Pension Privatization and Worker Behavior: Evidence from a Cohort-Based Reform*

Maximiliano Lauletta†
Federal Reserve Board

Marcelo Bérgolo‡
IECON-UDELAR & IZA

August 27, 2025

Abstract

We provide the first causal evidence of the effect of pension privatization on workers' behavior. We study a reform in Uruguay that, starting from a pay-as-you-go system with defined benefits, introduced a defined contribution component with retirement accounts. For identification, we leverage a cohort-based discontinuity in the introduction of the new system with regression discontinuity analyses. Using rich administrative and census data, we find that privatization increased older workers' employment, particularly among low-wealth individuals and those with poor health, reducing rates of early retirement. Privatization also induced workers to report higher early career earnings, with evidence that this is driven by reduced tax evasion.

JEL Codes: H55, J22, J26, J32, H26

*First version: September 2021. We are grateful to Emmanuel Saez, Edward Miguel, Alan Auerbach, and Ricardo Pérez Truglia for their constant support throughout this project. We also thank George Bulman, Federico Echenique, Álvaro Forteza, Marco González Navarro, Hilary Hoynes, Supreet Kaur, Patrick Kline, Ianina Rossi, Damián Vergara, Gabriel Zucman, and seminar participants at UC Berkeley, UC Los Angeles, UC Santa Cruz, the George Washington University, the Federal Reserve Board, Nova SBE, Universidad de San Andrés, the IIPF Annual Congress, the EEA Annual Meeting, the Zurich Conference on Public Finance in Developing Countries, and the RIDGE Workshop on Public Economics for helpful comments and suggestions. Lauletta gratefully acknowledges funding from the Center for Equitable Growth at UC Berkeley. Any mistake or omission is our responsibility. The analysis and conclusions set forth here are those of the authors and do not indicate concurrence by other members of the Federal Reserve Board research staff, the Board of Governors, or the Federal Reserve System.

†Corresponding author, e-mail: maximiliano.lauletta@frb.gov.

‡E-mail: marcelo.bergolo@fcea.edu.uy.

1. Introduction

Pension systems are central to modern social insurance, with ongoing debates about their design. Discussions typically revolve around whether systems should be funded or unfunded, public or private, and based on defined benefit (DB) or defined contribution (DC) formulas. A prominent reform proposal is the shift from government-run pay-as-you-go (PAYG) DB systems to funded retirement accounts—commonly referred to as *privatization*. This typically involves financing pensions through individual savings and replacing DB with DC formulas.[1] Advocates argue that this shift improves financial sustainability, while boosting labor supply and economic efficiency by strengthening the link between contributions and benefits (Lindbeck and Persson, 2003; Auerbach and Kotlikoff, 1985; Feldstein, 1995). Even though multiple countries have undergone privatizations following such recommendations, empirical evidence remains limited.

In this paper, we contribute to this discussion by studying the effects of the partial privatization of the pension system on workers' employment and earnings trajectories. We study a reform in Uruguay that, starting from an exclusively unfunded DB public system, introduced an individual capitalization component in which a fraction of workers' contributions goes to a retirement account managed by a pension fund, while the rest is used to fund public pensions.[2] Pensions then consist of a government-provided pension based on average earnings over the last 10 years of employment and an annuity from accumulated savings.[3] The new system applied to workers under 40 at the time of the reform, while older cohorts remained in the original system, creating a sharp date of birth-based cutoff that underpins our identification strategy. Using rich administrative and census records, we leverage the cohort-based discontinuity in the introduction of the mixed system with a Regression Discontinuity Design (RDD), comparing the trajectories of individuals born within days of the cutoff over the course of two decades.

We first analyze employment responses. Employment rates are similar across systems during workers' 40s and early 50s, but diverged closer to retirement: by age 57, workers in the mixed system are about 4 percentage points more likely to be employed. Census data

[1] Note that "privatization" in the pension reform context does not necessarily imply that the system is privately run, many countries that have undergone such reforms have often had a government-run pension fund that workers can choose.

[2] The share of the contributions that is allocated to the retirement fund varies depending on the level of earnings and choices that workers can make within the system. We describe the system in more detail in section 2.

