Pension Incentives and Formal-Sector Labor Supply: Evidence from Colombia

Oscar Becerra*†

October 31, 2015

Abstract

I analyze how future pension benefits affect the labor supply in economies with an informal sector. For workers, a formal-sector job provides long-run gains, as it increases their likelihood of getting pension benefits in the future. If workers take into account those gains when they search for formal-sector jobs, the pension system affects the formal-sector labor supply. I estimate the causal effect of pension-related incentives on formal-sector labor supply, using a cohort-based reform in Colombia. I show that a change in future pension benefits generates a large shift in the labor supply between the formal and informal sector, and does not affect the labor force participation. Consistent with the predictions of a model with a pension system and informal job opportunities, the average effect of pension incentives on formal-sector labor supply is heterogeneous. The effect concentrates among workers for whom the minimum qualifying conditions are binding and workers with a higher expected pension wealth. The results suggest that pension reforms might create large efficiency costs, which should be taken into account when designing pension programs.

*University of British Columbia. obecerra@alumni.ubc.ca

†I would like to thank Joshua Gottlieb, Thomas Lemieux, Florian Hoffman, Kevin Milligan, Nicole Fortin, David Green, participants at the CEA conference, workshop participants at UBC and the UBC Public Finance Reading Group for helpful comments and valuable discussion. I also 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 the view of the author and does not necessarily reflect those of DANE or Colpensiones.

1 Introduction

How do future pension benefits determine workers' choices in developing economies? Workers in developing economies can respond to public policies by changing their search for jobs in the formal or the informal sector. If workers respond to changes in their expected pension benefits by changing their efforts of searching for formal-sector jobs, reforms to the pension system affect the formal-sector labor supply and create efficiency costs.

The analysis of the behavioral responses along the formal-informal margin is important for the design of retirement policies in Latin American countries. In Latin America, approximately 50 percent of the workers do not contribute to the pension system. Since most of them have low incomes, the lack of pension contributions exacerbates the income inequality after retirement (Frölich et al., 2014). In response to the concerns about low coverage and high inequality in benefits, Latin American policymakers have implemented major pension reforms. The reforms have included changes in funding sources, changes in qualifying conditions for receiving a pension, and the introduction of pension assistance programs.[1] One cost of such reforms, though, is that they may reduce the expected gains from retirement contributions, reducing the incentive to search for formal-sector jobs.

Despite its policy importance, 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 the worker's expected pension benefits also influence the worker's current labor choices. Therefore, there is little useful variation to identify the causal effect of pension incentives on formal-sector labor supply (Liebman et al., 2009).

In this paper, I estimate the causal effect of pension incentives on formal-sector labor supply. I overcome the identification problems by using the changes in the structure of the Colombian pension system. In 1993, the Colombian government increased the pension

[1] Ten countries in the region implemented large 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-contributive pension programs is in Bosch et al. (2013).

contribution rate and the qualifying conditions for receiving a pension in the defined-benefit system. However, the reform kept the pre-reform qualifying conditions for eligible men born before April 1954 and women born before April 1959. Compared with younger workers, eligible workers 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). Thus, the reform caused a permanent change in the long-run value of a formal-sector job depending on the worker's birth date.

The difference in qualifying conditions by birth date provides a source of exogenous variation to estimate the causal effect of pension incentives on formal-sector labor supply. To estimate this effect, I use a regression discontinuity design 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 no other economic or institutional factor can explain a change in the outcome at the cutoff, the difference is an estimate of the causal effect of pension incentives on formal-sector labor supply.

To understand the impact of the changes in the 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 search behavior proposed by Chetty (2006) for unemployment insurance and adapted by Gerard and Gonzaga (2013) to include informal labor markets. I modify the model to incorporate a defined-benefit pension system, in which 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, a formal-sector job provides long-run gains because it increases the worker's likelihood of getting pension benefits in the future. Within an age group, the long-run gains from searching for a formal-sector job are a nonlinear function of the years of contribution. The higher gains concentrate among workers who are close to the vesting period, as an additional period in the formal sector increases their probability of securing a pension in the future.

The comparative statics of the model shows that the effect of an increase in the minimum qualifying conditions on labor supply are heterogeneous, and its sign is ambiguous. The direction of the response depends on the years of contributions relative to the new vesting period. On the one hand, workers who are a long way from reaching the new vesting period reduce their formal-sector labor supply, as it is unlikely that they can reach it. On the other hand, workers who are close to reach the new vesting period increase their formal-sector labor supply to secure their pension benefits. The magnitude of the effect also depends on the worker's age and his opportunities of finding a formal-sector job.

I present three main empirical findings. First, there is a sizable and significant response of the formal-sector labor supply to changes in the pension incentives. The effect concentrates among men. For 2005, the average effect of easier qualifying conditions on salaried-formal labor supply for men is 16 percent. For 2011, the average effect is negative 7 percent. This change is consistent with the model insights. Compared to younger workers, eligible workers have less incentive to contribute because they already met the vesting period. In addition, I find little evidence that the increase in the formal-sector labor supply is offset by a reduction in wages.

Second, the increase 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 the eligibility for easier qualifying conditions on self-employment (mostly informal) is negative and of the similar magnitude to the effect for salaried-formal employment. The estimated effect is not significant for labor force participation. The results are similar to those at Almeida and Carneiro (2012) who find that higher mandated benefits with no wage adjustment generate incentives for self-employed workers to switch to salaried-formal jobs.

Third, the response of the formal-sector labor supply to pension incentives is heterogeneous, and depends on labor market opportunities. I analyze the response for groups with different incentives to work in the formal sector (e.g. education). For the education analysis, workers with primary and post-secondary education are less responsive to pension incentives than workers with secondary education. The result is consistent with the model predictions. I obtain similar conclusions for other subsamples such as regions and married workers. Using the variation by region and education and assumptions about earnings and probability of contributions, I estimate an elasticity of the formal-sector labor supply with respect to the net pension contribution rate of 1.8. This estimate is likely a lower bound of the actual elasticity, suggesting large behavioral responses.

The paper is organized as follows: Section 2 describes the Colombian pension system and the labor market institutions. Section 3 presents the conceptual framework that provides insights about the expected sign and sources of heterogeneity in the results. Section 4 discusses the identification strategy and the data sources. Section 5 reports the estimation results. Finally, Section 6 concludes.

2 Institutional background

2.1 The Colombian pension system

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

For young workers and new entrants, the GPS has two pension systems. All public and private workers must contribute to one system, and their choice determines their pension eligibility and benefits.[2] The first system is the social insurance system, in which workers 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

[2]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).

of the worker's wage and the minimum wage, while the eligibility is based on the worker's age and the years of contribution. The second system is the individual account system, in which workers 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 financial 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. The eligibility for the guaranteed minimum pension is based on the worker's age and length of contributions.

For all other workers, the GPS is the transition system, a defined-benefit system managed by Colpensiones. The transition system keeps the pre-reform eligibility and benefits for eligible workers. The eligibility criteria for the transition system were based on age and contributions by the time the reform took effect (April 1, 1994). Originally, three groups of workers were eligible for the transition system: men born before April 1954 and women born before April 1959 who had contributed to the pension system, and younger workers who had contributed to the system by at least 750 weeks. A reform in 2005 required that eligible workers had contributed more than 750 weeks by July 2005 and met the qualifying conditions by 2014.[3]

Table 1 and Figure 1 summarize the main characteristics of the GPS. Workers in the three systems face the same contribution rate (16 percent) 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 earlier age (55 years for women and 60 years for men instead of 57 and 62). Moreover, the transition system has a higher nominal replacement rate than the social insurance system (Figure 1). However, this

[3]In the rest of the paper, I focus on the eligibility criteria based on the birth date of the worker. The criterion based on 750 weeks by 1993 has limited effect on younger workers, as the requirement implies that men younger than 40 and women younger than 35 would have worked by at least 15 years in the formal sector. Moreover, after the 2005 reform, men born after 1954 and women born after 1959 became not eligible for the transition system, even though they could have met the original eligibility criteria.

difference is less important because the minimum pension guarantee implies that low wage workers face the same effective replacement rate. This is a relevant feature of the system, as 90 percent of workers in the GPS report earnings between one and two minimum wages (Table 1).

Relevance of the minimum qualifying conditions The importance of differences in the minimum qualifying conditions for a pension depends on whether the workers take into account these conditions to make their retirement decisions. Figure 2 shows that most workers claim their pension benefits as soon they meet the requirements. Based on statistics from Colpensiones, the figure displays the distribution by age, weeks of contribution and gender for non retired population who have 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 shows a clear discontinuity around the minimum retirement age only 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. Poor population is eligible for coverage of the public non-contributory healthcare system, a subsidy for old-age population living in extreme poverty (about 20 percent of the minimum wage), and conditional cash transfers for families with children in schooling age. Except for the eligibility for the public healthcare system, the eligibility for other social programs does not depend on whether the worker has a formal-sector job. More importantly, no program depends on the eligibility for the transition system. As a result, the interactions with the assistance programs do not offset the cohort differences induced by the eligibility for the transition system.

2.2 Labor market institutions

The Colombian government mandates that all employers provide benefits to their employees, and that self-employed workers pay the contributions to the pension and contributive healthcare systems. Thus, workers covered by the 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 insurance through the pension and the contributive family healthcare systems, provide paid vacations (two weeks per year), severance payments (one extra 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 benefits of paying their contributions are related to the insurance provided by the pension and the contributive family healthcare systems.[4]

Second, formal-sector workers earn higher wages. As La Porta and Shleifer (2014) show, formal firms tend to be more productive and pay higher wages than informal firms, and formal self-employed tend to be more educated. Using data from the household surveys (described in the Section 4.1), Table 2 presents the average wage and distribution of urban workers aged 20 to 65 that work at least 30 hours per week. I define a formal worker as a worker who is contributing to the pension system and is covered by the contributive healthcare system.[5] The average wage gap is 75 percent for salaried workers and 100 percent

[4] The minimum contribution for the pension and the contributive healthcare systems are 16 and 12 percent of the minimum wage.

