Additionally, the binding minimum wage implies that the tax incidence –for a significant portion of Colombian workers– is fully borne by employers. In their comparison of wage distribution for Latin American economies, Maloney and Nuñez (2004) show that Colombia has a particularly binding minimum wage. When compared to seven Latin American economies,[24] the wage distribution in Colombia exhibits the second highest minimum wage- to median-wage ratio, the lowest standard deviation, and the highest skewness coefficient. Taken together, the results presented by Maloney and Nuñez (2004)indicate that the distribution of wages in Colombia is concentrated around the minimum wage and it is more concentrated than other Latin American economies. The binding minimum wage is confirmed by the results presented Table 3.2 (discussed in detail in Section 3.4), which shows that about 40 percent of formal-sector workers earn the minimum wage.

A third factor preventing wages from adjusting to changes in payroll taxes is the large size of the informal sector. The informal sector so characteristic of the Colombian economy is explained by a number of factors. Firms are inclined to operate informally given weak enforcement of registration requirements, large differences in costs of labor between the formal and informal sector, and a low valuation given to the benefit of operating in the formal sector (Santa María et al., 2009). As Mondragón-Vélez, Peña, and

[24] Argentina, Bolivia, Brazil, Chile, Honduras, Mexico and Uruguay.

Wills (2010) show, employment in the informal sector accounts for between 50 to 60 percent of total employment in Colombia, mostly less-educated people working as a self-employed or as a salaried-worker in a small firm. These patterns are similar to those found in other Latin American countries (Perry et al., 2007).

3.3.2 First Job Act

In order to encourage the generation of formal-sector jobs, the Colombian government enacted the First Job Act (Law 1429 of 2010). The Act had two objectives: to increase formal-sector employment of workers facing difficulties in finding formal-sector jobs, and to increase the registration rate of small firms.

The first component of the Act provided tax credits to existing firms for hiring workers under the age of 28. Starting in January 2011, employers could deduct from their corporate taxes 11 percentage points of the payroll taxes paid for these new workers. The tax credits were temporary, allowing the employer to claim the benefit for up to two years. The deducted contributions correspond to payroll taxes used to fund public goods (SENA, ICBF, and the public health care system) and Family Benefits funds. A worker's eligibility was based on the age at which employment commenced. For example, if the worker was hired when he was 27 years and 11 months old and continued working in the same firm, the firm would still be able to claim the tax credits.[25] Because the intention of the Act was to encourage the creation of new jobs, eligibility for tax credits was conditional on the firm increasing its total payroll by the end of the year.

The second component of the Act provided incentives for new firms employing up to 50 workers. The Act defined new firms as those registered after

[25]The Act included other groups of eligible workers: women 40 and above without a formal job in the last 12 months; people with disabilities; heads of households eligible for social assistance programs; low-wage workers, up to 1.5 times the minimum wage, who had not worked in the formal sector; refugees; demobilized guerrilla soldiers, and paramilitary members. When a worker met more than one of the eligibility criteria, the exemptions apply once. Due to the lack of information needed in order to identify these workers, I restrict my analysis to workers under the age of 28.

January 2011, and did not distinguish between new entrants and existing unregistered firms. New registered firms were exempt from paying corporate taxes (33 percent), along with 11 percentage points of the payroll taxes of all their workers (SENA, ICBF, Family Benefits, and public health care), and registration fees. The exemptions were temporary, allowing full exemption for the first two years and partial exemptions for the following three years.

The prospect of a reduction in payroll taxes resulted in a positive shock in the demand for workers under the age of 28. After the Act was enacted, the total labor cost (wage plus payroll taxes) of eligible workers declined by 11 percent. In contrast, the Act had a limited effect on labor supply because the reduction in payroll taxes had no effect on the benefits that workers received. Most of the reduction in payroll taxes was associated with contributions used to fund public goods. In addition, given that firms were still required to contribute to the Family Benefits funds, the workers continue receiving the benefits from the contributions to the Family Benefits funds.

Given my focus on identifying the effects of payroll taxes on formal-sector employment, I restrict my analysis to the effect of the reduction of payroll taxes for workers below the age of 28. Although the First Job Act reduced payroll taxes for new registered firms, it also reduced corporate taxes and registration fees. As a result, I am not able to distinguish the causal effect of the reduction in payroll taxes on employment from the effect of reductions in the corporate tax rate and registration fees on employment for new entrants.

3.4 Empirical strategy

3.4.1 Data and sample selection

To investigate the effects of payroll taxes on the formal-sector labor market, I use the PILA dataset covering the period between 2010 and 2012. The PILA is an employer-employee dataset obtained from the system used to collect payroll taxes. From the taxes reported in Table 3.1, all but the severance savings contributions are collected monthly through the PILA system. Because formal-sector employers pay payroll taxes and mandated contributions, the dataset collects information for the universe of formal-sector workers.

The PILA dataset includes information with respect to the type of employer (public or private), the type of worker (independent or employee), days worked (typically 30 per month), job location, and the worker's wage, gender, and date of birth.[26] Although the dataset covers all formal-sector employment, there are problems with some of the identifiers for employers and employees. To avoid false transitions, I fill in job spells in cases where a match is missing in a month but the dataset records the same match within a three-month window. In addition, I drop employers with extremely large variations in the number of workers. After processing, the PILA dataset contains 3.6 million private-sector employer-employee matches per month.

The sample selection is based on characteristics of both the employers and the employees, and the conditions of the First Job Act. With respect to employers, I restrict the sample to employers appearing in the entire sample (January 2010 to December 2012). Because these employers were registered before January 2011, they became eligible for tax credits only when hiring new workers below the age of 28. These employers are likely to be more stable and less likely to be affected by the entry effects associated with the provision of additional benefits as set out by the Act. The final sample contains 126,855 employers, which account for 57 percent of private-sector employers and 87 percent of formal-sector employment between 2010 and 2012.

With respect to employees, I restrict the sample to workers aged 26 to 29, hired between February 2010 and December 2012. I focus on new formal-sector workers, defined as workers reported by the first time in a formal-sector firm. Since the Act reduced payroll taxes for new workers only, workers younger than 28 who did not change their job after January 2011 were not eligible for the tax credit. Regarding the selection of workers by age group, I also restrict the sample to those aged 26 to 29 to mitigate concerns about systematic differences in the time trends of formal-sector employment by age or cohort (e.g., college enrollment decisions for workers 25 and younger).

[26]Date of birth and gender were added by the Ministry of Health based on the employee ID number. About 0.4 percent of the sample do not have information about these variables.

Table 3.2 presents summary statistics for the selected sample of employers. The table shows the average employment composition and wages by age group for all workers and for new workers, between February 2010 and December 2012. Table 3.2 shows that workers aged 26 to 29 share similar characteristics, while exhibiting noticeable differences with other age groups. In particular, relative to the total population,[27] workers aged 26 to 29 have the highest participation on both total formal-sector employment (22 percent) and entry into the formal-sector (1 percent). The result is consistent with previous evidence for Latin American economies, where young workers enter into the labor market in salaried-informal jobs, switch in their adulthood to salaried-formal jobs and, after accumulating experience, switch to informal self-employment (Perry et al., 2007). Despite noticeable differences in age composition, the distribution of employment in other demographic characteristics is similar across age groups. Except for the higher prevalence of men older than 30, the fraction of men and the regional distribution are similar across age groups.

The bottom part of Table 3.2 shows summary statistics for the wage distribution of formal-sector workers. To reduce the effect of extreme observations on the summary statistics, I trim the top and bottom 1 percent wages based on the monthly wage distribution. The first noticeable characteristic of the wage distribution is the binding minimum wage, especially for new workers. About 36 percent of workers older than 25 earn exactly the minimum wage and this fraction reaches 50 percent for the group of new workers. The average and median wage show a positive gradient over age, yet the median wage is just 20 percent larger than the minimum wage. Comparing the wage distribution of new workers with the wage distribution for all workers, the wage distribution of new workers is more concentrated towards the minimum wage. On average, new workers earn less, and their wage distribution exhibits a lower standard deviation. These differences suggest that less-skilled workers are more prevalent in the new workers group than in the full sample.

[27] To compute the fraction of total population, I used the total urban population by age from the Colombian Census of 2005. I describe the 2005 Census in Chapter 1.

In summary, workers aged 26 to 29 are at their prime age for working in the formal-sector. In spite of the fact that new hires share similar demographic characteristics with workers with more tenure in the firm, their wages tend to be located near the minimum wage. The importance of the minimum wage in shaping the wage distribution is remarkable, and this may play a key role in preventing taxes from passing through wages even after the reduction of payroll taxes.

3.4.2 Identification strategy

To measure the causal effect of changes in payroll taxes on formal-sector employment and wages, I use the variation across age and time induced by the First Job Act to implement two identification strategies. In both strategies, I analyze the behavior of employment and wages of workers younger than 28 relative to the behavior of workers 28 and older. I focus on labor market indicators for workers in the month that they were hired, so I am estimating the effects of payroll taxes on employment generation and hiring wages.

The first strategy is a regression discontinuity design (RD). I identify the employment and wage effect of payroll taxes by comparing those indicators for new workers just below and just above 28. I group workers aged 26 to 29 hired after January 2011 by age group (a) and month of entry (m_0) and run regressions of the form

$$y_{a,m_0} = \alpha_0 + \rho^{RD} I\left(a < 0\right) + \sum_{k=1}^{2} (\alpha_k + \tilde{\alpha}_k I\left(a < 0\right)) \cdot a^k + u_{a,m_0}. \tag{3.8}$$

In equation (3.8), y_{a,m_0} is the labor market outcome for age group a and month of entry m_0, $a = age - 28$ is normalized age (in quarters), and $I\left(\cdot\right)$ is the indicator function. The polynomial $\sum_{k=1}^{2}(\alpha_k a^k + \tilde{\alpha}_k a^k I\left(a < 0\right))$ accounts for the relationship between the labor market indicators and age. Under the identification assumption that determinants of formal-sector employment evolve smoothly around the eligibility threshold, ρ^{RD} is the causal effect of the reduction of payroll taxes on the outcome of interest (Imbens and Lemieux, 2008).

An additional identification assumption is required for the causal effect of the reduction of payroll taxes on employment. Because I have the universe of formal-sector workers, estimates of equation (3.8) are based on counts of formal-sector workers instead of measures of formal-sector employment relative to the entire population. As a result, the identification strategy relies on the assumption that the density of the overall population by age evolves smoothly around the eligibility threshold. If so, the estimates based on the density of formal-sector employees identify the changes in formal-sector employment caused by the reduction of payroll taxes instead of a change in the population by age.

The second identification strategy is a differences-in-differences design (DD). I identify the employment and wage effect of payroll taxes by comparing these indicators for new workers younger and older than 28 who started working in a formal-sector firm before and after the entry of the First Job Act. Using information for workers entering formal-sector firms by age group (a) and month of entry (m_0), I run regressions of the form[28]

$$y_{a,m_0} = \beta_a + \beta_{m_0} + \rho^{DD} I\left(a < 0\right) \cdot I\left(m_0 \geq 2011:01\right) + v_{a,m_0} \tag{3.9}$$

where the variables are defined in equation (3.8). Under the identification assumption that determinants of formal-sector employment do not change differentially across the treatment and control groups around the reform, ρ^{DD} is the causal effect of the reduction in payroll taxes on formal-sector employment for workers younger than 28.