[3] These short "windows" of pre-retirement earnings to calculate pension benefits are common in low- and middle-income countries and some specific systems in high-income countries (such as some public-sector workers). In Latin America, several countries use a 10-year window for benefits calculation (such as Argentina, Colombia, and Uruguay), while others use shorter windows (such as Peru and Paraguay). Pension systems for civil servants in some African countries use for reference the very last salary (Stewart and Yermo, 2009).

show this is largely due to delayed retirement, consistent with stronger incentives to remain employed in a capitalization system, where annuity values rise significantly with postponed retirement. Effects are concentrated among low-socioeconomic status workers and those with mild health problems—groups more likely to retire early (Leijten et al., 2015; Van Rijn et al., 2014). Since these mild health problems are unlikely to qualify for permanent disability pensions, the results likely reflect regular early retirement choices. Interestingly, this contrasts findings in Manoli and Weber (2016), where unhealthy individuals are less likely to respond to financial incentives. In our scenario, unhealthy individuals are *more* likely to respond to pension incentives.

Turning to earnings, workers in the new system reported earnings 15 to 20 percent higher immediately after the reform, with effects fading by their mid-to-late 50s. This is in line with the theory that tighter contribution-benefit links can incentivize labor supply (Auerbach and Kotlikoff, 1985; Kotlikoff, 1996; French et al., 2022), though the magnitude suggests implausibly large elasticities. Motivated by this, and survey evidence on widespread income underreporting (Bergolo and Cruces, 2014; Feinmann et al., 2024), we assess whether this is a real labor supply response or driven by reduced evasion.

Multiple pieces of evidence point to reduced underreporting. First, we find no effect on reported days or hours worked, both measures of real labor supply. Second, we find no effect on earnings in the public sector, where underreporting is minimal. Third, the effect is larger for firm owners and the self-employed, who have more scope to underreport earnings. Fourth, the effect is highly concentrated in sectors high in informality and evasion, as measured by household survey data. Finally, multiple measures of socioeconomic status and consumption from census data are perfectly balanced despite the two groups having very different reported earnings for over a decade, suggesting the effect is primarily an evasion response. Together, these suggest that privatization reduced evasion by strengthening the contribution-benefit link.

These findings align with standard economic models. Intuitively, the DC component rewards delaying retirement by increasing annuities relative to the minor adjustments of the DB formula. In the unfunded DB system, contributions early in the career are disconnected from future benefits—encouraging evasion—since only the last 10 years matter. The mixed system reduces this distortion by linking all contributions to benefits. As DB-only workers enter the pension calculation window, their incentives to evade fall, closing the earnings gap. A simple conceptual framework and calibrated simulations illustrate the incentives of such a reform. In our simulations, the reform increased the marginal incentive to earn by over 20% early on, due to the effect of earnings in future pension benefits, suggesting a back-of-the-envelope elasticity of earnings relative to the marginal incentive to earn of around 0.7 to

0.9.

Our paper contributes to several literatures. First, we provide rare causal evidence on pension privatization, a topic largely studied via general equilibrium models (Auerbach and Kotlikoff, 1985; Feldstein, 1995; Kotlikoff, 1996; Edwards and Edwards, 2002; Nishiyama and Smetters, 2007; Hosseini and Shourideh, 2019). While these models motivated widespread reform (Orenstein, 2013), credible empirical evidence has been lacking, potentially due to uniform national reforms (e.g., Chile) or dual systems with voluntary choice (e.g., Colombia, Argentina). Uruguay's cohort-based cutoff and rich administrative data provide an ideal quasi-experimental setting to study privatization and its effects on labor supply over the life cycle. We document behavioral responses consistent with theory, in particular with incentives to reduce tax evasion (consistent with suggestive evidence in Kumler et al., 2020).

Second, we contribute to the empirical literature on labor supply responses to pension incentives (Gelber et al., 2016; Liebman et al., 2009; Manoli and Weber, 2016; Fetter and Lockwood, 2018; Brown, 2013; Hernæs et al., 2016; Coile and Gruber, 2007), extending the focus from benefit levels to system structure.[4] By studying outcomes decades before retirement, we show that structural reform to pension incentives can affect behavior even when retirement is distant—despite potential behavioral biases like exponential growth bias or present bias (Goda et al., 2019) and poor understanding of pension rules (Liebman and Luttmer, 2012, 2015). Our findings suggest that privatization can reshape reporting behavior early in life, consistent with DC incentives.

A recent related paper by French et al. (2022) studies the effects of switching from an unfunded DB system to a Notional Defined Contributions system on labor supply in Poland, finding a significant reduction in employment rates at ages 51-54.[5] A calibrated life-cycle model suggests that the reform slightly reduced overall labor supply due to disincentives later in life more-than-compensating stronger labor supply incentives early in life. Our paper differs significantly, not only by studying a reform that introduced actual retirement accounts, but also by studying responses throughout the life-cycle without need of a calibrated model and finding different patterns. Contrary to French et al. (2022), we find that the transition from DB to a mixed system induced an overall *increase* in lifetime labor supply, both in later-in-life employment and early-in-life reported earnings, and we highlight the importance of evasion as a key margin of response to pension incentives.