[5] The coverage of pension and contributive healthcare systems is a widely used measure of formal employment (Perry et al., 2007). Because workers are not subject to penalties for being uncovered, it is unlikely they misreport their coverage status. Moreover, the formal employment indicators are consistent with aggregate

for self-employed workers. The wage gap is positive regardless of the level of education and the type of employment. Nonetheless, as Figure 3 points out, the wage for a large fraction of informal-sector workers is above the minimum wage.

Despite the gains from working in the formal sector, other supply and demand factors prevent workers from searching for formal-sector jobs. On the supply side, workers may find it optimal to work in the informal sector (Maloney, 2004). Low valuation or substitutes for 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 demand side, high regulation costs coupled with weak enforcement cause that firms find it optimal to operate informally, reducing the worker's opportunities of finding a formal-sector job. The effect of regulation on the labor market is evident for the minimum wage, as it is binding for a large fraction of formal-sector workers (Figure 3).

3 Pension incentives and formal-sector labor supply

To understand the incentives that workers face when making their labor supply decisions, I present a model that characterizes the workers' decisions about retirement and formal-sector participation. In the 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.

Assume a representative worker who lives during $T - a_0$ periods, indexed by $a = a_0, \ldots, T$. Every period, the worker chooses whether he retires leaving the labor market for good. If he retires and is eligible for retirement benefits, he receives a fraction b of the wage in the formal sector w^f and other benefits valued θ^r (e.g. healthcare). If he retires and is not eligible for retirement benefits, he gets zero income. Thus, conditional on retirement, the worker's earnings at age a are $e_a\left(\tau_{a-1}\right) bw^f$, where $e_a\left(\tau_{a-1}\right)$ is an indicator variable of the worker's eligibility for pension and 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 has to

statistics obtained from administrative data.

work for at least τ^* periods in the formal sector and be at least R periods old, and thus $e_a(\tau_{a-1}) = 1\{\tau_{a-1} \geq \tau^*\} 1\{a \geq R\}$.

If the worker does not retire, he draws a random cost of searching for a formal-sector job $\psi_a \overset{i.i.d}{\sim} G(\psi)$, where $G(\cdot)$ is defined over the range $[0, \infty)$. ψ_a is measured in utility units and it is used as a catch-all variable summarizing the searching costs and preferences for formal-sector jobs. After drawing ψ_a, he decides between searching for a job in the formal sector and working in the informal sector. If the worker works in the formal sector, he receives a wage w^f and mandated benefits valued θ^f, incurs in the search cost ψ_a, and pays a pension tax rate t^{nom}. If he decides to work in the informal sector, he receives a wage w^i (I assume that $w^i \leq w^f(1 - t^{nom})$ and $\theta^r \geq \theta^f$). In addition, working in the formal sector increases the worker's periods of contribution to the pension system by one period. As a result, $\tau_a = \tau_{a-1} + d_a$, where d_a is an indicator variable of whether the worker searches for a formal-sector job. At the end of the period, the worker loses his job with probability one.[6]

I assume that workers do not save, and so the worker's consumption per period is equal to his income. Let r_a denote an indicator variable of whether the worker retires at the beginning of period a. Given τ_{a-1}, the problem for the worker is

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

where

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

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

[6] The model implications are robust to a separation rate lower than one.

The model encompasses the two common views about the determinants of the decision to work in the informal sector (Gerard and Gonzaga, 2013). First, workers may not search for formal-sector jobs because the perceived gains from searching 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 not search for formal-sector jobs 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 heavier tails. When $G(\psi)$ has heavy tails, it is likely to draw a realization of ψ high enough to offset the gains from searching. It is common that these two forces interact and reinforce each other. For example, workers with narrower wage gaps may also face higher search costs, reducing even further the likelihood that a worker searches for a formal-sector job.

3.1 Retirement and formal-sector participation decisions

The optimal 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. First, the worker finds the optimal plan for searching for a formal-sector job and the value function conditional on labor force participation, $v_a^w(\tau_{a-1})$. Second, the worker compares the value function from working with the value function from retiring, and determines the optimal retirement decision rule.

First, 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 searching are greater than the costs. Thus, the worker searches for a formal-sector job (sets $d_a = 1$) if

$$\bar{u}_a(\tau_{a-1}) = \tilde{u} + \beta \Delta v_{a+1}(\tau_{a-1} + 1) \geq \psi_a. \tag{2}$$

where $\tilde{u} = u\left(w^f\left(1 - t^{nom}\right)\right) + \theta^f - u\left(w^i\right)$ and $\Delta v_{a+1}\left(\tau_{a-1}+1\right) = v_{a+1}\left(\tau_{a-1}+1\right) - v_{a+1}\left(\tau_{a-1}\right)$. In equation (2), $\bar{u}_a\left(\tau_{a-1}\right)$ summarizes the gains from searching for a formal-sector job. The first term represents the short-run gains from searching, the utility gains determined by the differences in wages in both sectors plus the net valuation of the mandated benefits. The second term represents the long-term gains from searching, which accounts for the effect that one additional period of working in the formal sector has on the likelihood that the worker gets pension benefits.

From equation (2), the ex ante probability that a worker works in the formal sector is $P\left(d_a = 1 \mid \tau_{a-1}\right) = G\left(\bar{u}_a\left(\tau_{a-1}\right)\right)$, and the value function is

$$v_a^w\left(\tau_{a-1}\right) = u\left(w^i\right) + v_{a+1}\left(\tau_{a-1}\right) + G\left(\bar{u}_a\left(\tau_{a-1}\right)\right)\mathbb{E}\left(\bar{u}_a\left(\tau_{a-1}\right) - \psi_a \middle|\, \psi_a \leq \bar{u}_a\left(\tau_{a-1}\right)\right). \quad (3)$$

Second, the worker retires if the value function for retirement is greater than the value function conditional on labor force participation (equation (3)). Thus, the worker retires (sets $r_a = 1$) if

$$v_a^r\left(\tau_{a-1}\right) \geq v_a^w\left(\tau_{a-1}\right). \quad (4)$$

3.2 Model implications

The model has four useful implications to understand the empirical results of the paper. The first three implications are discussed in detail in the Appendix.

First, when the replacement rate is equal to one, the worker retires as soon as he meets the qualifying conditions. The retiree receives as pension benefit the wage in the formal sector 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 2.1, where workers retire as soon as they meet the age and weeks requirements. Values of b lower than one may delay the retirement decision, depending on the value function conditional on working.[7]

Second, the long-run gains from searching for a formal-sector job are heterogeneous and depend on the worker's employment history (τ_{a-1}). The intensity of the worker's search for formal-sector jobs depends on the likelihood of getting retirement benefits. Workers who cannot accumulate enough years to meet the vesting period have no chance of receiving pension benefits, and so $\Delta v_{a+1}(\tau_{a-1}+1)=0$. Workers who meet the vesting period do not have any extra long-run gains from working an extra period in the formal-sector, and so $\Delta v_{a+1}(\tau_{a-1}+1)=0$. For the rest of workers, the long-run gains from working in the formal sector are positive. Because the probability of searching for formal-sector jobs is increasing in $\Delta v_{a+1}(\tau_{a-1}+1)$, the result implies that workers with positive long-run gains from searching search more actively for formal-sector jobs. Nonetheless, workers with no long-run gains from searching still search for formal-sector jobs, but their search is motivated by short-run gains only.

Third, a change in the minimum retirement age R or the vesting period τ^* affects the search for formal-sector jobs of younger workers. The effect of an increase in R on the formal-sector labor supply is negative, as it reduces the long-run gains from searching for formal-sector jobs. In contrast, the effect of an increase in τ^* on the formal-sector labor supply is ambiguous. On the one hand, workers who are close to meet the vesting period increase their search efforts to meet the new vesting period. On the other hand, workers with a low number of years of contribution reduce their search efforts because it is unlikely that they meet the new requirements.

Fourth, 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 heavy tails) and labor markets with high formality rates (a high value of $\tilde{u}$ and a search cost

[7] An alternative explanation for this fact 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. I also consider a version of the model in which the minimum retirement age is exogenous, and the implications of the model hold.

distribution with light tails). When $P(d_a = 1 \mid \tau_{a-1}) \to 0$, workers cannot reach the vesting period and the long-run gains from searching are zero. When $P(d_a = 1 \mid \tau_{a-1}) \to 1$, workers always reach the vesting period and the long-run gains from searching formal-sector jobs are zero. The effect of changes in the qualifying conditions concentrates among workers who struggle for reaching the vesting period, but they still could reach it.

Figure 4 shows the probability of searching for a formal-sector job by years of contribution for two simulated cohorts observed at age $a = 50$. Both cohorts face the same labor market opportunities, but different defined-benefit pension plans. One plan assumes $R = 60$, $\tau^* = 20$ and $b = 1$ and the other assumes $R = 62$, $\tau^* = 25$ and $b = 1$.[8] Since $b = 1$ in both systems, workers retire as soon as they meet the requirements. For both cohorts, the probability of searching for a formal-sector job is higher on years right below the vesting period, as the long-run gains from searching are concentrated on those years. The increase in the vesting period shifts the long-run gains from searching to the right, while the increase in the minimum retirement age reduces them. The change in the long-run gains from searching shifts the probability of searching for formal-sector jobs.

The difference in the probability of searching given the change in the minimum qualifying conditions is presented in the bottom panel of Figure 4. A negative value of the difference implies that workers under the easier qualifying conditions search more actively for formal-sector jobs, and vice versa. The increase in the minimum qualifying conditions has two types of effects on formal employment. On one hand, harder qualifying conditions discourage workers with a low number of years of contributions, because it is more difficult for them to reach the vesting period. On the other hand, harder qualifying conditions encourages workers close to the new vesting period to search for formal-sector jobs, because they have to contribute additional periods to reach the new vesting period. Thus, the effect of changes in minimum qualifying conditions on formal-sector labor supply is heterogeneous. The sign of the effect is ambiguous and depends on the worker's employment history.