I use three indicators to estimate employment and wage effects of the reduction of payroll taxes caused by the entry of the First Job Act. Given that I only had access to information about formal-sector workers, I estimate formal-sector employment effects based on the density of new workers by age per month of entry. To compute the density, I count the number of workers hired in month m_0 by age group, normalized by the total number of workers aged 26 to 29 hired in month m_0. To measure formal-sector wage effects, I

[28] I restrict the sample to workers in the month they enter into the formal-sector firm (m_0). Because of that, each new formal-sector match appears only once in the sample.

use the average log wage and the fraction of new workers with wage equal to the minimum wage by age-month cell. For the wage regressions, I weight each average by the number of observations per cell.

A limitation of using the density of new entrants as dependent variable is that ρ^{RD} and ρ^{DD} do not allow a straightforward interpretation. The density of new entrants is computed as the fraction of workers of a given age group relative to the total number of entrants aged 26 to 29. As a result, the estimated effects ρ^{RD} and ρ^{DD} are relative to the total formal-sector employment for this particular group. To estimate employment effects, I use relative measures using estimates of equations (3.8) and (3.9). For the RD strategy, I use the test statistic proposed by McCrary (2008) and compute the employment effect as the log difference of the density just below and just above 28. Using estimates from equation (3.8), I estimate the relative effect on formal-sector employment as

$$\gamma^{RD} = \log\left(\alpha_0 + \rho^{RD}\right) - \log\left(\alpha_0\right). \tag{3.10}$$

For the DD strategy, I normalize the average difference in the density of formal-sector employment by the density of workers aged 28 in December 2010 (one month before the First Job Act took place). Using estimates from equation (3.9), I estimate the employment effect of the reduction of payroll taxes as

$$\gamma^{DD} = \frac{\rho^{DD}}{\beta_{28} + \beta_{2012:12}}. \tag{3.11}$$

In both cases, I compute the standard errors of the employment effects by using the delta method.

Because the minimum wage is binding for a large fraction of new workers, the expected sign for the employment effect of the reduction of payroll taxes is positive, but the expected sign for the wage effect is ambiguous. To see this, note that if the workers productivity is not constant, a reduction of payroll taxes has two effects on workers with productivity close to the minimum wage (Kramarz and Philippon, 2001). On one hand, some workers who would have entered earning the minimum wage will receive a higher wage. This effect reduces the fraction of workers at the minimum wage and increases the average wage. On the other hand, the reduction allows the entry of new minimum wage workers who were not productive enough to work at the original labor cost level. This effect increases the fraction of workers at the minimum wage and reduces the average wage. Thus, the overall effect of the reduction of the payroll taxes on wages around the minimum wage depends on the relative magnitude of these offsetting effects.

3.5 Estimation results

3.5.1 Identification checks

The regression discontinuity identification strategy relies on the assumption that the unobserved determinants of formal-sector employment and wages evolve smoothly around the eligibility threshold. This assumption could be undermined if the estimated effect of the policy could be confounded by changes in other covariates that might influence the outcome (Imbens and Lemieux, 2008).

To test the extent in which other observable characteristics may change around the eligibility threshold, I analyze labor market outcomes and observable characteristics of workers entering into formal-sector firms between February and December 2010. Because the First Job Act took place in January 2010, the distribution and characteristics of entrants should not be affected by the fact that the worker is under the age of 28.

Table 3.3 presents estimates of equation (3.9) for observable characteristics of the new entrants (panel A) and the formal-sector employment and wages in 2010 (panel B). Observable characteristics found in the PILA dataset are gender and whether the worker starts a new job in a firm with less than 10 workers.[29] Panel A presents regression discontinuity estimates using as dependent variables the fraction of male formal-sector workers and the fraction of formal-sector workers entering into a small firm (10 workers

[29]Since the Act could have differential effects by gender and firm size after its entry, it is not possible to implement this test for 2011 and 2012.

or less). The results show that there are no systematic differences in those observable categories between workers just below and just after the eligibility threshold of the First Job Act.

Panel B of Table 3.3 presents estimates of the employment and wage effects for 2010. Since the Act took place before in January 2011, estimates for 2010 provide evidence of differential changes in wages and employment associated with other factors different from the entry of the First Job Act. The estimation results show that there is no significant employment effect. However, there is a significant difference in the average wage around the age of 28 before the entry of the First Job Act. The significant effect is not robust to alternative estimators and control functions.

3.5.2 Baseline results

Table 3.4 reports estimates of the employment and wage effects of the reduction of payroll taxes for the sample of new workers aged 26 to 29. The table presents estimates of the employment effects and wage effects using regression discontinuity (RD) and differences-in-differences (DD) identification strategies. For each strategy, the columns present the effect of the reduction of payroll taxes on employment, on average wages, and on the fraction of workers earning the minimum wage. All estimates are presented in percentage points. I allow for correlated errors by age group over time by clustering the standard errors by age group (16 clusters).

Table 3.4 shows that the main adjustment to the reduction of payroll taxes for new workers younger than 28 was through employment rather than wages. In both strategies, I find a positive and significant effect on employment and a small and insignificant effect on the average wage and the fraction of new formal-sector workers earning the minimum wage. The estimates from the RD strategy indicate that, at the boundary, the reduction of payroll taxes increased employment for workers aged 28 by 1.14 percent. Similarly, the DD strategy indicates that the reduction of payroll taxes increased employment of workers younger than 28 by 3.38 percent. The RD results are robust to the specification of the control function, alternative estimators and bandwidth selections. Similarly, the DD estimation results are robust to the inclusion of time trends by age group to control for differential entry rates of younger workers to the formal-sector. The results of the robustness tests are presented in Appendix C.

The bottom part of Table 3.4 presents estimates of the effects of the First Job Act for new workers, in which I allow that the effect of the Act changes per year. The results show that firms responded slowly to the implementation of the reduction of payroll taxes, as the estimated employment effects were larger in 2012 than in 2011. The RD employment effects are estimated with low precision, and it is not possible to reject the null hypothesis that the employment effects per year are the same (p-value: 0.169). In contrast, the employment effect from the DD strategy is significant for both years, and significantly larger in 2012 (p-value for the difference between employment effects per year: 0.013). On average, the estimated employment effect based on the differences-in-differences strategy is about 4.5 percent in 2012.

Graphical evidence for the RD identification strategy is presented in Figure 3.2. The top panel presents the estimated density of formal-sector employment by age group in the year before and two years after the entry of the First Job Act. In 2010, there is no visible difference at the density of employment around the age of 28, however there is a positive and significant effect after the entry of the First Job Act. The density of employment in 2011-12 tends to be more volatile for workers above the age of 28, which is reflected on the wider confidence bands after the entry of the policy. With respect to average wages, there is no observable difference between wages around the discontinuity threshold, that in fact is observed in 2010. Although this difference is not robust to the specification of the control function, the result casts doubts on the RD identification strategy.

To further investigate the differences in the employment effects over time, I compute the time trend of the estimated employment effects for the DD strategy. Figure 3.3 presents estimates of γ^{DD} in which I allow that the employment effect changes by semester, using the second semester of 2010 as base category. Specifically, I estimate the equation

$$y_{a,m_0} = \beta_a + \beta_{m_0} + \rho^{DD}_{h[m_0]} I\left(a < 0\right) \cdot I\left(m_0 \geq 2011:01\right) + v_{a,m_0} \tag{3.12}$$

where $h\left[m_0\right]$ stands for the semester associated with month of entry m_0. Based on estimates of $\rho^{DD}_{h[m_0]}$, I compute the employment effects following equation (3.11). The employment effect takes larger values only after the second semester of 2011, consistent with the idea that it took some time for the firms to implement the First Job Act. Moreover, the estimated effect for the first half of 2010 suggests that, prior to the entry of the First Job Act there were no significant differences in the trends for workers in the treatment and control group. This result provides evidence to support the common-trend assumption required for the differences-in-differences identification strategy.

In light of the discussion presented in Section 3.2, the estimation results suggest that effect of the reduction in payroll taxes was not passed-through higher wages. Taking the estimated employment effect from the DD strategy $\left(\hat{\gamma}^{DD} = 0.0338\right)$, the reduction in payroll taxes ($dt = -0.11$), and the employer payroll tax rate ($t \approx 0.42$),[30] equation (3.4) implies an elasticity of formal-sector labor demand of $\eta^f = -0.44$. This is likely a lower bound of the actual elasticity, as not all firms can claim the tax credit (so dt may be closer to zero). The elasticity is in the middle range of previous estimates found in the literature, which vary between -0.65 and -0.3 (Arango Thomas, Gómez, and Posada, 2009; Bell, 1997; Cardenas and Bernal, 2004; Roberts and Skoufias, 1997). In contrast to these previous studies, the estimated elasticity does not rely on specific functional forms for labor demand, but on the zero wage effect of the reduction of payroll taxes.

Differences between identification strategies

Although both identification strategies find positive and significant employment effects and no wage effects, the magnitude of the RD employment

[30] I assume $t = 0.42$, as it is the middle point of the employer payroll tax rate (see Table 3.1).

effects are smaller than the DD employment effects. Those differences may be explained by the limited external validity of the RD identification strategy under heterogeneous treatment effects. As Imbens and Lemieux (2008) point out, regression discontinuity designs provide the average effect for the subpopulation located at the boundary (in this case, new hires aged 28). However, if the treatment effect is heterogeneous, the conclusions drawn from this strategy cannot be extrapolated to other subpopulations.

To test the role of heterogeneity in driving the differences between identification strategies, I estimate employment effects by age group using a differences-in-differences strategy. I estimate employment effects based on equation (3.11), allowing γ^{DD} to change by age group (grouped in 6-month bins), and using new hires aged 28 as the base group. Figure 3.4 presents the estimates by age group. Under the homogeneity hypothesis, the coefficients located to the left of 28 should be equal. However, the estimate for the group just below 28 is smaller than the average effects for younger workers, and it is significantly different from the other average effects (p-value:$<$0.01). Moreover, as Table C.2 in the Appendix shows, the aggregate results are robust to the inclusion of time trends by age group, suggesting that the effect is not the result of differential time trends between young and old workers. Thus, the results indicate that employment effects were larger for younger workers. This result may explain part of the difference between the estimated employment effects.

Figure 3.4 provides additional information to test the robustness of the DD identification strategy. Since the First Job Act affected new workers younger than 28, the estimated effects for older age groups should not be affected for the entry of the First Job Act. The coefficients above 28 are not significantly different from zero (p-value 0.193), which provides supporting evidence of the robustness of the DD results.

3.5.3 Heterogeneity analysis

Previous section shows that the effect of the reduction in payroll taxes for younger workers exhibits heterogeneous effects over time and over age groups.

Next, I analyze whether the effects is heterogeneous with respect to other observable characteristics available in the data. To take into account the average effects over the whole group of workers younger than 28, I present results for the differences-in-differences identification strategy.

I estimate the response of the formal-sector labor market to changes in payroll taxes along four dimensions. First, I look at whether the significant employment effect reported before corresponds to an actual increase of employment, or whether it corresponds to a reallocation from younger workers to new formal-sector jobs. I test this by grouping new workers on the basis of their presence in the PILA dataset before (regardless of the type of job). I then estimate the employment and wage effects for workers with and without formal-sector experience.

Panel A of Table 3.5 presents the estimated employment and wage effects obtained by using the differences-in-differences identification strategy. Although the employment effects are significant for workers with and without formal-sector experience, the larger employment effects are for new formal-sector workers. The estimated employment effect for workers with and without previous experience in the formal sector are 3.6 and 7.5 percent, and the difference is statistically significant (p-value: 0.014). Contrary to the aggregate results, the average effect on starting wages for new formal-sector workers is positive and significant. However, the estimated increase in wages is relatively small compared to the employment effects (0.78 percent).