Third, we contribute to the literature on tax evasion and compliance. While prior work focuses on firms (e.g. Pomeranz, 2015; Naritomi, 2019; Bachas and Soto, 2021; Brockmeyer and

[4] Also related are papers on savings and retirement spending (Attanasio and Rohwedder, 2003; Attanasio and Brugiavini, 2003; Lachowska and Myck, 2018).

[5] The Polish reform allowed workers to opt-in for saving part of their contributions in actual retirement accounts, but very few workers around the cutoff date of birth opted for this option (French et al., 2022).

Sáenz Somarriba, 2025) or individual enforcement (e.g. Londoño-Vélez and Ávila-Mahecha, 2021), labor earnings underreporting has received less attention. Yet, it is widespread and sizable (Feinmann et al., 2024; Bergolo and Cruces, 2014; Gavoille and Zasova, 2023), even in high-income countries (Saez, 2010; Bjørneby et al., 2021). Our evidence shows that pension incentives significantly affect reporting among both employees and the self-employed. The earnings response we observe is substantially larger than the one documented by Dean et al. (2022) for self-employed workers nearing the DB calculation window. Our effects appear decades earlier and extend to all worker types, including dependent employees.

Finally, the Uruguayan case offers a valuable middle ground between external validity and identification. The reform closely resembles other partial privatizations globally—such as in Argentina, Colombia, and even proposed models in the U.S.—and Uruguay shares key labor market features with other Latin American countries, including informality rates of 20–35%. Although the complexity behind pension privatization precludes from perfectly isolating the factors that drive the responses, our findings carry broad policy relevance for understanding how such reforms can affect labor supply and tax evasion.

The rest of the paper is structured as follows. Section 2 presents the institutional context. Section 3 introduces a conceptual framework for analyzing workers' responses. Section 4 presents the data. Section 5 presents the econometric strategy. Section 6 presents the main results. Section 7 concludes.

2. Institutional context

This section outlines the institutional context. We first briefly describe the context of Uruguay and the pension system in place prior to the reform. We then describe the main aspects of the reform and the key differences between the two systems we compare. A summary of the key institutional features can be found in table 1.

Uruguay is an upper-middle-income country with 3.5 million people and a 2018 GDP per capita of $18,000. The country has a comprehensive contributory social security system for formal workers, including retirement, health, and unemployment benefits, managed by the Social Security Agency (*Banco de Previsión Social*). As in much of Latin America, informality and payroll tax non-compliance have historically been high, though both have declined in recent years.

Prior to 1996, the pension system was a pay-as-you-go defined benefit (DB) scheme, with pensions funded by payroll taxes and calculated over the last 5 years of earnings. Concerns about its sustainability led to a reform via Law 16,713, which introduced a two-pillar system

combining the DB scheme with a funded defined contribution (DC) component.[6] Workers contribute 15% of their salary, split between the SSA and individual retirement accounts, depending on income thresholds and contribution options. The default allocates contributions below the first threshold to the DB pillar, while income above it goes to retirement accounts. An alternative option (known as Article 8) spreads contributions evenly below the first threshold and is chosen by about 75% of workers (CESS, 2021). Contributions above the third threshold are voluntary, and workers can choose among several pension funds to manage their retirement savings.

In the new system, pensions consist of two parts: (i) a government-provided DB pension and (ii) an annuity based on funds accumulated in individual retirement accounts. The DB pension is calculated as a replacement rate applied to a "contributory salary," which is the average indexed earnings over the last 10 years of employment. For those opting into the Article 8 option, this salary is reduced to 75% of the default, despite contributing only half as much—effectively providing a subsidy to the DB part. The statutory replacement rate starts at 45% for individuals with 30 years of contributions and increases with additional years and delayed retirement, up to a maximum of 82%. The DC component is determined by the accumulated savings and life expectancy at retirement, with pension funds transferring balances to a government-run insurer to provide annuities, and includes provisions for early retirement in certain sectors and for permanent disability.[7] There is a maximum and minimum pension for the DB part.

To gradually roll the new pension system in, workers aged less than 40 at the time the law entered into effect would be switched to the new mixed system, while those aged 40 or more would remain in a *transition system* that retained the PAYG-DB nature of the original system. For workers left in the transition system, their pension is determined under the same rules as the DB part of the workers in the new system and their contribution rate is still 15%, with the difference that these contributions only go to the public PAYG-DB system based on all of their labor earnings, and the maximum pension is capped at a higher level to compensate for the fact that they do not receive a DC pension. This discontinuity implied that individuals born up until April 1st 1956 remained by default in a PAYG-DB system, while those born after entered the new mixed system–meaning people born just hours apart faced different pension regimes. Exceptions applied to older workers with no prior formal employment or those who opted into the new system within six months of the reform.