[8] In the simulation, I assume that the worker works from $a_0 = 20$ up to $T = 75$ years and his 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 $\beta = \frac{1}{1.05}$.

3.3 Labor demand

The labor supply model presented above shows that pension incentives affect the workers' search for formal-sector jobs. In general equilibrium, though, changes in the long-run gains from searching should be offset by changes in wages. Workers would be willing to give up part of their wage to receive the extra expected pension benefits, and vice versa (Summers, 1989).

Therefore, in equilibrium, the wage gap should exhibit an inverse pattern to the observed in Figure 4. The changes in qualifying conditions would be reflected on wages, leaving formal employment unchanged. However, this result requires that the extra benefits can be passed to workers through lower wages, which in turn is determined by the wage setting process (Saez et al., 2012). Institutional factors such as minimum wage laws, search processes based on posted earnings, unobservable employment history, and pay fairness norms may prevent firms from setting differential wages among workers. If firms cannot set different wages among similar workers, the general equilibrium effect of the extra search for a formal-sector job narrows the wage gap for all workers. In such equilibrium with spillover effects, the comparative statics result exhibits the same patterns presented above.

4 Data and Empirical Approach

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 long-form questionnaire of the Colombian Census of 2005, a cross-sectional dataset including information of labor market outcomes, pension and healthcare coverage, and demographic and household characteristics. The second source is the PILA dataset of 2011, an administrative dataset that collects the information of all workers and earnings in the formal sector.

Despite of the compromises that the combination of two datasets may imply, I include both for the estimation results. I can analyze factors with the Census dataset that I cannot analyze with the PILA dataset, and vice versa. In particular, the Census dataset does not include information about worker's earnings, while the PILA dataset does not include information about informal employment or demographic characteristics. Using both datasets I can study the response of formal employment to changes in the pension incentives as the workers age.

A limitation of the Census and PILA datasets is that none of them includes the worker's employment history. I complement the information of the Census and PILA datasets by adding statistics of the years of contributions from the Colombian household surveys.

4.1.1 Colombian Census (2005)

The long-form questionnaire of the Colombian Census of 2005 contains information from 2 million households and 9.7 million people, approximately 20 percent of the households. The dataset includes birth date (in months), demographic information, type of employment, contributions to the pension system and coverage of the healthcare system. The information about birth date is highly reliable, as the interviews were carried out in-person and the interviewer verified the birth date from the respondent's identification card. If the person did not have identification card, the birth date was provided by the respondent or inferred based on the reported age. In the long-form questionnaire dataset, the birth date of 92 percent of the urban population was based on the identification card.

The sample used in the paper is based on people living in urban areas, with known birth date and born up to four years before or after the date of eligibility for the transition system (March 1950 to March 1958 for men and March 1955 to March 1963 for women). The final samples sizes are 128,531 for men and 178,333 for women.

4.1.2 PILA dataset (2011)

The PILA dataset is a new dataset that collects information of the system used by firms and independent workers to pay for all the mandated benefits. Since formal workers must be covered by the mandated benefits, the dataset collects the information of all formal workers. The dataset includes identifiers for employer and employee, basic wage, job location, gender and birth date (in days) of the employee. The birth date and gender are added by the Ministry of Health, based on the employee ID number. The dataset also includes information on whether the firm is public or privately owned, (scrambled) economic sector, and type of worker (independent or employee). Counting all types of employers and employees, the dataset covers about eight million employer-employees pairs per month.

Out of the entire dataset, I keep information of 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 whenever an employer-employee match is missing and the dataset records the same match up to three months before and after. In addition, I drop employees appearing only once.

The sample used in the paper is based on all workers 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).

4.1.3 Household Surveys (2006-2011)

The Colombian Household survey is the official source of the employment statistics in Colombia. After a large methodological change in 2006, the dataset includes information about the person's birth date (in months) and coverage of the pension and contributive healthcare system. The surveys also contain information about the worker's earnings and the number of years of contributions (conditional on contributing). The main limitation of the household surveys for this study is their small sample size for the cohort of interest. The number of observations by month of birth in a year is about 200 people, and just about 40 people report information about years of contributions.

The sample used in the paper is based on all 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.

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 estimate regressions of the effect of easier qualifying conditions on labor market outcomes. It provides evidence of the response of the formal-sector employment to the pension incentives, without making further assumptions about the earnings process and expectations. In the second stage, I use additional assumptions to recover an estimate of the incentive effect of pensions on formal-sector labor supply.

4.2.1 Effect of Easier Qualifying Conditions on Labor Market Outcomes

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

$$Y_i = \beta + \rho 1_{\{\tilde{a}_i \geq 0\}} + \delta_1 \left(\tilde{a}_i\right) 1_{\{\tilde{a}_i < 0\}} + \delta_2 \left(\tilde{a}_i\right) 1_{\{\tilde{a}_i \geq 0\}} + \varepsilon_i \tag{5}$$

where Y_i is an indicator of the formal-sector labor supply, $\tilde{a}_i$ is the normalized age of the person (so $\tilde{a}_i = 0$ corresponds to the age at the eligibility threshold), and $\delta_1 \left(\tilde{a}_i\right)$ and $\delta_2 \left(\tilde{a}_i\right)$ are control functions, such that $\delta_1 (0) = \delta_2 (0) = 0$. Based on the reported birth date, the relevant cutoff for eligibility for the transition system is March 1954 for men and March 1959 for women. Workers born before those dates were eligible for retirement benefits with 1,000 weeks of contributions and at age 55 (women) and 60 (men), while workers born afterwards have to retire with up to 300 more weeks of contributions and two years older.

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 is interpreted as the average effect of the eligibility for easier qualifying conditions on the formal-sector labor supply, defined as

$$\rho = \lim_{c \downarrow 0} \mathbb{E}\left(Y_i \middle| \tilde{a}_i = c\right) - \lim_{c \uparrow 0} \mathbb{E}\left(Y_i \middle| \tilde{a}_i = c\right). \tag{6}$$

(Imbens and Lemieux, 2008). However, as discussed in Section 3, the incentive to search for formal-sector jobs depends on the worker's age (a) and years of contribution (τ_{a-1}). Therefore, ρ corresponds to the weighted average effect by years of contribution, i.e.

$$\rho = \int_{\tau_{a-1}} \left(\lim_{c \downarrow 0} \mathbb{E}\left(Y_i \middle| \tilde{a}_i = c, \tau'\right) - \lim_{c \uparrow 0} \mathbb{E}\left(Y_i \middle| \tilde{a}_i = c, \tau'\right) \right) dF_a\left(\tau'\right), \tag{7}$$

where $F_a(\tau)$ 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 expected sign of ρ is ambiguous. As discussed in Section 3, the expected response for workers who are a long way from reaching the new vesting period is positive, while it is negative for workers near the new vesting period. The sign of the average effect depends on the specific distribution of the number of years of contribution.

Although the distribution of the years of contribution is not observed in the data, the analysis in Section 3 provides useful insights about the expected sign and magnitude of ρ. First, ρ should decrease with the worker's age, as the distribution of τ moves to the right as workers age, weighting more the negative part of the effect. Second, ρ should be larger (in absolute value) for groups of workers for whom $F_a(\tau)$ is located on middle-range values of the years of contribution. For workers with a low probability of finding a formal-sector job, the estimated average effect should be small, as they have low long-run gains from searching and a right-skewed distribution of years of contribution. A similar argument explains 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 the measure.

To estimate equation (5), I run regressions separately by gender, as the cohorts affected by the reform are different. I cluster the standard errors by age in months to account for the 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 the control function and bandwidth.

4.2.2 Labor Supply Elasticity with Respect to the Effective Pension Tax Rate

Following the convention in public economics literature, I measure the incentive effects of pension benefits on the labor supply by using the elasticity of labor supply with respect to the net effective pension tax rate (Liebman et al., 2009). I estimate the elasticity with respect to the effective pension rate along the formal-informal margin, defined as

$$\sigma = \frac{d \ln L_a^f}{d \ln \left(1 - t_a^{eff}\right)}. \tag{8}$$

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,

$$t_a^{eff} = 0.04 - \frac{\beta \Delta \mathbb{E} PW_{a+1}(\tau+1)}{w}.$$

In the definition of t_a^{eff}, $\mathbb{E}PW_a(\tau)$ stands for the expected pension wealth at age a and τ years of contribution, $\Delta \mathbb{E}PW_{a+1}(\tau+1) = \mathbb{E}PW_{a+1}(\tau+1) - \mathbb{E}PW_{a+1}(\tau)$, and w is the worker's wage. A detailed discussion of the procedure used to compute the effective tax rate is presented in Section 5.2.1.

The effective pension tax rate along the formal-informal margin measures the net gains from working the current period in the formal sector. It is the difference between the nominal 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.[9] Based on the results of Section 3, the 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 obtain an estimate of σ, I split the samples in groups characterized by different propensities to work in the formal sector (e.g. education and region). For each group (denoted by X), I compute both the estimate of the average response of the formal-sector employment $\left(\Delta \ln L_{aX}^{f}\right)$ and the average change at the discontinuity of the effective pension rate $\left(\Delta \ln \left(1 - t_{aX}^{eff}\right)\right)$. Then, I estimate σ by running the regression

$$\Delta \ln L_{aX}^{f} = \alpha_0 + \sigma \Delta \ln \left(1 - t_{aX}^{eff}\right) + \varepsilon_X. \tag{9}$$

In equation (9), 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.

5 Estimation results

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 violated for at least two reasons. First, workers who were likely to work in the formal sector could manipulate their birth date to become eligible for the program when they are not (McCrary, 2008). Second, the estimated effect of the policy could be confounded by

[9] The nominal 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 nominal tax rate.

changes in other covariates that effectively determines the outcome (Imbens and Lemieux, 2008). In this section, I assess these two potential threats to the identification.