Second, I examine the effects of the reduction of payroll taxes on employment and wages by gender. Men and women have different labor force and formal-sector participation patterns over their life cycle (Perry et al., 2007). Although men and women are in their prime age for both types of participation, fertility and household decisions might differentially affect the time trends for men and women, which may confound my results.

Panel B of Table 3.5 presents the estimation results by gender. Both groups exhibit positive and significant employment effects, while the wage effects are small and estimated with low precision. The employment effects are higher for men than for women, 4.7 versus 2.7 percent, yet their difference is only statistically significant at 10 percent level (p-value: 0.072).

Third, I analyze the effects of the reduction of payroll taxes on employment and wages by region. Geographical variation has received increasing attention in the literature concerned with determinants and consequences of the informal sector (Almeida and Carneiro, 2012; Gerard and Gonzaga, 2014). Economies with low levels of economic development tend to exhibit larger informal sectors (La Porta and Shleifer, 2014). Regions characterized by lower economic development tend towards smaller and less productive firms, with lower levels of enforcement of labor regulation.

Panel C of Table 3.5 presents the estimation results by region. I split the sample between developed regions, comparing the largest industrial regions of the country (Bogota/Cundinamarca, Antioquia, and Valle) versus the rest of the country.[31] The estimated employment effects are large and significant for both regions. Employment effects are larger for developing regions, suggesting that formal-sector labor demand is more elastic in less developed regions. Nonetheless, the standard errors of the estimates are large enough to fail to reject the null of equality of employment effects (p-value: 0.13).

Finally, I analyze the effects of the reduction of payroll taxes on employment and wages by firm size. Small firms tend to hire low-skilled workers and pay lower wages. Taking into account the binding minimum wage in Colombia, it is more likely that the wage distribution for workers in small firms is more concentrated around the minimum wage. Figure 3.5 shows that this is certainly the case. The figure presents the average fraction of workers earning the minimum wage by firm size. To maintain consistency in the composition of the sample over time, I define firm size based on the number of workers reported in January 2010. Although the minimum wage is binding for firms of all sizes, small formal-sector firms are more likely to pay the minimum wage than larger firms. About 60 percent of workers in firms with 1 to 10 workers earn the minimum wage, and this figure drops to 35 percent for firms with more than 100 workers. A similar pattern is observed

[31] According to official reports from the Colombian Statistics Office (DANE), developed regions account for 60 percent of the total GDP and about 45 percent of the total population. Moreover, as I show in Chapter 1, developed regions exhibit larger formal-sector employment rates than developing regions.

in the distribution of hiring wages, where about 70 percent of new workers in firms with 1 to 10 workers earn the minimum wage. Thus, because the minimum wage is more binding for small firms, a reduction in payroll taxes should have larger effects for them.

Panel D of Table 3.5 presents the estimation results by firm size. The results suggest that the response of labor demand to changes in payroll taxes was similar across firm sizes. Estimates of employment effects oscillate between 3.0 and 3.5 percent, and I fail to reject the null hypothesis of equality of employment effects (p-value: 0.918). Wage effects are small and insignificant for all firms. Taken together, the reduction in payroll taxes had a widespread positive impact in the formal-sector labor demand.

3.6 Final remarks

In this chapter, I analyze the response of formal-sector labor demand to payroll taxes in an economy with wage rigidity. In developing economies, payroll taxes may have large distortionary effects, given the likelihood that their institutional characteristics will have employers bearing all the incidence of the payroll tax.

In particular, I analyze the incidence of payroll taxes in the formal-sector in Colombia. Colombia is an example of an economy with labor market institutions that prevents payroll taxes from passing-through wages. On one hand, it has a strong wage rigidity. It exhibits one of the most binding minimum wages in the region and half of the labor force works in the informal sector. On the other hand, the majority of the payroll tax system has a low tax-benefit link, which leaves workers less willing to give up part of their wage in exchange for access to the benefits from payroll taxes.

To estimate the incidence of payroll taxes on the Colombian formal sector I use the First Job Act. Starting in 2011, the Act reduced payroll taxes for new workers under the age of 28. The Act has two useful aspects for the identification of the incidence of payroll taxes. First, it modified payroll taxes for only a subpopulation of workers, which allows the identification of employment and wage effects by using group variation over time. Second, the Act reduced taxes that did not provide a direct benefit for workers, and thus the variation induced by the Act can be interpreted as a shock in the formal-sector labor demand.

I estimate the payroll tax incidence by applying two identification strategies (regression discontinuity and differences-in-differences) to a new source of administrative data for the formal sector. I estimate effects of the reduction of payroll taxes on both formal-sector employment and wages. Consistent with the idea that the incidence of payroll taxes is borne by employers, I find that the reduction of payroll taxes increased formal-sector demand for young workers by 3.38 percent and no significant effect on wages. The estimated employment and wage effects are consistent across different specifications and subsamples. The estimated impacts are similar across firms of all sizes, and are concentrated more in workers with no previous experience in the formal sector, in male workers, and in workers living in less developed regions.

Using the estimates from the differences-in-differences strategy and the change in payroll taxes, I find that the implied elasticity of demand in the formal sector is -0.44. Because the wage distribution of new workers is concentrated around the minimum wage, the elasticity is informative with respect to the labor demand response around the minimum wage. Thus, the estimation results indicate that an increase of 10 percent of the minimum wage reduces formal-sector employment by 4.4 percent. This implication is country-specific, as the Colombian minimum wage is binding for a large fraction of formal-sector workers, which makes likely that the results are driven by changes in the formal-sector labor demand only.

The results show that changes in payroll taxes are an effective policy tool for generating formal-sector employment when the institutional arrangement prevents labor market from passing-through payroll taxes to wages. The generalization of these results is not straightforward, though, because such generalization would depend on the particular type of rigidity affecting the labor market. Nonetheless, this paper shows the importance of understanding the wage-setting process in order to better measure the extent and efficacy of labor market policies.

Table 3.1: Payroll taxes in Colombia, 2010

% of monthly wage	Total	Employer tax rate	Employee tax rate	Benefits for Worker	Benefits for Other
A. Insurance					
Health care	12.5	8.5	4.0	10.5	2.0
Workplace safety	0.4-8.7	0.4-8.7	–	0.4-8.7	–
Pension benefits	16.0	12.0	4.0	16.0	–
Severance savings	8.1	8.1	–	8.1	–
B. Family Benefits funds					
Family benefits	4.0	4.0	–	4.0	–
C. Public goods					
SENA/ICBF	5.0	5.0	–	–	5.0
Total	46.0-54.3	38.0-46.3	8.0	39.0-47.3	7.0

Notes: This Table presents a summary of the payroll taxes paid by Colombian firms and workers in the formal sector. It shows the employer and employee payroll tax rates, and the distribution of the rate between services provided to the worker and the financing of public goods. SENA is a public education institution with a focus on technical programs and training, ICBF is the government agency responsible for providing child protection and family services, and Family Benefits funds are non-profit organizations responsible for providing benefits to workers, such as child allowances, access to recreation facilities, and subsidies for housing.

Table 3.2: Summary statistics, average 2010-2012

	All workers				New workers			
Age group (years)	20-25	26-27	28-29	30-60	20-25	26-27	28-29	30-60
Workers/month (Thousands)	503.0	226.9	237.7	2,069.8	40.6	11.9	10.7	57.7
% of total	16.6	7.5	7.8	68.1	33.6	9.8	8.9	47.8
% of Population	15.0	21.6	22.6	16.9	1.21	1.13	1.02	0.47
Demographic Characteristics (% of total by age group)								
Men	55.4	55.9	56.2	60.2	56.4	56.9	57.3	61.9
Developed regions	70.6	69.2	69.0	68.7	68.8	66.7	66.4	65.8
Workers by Employer's Size in January 2010 (% of total by age group)								
1-10 workers (97,526 Firms)	9.4	9.4	9.5	12.6	7.9	8.1	8.2	9.9
11-50 (21,478 Firms)	14.3	14.5	14.5	15.8	12.6	13.1	13.3	14.8
51-100 (3,506 Firms)	7.6	8.1	8.2	8.6	7.0	7.6	7.9	8.8
101-1000 (3,983 Firms)	32.3	34.7	35.5	36.6	30.0	32.4	33.4	35.5
1000+ (362 Firms)	36.3	33.4	32.4	26.3	42.5	38.7	37.1	31.1
Wage distribution (Minimum wage = 1)								
Average	1.34	1.66	1.80	2.03	1.25	1.48	1.55	1.58
Standard Deviation	0.73	1.22	1.49	1.99	0.64	1.06	1.26	1.47
Median	1.02	1.13	1.17	1.18	1.00	1.00	1.00	1.00
Min. wage earners (%)	46.9	37.6	35.9	37.0	59.8	51.5	50.8	54.7

Notes: This Table presents summary statistics from the PILA dataset between February 2010 and November 2012. It compares the distribution of employment and wages across age groups (columns) for all workers in the firms included in the sample, as well as the subset of new workers. For the wage distribution statistics, I trim the top and bottom one percent of observations based on the monthly distribution of wages. "Min. wage earners" refers to the fraction of workers earning the monthly minimum wage.

Table 3.3: Balance tests, 2010

A: RD results for predetermined variables, 2010

	Men	Small firm worker
Under the age of 28	0.20	0.06
	[0.31]	[0.13]
Number of cells	176	176
Observations	241,368	241,368

B: Employment and wage effects for 2010

	Employment	Wage	Min. Wage
Under the age of 28	1.10	0.61	-0.46
	[1.24]	[0.27]**	[0.32]
Number of cells	176	176	176
Observations	241,368	236,240	236,240

Notes: This Table investigates whether there are factors that affect the validity of the identification assumptions required for the regression discontinuity (RD) design described in Section 3.4.2. Each cell reports an RD estimate (escalated by 100) based on a separate regression of a variable observed in 2010 (the year before the entry of the First Job Act). The regressions include a quadratic polynomial on age and its interaction with a treatment indicator for being under the age of 28 as independent variables (See equation (3.8)). Panel A presents estimates for the composition by gender and firm size between workers just below and just after the eligibility threshold of the First Job Act (28 years). Panel B presents RD estimates for employment and wage effects. The estimated effect for wages is significant, but it is not robust to the specification of the control function. Standard errors clustered by age group in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.4: Estimation results, Regression Discontinuity (RD) and Differences-in-Differences (DD) identification strategies

	RD strategy			DD strategy		
	Employment	Wages	Min. Wage	Employment	Wages	Min. Wage
Panel A. Average effect 2011-2012						
Lower payroll taxes	1.14	0.02	-0.13	3.38	-0.23	0.07
	[0.57]*	[0.36]	[0.17]	[0.81]***	[0.19]	[0.23]
Panel B. Effects by year						
2011	0.32	-0.41	-0.04	2.20	0.01	-0.09
	[0.70]	[0.64]	[0.46]	[1.09]*	[0.25]	[0.28]
2012	1.91	0.39	-0.22	4.48	-0.46	0.22
	[0.88]**	[0.26]	[0.21]	[0.70]***	[0.22]*	[0.25]
Number of cells	384	384	384	560	560	560
Observations	549,171	539,220	539,220	790,539	775,460	775,460