I test the manipulation hypothesis by estimating the density of the total population by age group above and below the eligibility thresholds, and implementing the test statistic proposed by McCrary (2008).[10] The results are presented in the top panel of Table 3. The manipulation hypothesis implies that the estimated difference should be positive, as younger workers change their documentation to become eligible for the transition system. However, the sign of the effect is the contrary to the expected in the manipulation hypothesis for men and women, and it is significant for women. The results for women raise concerns about other factors that may be driving their results.

To test further the potential 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 5. The t-statistics exhibit two-year cyclical patterns over time, where the large (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 of men. The result suggests that other time trends different than the eligibility for the transition system are driving the changes in the density for women.

In addition to the manipulation tests, I look for discontinuities in other observable variables that can explain the worker's labor supply choice. The variables considered are indicators for whether the person has a High School diploma or less, whether the person reports any disability, and whether the person identifies himself as a member of an ethnic group (black or indigenous). These variables are correlated with the likelihood that a person has a formal-sector job, but they are predetermined by the time the policy change took place. Thus, significant differences in other 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 3 presents the estimation results, and they show no significant

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

differences for all three indicators for both men and women.

Taken together, the results in Table 3 and Figure 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 age for women is not continuous around the eligibility threshold, the placebo test suggests that the change is caused by other time trends different than changes in the pension eligibility. Nonetheless, the interpretation of the results for women must take into account this caveat.

5.2 Results

The regression discontinuity estimates for 2005 and 2011 are presented in Table 4 and in Figure 6.

The top panel of Table 4 presents regression discontinuity estimates of the effect of easier qualifying conditions on salaried-formal employment for 2005. I use as dependent variable an indicator of whether the person works as salaried-formal worker.[11] Thus, the estimated effect is the average effect of easier qualifying conditions on salaried-formal employment rate. This specification is my preferred specification because it is more robust to other changes in population (unrelated to worker's self-selection), which is a particular concern for women. The middle and bottom panel of Table 4 show the regression discontinuity estimates using as dependent variable the log of the number of salaried-formal workers for 2005 and 2011. In all regressions, I use a quadratic polynomial in age as a control function to account for potential non linearities in the formal-employment rate, and use a bandwidth 48 months for 2005 and 730 days for 2011.[12]

The results in Table 4 show that Colombian workers actively responded to changes in the pension incentives. For men, the estimated effects are significant and change over time. In 2005, the average effect of being eligible for easier qualifying conditions increased the

[11] 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 contributive healthcare system.

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

salaried-formal employment age by 3.1 percentage points (relative to a basis of 18 percentage points). The effect is confirmed by an specification using as dependent variable the number of salaried-formal workers for 2005 (panel B of Table 4). The advantage of this specification is that it can also be implemented with the 2011 data. The regression discontinuity estimates show that the increase in the number of salaried-formal workers at the discontinuity is 15.8 percent. In 2011, the estimated average effect on the salaried-formal employment for men is negative and significant, implying a reduction 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).

The results for men are consistent with the framework presented above, in which the average effect of easier qualifying conditions depends on the distribution of the number of years of contribution. When men were 51, nine years away from the minimum retirement age, the population eligible for the transition system were more likely to work as a salaried-formal worker. The result reverses when workers were 57. Table 4 also includes 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.[13] The distribution of years of contribution shifts to the right from 2006 to 2011. Thus, the average effect for 2011 should be less than the average effect for 2005, as there are more eligible workers with negative incentives to search for formal-sector jobs.

The results for women are intriguing. The 2005 estimates do not show any sizable or significant response. For 2011, though, Table A.2 shows significant results, depending on the specification. Since the 2011 results do not account for changes in the population, it is not possible to disentangle the potential effect of the changes in the policy and the documented changes in the total population around the discontinuity. An explanation for the women's non-response is that the transition system required that workers had contributed to the pension system by 1994. This condition immediately limits the applicability of the reform

[13]The distribution is conditional on making contributions.

for women because of their relatively low labor force participation (62 percent from 1984 to 1993).[14]

Labor Demand. To address the labor demand response to changes in pension incentives, I estimate the effect of the eligibility for easier qualifying conditions on the wages of formal-sector workers for 2011. If firms are able to set different wages between workers, wages offset part of the long-run gains from searching for formal-sector jobs. Therefore, the expected sign of the average effect eligibility for easier qualifying condition 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 5. The average effect of easier qualifying conditions on formal-sector wages is about 3 percent for men and is not significant for women (columns (1) and (5)).

To understand the sources of the aggregate results, I estimate the average effect of easier qualifying conditions on formal-sector wages and employment by wage range. The results are presented in Table 5. Consistent with the analytical framework, low-wage men are the most responsive to changes on 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 extra long-run gains from searching for formal-sector jobs once they met the requirements. The response for women is not significant.

The top panel of Table 5 presents the average effect of easier qualifying conditions on formal-sector wages by wage range. For men, the estimated effects are small and not significant. The difference in the aggregate results is driven by a composition effect, as the percentage of workers earning the minimum wage is larger for younger workers (panel B of Table 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 impact of the policy

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

change on wages was limited.

Nonetheless, the results on wages presented in Table 5 do not imply that changes in pension benefits cannot be 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 the pension incentives on the composition of the labor force. Based on information from 2005, I run versions of equation (5) using as dependent variable indicators for labor force participation, self-employment, and salaried-informal employment.

The estimation results are presented in Table 6, while the graphical evidence for men is presented in Figure 6. The increase in the salaried-formal employment rate for men is associated with reductions in informal-sector jobs, in particular self-employed jobs. For men, the response of self-employment rate is of the same magnitude but opposite sign than the response for salaried-formal employment. In contrast, neither salaried-informal employment nor labor force participation shows a significant response. For women, there is no significant response either in labor force participation or formal employment. The reallocation result has been observed in previous literature of mandated benefits and formal labor supply (Almeida and Carneiro, 2012).

Heterogeneity and Additional Robustness. In this section, I analyze the differential effect of easier qualifying conditions on the 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 about the mechanisms driving the aggregate results.

I estimate the response of the formal-sector labor supply to changes in pension incentives on three dimensions: educational attainment, household composition (e.g. presence of a spouse in the household), and region. I present the results for men, as the results for women are not significant and may be affected by changes in the distribution by age. Because of data limitations, I present the results for educational attainment and household characteristics for 2005 and the regional results for 2005 and 2011.

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

The estimation results show that the effect of easier qualifying conditions concentrates among workers with secondary school (Table 7).[15] For workers with secondary schooling, the eligibility for easier qualifying conditions increased the salaried-formal employment rate by 10 percentage points (on a 21 percent basis). In contrast, the estimated effects for workers with primary or post-secondary schooling are smaller and not significant. The additional columns of Table 7 show the average level of salaried-formal employment rates by educational attainment. Consistent with the theoretical framework, workers with low or high informality rates are less responsive to changes in pension benefits.

The second dimension is based on the composition of the household. I analyze the response of workers in households with different incentives to search for formal-sector jobs. The groups are married men and men living in a household with only one member in the labor force. These two groups should respond more actively to the eligibility for the transition system. First, men tend to get married to younger women (the median difference is 5 years). Because 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 in this part of the analysis is that these variables are endogenous to the eligibility for the transition system. However, I find no evidence that household structure changes due to the

[15]Implicitly, I am assuming that workers do not change their schooling as response of 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 Table 3.

eligibility for easier qualifying conditions (Table 8).

Table 9 reports the results for the considered samples. The RD estimates indicate that the effect of pension incentives varies depending on the household characteristics. The response is concentrated on married men, and households in which there is only one member in the labor force.

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

Table 10 reports the regression discontinuity estimates by region for 2005 and 2011. I group workers according to the departments (provinces) with the highest GDP per capita excluding oil. The developed departments are Bogota-Cundinamarca, Antioquia, and Valle and the developing departments are the rest of the country. The developed regions concentrate about 60 percent of the total GDP in 2005 and 45 percent of total population. The average response to changes in the pension benefits is large and significant for the developed regions, where most of the formal employment is.

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

5.2.1 Elasticity of Formal-sector labor Supply to the Effective Pension Tax Rate

To compute the elasticity of the formal-sector labor supply to the effective pension tax rate (σ), I implement a two-stage procedure. In the first stage, I compute the average change in the effective pension tax rate at the discontinuity by selected samples. In the second stage, I regress the estimates of the changes in employment against the changes in the effective pension tax rate to recover the elasticity.

For the first stage of the estimation of σ, I compute the effective pension tax rate for selected groups of workers. I use the groups analysis to compare the pension incentives of workers with different propensities to work in the formal sector. The final groups are based on combinations of region and educational attainment for 2005 and region and wage range for 2011 (12 groups).[16] For each group in the sample (denoted by X), I compute the average change in the effective pension tax rate at the discontinuity in three steps. First, 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 retires as soon as he meets the retirement conditions and enjoys 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 asks for the refund of his contributions up to date assuming an average contribution rate of 10 percent.[17] Otherwise, the worker works an additional period in the formal sector with probability $p_X(a)$.

Second, I compute the change in the log effective tax rate at the discontinuity as

$$\begin{aligned}\Delta \ln\left(1 - t^{eff}_{a\tau' X}\right) &= \ln\left(1 - 0.04 + \frac{\beta \Delta \mathbb{E}_X PW^{T}_{a+1}(\tau'+1)}{w_X}\right) \\ &\quad - \ln\left(1 - 0.04 + \frac{\beta \Delta \mathbb{E}_X PW^{SI}_{a+1}(\tau'+1)}{w_X}\right).\end{aligned}$$

where the superscripts T and SI denote that the expected pension wealth is computed using the conditions of the transition and the social insurance systems. Because the estimates of the changes in employment are observed for men in 2005 and 2011, I keep the change in the

[16]For 2005, the selected groups are developed and developing regions by primary, secondary and post-secondary education (6 groups). For 2011, the selected groups are developed and developing regions by wage range (1, 1-2, 2+ minimum wage, 6 groups).

[17]I choose 10 percent instead of 16 percent because the pension contribution rate before 1994 was 6.5 percent of the worker's wage.

log effective tax rate for workers with age 51 and 57 (the age of the eligible men at the cutoff in 2005 and 2011).

Third, 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 effective pension tax rate along the formal-informal margin as

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

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. Finally, I assume that w_X is constant over time and set it to twice the minimum wage for skilled workers and to the minimum wage for the other groups.[18]

For the second stage of the estimation of σ, I regress the changes in the formal employment rate at the discontinuity on the change of the effective pension tax rate. Figure 8 displays a scatterplot with the average changes in log employment (vertical axis) and in effective tax pension rate (horizontal axis) at the discontinuity. The reported values for 2005 and 2011 are represented by triangles and circles. Consistent with the predictions of the model, workers with higher pension incentives along the formal-informal margin also exhibit higher responses in the formal-sector labor supply. A linear regression fitted for these points yields an estimated elasticity of $\sigma = 1.77$. The estimated elasticity is slightly larger than the values of the same regression restricted to cross through the origin ($\sigma = 1.66$) and the median value of the elasticities by group ($\sigma = 1.37$). 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. First, the estimates of changes in the formal-sector supply do not account for spillover effects. Second, because of the definition of the transition system a fraction of the population could not take up

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

the benefits (Section 2.1). Third, $\Delta \ln \left(1 - t_{aX}^{eff}\right)$ may be an upper bound of the actual change in the effective pension tax rate 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 the utility function of workers is concave instead of linear (Stock and Wise, 1990),

6 Conclusion

In this paper, I show that workers take into account their future pension benefits for making their labor supply decisions. Using the Colombian pension system, I show that a change in future pension benefits generates a large shift in the labor supply between the formal and informal sector. 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 of finding formal-sector jobs. I also obtain an elasticity of formal-sector labor supply with respect to the effective pension tax rate of 1.8.

Although the estimation results cannot be generalized to other cohorts or countries, the results suggest that the behavioral response to pension incentives may be large. The behavioral response should be taken into account in the design of pension programs, as it may imply large efficiency costs. In particular, pension programs that reduce the value of the expected pension benefits have a negative effect on the formal-sector labor supply. From a fiscal perspective, the effect of such programs is twofold. On the revenue side, they reduce the revenue from contributions to the pension system, as less workers contribute. On the expenditure side, they increase the future expenditure in assistance programs, as more retirees would claim non contributive pension benefits.

Nevertheless, a complete evaluation of pension programs must account for other factors that may mitigate their efficiency costs. First, the welfare gains from the insurance against consumption loses after retirement may be significant. Second, the overall effect of pension programs depends on what population is affected. For example, non contributive pension programs for workers with low opportunities of finding formal-sector jobs could be welfare-enhancing. For those workers, the behavioral response is small and the extra gains from insurance may be large. An analysis of the efficiency costs and welfare consequences of retirement policies are additional topics for future research.

References

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

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

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

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.

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, 2013.

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.

Rafael La Porta and Andrei Shleifer. Informality and development. *Journal of Economic Perspectives*, 28(3):109–26, 2014. doi: 10.1257/jep.28.3.109.

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 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.

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.

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, 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.

Table 1: General Pension System characteristics

	Transition (DB)	**Social Insurance (DB)**	**Individual Account (DC)**
Managed by	Colpensiones	Colpensiones	Private pension funds
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.
Replacement rate	Function of length of contributions. From 65% to 85% (See Figure 1)	Function of length of contributions and wage. From 65% to 85% (See Figure 1).	It depends only on the accrued capital
Contribution	16% of wage – 11.5% contribution, 4.5% for administrative fees and insurance		
Employer payment	Salaried workers: 3/4 employer - 1/4 employee. Self-employed 100%		
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		Accrued capital + interest
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 et al. (2010), Llano et al. (2013) and texts of the reforms.

Table 2: 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 contributive healthcare system. Source: Colombian Household Surveys, 2011

Table 3: Identification checks, 2005

A: McCrary's Density Test

	Men	**Women**
Test Statistic	-0.023	-0.077
(Bandwidth 48 months)	[0.033]	[0.025]***
Observations	125,878	174,569

B: Balance Tests (estimates scaled up by 100)

	Men	**Women**
High School or less indicator	-1.05	-1.34
(Bandwidth 48 months)	[1.24]	[1.08]
Disability indicator	0.51	-0.16
(Bandwidth 48 months)	[0.98]	[0.67]
Ethnical Minority indicator	0.27	-0.82
(Bandwidth 48 months)	[0.77]	[0.63]
Observations	76,730	107,823

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 4.2.1. 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 age. 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 trend on age and its interaction with the eligibility for the transition system as independent variables (See equation (5)). 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 age (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 4: RD estimation results, 2005 and 2011

A: 2005 Results – Dependent variable: Salaried-formal indicator

	RD estimates		Distribution of years of contribution (%)					
	Men	**Women**	**Men**			**Women**		
			0-10	**11-20**	**21+**	**0-10**	**11-20**	**21+**
Easier qualifying conditions	3.13	0.40	20.2	39.8	40.0	27.6	36.7	35.7
(Bandwidth 48 months)	[1.23]**	[1.04]						
Ave. Dep. Variable (%)	18.1	15.7						
Observations	128,531	178,333						

B: 2005 Results – Dependent variable: Log salaried-formal workers by age in months

	RD estimates		Distribution of years of contribution (%)					
	Men	**Women**	**Men**			**Women**		
			0-10	**11-20**	**21+**	**0-10**	**11-20**	**21+**
Easier qualifying conditions	15.75	-6.79	20.2	39.8	40.0	27.6	36.7	35.7
(Bandwidth 48 months)	[8.51]*	[7.56]						
Observations	15,252	20,536						

C: 2011 Results – Dependent variable: Log salaried-formal workers by age in days

	RD estimates		Distribution of years of contribution (%)					
	Men	**Women**	**Men**			**Women**		
			0-10	**11-20**	**21+**	**0-10**	**11-20**	**21+**
Easier qualifying conditions	-6.80	-2.04	13.0	31.6	55.5	20.8	37.6	41.5
(Bandwidth 730 days)	[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 versus a quadratic polynomial on age and its interaction with the eligibility for the transition system as independent variables (See equation (5)). Panel A includes all 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 contributive healthcare 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 age (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 5: RD estimation results for wages in the formal sector, 2011

A: RD Estimates for log wages (estimates scaled up by 100)

	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)**
Easier qualifying conditions	3.11	–	0.25	1.51	-1.16	–	-0.23	-1.22
(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)

	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)**
Easier qualifying conditions	-7.85	-12.63	-7.82	-0.02	-1.74	-1.70	0.96	-3.80
(Bandwidth 730 days)	[2.53]***	[3.26]***	[2.59]***	[3.41]	[2.17]	[3.29]	[3.40]	[3.48]
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 versus a quadratic polynomial on age and its interaction with the eligibility for the transition system as independent variables (See equation (5)). 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 definition, 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 age (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 6: Estimation results for other labor market outcomes, 2005

A: RD estimates for other labor market outcomes – Men, 2005

	Participation	**Salaried formal**	**Salaried informal**	**Self employed**
Least Squares estimator (all estimates scaled up by 100)				
Easier qualifying conditions	-1.34	3.13	-0.65	-2.51
(Bandwidth 48 months)	[1.41]	[1.23]**	[1.51]	[1.25]**
Observations	128,531	128,531	128,531	128,531
Ave. Dep. Variable (%)	78.2	18.1	27.3	25.7

B: RD estimates for other labor market outcomes – Women, 2005

	Participation	**Salaried formal**	**Salaried informal**	**Self employed**
Least Squares estimator (all estimates scaled up by 100)				
Easier qualifying conditions	1.11	0.4	0.74	-0.54
(Bandwidth 48 months)	[1.73]	[1.04]	[1.33]	[0.91]
Observations	178,333	178,333	178,333	178,333
Ave. Dep. Variable (%)	49.6	15.7	18.9	10.8

Notes: Each cell reports an RD estimate based on a separate regression of a different labor market indicator versus a quadratic polynomial of age and its interaction with the eligibility for the transition system as independent variables (See equation (5)). The columns labeled salaried-formal presents the baseline RD estimates presented in Table 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 age (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 7: Estimation results by educational attainment – Men, 2005

A: 2005 Results - Dependent variable: Salaried-formal employment indicator

	RD Estimate	**Salaried-formal emp. rate**	**Dist. of years of contribution (%)**		
			0-10	**11-20**	**21+**
Primary	0.34	10.6	29.3	35.6	35.1
(Bandwidth 48 months)	[1.55]				
Secondary	10.12	21.2	26.3	40.8	33.0
(Bandwidth 48 months)	[3.40]***				
Post-Secondary	1.71	38.4	8.0	41.5	50.5
(Bandwidth 48 months)	[2.79]				
Observations	128,531				
Ave. Dep. Variable (%)	18.1				

Notes: The first column reports RD estimates based on a separate regression of the salaried-formal employment indicator versus a quadratic trend on age and its interaction with the eligibility for the transition system and educational attainment. The cells report the average effect of the eligibility for easier 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 age (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 8: Estimation results for household composition indicators – Men, 2005

Dependent variable	**Married**	**Only worker in HH**
Easier qualifying conditions	-0.37	1.51
(Bandwidth 48 months)	[1.40]	[1.52]
Observations	109,764	106,387
Ave. Dep. Variable (%)	83.1	38.7

Notes: Each cell reports an RD estimate based on a separate regression of the household composition indicator versus versus a quadratic trend on age and its interaction with the eligibility for the transition system (See equation (5)). 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 age (in months) in brackets. * p<0.1, ** p<0.05, *** p<0.01.