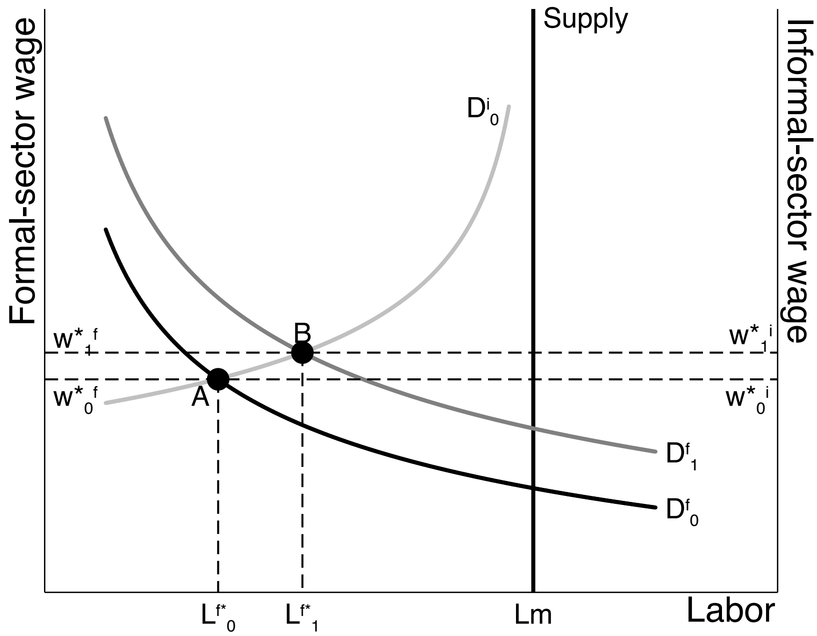
Notes: This Table presents the estimation results for the employment and wage effects (in percentage points) for the sample of workers entering into formal-sector firms between February 2010 and December 2012. Each cell represents an estimated employment or wage effect obtained by using regression discontinuity (RD) and differences-in-differences (DD) identification strategies (See Section 3.4 for details). "Min. Wage" refers to the average effect on the fraction of workers earning the minimum wage. Panel A displays the estimation results assuming that the average effect of the policy is constant; Panel B allows that the estimated effects vary by year. Standard errors clustered by age group in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 3.5: Estimation results for subsamples

	Employment	Wages	Min. Wage
A. Effects by previous experience in the formal sector			
Previous experience	3.64	-0.06	0.05
	[0.98]***	[0.29]	[0.30]
No previous experience	7.46	0.78	-0.46
	[1.49]***	[0.29]**	[0.38]
B. Effects by gender			
Men	4.23	-0.59	0.06
	[1.19]***	[0.31]*	[0.31]
Women	2.22	0.30	0.04
	[0.51]***	[0.31]	[0.39]
C. Effects by region			
Developed	2.73	-0.41	0.30
	[0.93]**	[0.25]	[0.35]
Developing	4.72	0.17	-0.47
	[1.14]***	[0.26]	[0.35]
D. Effects by firm size			
1–10 workers	3.27	-0.18	0.85
	[1.55]*	[0.51]	[0.71]
11–100	3.03	-0.26	-0.49
	[1.11]**	[0.45]	[0.48]
100+	3.47	-0.21	0.11
	[0.83]***	[0.25]	[0.30]
Observations	790,539	775,460	775,460

Notes: This Table investigates whether the effect of the First Job Act on employment and wages is heterogeneous by observable characteristics of firms and workers. Each cell represents an estimated employment or wage effect (in percentage points) obtained by using a differences-in-differences (DD) identification strategy on the selected sample. "Min. Wage" refers to the average effect on the fraction of workers earning the minimum wage. Panel A analyzes whether the employment effect found in Table 3.4 is the result of the entry of new formal-sector workers, or the result of reallocating younger workers who were at one time working in the formal sector. Panel B shows the results by gender. Panel C shows the results by region. Panel D shows the results by firm size, defined as the number of workers employed by the firm at the beginning of the sample (January 2010). Standard errors clustered by age group in brackets. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

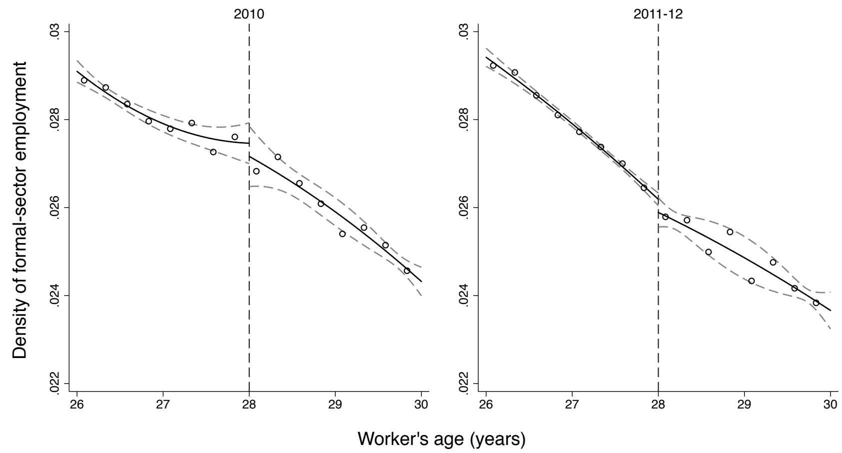
Figure 3.1: Effect of a reduction of payroll taxes in an economy with an informal sector



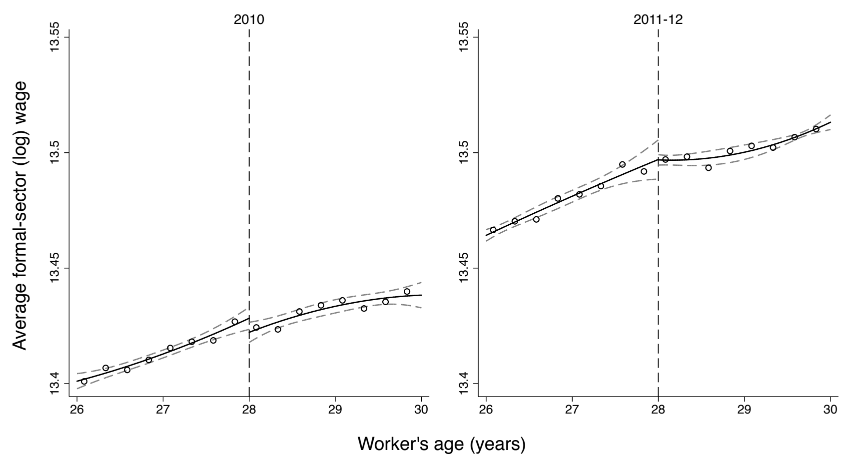
Notes: This Figure shows an example of the equilibrium effect of a reduction of payroll taxes in an economy with an informal sector. I assume that the aggregate labor supply is inelastic and equal to L_m, and that a worker always gets a job in either the formal or the informal sector. The formal-sector labor demand is represented by the curve D_0^f (drawn from left to right), while the informal-sector labor demand is represented by the curve D_0^i (drawn from right to left, starting at L_m). The initial equilibrium is denoted by the point A, where the wage received by workers in both sectors is the same $\left(w_0^{*i} = w_0^{*f}\right)$. A reduction of payroll taxes shifts formal-sector labor demand curve to the right by $\frac{\eta^f}{1+t}dt$ to D_1^f, where η^f stands for the elasticity of formal-sector labor demand and t is the employer payroll tax. In the new equilibrium, the reduction of payroll taxes increases wages in both sectors and reallocates employment from the informal to the formal sector (point B). Overall, the effect on formal-sector employment caused by the reduction of the payroll taxes $\left(L_1^{f*} - L_0^{f*}\right)$ is smaller than that observed in a case of a binding minimum wage $\left(\frac{\eta^f}{1+t}dt\right)$, but larger than that observed in a case with full pass-through from taxes to wages and no informal sector (0).

Figure 3.2: Formal-sector employment and wages by age, 2010–2012

(a) Density of formal-sector employment (new workers)

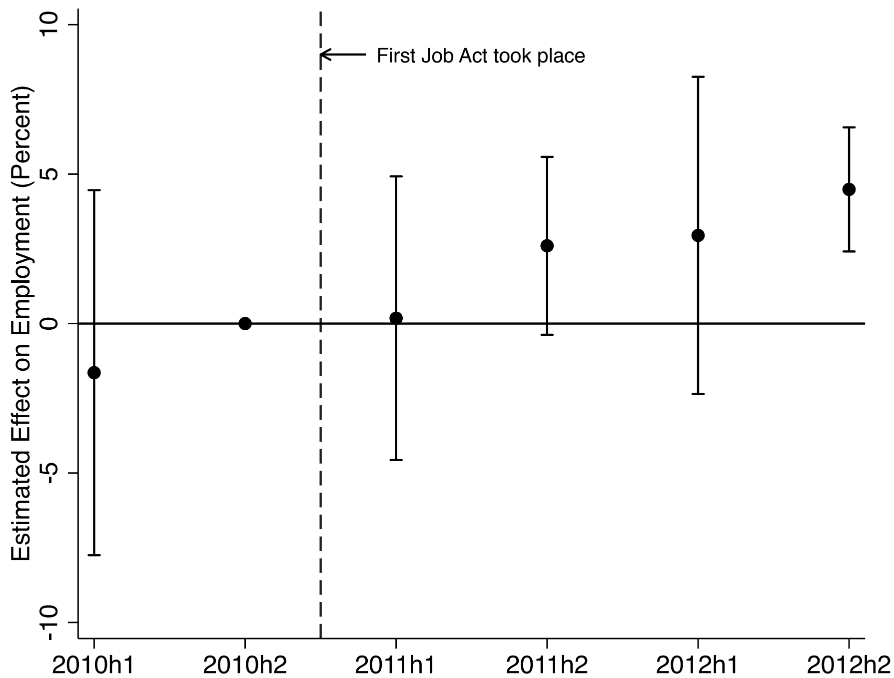


(b) Average log-wages (new workers)



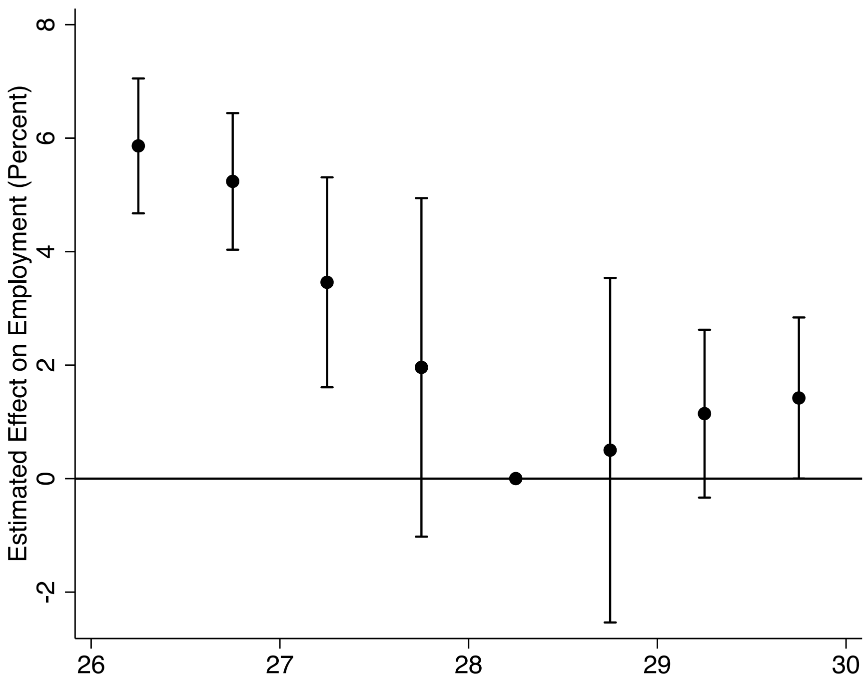
Notes: This Figure shows the differences in the density of employment and the average wages by age before (2010) and after (2011-12) the entry of the First Job Act. Confidence bands (95 percent) were computed by using standard errors clustered by age group (in quarters).