Table 9: RD results for subsamples based on household characteristics – Men, 2005

A: Estimates for men with spouse in the household

	RD Estimate	**Salaried-formal emp. rate**	**Dist. of years of contribution (%)**		
			0-10	**11-20**	**21+**
No Spouse	-1.11	21.8	11.5	62.5	26.0
(Bandwidth 48 months)	[3.63]				
Spouse	4.63	18.3	21.6	36.6	41.8
(Bandwidth 48 months)	[1.55]***				
Observations	109,764				
Ave. Dep. Variable (%)	19.6				

B: Estimates for men in households with one or more members in the labor force

	RD Estimate	**Salaried-formal emp. rate**	**Dist. of years of contribution (%)**		
			0-10	**11-20**	**21+**
More than one member in LF	2.10	16.4	21.7	36.1	42.2
(Bandwidth 48 months)	[1.82]				
One member in labor force	5.81	20.2	15.9	50.7	33.4
(Bandwidth 48 months)	[1.90]***				
Observations	106,387				
Ave. Dep. Variable (%)	20.4				

Notes: The first column of the Table reports an RD estimate based on a separate regression of the salaried-formal employment indicator versus a quadratic trend on age and its interaction with the eligibility for the transition system by household characteristics (See equation (5)). 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 age (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 10: Estimation results by region – Men, 2005 and 2011

A: 2005 Results - Dependent variable: Salaried-formal employment indicator

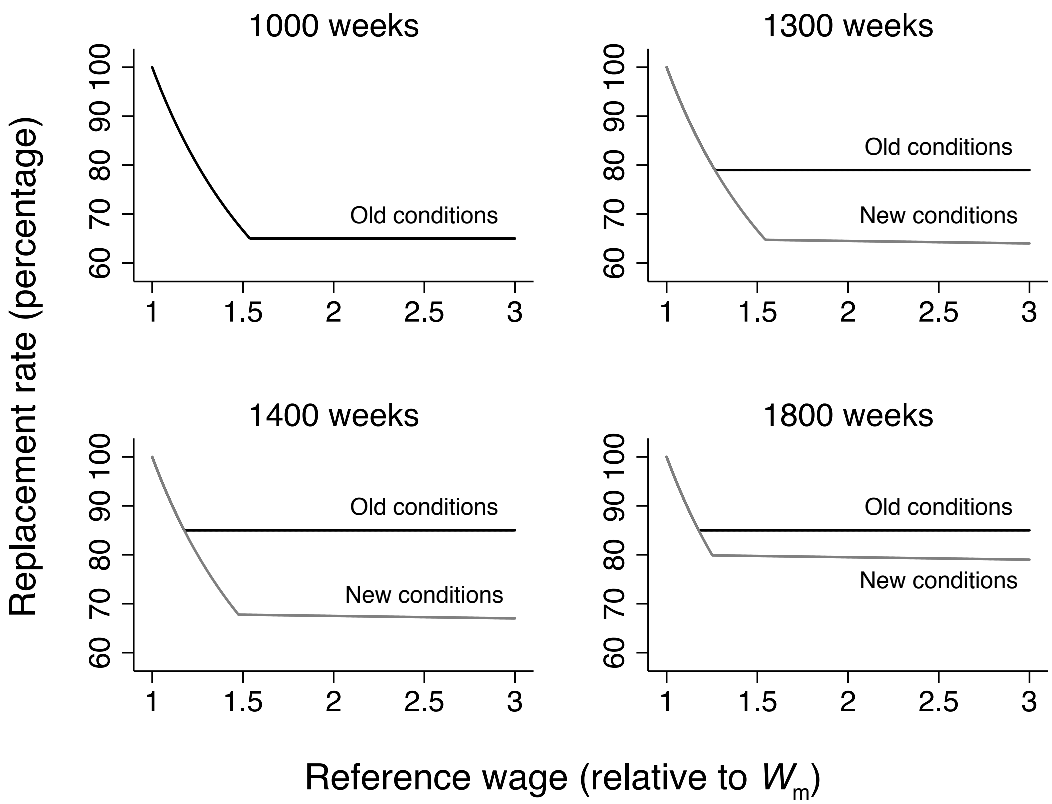
	RD Estimate	**Salaried-formal emp. rate**	**Dist. of years of contribution (%)**		
			0-10	**11-20**	**21+**
Developing regions	1.16	13.3	16.7	38.9	44.4
	[1.63]				
Developed regions	4.57	21.2	22.4	40.3	37.3
	[1.65]***				
Observations	128,531				
Ave. Dep. Variable (%)	18.1				

B: 2011 Results - Dependent variable: log number of workers

	RD Estimate	**Salaried-formal emp. rate**	**Dist. of years of contribution (%)**		
			0-10	**11-20**	**21+**
Developing regions	-1.56	18.1	14.0	38.4	47.7
	[5.68]				
Developed regions	-8.23	23.3	12.7	29.6	57.7
	[2.38]***				
Observations	964,558				
Ave. Dep. Variable (%)	20.4				

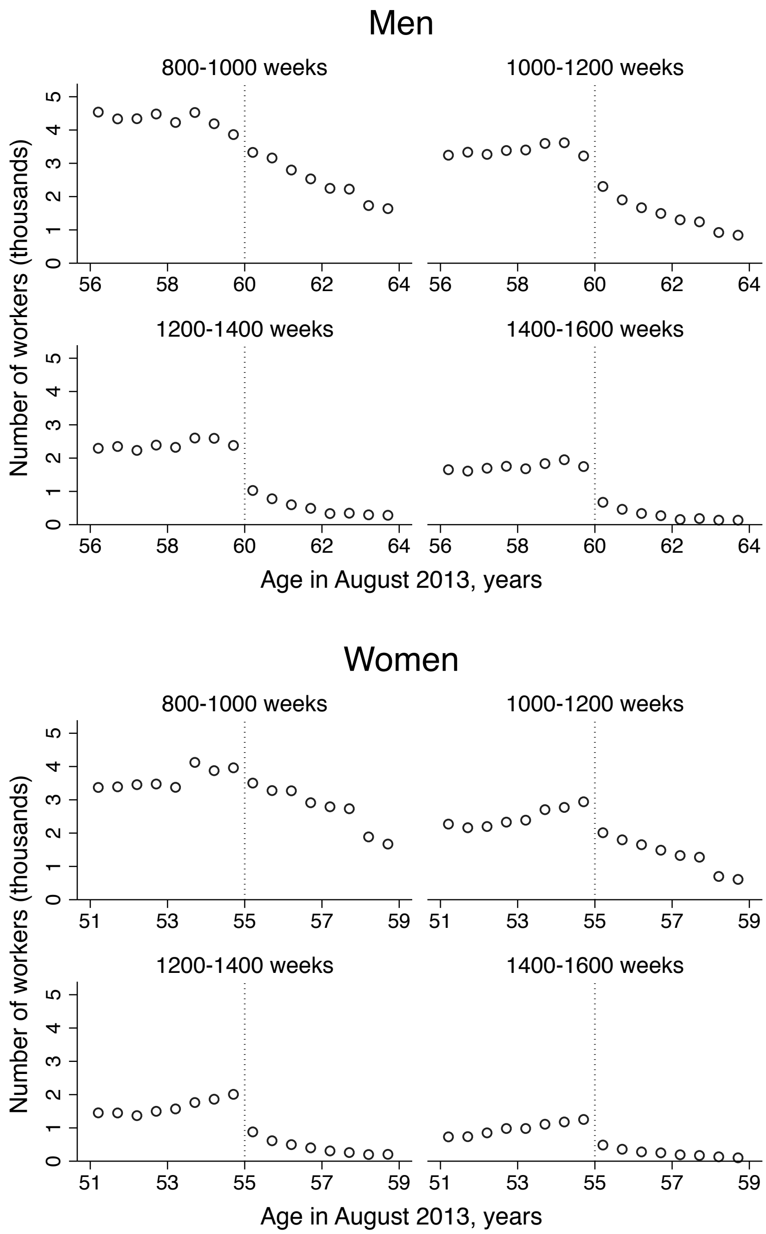
Notes: The first column reports RD estimates based on a separate regression of a salaried-formal employment variable versus a quadratic trend on age and its interaction with the eligibility for the transition system by region (See equation (5)). The top panel presents results by region for 2005, while 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 age (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: Replacement rate for the defined-benefit systems by weeks of contributions



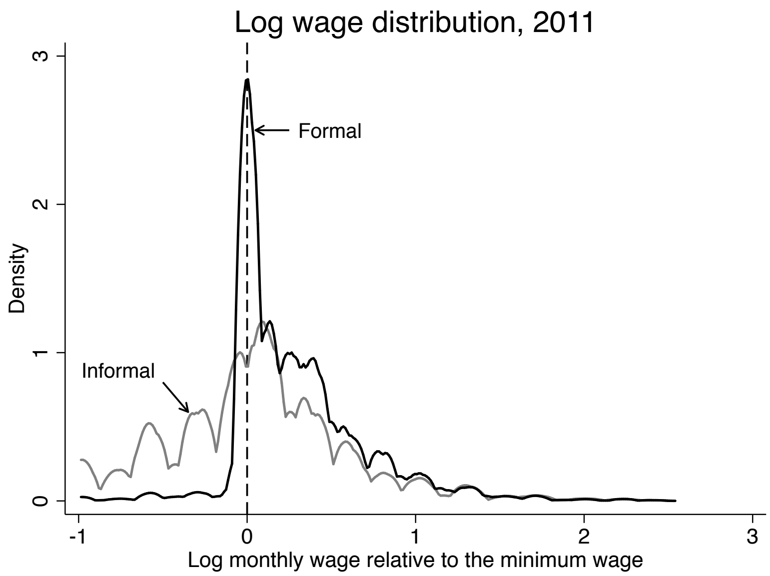
Notes: The figure displays the replacement rates as percentage of the reference wage for the social insurance (gray line) and the transition (black line) systems. Each panel represents the particular value of the formula defining the replacement rate by weeks of contributions. For the insurance system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

Figure 2: Distribution of workers by age and number of weeks



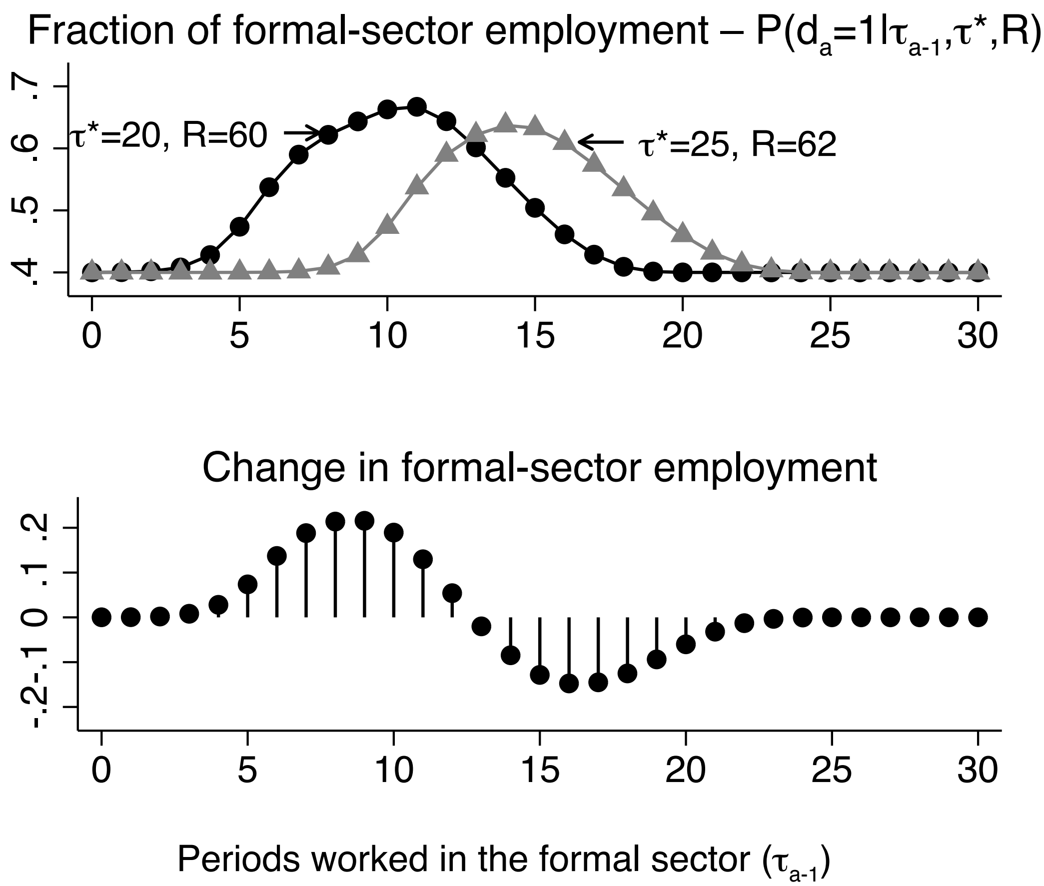
Notes: The figure presents the distribution by age, weeks of contribution and gender for non retired workers who have made contributions to the public pension system throughout their lifetime up to December 2013, based on Colpensiones administrative data. Once the workers claim the pension benefits they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

Figure 3: 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 workers are defined as workers who contributed to the pension and are covered by the contributory healthcare system.

Figure 4: Effects of changes of the minimum qualifying conditions for retirement on the probability of searching for a formal-sector job, worker age 50

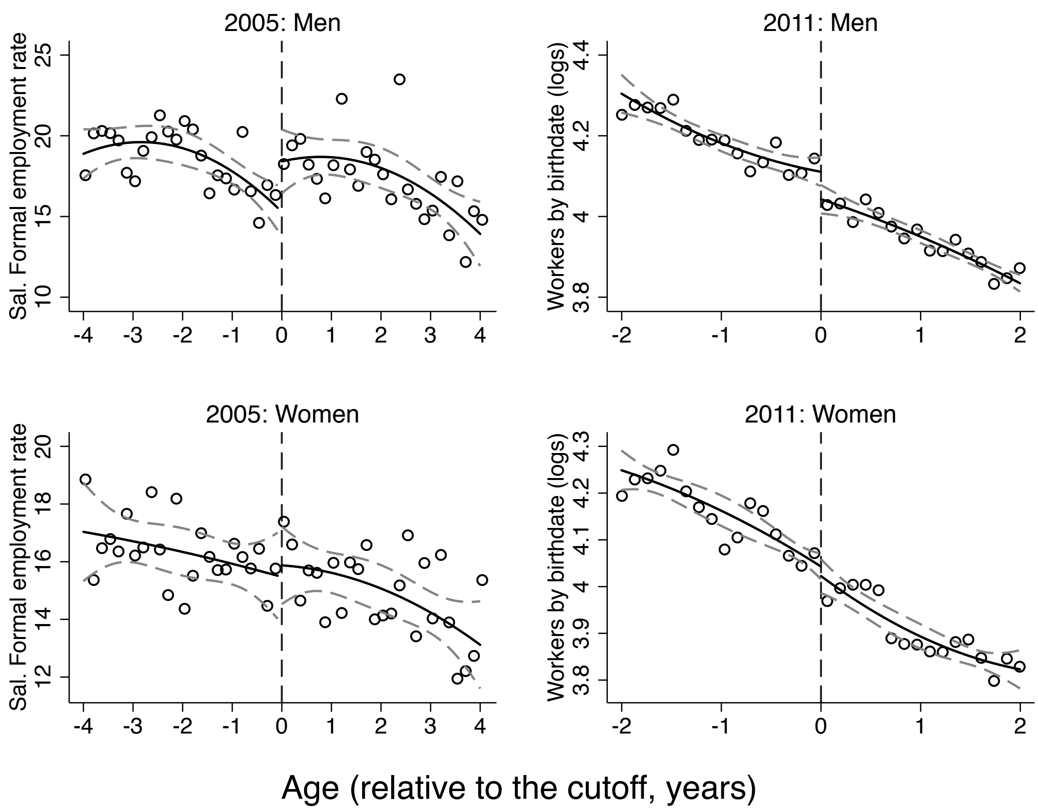


Notes: The figure presents a simulated scenario of the probability of searching for a formal-sector job as a function of the years of contribution of workers of age $a = 50$, given different qualifying conditions for retirement benefits. I assume the worker lives from $a_0 = 20$ until $T = 75$ years and his utility is linear, the wages in the formal and informal sector are $w^f\left(1 - t^{nom}\right) = 1.2$ and $w^i = 1$, $\theta^r = \theta^f = 0$, the random search costs follow a uniform distribution $\psi \sim \mathcal{U}\left(0, 0.5\right)$, $\beta = \frac{1}{1.05}$, and two different pension systems. The first pension system is given by $R = 60$, $\tau^* = 20$ and $b = 1$ and the second system is given by $R = 62$, $\tau^* = 25$ and $b = 1$. The top panel of the figure shows the probability of searching for a formal-sector job for two cohorts with age $a = 50$ but the two different pension systems. Within a cohort that is not yet retired, increasing the minimum retirement age and the vesting period changes the incentives to search for formal-sector jobs, as a function of the experience in the formal sector, τ_{a-1}. The probability of searching as a function of the years of contribution in the formal sector shifts rightwards, as the relevant vesting period shifts and the "hump" is smaller, as the minimum retirement age increases. The bottom panel displays the difference between the probability of searching with $R = 60$ and $\tau^* = 20$ and the same probability with $R = 62$ and $\tau^* = 25$. In the figure, a negative value implies that the workers are discouraged to search for a formal-sector job with the more difficult qualifying conditions, and vice versa.

Figure 5: Rolling t-statistics for testing the manipulation in birth date, 2005

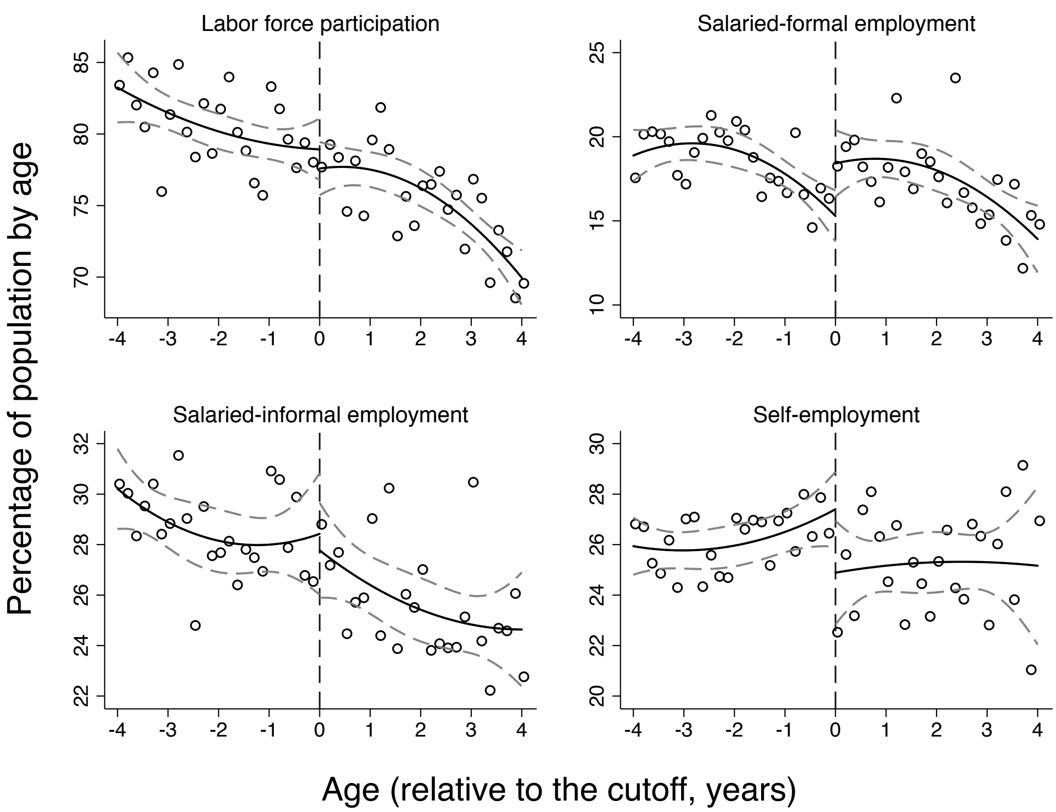
Notes: The figure displays the t-statistics using the test proposed by McCrary (2008) for testing discontinuities in the density of the running variable in the regression discontinuity setup. Each panel represents the value of the t-statistics changing the cutoff point, where the vertical dashed lines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