Figure 3.3: Estimated employment effects by Semester, 2010–2012



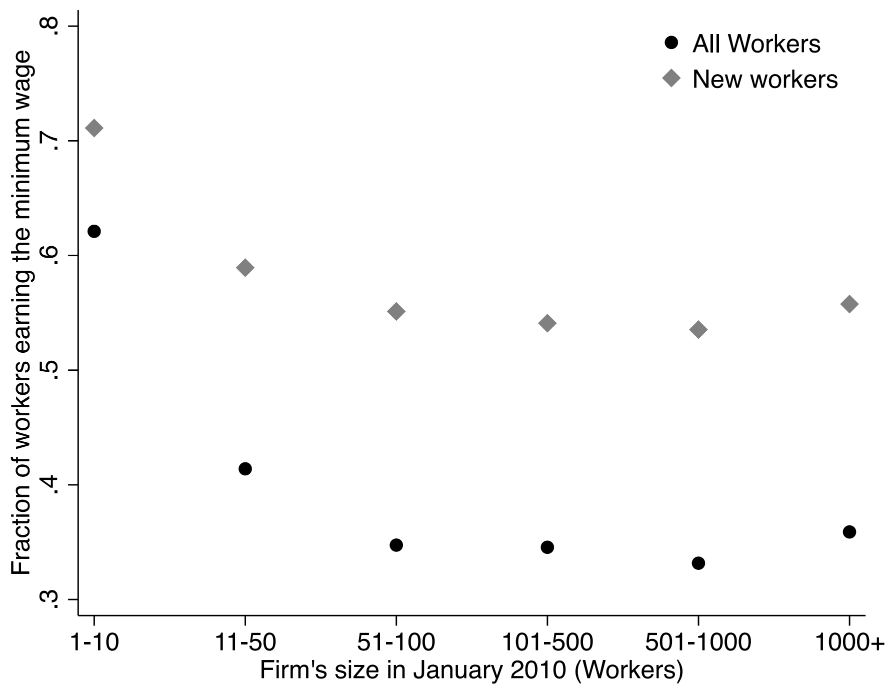
Notes: This Figure investigates whether the employment effect of the First Job Act on employment changes over time. Each point in the graph represents the estimated employment effect (in percentage points), estimated using a differences-in-differences strategy in which I allow that the employment effect varies by semester. I use as a base category the second semester of 2010, so the reported coefficients are relative to the difference between the density of entrants younger and older than 28 in that date. Confidence bands are 95 percent confidence intervals using standard errors clustered by age group (in quarters).

Figure 3.4: Estimated employment effects by Age Group



Notes: This Figure investigates whether the employment effect of the First Job Act is heterogeneous by age. Each point in the graph represents the estimated employment effect (in percentage points) estimated using a differences-in-differences strategy in which I allow that the employment effect varies by age, grouped in 6-month bins. I use as base category the group aged 28 to 28.5 years, so the reported coefficients are relative to the difference between employment for that group. Confidence bands are 95 percent confidence intervals using standard errors clustered by age group (in quarters).

Figure 3.5: Fraction of workers earning the minimum wage by firm size, 2010-2012



Notes: This Figure investigates whether minimum wage is more binding for smaller firms. The figure displays the fraction of workers earning the minimum wage between February 2010 and November 2012, according to the number of workers the employer had in January 2010. The figure presents the fraction of minimum wage earners for all workers (black dots) and new hires (grey dots).

Bibliography

James Albrecht, Lucas Navarro, and Susan Vroman. The Effects of Labour Market Policies in an Economy with an Informal Sector. *Economic Journal*, 119(539):1105–1129, 07 2009.

Rita Almeida and Pedro Carneiro. Enforcement of labor regulation and informality. *American Economic Journal: Applied Economics*, 4(3):64–89, July 2012.

Luis Arango Thomas, Monica Alexandra Gómez, and Carlos Posada. La demanda de trabajo formal en Colombia: determinantes e implicaciones de política. Borradores de Economia 563, Banco de la Republica, 2009.

Linda A. Bell. The impact of minimum wages in Mexico and Colombia. *Journal of Labor Economics*, 15(S3):S102–S135, 1997.

Mariano Bosch, Ángel Melguizo, and Carmen Pagés. *Better Pensions, Better Jobs: Towards Universal Coverage in Latin America and the Caribbean.* IADB, 2013.

Adriana Camacho, Emily Conover, and Alejandro Hoyos. Effects of colombia's social protection system on workers' choice between formal and informal employment. *The World Bank Economic Review*, 28(3):446–466, 2013.

Mauricio Cardenas and Raquel Bernal. Determinants of labor demand in Colombia. 1976-1996. In James J. Heckman and Carmen Pagés, editors, *Law and Employment: Lessons from Latin America and the Caribbean*, pages 229–272. University of Chicago Press, August 2004.

Eduardo Cavallo and Tomás Serebrisky, editors. *Saving for Development*. Springer Nature, 2016.

Raj Chetty. A general formula for the optimal level of social insurance. *Journal of Public Economics*, 90(10-11):1879–1901, 2006.

Ana Corbacho, Vicente Fretes Cibils, and Eduardo Lora, editors. *More than Revenue: Taxation as a development tool*. Springer, 2013.

Guillermo Cruces, Sebastian Galiani, and Susana Kidyba. Payroll taxes, wages and employment: Identification through policy changes. *Labour Economics*, 17(4):743 – 749, 2010.

Nada Eissa, Henrik Jacobsen Kleven, and Claus Thustrup Kreiner. Evaluation of four tax reforms in the United States: Labor supply and welfare effects for single mothers. *Journal of Public Economics*, 92(3 – 4):795 – 816, 2008.

Martin Feldstein and Jeffrey B. Liebman. Social security. In *Handbook of Public Economics*, volume 4, pages 2245 – 2324. Elsevier, 2002.

Markus Frölich, David Kaplan, Carmen Pagés, Jamele Rigolini, and David Robalino. *Social Insurance, Informality, and Labour Markets: How to Protect Workers While Creating Good Jobs*. Oxford University Press, 2014.

S. Galiani and F. Weinschelbaum. Modeling informality formally: households and firms. *Economic Inquiry*, 50(3):821–838, 2011.

Francois Gerard and Gustavo M. Gonzaga. Informal labor and the cost of social programs: Evidence from 15 years of unemployment insurance in Brazil. Unpublished Manuscript, 2014.

Jonathan Gruber. The incidence of payroll taxation: Evidence from Chile. *Journal of Labor Economics*, 15(3):S72–101, 1997.

Jonathan Gruber. Payroll taxation, employer mandates, and the labor market. *Employee benefits and labor markets in Canada and the United States*, page 183, 2000.

Jonathan Gruber and Alan B Krueger. The incidence of mandated employer-provided insurance: Lessons from workers' compensation insurance. In *Tax Policy and the Economy, Volume 5*, pages 111–144. The MIT Press, 1991.

Guido Imbens and Karthik Kalyanaraman. Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, 79(3):933–959, 2012.

Guido W Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, 2008.

Clement Joubert. Pension design with a large informal labor market: Evidence from Chile. *International Economic Review*, 56(2):673–694, 2015. ISSN 1468-2354.

Francis Kramarz and Thomas Philippon. The impact of differential payroll tax subsidies on minimum wage employment. *Journal of Public Economics*, 82(1):115 – 146, 2001.

Adriana Kugler and Maurice Kugler. Labor market effects of payroll taxes in developing countries: Evidence from Colombia. *Economic Development and Cultural Change*, 57(2):pp. 335–358, 2009.

Rafael La Porta and Andrei Shleifer. Informality and development. *Journal of Economic Perspectives*, 28(3):109–26, 2014.

David S Lee and David Card. Regression discontinuity inference with specification error. *Journal of Econometrics*, 142(2):655–674, 2008.

S. Levy. *Good intentions, bad outcomes: Social policy, informality, and economic growth in Mexico*. Brookings Inst Press, 2008.

Jeffrey B. Liebman, Erzo F.P. Luttmer, and David G. Seif. Labor supply responses to marginal social security benefits: Evidence from discontinuities. *Journal of Public Economics*, 93(11-12):1208–1223, 2009.

Jorge Llano, Jaime Cardona, Natalia Guevara, Gonzalo Casas, Camilo Arias, and Fernando Cardozo. Movilidad e Interacción entre Regímenes del Sistema General de Pensiones Colombiano. Technical report, Ministry of Finance, 2013.

William Maloney and Jairo Nuñez. Measuring the impact of minimum wages. evidence from Latin America. In *Law and Employment: Lessons from Latin America and the Caribbean*, pages 109–130. University of Chicago Press, 2004.

William F Maloney. Does informality imply segmentation in urban labor markets? evidence from sectoral transitions in Mexico. *World Bank Economic Review*, 13(2):275–302, May 1999.

William F Maloney. Informality revisited. *World Development*, 32(7):1159 – 1178, 2004. ISSN 0305-750X.

Justin McCrary. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, 142(2):698 – 714, 2008. The regression discontinuity design: Theory and applications.

Costas Meghir, Renata Narita, and Jean-Marc Robin. Wages and informality in developing countries. *American Economic Review*, 105(4):1509–46, 2015.

Camilo Mondragón-Vélez, Ximena Peña, and Daniel Wills. Labor market rigidities and informality in Colombia [with comment]. *Economía*, 11(1): 65–101, 2010.

OECD. *Taxing Wages*. OECD Publishing, 2016.

Guillermo E. Perry, William F. Maloney, Omar S. Arias, Pablo Fajnzylber, Andrew D. Mason, and Jaime Saavedra-Chanduvi. *Informality: Exit and exclusion*. World Bank Publications, 2007.

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

Mark J. Roberts and Emmanuel Skoufias. The long-run demand for skilled and unskilled labor in Colombian manufacturing plants. *The Review of Economics and Statistics*, 79(2):330–334, 1997.

Emmanuel Saez, Manos Matsaganis, and Panos Tsakloglou. Earnings determination and taxes: Evidence from a cohort-based payroll tax reform in Greece. *The Quarterly Journal of Economics*, 127(1):493–533, 2012.

Mauricio Santa María, Fabián García, and Ana Virgina Mujica. Los costos no laborales y el mercado laboral: impacto de la reforma de salud en Colombia. Working Paper Series 009186, Fedesarrollo, August 2009.

Mauricio Santa María, Roberto Steiner, Jorge Humberto Botero, Mariana Martinez, Natalia Millán, Maria Alejandra Arias, and Erika Schutt. El Sistema Pensional en Colombia: Retos y Alternativas para Aumentar la Cobertura (Informe Final). Technical report, Fedesarrollo, 2010.

James H. Stock and David A. Wise. Pensions, the option value of work, and retirement. *Econometrica*, 58(5):pp. 1151–1180, 1990.

Lawrence H. Summers. Some simple economics of mandated benefits. *The American Economic Review*, 79(2):pp. 177–183, 1989.

Gabriel Ulyssea. Regulation of entry, labor market institutions and the informal sector. *Journal of Development Economics*, 91(1):87–99, January 2010.

Appendix A

Appendix to Chapter 1

A.1 Model implications and robustness tests

The conditions characterizing the retirement and search decisions have useful implications to understand the empirical results of the paper. To simplify notation, let $u^f = u\left(w^f\left(1-t^{nom}\right)\right)$ and $u^i = u\left(w^i\right)$ denote the utility levels the worker receives when working in the formal and informal sector, $u^r = u\left(bw^f\right)$ the utility the worker gets when he retires and is eligible for pension benefits, and $u^0 = u\left(0\right)$ the baseline utility the worker receives when he retires but is not entitled to pension benefits. Thus, $\tilde{u} = u^f + \theta^f - u^i$ is the gap (in utility terms) between the formal an informal sector and $\bar{u}_a\left(\tau\right) = \tilde{u} + \beta\Delta v_{a+1}\left(\tau+1\right)$.

Most of the proofs use backward induction.

Proposition 1. *A replacement rate of $b = 1$ implies that the worker retires as soon as he meets the requirements.*

Proof. Assume $b = 1$ and $\tau^* \leq R < T$. In period T, the value function for retirement is given by

$$v_T^r\left(\tau_{T-1}\right) = \begin{cases} u^0 & \text{if } \tau_{T-1} < \tau^* \\ u^r + \theta^r & \text{if } \tau_{T-1} \geq \tau^* \end{cases}.$$

The value function if the worker continues working is

$$v_T^w\left(\tau_{T-1}\right) = u^i + G\left(\tilde{u}\right)\mathbb{E}\left(\tilde{u} - \psi_T \middle| \psi_T \leq \tilde{u}\right).$$

By assumption $u^0 < u^i$, and so the comparison of both value functions implies that the worker retires when $\tau_{T-1} \geq \tau^*$, and

$$v_T(\tau_{T-1}) = \begin{cases} v_T^w(\tau_{T-1}) & \text{if } \tau_{T-1} < \tau^* \\ v_T^r(\tau_{T-1}) & \text{if } \tau_{T-1} \geq \tau^* \end{cases}.$$

For period $T-1$, the value function conditional on retirement and working are equal to

$$v_{T-1}^r(\tau_{T-2}) = \begin{cases} u^0 + \beta v_T^r(\tau_{T-2}) & \text{if } \tau_{T-2} < \tau^* \\ u^r + \theta^r + \beta v_T^r(\tau_{T-2}) & \text{if } \tau_{T-2} \geq \tau^* \end{cases}.$$

$$\begin{aligned} v_{T-1}^w(\tau_{T-2}) = {} & u^i + \beta v_T(\tau_{T-2}) + \\ & G(\bar{u}_{T-1}(\tau_{T-2})) \times \\ & \mathbb{E}(\bar{u}_{T-1}(\tau_{T-2}) - \psi_{T-1} | \, \psi_{T-1} \leq \bar{u}_{T-1}(\tau_{T-2})). \end{aligned}$$

The second term of the latter equation is non negative, which implies that the worker does not retire when $\tau_{T-2} < \tau^*$. When $\tau_{T-2} \geq \tau^*$, rewrite $v_{T-1}^w(\tau_{T-2})$ as

$$\begin{aligned} v_{T-1}^w(\tau_{T-2}) = {} & (1 - G(\bar{u}_{T-1}(\tau_{T-2})))\left(u^i + \beta v_T(\tau_{T-2})\right) \\ & + G(\bar{u}_{T-1}(\tau_{T-2}))\left(u^f + \theta^f + \beta v_T(\tau_{T-2} + 1)\right) \\ & - G(\bar{u}_{T-1}(\tau_{T-2}))\,\mathbb{E}(\psi_{T-1} | \, \psi_{T-1} \leq \bar{u}_{T-1}(\tau_{T-2})) \end{aligned}$$

which is strictly less than $v_{T-1}^r(\tau_{T-2})$, and therefore the worker retires if $\tau_{T-2} \geq \tau^*$. A similar analysis applies for $a = R, R+1, \ldots, T-2$.

For $a \leq R-1$, the worker cannot claim pension benefits even if $\tau_{a-1} \geq \tau^*$. The value function if he retires is $v_a^r(\tau_{a-1}) = u^0 + \beta v_a^r(\tau_{a-1})$, which is less than $v_a^w(\tau_{a-1})$. As a result, he does not retire before period R. □

Proposition 2. *The intensity of the search for formal-sector jobs depends on the likelihood of getting retirement benefits.*

Proof. For simplicity, assume $b = 1$ so the worker retires as soon as he mets the requirements. The proof of the proposition has two parts. First, I show that for workers for whom $\tau \geq \tau^*$ or $\tau + (T - a + 1) < \tau^*$, $\Delta v_a(\tau + 1) = 0$ and therefore $\bar{u}_a(\tau) = \tilde{u}$. Second, I show that $\Delta v_a(\tau + 1) \geq 0$ for all other values of τ, and so $\bar{u}_a(\tau) \geq \tilde{u}$ for all τ. As a result, workers who still have a chance of meeting the minimum requirement conditions are the ones who search more actively for formal-sector jobs.

First, there are two cases in which the accrual value of a period worked in the formal sector is zero: (i) when workers are vested ($\tau_{a-1} \geq \tau^*$) and when workers do not have enough periods to reach τ^* ($\tau_{a-1} + (T - a + 1) < \tau^*$).

For the first part of the proof, note that when $a \geq R$ and $\tau \geq \tau^*$, the optimal retirement decision implies that $v_a(\tau + 1) = v_a(\tau) = v_a^r(\tau)$ for $a = R, \ldots T - 1$. For the case $a = R - 1$ and $\tau \geq \tau^*$, condition (1.2) implies that $u_{R-1}(\tau) = \tilde{u}$, and therefore the accrual value of an additional period worked in the formal sector is

$$\Delta v_{R-1}(\tau + 1) = (1 - G(\tilde{u}))\,\beta \Delta v_R(\tau + 1) + G(\tilde{u})\,\beta \Delta v_R(\tau + 2) = 0.$$

The same argument can be extended for $a = \tau^*, \ldots, R - 2$.

When $\tau_{T-1}+1 < \tau^*$, $\Delta v_T(\tau_{T-1} + 1) = 0$ and $\bar{u}_{T-1}(\tau) = \tilde{u}$ for $\tau+1 < \tau^*$. Using backward induction, the result follows.

Second, for $\tau \in [\tau^* - (T - a + 1), \tau^* - 1]$, $\Delta v_a(\tau + 1) \geq 0$. To see this, note first that from the definition of $v_T(\tau_{T-1})$ presented above, $v_T(\tau + 1) \geq v_T(\tau)$ for all τ. For any other period $a < T$, assume $\bar{u}_a(\tau + 1) \geq \bar{u}_a(\tau)$, and rewrite the first difference of the value function as

$$\begin{aligned}
\Delta v_a(\tau + 1) = &(1 - G(\bar{u}_a(\tau)))\,\beta \Delta v_{a+1}(\tau + 1) + \\
&G(\bar{u}_a(\tau + 1))\,\beta \Delta v_{a+1}(\tau + 2) + \\
&(G(\bar{u}_a(\tau + 1)) - G(\bar{u}_a(\tau))) \times \\
&\mathbb{E}(\tilde{u} - \psi_a |\, \bar{u}_a(\tau) \leq \psi_a \leq \bar{u}_a(\tau + 1)) \\
\geq &(1 - G(\bar{u}_a(\tau)))\,\beta \Delta v_{a+1}(\tau + 1) + G(\bar{u}_a(\tau))\,\beta \Delta v_{a+1}(\tau + 2) \\
\geq &0
\end{aligned}$$

and thus $\Delta v_a(\tau + 1) \geq 0$. A similar argument can be used to show the result when $\bar{u}_a(\tau+1) \leq \bar{u}_a(\tau)$. $\square$

Proposition 3. *Assume $b = 1$. Holding all other variables constant, a change in the minimum retirement age R affects the incentives to search for formal-sector jobs. The effect is ambiguous and depends on a and τ_{a-1}.*

Proof. Assume an increase in the minimum age of retirement from R to R'. To characterize the full set of cases, assume that $R' - R \geq 3$. Since b equals one, workers retire as soon as they meet the requirements, and therefore a change in the minimum age of retirement affects the incentives to search for formal-sector jobs.[32]

Let $v_a(\tau)$ and $v'_a(\tau)$ denote the value functions for the workers under R and R'. The effect of a change in the minimum retirement age depends on the worker's age a. For workers with $a \geq R'$, there is no labor supply response, as a change in the retirement age does not change their retirement behavior. Therefore, $v_a(\tau) = v'_a(\tau)$ for all τ and $a = R', \ldots, T$.

For $R \geq a > R'$, the effect of changes in the retirement age the effect is ambiguous. To see this, start with $a = R' - 1$. Since $\Delta v_{R'}(\tau+1) = \Delta v'_{R'}(\tau+1)$, the long-run gains from retirement do not change for this group, and so $\bar{u}_{R'-1}(\tau) = \bar{u}'_{R'-1}(\tau)$. However, workers with $\tau \geq \tau^*$ are no longer eligible to retire, and so they search for formal-sector jobs. As a result, $v_{R'-1}(\tau) = v'_{R'-1}(\tau)$ for $\tau < \tau^*$ and $v_{R'-1}(\tau) \geq v'_{R'-1}(\tau)$ for $\tau \geq \tau^*$ (otherwise the retirement decision would not have been optimal). For this age group, there is an increase in formal-sector employment, as the workers who are not longer eligible to retire search for formal-sector jobs – driven by short-run gains only. A direct implication of the definition of $v'_{R'-1}(\tau)$ is that $\Delta v_{R'-1}(\tau+1) = \Delta v'_{R'-1}(\tau+1)$ for $\tau \neq \tau^* - 1$ and $\Delta v_{R'-1}(\tau^*) \geq \Delta v'{}_{R'-1}(\tau^*)$.

For $a = R' - 2$, the change in the incentives to search for formal-sector jobs depends on τ. The definitions of $v_{R'-1}(\tau)$ and $v'_{R'-1}(\tau)$ imply that

[32] The assumption $b = 1$ is not a necessary condition for the proof. In order that a change in the minimum retirement age generates changes in the incentives to search for formal-sector jobs, it is necessary that at least a group of workers finds optimal to retire in an age R^* such that $R \leq R^* < R'$. Otherwise, the minimum retirement age is not binding and workers do not respond to the change.

$\bar{u}_{R'-2}(\tau) = \bar{u}'_{R'-2}(\tau)$ for $\tau \neq \tau^* - 1$ and $\bar{u}_{R'-2}(\tau^* - 1) \geq \bar{u}'_{R'-2}(\tau^* - 1)$. Thus, the effect of a change in the minimum retirement age on the formal-sector labor supply response for is ambiguous, as there is a group that is not affected by the measure (those with $\tau \leq \tau^* - 2$), a group that reduces its searching for formal-sector jobs ($\tau = \tau^* - 1$), and a group that increases their formal-labor supply, as they would have retired under the previous conditions ($\tau \geq \tau^*$). Using the definition of $v_a(\tau)$, $v_{R'-2}(\tau) = v'_{R'-2}(\tau)$ for $\tau < \tau^* - 1$ and $v_{R'-2}(\tau) > v'_{R'-2}(\tau)$ for $\tau \geq \tau^* - 1$.