Figure 6: RD estimation results, 2005 and 2011



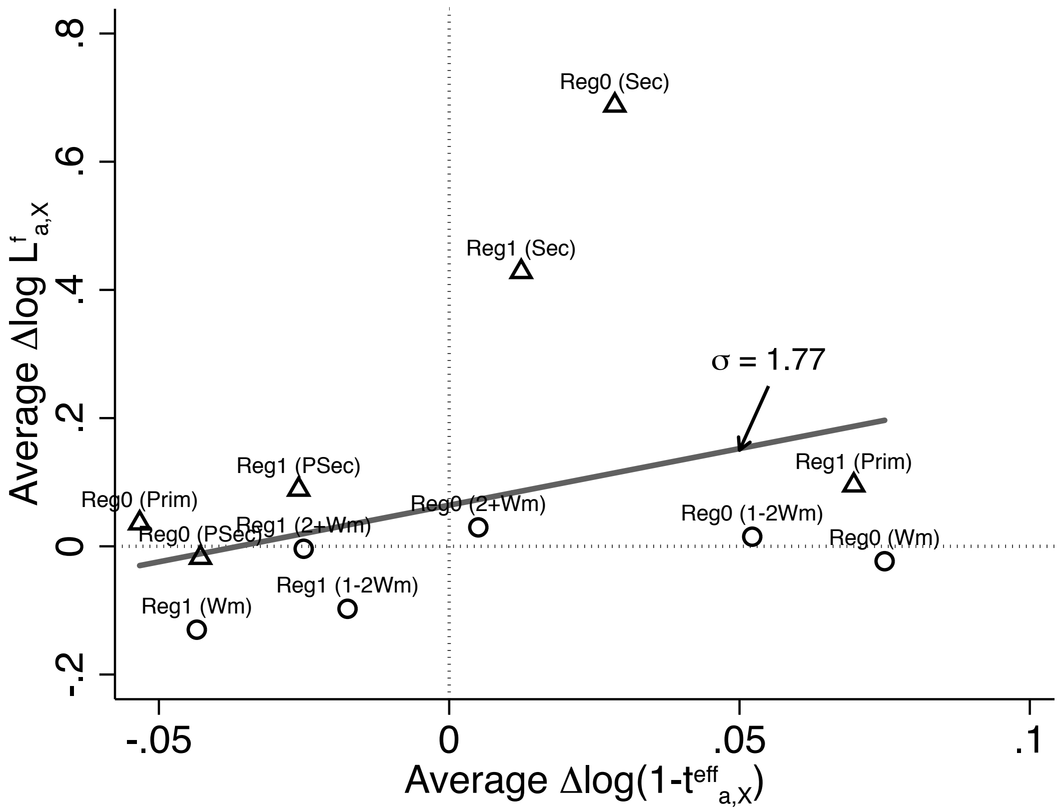
Notes: The figure presents the salaried-formal employment indicators by gender and age. Each point represents the 2-month average of the salaried-formal employment rate by month in 2005 and the 47-days average of the log number of workers by age in 2011. The regression estimates on the graphs are based on the estimates reported in panels A and C of Table 4. Confidence bands are computed with standard errors clustered by age (in months).

Figure 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 age. 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 6) of the Colombian Census long-form questionnaire dataset. Confidence bands are computed with standard errors clustered by age (in months).

Figure 8: Elasticity of the formal-sector labor supply to changes in the effective pension rate



Notes: The Figure displays the average change in the log effective tax rate (horizontal axis), computed in Section 5.2.1, and the average change in the salaried-formal labor supply (vertical axis), derived from the results obtained in Section 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 effective pension rate along the formal-informal margin.

A Appendix

A.1 Model implications

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

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

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

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

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

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 $v_{T-1}^{w}\left(\tau_{T-2}\right)$ as

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

which is strictly less than $v_{T-1}^{r}\left(\tau_{T-2}\right)$, 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 $v_{a}^{r}\left(\tau_{a-1}\right)=u^{0}+\beta v_{a}^{r}\left(\tau_{a-1}\right)$, which is less than $v_{a}^{w}\left(\tau_{a-1}\right)$. 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 $\tau+(T-a+1)<\tau^{*}$, $\Delta v_{a}\left(\tau+1\right)=0$ and therefore $\bar{u}_{a}\left(\tau\right)=\tilde{u}$. Second, I show that $\Delta v_{a}\left(\tau+1\right)\geq0$ for all other values of τ, and so $\bar{u}_{a}\left(\tau\right)\geq\tilde{u}$ 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 (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 values of $\tau \in [\tau^* - (T - a + 1), \tau^* - 1]$, $\Delta v_a(\tau + 1) \geq 0$. 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

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

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. Consider 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.[19]

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, consider first the age $a = R' - 1$. 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 $\bar{u}_{R'-2}(\tau) = \bar{u}'_{R'-2}(\tau)$ 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

[19] 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.

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

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

and

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

For $a = R' - 3$, the change in the minimum retirement age have the same type of composition effects than 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. In addition, $\Delta v_{R'-3}(\tau + 1) = \Delta v'_{R'-3}(\tau + 1)$ for $\tau \notin \{\tau^* - 3, \tau^* - 2, \tau^* - 1\}$ and $\Delta v_{R'-3}(\tau + 1) \geq \Delta v'_{R'-3}(\tau + 1)$ 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 than the case of $a = R' - 2$, and for $\tau = \tau^* - 2$,

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

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 same arguments presented in propositions 1 and 2, solving the model by backward induction yields the same type of policy function than before, but it uses τ' as a reference point instead of τ^*, and so 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 and retire goes down. $\square$

Table A.1: Robustness test, 2005

	Men				Women			
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
A: Least squares estimator (all estimates scaled up by 100)								
Linear control function	2.75	2.43	2.84	2.90	0.71	0.62	0.49	0.72
(Bandwidth 48 months)	[0.92]***	[0.88]***	[0.84]***	[0.93]***	[0.71]	[0.70]	[0.70]	[0.70]
Quadratic control function	3.13	2.87	2.99	3.20	0.4	0.28	0.32	0.34
(Bandwidth 48 months)	[1.23]**	[1.29]**	[1.09]***	[1.26]**	[1.04]	[0.99]	[1.05]	[1.03]
B: Probit estimator (all estimates scaled up by 100)								
Linear control function	2.79	2.61	2.63	3.25	0.73	0.64	0.44	0.86
(Bandwidth 48 months)	[0.94]***	[0.97]***	[0.77]***	[1.05]***	[0.71]	[0.70]	[0.59]	[0.78]
Quadratic control function	2.99	2.93	2.56	3.39	0.38	0.27	0.34	0.39
(Bandwidth 48 months)	[1.20]**	[1.34]**	[0.95]***	[1.38]**	[1.04]	[0.98]	[0.85]	[1.14]
C: Logit estimator (all estimates scaled up by 100)								
Linear control function	2.78	2.64	2.44	3.35	0.73	0.63	0.38	0.85
(Bandwidth 48 months)	[0.94]***	[0.99]***	[0.73]***	[1.08]***	[0.71]	[0.70]	[0.54]	[0.81]
Quadratic control function	2.97	2.94	2.38	3.48	0.38	0.26	0.23	0.37
(Bandwidth 48 months)	[1.20]**	[1.36]**	[0.90]***	[1.41]**	[1.04]	[0.98]	[0.79]	[1.18]
D: Local linear estimator (all estimates scaled up by 100)								
Local linear	2.68	–	–	–	1.04	–	–	–
(Bandwidth 24 months)	[1.08]**				[0.95]			
Local linear	3.01	–	–	–	0.63	–	–	–
(Bandwidth 36 months)	[1.00]***				[0.84]			
Observations	128,531	128,531	128,531	128,531	178,333	178,333	178,333	178,333
Average dep. Variable	18.1	18.1	18.1	18.1	15.7	15.7	15.7	15.7
Fixed effects		Month of birth	HS or less	Region		Month of birth	HS or less	Region

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 contributive healthcare system versus a polynomial of age and its interaction with the eligibility for the transition system (See equation (5)). 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 age (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 (all 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	-6.80	-5.85	-6.80	-2.04	-1.92	-2.04
(Bandwidth 730 days)	[2.39]***	[2.04]***	[2.39]***	[2.17]	[1.58]	[2.17]
B: Local linear estimator (all estimates scaled up by 100)						
Local Linear	-8.4	–	–	-3.65	–	–
(Bandwidth 360 days)	[2.24]***			[1.62]**		
Local Linear	-6.75	–	–	-3.86	–	–
(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 contributive healthcare system versus a polynomial of age and its interaction with the eligibility for the transition system (See equation (5)). 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 age (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

Estimator	**Men**				**Women**			
	Salaried formal (pension)	**Formal worker**	**Formal worker (ISS)**	**Pension**	**Salaried formal (pension)**	**Formal worker**	**Formal worker (ISS)**	**Pension**
	(1)	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Least Squares estimator (all estimates scaled up by 100)								
Quadratic control function	2.92	3.22	2.16	3.88	0.37	0.54	0.14	0.26
(Bandwidth 48 months)	[1.36]**	[1.31]**	[0.87]**	[1.27]***	[1.10]	[0.91]	[0.50]	[1.04]
Observations	128,531	128,531	128,531	128,531	178,333	178,333	178,333	178,333
Ave. Dep. Variable (%)	19.0	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 trend on age and its interaction with the eligibility for the transition system as independent variables (See equation (5)). The definitions used are (i) salaried-formal employment based on contributions to the pension system, regardless of coverage of the contributive healthcare system; (ii) formal employment for all workers contributing to the pension and covered by the contributive healthcare system, regardless of type of employment; (iii) formal employment for workers contributing to the pension system and covered by the contributive healthcare system managed by the ISS; 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 age (in months) in brackets. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.