Finally, $\Delta v_{R'-2}(\tau + 1) = \Delta v'_{R'-2}(\tau + 1)$ for $\tau \notin \{\tau^* - 2, \tau^* - 1\}$ and $\Delta v_{R'-2}(\tau + 1) \geq \Delta v'_{R'-2}(\tau + 1)$ for $\tau \in \{\tau^* - 2, \tau^* - 1\}$ since

$$\begin{aligned}
\overline{\Delta v_{R'-2}(\tau^* - 1)} &= \Delta v_{R'-2}(\tau^* - 1) - \Delta v'_{R'-2}(\tau^* - 1) \\
&= \beta\left(v_{R'-2}(\tau^* - 1) - v_{R'-2}(\tau^* - 2)\right) \\
&\quad - \beta\left(v'_{R'-2}(\tau^* - 1) - v'_{R'-2}(\tau^* - 2)\right) \\
&= \beta\left(v_{R'-2}(\tau^* - 1) - v'_{R'-2}(\tau^* - 1)\right) \\
&\geq 0
\end{aligned}$$

and

$$\begin{aligned}
\overline{\Delta v_{R'-2}(\tau^*)} &= \Delta v_{R'-2}(\tau^*) - \Delta v'_{R'-2}(\tau^*) \\
&= \left(1 - G\left(\bar{u}_{R'-2}(\tau^* - 1)\right)\right)\beta \Delta v_{R'-1}(\tau^*) \\
&\quad - \left(1 - G\left(\bar{u}'_{R'-2}(\tau^* - 1)\right)\right)\beta \Delta v'_{R'-1}(\tau^*) \\
&\quad - \left(G\left(\bar{u}_{R'-2}(\tau^* - 1)\right) - G\left(\bar{u}'_{R'-2}(\tau^* - 1)\right)\right) \times \\
&\quad \mathbb{E}\left(\tilde{u} - \psi_{R'-2} \middle| \bar{u}'_{R'-2}(\tau^* - 1) \leq \psi_{R'-2} \leq \bar{u}_{R'-2}(\tau^* - 1)\right) \\
&\geq \left(1 - G\left(\bar{u}_{R'-2}(\tau^* - 1)\right)\right)\beta\left(\Delta v_{R'-1}(\tau^*) - \Delta v'_{R'-1}(\tau^*)\right) \\
&\geq 0.
\end{aligned}$$

For $a = R' - 3$, the change in the minimum retirement age have the same type of composition effects as those for $a = R' - 2$. In this case, the workers that exhibit a reduction in their formal-sector labor supply are those with $\tau \in \{\tau^* - 3, \tau^* - 2, \tau^* - 1\}$. Again, $v_{R'-3}(\tau) = v'_{R'-3}(\tau)$ for $\tau < \tau^* - 2$ and $v_{R'-3}(\tau) \geq v'_{R'-3}(\tau)$ otherwise. Moreover, $\Delta v_{R'-3}(\tau + 1) = \Delta v'_{R'-3}(\tau + 1)$

for $\tau \notin \{\tau^* - 3, \tau^* - 2, \tau^* - 1\}$ while $\Delta v_{R'-3}(\tau + 1) \geq \Delta v'_{R'-3}(\tau + 1)$ for $\tau \in \{\tau^* - 3, \tau^* - 2, \tau^* - 1\}$. The proof of $\Delta v_{R'-3}(\tau + 1) \geq \Delta v'_{R'-3}(\tau + 1)$ for $\tau \in \{\tau^* - 3, \tau^* - 1\}$ follows the same steps as those for $a = R' - 2$. For $\tau = \tau^* - 2$, difference between the value functions is equal to

$$\begin{aligned}
\overline{\Delta v_{R'-3}(\tau^* - 1)} &= \Delta v_{R'-3}(\tau^* - 1) - \Delta v'_{R'-3}(\tau^* - 1) \\
&= (1 - G(\bar{u}_{R'-3}(\tau^* - 2)))\, \beta \Delta v_{R'-2}(\tau^* - 1) \\
&\quad - \left(1 - G\left(\bar{u}'_{R'-3}(\tau^* - 2)\right)\right) \beta \Delta v'_{R'-2}(\tau^* - 1) \\
&\quad + G(\bar{u}_{R'-3}(\tau^* - 1))\, \beta \Delta v_{R'-2}(\tau^*) \\
&\quad - G\left(\bar{u}'_{R'-3}(\tau^* - 1)\right) \beta \Delta v'_{R'-2}(\tau^*) \\
&\quad + \left(G(\bar{u}_{R'-3}(\tau^* - 1)) - G\left(\bar{u}'_{R'-3}(\tau^* - 1)\right)\right) \times \\
&\quad \mathbb{E}\left(\tilde{u} - \psi_{R'-3} \middle| \bar{u}'_{R'-3}(\tau^* - 1) \leq \psi_{R'-3} \leq \bar{u}_{R'-3}(\tau^* - 1)\right) \\
&\quad - \left(G(\bar{u}_{R'-3}(\tau^* - 2)) - G\left(\bar{u}'_{R'-3}(\tau^* - 2)\right)\right) \times \\
&\quad \mathbb{E}\left(\tilde{u} - \psi_{R'-3} \middle| \bar{u}'_{R'-3}(\tau^* - 2) \leq \psi_{R'-3} \leq \bar{u}_{R'-3}(\tau^* - 2)\right) \\
&\geq (1 - G(\bar{u}_{R'-3}(\tau^* - 2))) \times \\
&\quad \beta\left(\Delta v_{R'-2}(\tau^* - 1) - \Delta v'_{R'-2}(\tau^* - 1)\right) \\
&\quad + G\left(\bar{u}'_{R'-3}(\tau^* - 1)\right) \beta\left(\Delta v_{R'-2}(\tau^*) - \Delta v'_{R'-2}(\tau^*)\right) \\
&\geq 0.
\end{aligned}$$

Using backward induction, the implications above apply for all age groups $a = \{R, \ldots, R' - 3\}$.

For $a < R$, the effect of a change in the minimum age for retirement is a reduction of the formal-sector labor supply. In this case, the searching efforts of two types of workers are not affected by the change in R: workers who are too far to retire before R' ($\tau_{a-1} + R' - a < \tau^*$) and those who already met the vesting period ($\tau_{a-1} \geq \tau^*$). For all other workers, the change in R reduces their labor supply (the proof is similar to the presented above). Thus, for $a < R$, $v_{a+1}(\tau) = v'_{a+1}(\tau)$ for $\tau < \tau^* + (R' - a + 1)$ and $v_{a+1}(\tau) > v'_{a+1}(\tau)$ otherwise. In addition, $\Delta v_{a+1}(\tau + 1) \geq \Delta v'_{a+1}(\tau + 1)$ for $\tau \in \{\tau^* - (R' - a), \ldots, \tau^* - 1\}$ and $\Delta v_{a+1}(\tau + 1) = \Delta v'_{a+1}(\tau + 1)$ otherwise. $\square$

Proposition 4. *Assume $b = 1$. Holding all other variables constant, a change in the vesting period τ^* affects the incentives to search for formal-sector jobs. The effect is ambiguous and depends on a and τ_{a-1}.*

Proof. Given an increase of the vesting period from τ^* to τ', the optimal retirement and searching policy change. Using the arguments presented in propositions 1 and 2, solving the model by backward induction yields a type of policy like the one presented before, but it uses τ' as a reference point instead of τ^*. Thus, the policy function shifts rightwards. The shift generates two types of changes within each cohort. Workers with $\tau_{a-1} \in \{\tau^*, \ldots, \tau'\}$ increase their searching efforts, as they are not vested yet, and workers with low values of τ_{a-1} tend to reduce their efforts, as the probability of reaching the vesting period goes down. □

Table A.1: Robustness test, 2005

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
A: Least Squares estimator (estimates scaled up by 100)						
Linear control function	-2.79	-2.61	-2.87	-0.84	-0.74	-0.56
(Bandwidth 48 months)	[0.94]***	[0.91]***	[0.85]***	[0.70]	[0.70]	[0.70]
Quadratic control function	-2.62	-2.36	-2.52	-0.18	-0.13	-0.15
(Bandwidth 48 months)	[1.28]**	[1.34]*	[1.14]**	[1.03]	[1.00]	[1.05]
B: Logit estimator (estimates scaled up by 100)						
Linear control function	-2.82	-2.84	-2.48	-0.87	-0.77	-0.44
(Bandwidth 48 months)	[0.96]***	[1.04]***	[0.74]***	[0.71]	[0.71]	[0.54]
Quadratic control function	-2.48	-2.39	-2.00	-0.15	-0.10	-0.10
(Bandwidth 48 months)	[1.27]*	[1.42]*	[0.95]**	[1.04]	[1.00]	[0.80]
C: Local linear estimator (estimates scaled up by 100)						
Local linear	-2.64	–	–	-1.10	–	–
(Bandwidth 24 months)	[1.16]**			[0.94]		
Local linear	-2.80	–	–	-0.68	–	–
(Bandwidth 36 months)	[1.03]***			[0.83]		
Observations	129,061	129,061	129,061	178,990	178,990	178,990
Mean dep. variable (%)	18.1	18.1	18.1	15.7	15.7	15.7
Fixed effects		Month of birth	High School or less		Month of birth	High School or less

Notes: Each cell reports an RD estimate based on a separate regression of an indicator variable of whether the person is a salaried worker contributing to the pension system and covered by the contributory health care system versus a polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). Columns (1) and (5) present the baseline regressions without any fixed effects, while the other columns include fixed effects to test the sensitivity of the results. The included fixed effects are based on categories of month of birth, educational attainment, and region. Regressions were computed using the Colombian Census long-form questionnaire dataset (2005). Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.2: Robustness test, 2011

	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
A: Least Squares estimator (estimates scaled up by 100)						
Linear control function	4.32	3.36	4.32	5.10	4.99	5.10
(Bandwidth 730 days)	[1.92]**	[1.43]**	[1.92]**	[1.98]**	[1.60]***	[1.98]**
Quadratic control function	6.79	5.84	6.79	2.03	1.89	2.03
(Bandwidth 730 days)	[2.39]***	[2.04]***	[2.39]***	[2.17]	[1.58]	[2.17]
B: Local linear estimator (estimates scaled up by 100)						
Local linear	8.39	–	–	3.70	–	–
(Bandwidth 360 days)	[2.24]***			[1.63]**		
Local linear	6.74	–	–	3.85	–	–
(Bandwidth 540 days)	[2.05]***			[1.69]**		
Observations	964,558	964,558	964,558	927,691	927,691	927,691
Fixed effects		Month of birth	Month of contribution		Month of birth	Month of contribution

Notes: Each cell reports an RD estimate based on a separate regression of the log number of salaried formal workers contributing to the pension system and the contributory health care system versus a polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). Columns (1) and (4) present the baseline regressions without any fixed effects, while the other columns include fixed effects to test the sensitivity of the results. The included fixed effects are based on categories of month of birth and month of contribution. Regressions were computed using the PILA dataset (2011). Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table A.3: Estimation results with alternative definitions of formal employment,2005

	Men				Women			
	Salaried-formal (pension)	Formal worker	Formal worker (Colp)	Pension (all)	Salaried-formal (pension)	Formal worker	Formal worker (Colp)	Pension (all)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Least Squares estimator (estimates scaled up by 100)								
Harder qualifying conditions	-2.57	-2.71	-2.17	-3.49	-0.16	-0.36	0.02	-0.16
(Bandwidth 48 months)	[1.42]*	[1.36]**	[0.90]**	[1.29]***	[1.08]	[0.91]	[0.50]	[1.02]
Observations	129,061	129,061	129,061	129,061	178,990	178,990	178,990	178,990
Mean dep. variable (%)	19.1	22.0	5.5	25.8	16.3	17.6	3.4	20.9

Notes: Each cell reports an RD estimate based on a separate regression of formal employment versus a quadratic polynomial on date of birth and its interaction with a dummy for being born after March-54 (men) and March-59 (women) as independent variables (See equation (1.6)). The definitions used are (i) salaried-formal employment based on contributions to the pension system, regardless of coverage of the contributory health care system; (ii) formal employment for all workers contributing to the pension and covered by the contributory health care system, regardless of type of employment; (iii) formal employment for workers contributing to the pension system and covered by the contributory health care system managed by Colpensiones; and (iv) an indicator for all people contributing to the pension system, regardless of their labor force participation. Regressions were estimated using the Colombian Census long-form questionnaire dataset. Standard errors clustered by date of birth (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Appendix B

Appendix to Chapter 2

B.1 Formal-sector labor supply of the young worker

In this appendix, I show that the properties of the function $\bar{u}_a(\tau_{a-1}, A_{a-1})$ described in Section 2.3 also hold for the young worker ($a = 1$).

Given an initial endowment of formal-sector experience τ_0 and assets A_0, the problem for the young worker is

$$v_1(\tau_0, A_0) = \max_{c_1, A_1, h_1} u(c_1) - \psi_1 h_1 + \beta \mathbb{E} v_2(\tau_1, A_1) \tag{B.1}$$

subject to

$$A_1 = (1+r) A_0 + w_1^i + h_1 \left(w_1^f (1 - t_c - t_p) - w_1^i \right) - c_1 \tag{B.2}$$

$$\tau_1 = \tau_0 + h_1 \tag{B.3}$$

where the expected value function $\mathbb{E}v_2(\tau_1, A_1)$ is defined in equation (2.17).

The optimal consumption and formal-sector labor supply plan for the young worker does not have an analytical solution. However, conditional on a formal-sector labor supply decision being made, the solution is characterized by the system of equations

$$u'(c_1) = \mathbb{E}u'(c_2(\tau_1, A_1)) \tag{B.4}$$

$$A_1 = (1+r) A_0 + w_1^i + h_1 \left(w_1^f (1 - t_c - t_p) - w_1^i \right) - c_1 \tag{B.5}$$

$$\tau_1 = \tau_0 + h_1. \tag{B.6}$$

To find the utility gains from working in the formal sector, I first describe the response of the optimal consumption plan and the value function with to changes in h_1. Keeping h_1 constant, there exists a consumption plan that satisfies equations (B.4) and (B.5). Implicit differentiation on equation (B.4) yields

$$\frac{\partial c_1}{\partial A_1} = \frac{\Psi_{A_1}}{u''\left(c_1\right)} \tag{B.7}$$

where

$$\begin{aligned}\Psi_{A_1} = &\left(\frac{1+r}{1+\beta}\right)\mathbb{E}u''\left(c_2\right) + \\ &(1+r)\, g\left(\bar{u}_2\left(\tau_1, A_1\right)\right)\left(u'\left(c_2^f\right) - u'\left(c_2^i\right)\right)^2.\end{aligned} \tag{B.8}$$

In equation (B.8), $g\left(\psi\right)$ is the density function of the random variable ψ. Using the properties of the utility function, equation (B.7) implies that in the space (A_1, c_1), the sign of the expression in equation (B.7) is ambiguous. The first term of Ψ_{A_1} accounts for the effect that an increase in savings has on future consumption, while the second term accounts for the effect that an increase in savings has on the ex-ante probability of working in the formal sector in the future. In what follows, I assume that the effect of assets on the probability of working in the formal sector is small enough, such that $\Psi_{A_1} \leq 0$. This assumption rules out equilibria where the worker reduces his current consumption to reduce his sensitivity to the gains from working in the formal-sector in the future. Thus, assuming $\Psi_{A_1} \leq 0$, the equilibrium is unique, as the locus defined in equation (B.4) cuts the locus (B.5) from below in the (A_1, c_1) space.

Let c_1^f, A_1^f, c_1^i, and A_1^i denote the consumption and saving plans conditional on working in the formal and the informal sector, respectively. Then, the worker chooses to work in the formal sector if

$$\begin{aligned}\bar{u}_1\left(\tau_0, A_0\right) &= u\left(c_1^f\right) + \beta\mathbb{E}v_2\left(\tau_0+1, A_0^f\right) - \left(u\left(c_1^i\right) + \beta\mathbb{E}v_2\left(\tau_0, A_0^i\right)\right) \\ &= \tilde{v}_1^f\left(\tau_0, A_0\right) - \tilde{v}_1^i\left(\tau_0, A_0\right) \geq \psi_1,\end{aligned} \tag{B.9}$$

and the ex-ante probability of working in the formal sector is equal to

$$P\left(h_1 = 1 \middle|\, \tau_0, A_0\right) = G\left(\bar{u}_1\left(\tau_0, A_0\right)\right). \tag{B.10}$$

Next, I show that the value function (without utility shocks) is increasing in h_1, and therefore $\bar{u}_1\left(\tau_0, A_0\right)$ is always non-negative. Define $\tilde{v}_1$ as $\tilde{v}_1\left(\tau_0, A_0\right) = u\left(c_1\right) + \mathbb{E}v_2\left(\tau_1, A_1\right)$ evaluated at the optimal consumption plan given an arbitrary value of h_1. Then, the partial derivative of $\tilde{v}_1\left(\tau_0, A_0\right)$ with respect to h_1 is

$$\frac{\partial \tilde{v}_1\left(\tau_0, A_0\right)}{\partial h_1} = u'\left(c_1\right)\left(w_1^f\left(1 - t_c - t_p\right) - w_1^i\right) + \beta^2 \mathbb{E}\left(u'\left(c_2\right) B'\left(\tau_1\right)\right) \tag{B.11}$$

where $\mathbb{E}\left(u'\left(c_2\right) B'\left(\tau_1\right)\right)$ is defined as

$$\begin{aligned}\mathbb{E}\left(u'\left(c_2\right) B'\left(\tau_1\right)\right) = &\left(1 - G\left(\bar{u}_2\left(\tau_1, A_1\right)\right)\right) u'\left(c_2^i\right) B'\left(\tau_1\right) + \\ &G\left(\bar{u}_2\left(\tau_1, A_1\right)\right) u'\left(c_2^f\right) B'\left(\tau_1 + 1\right).\end{aligned} \tag{B.12}$$

Due to the assumptions imposed on the utility function, the wage gap, and the pension plan, (B.11) is a non-decreasing function of h_1. Therefore $\bar{u}_1\left(\tau_0, A_0\right)$ is always non-negative.

Finally, I show two more properties of the threshold $\bar{u}_1\left(\tau_0, A_0\right)$: it is decreasing in the level of assets and its relationship with the formal-sector experience is ambiguous. First, the derivative of $\bar{u}_1\left(\tau_0, A_0\right)$ with respect to A_0 is

$$\frac{\partial \bar{u}_1\left(\tau_0, A_0\right)}{\partial A_0} = (1 + r)\left(u'\left(c_1^f\right) - u'\left(c_1^i\right)\right). \tag{B.13}$$

The concavity of the utility function guarantees that the right hand side of (B.13) is non-positive as long as $c_1^f \geq c_1^i$. Using equations (B.4) and (B.5), the partial derivative of c_1 with respect to h_1 around the optimal consumption plan is

$$\frac{\partial c_1}{\partial h_1} = \frac{\Psi_{A_1}\left(w_1^f\left(1 - t_c - t_p\right) - w_1^i\right) + \Psi_{\tau_1}}{u''\left(c_1\right) + \Psi_{A_1}}, \tag{B.14}$$

where Ψ_{τ_1} is the partial derivative of the expected marginal utility with respect to τ_1 defined as

$$\begin{aligned}\Psi_{\tau_1} = \frac{\beta}{1+\beta}&\mathbb{E}\left(u''\left(c_2\right)B'\left(\tau_1\right)\right) + \\ &\beta g\left(\bar{u}_2\left(\tau_1, A_1\right)\right)\left(u'\left(c_2^f\right) - u'\left(c_2^i\right)\right) \times \\ &\left(u'\left(c_2^f\right)B'\left(\tau_1+1\right) - u'\left(c_2^i\right)B'\left(\tau_1\right)\right),\end{aligned} \tag{B.15}$$

with $\mathbb{E}\left(u''\left(c_2\right)B'\left(\tau_1\right)\right)$ defined in a similar way than (B.12). As in the case of Ψ_{A_1}, I assume that the effect of τ_1 on $G\left(\bar{u}_2\left(\tau_1, A_1\right)\right)$ is small enough such that $\Psi_{\tau_1} \leq 0$. As a result, the optimal consumption plan is a non-decreasing function of h_1, and therefore the derivative defined in (B.13) is negative.

Second, the derivative of $\bar{u}_1\left(\tau_0, A_0\right)$ with respect to τ_0 is

$$\begin{aligned}\frac{\partial \bar{u}_1\left(\tau_0, A_0\right)}{\partial \tau_0} &= \beta^2\mathbb{E}\left(u'\left(c_1\left(\tau_0+1, A_1^f\right)\right)B'\left(\tau_0+1\right)\right) \\ &\quad - \beta^2\mathbb{E}\left(u'\left(c_1\left(\tau_0, A_1^i\right)\right)B'\left(\tau_0\right)\right).\end{aligned} \tag{B.16}$$

In this case, the sign of (B.16) is ambiguous, as it depends on the curvature of the utility function and the concavity or convexity of the pension plan.

Appendix C

Appendix to Chapter 3

C.1 Robustness test

In this Appendix, I show that the estimation results presented in Table 3.4 are robust to alternative specifications. In particular, I show that the RD results are robust to the specification of control functions, estimators and bandwidth selections (Table C.1). Similarly, Table C.2 shows that the DD estimation results are robust to the inclusion of time trends by age group to control for differential entry rates of younger workers to the formal sector.

Table C.1: Regression Discontinuity robustness test, 2011-12

	Employment			Wage			Fraction at min. wage		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
A: Least Squares estimator (estimates scaled up by 100)									
Linear control function	1.25	1.24	1.55	0.34	0.26	0.36	-0.04	-0.03	-0.11
(Bandwidth 24 months)	[0.55]**	[0.56]**	[0.60]**	[0.29]	[0.36]	[0.31]	[0.16]	[0.16]	[0.17]
Quadratic control function	1.14	1.13	1.42	0.02	-0.04	0.17	-0.13	-0.10	-0.26
(Bandwidth 24 months)	[0.57]*	[0.57]*	[0.56]**	[0.36]	[0.28]	[0.29]	[0.17]	[0.18]	[0.16]
B: Local linear estimator (estimates scaled up by 100)									
Local linear	1.10	–	–	-0.18	–	–	0.07	–	–
(Bandwidth 12 months)	[0.35]**			[0.25]			[0.04]		
Local linear	1.21	–	–	0.22	–	–	-0.08	–	–
(Bandwidth 24 months)	[0.43]**			[0.30]			[0.11]		
Observations	549,171	549,171	549,171	539,220	539,220	539,220	539,220	539,220	539,220
Fixed effects		Quarter	Region		Quarter	Region		Quarter	Region

Notes: Each cell reports an estimate (escalated by 100) based on a separate regression of employment or wage as dependent variable. As independent variables, I include a polynomial on age (in quarters) and its interaction with a treatment indicator of whether the worker is under the age of 28 (See Section 3.4 for details). Columns (1), (4) and (7) test the robustness of the results to the specification of the control function and estimators, while other columns include fixed effects in the regressions to test the sensitivity of the results. The fixed effects are based on categories of quarters of entry and job location. Standard errors clustered by age (in quarters) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Table C.2: Estimation results, DD including time trends by age group

	Employment	Wages	Min. Wage
Lower payroll taxes	3.36	-0.17	0.04
	[0.88]***	[0.19]	[0.24]
Number of cells	560	560	560
Observations	790,539	775,460	775,460

Notes: This Table investigates whether the Differences-in-Differences (DD) estimation results are robust to the introduction of differential trends by age group. Each cell presents DD estimates for the employment and wage effects (escalated by 100) for the sample of workers entering into formal-sector firms between February 2010 and December 2012 (See Section 3.4 for details). In each regression, I estimate the DD regression including differential time trends by age group. "Min. Wage" refers to the average effect on the fraction of workers earning the minimum wage.* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.