The disability insurance system is integrated with pensions for both systems. In the case

[6] The reform also gradually raised the retirement age for women from 55 to 60 (for all systems).

[7] For example, similar to other settings, teachers in the education sector get 4 years of contributions per every 3 years worked. There are other activities that are deemed dangerous or too taxing that also have similar clauses. This is independent of being in the mixed system or the PAYG-DB system.

of the mixed system, the DB component is calculated by a replacement rate of 65% over the earnings that contributed to the DB system, and an additional 45% is provided to cover for the DC part via a mandated insurance scheme based on earnings that contributed to the DC system. For the case of the unfunded DB system, the disability pension is calculated using an equivalent formula but is funded exclusively via payroll taxes. Workers who are disabled need to go through a strict government medical evaluation to determine their disability status, and have to be deemed completely unable to work. Thus, disability pensions constitute a small minority of all pensions, and the mild health problems listed in the census data that we study are unlikely to qualify for this benefit.

3. Conceptual framework

This section presents a simple framework to understand how pension privatization can affect worker behavior. We focus on intuitive insights based on simplified pension formulas and a continuous time framework.[8] We assume workers live for T periods, and earn income $y(t)$ in each time period t until retirement at age R. We denote retirement systems with S, with M and DB representing the mixed and PAYG-DB systems, respectively. The DB pension benefits are calculated as a replacement rate $\rho^S(R)$ applied to average earnings over the last L years (we assume $\rho^M < \rho^{DB}$, as is the case in the real formulas), which increases with the retirement age. In the mixed system, a share γ of contributions τ^{SSC} is saved in a fund that grows at rate r and is paid out as an annuity over retirement.[9] Let $\mathbb{1}\{S = M\}$ be an indicator equal to 1 for the mixed system, the monthly pension benefits B^S for each system S can then be written as:

$$B^S = \rho^S(R)\frac{1}{L}\int_{R-L}^{R} y(t)dt + \mathbb{1}\{S = M\}\left(\frac{1}{T-R}\tau^{SSC}\gamma\int_0^R y(t)e^{r(R-t)}dt\right)$$

This formula highlights key differences in incentives. As for retirement timing incentives, in the PAYG-DB system delaying retirement slightly raises the replacement rate ρ. In contrast, the mixed system's annuity increases significantly with delayed retirement, since savings grow and are spread over fewer years, which strongly discourages early retirement.

[8] This discussion focuses on the switch from a DB to a DC formula, abstracting away from additional factors that can affect incentives after a privatization (e.g. aversion to financial markets, funding of retirement accounts, investment decisions, etc). These factors can all affect the incentives of a pension privatization but our setting does not allow us to tease out the differential impact of each, exceeding the scope of this paper.

[9] For simplicity, this framework abstracts from any uncertainty and specificities in the allocation of contributions in the mixed system (e.g. the income thresholds that affect how contributions are distributed between systems and the Article 8 choice).

Given an earnings trajectory, the effective replacement rate increases more with the retirement age in the mixed system than in the PAYG-DB system due to the strong effect that the retirement age has on the annuity from retirement savings. Thus, we expect workers in the mixed system to be more likely to be employed later-on.

We can pin down this intuition by analyzing the net return to remaining employed ($MNRR$), similar to the net return to work from French et al. (2022). In particular, let τ^{PIT} represent the marginal tax rate (from income taxes and other contributions) and τ^{SSC} on earnings $y(t)$, the marginal net return to remaining employed ($MNRR$) is given by:

$$MNRR^S(R) = \left(1 - \tau^{PIT} - \tau^{SSC}\right) y(R) \; - \; B^S(R) \; + \; (T-R)\frac{\partial PDV^S}{\partial R}$$

The worker gets one more period of net labor income, gives up a period of pension income, and increases the present discounted value of the future pension benefits they will receive as pension (PDV). The benefits received and the effect on future pension benefits depends on the system. Assuming for expositional simplicity that future flows are discounted at a zero rate, define $\bar{y}$ as the average wage in the last L periods of employment, we can get that the difference in the $MNRR$ between the mixed system and the DB system is given by:

$$\begin{aligned} MNRR^M(R) - MNRR^{DB}(R) = &-\left(\rho^M(R) - \rho^{DB}(R)\right)\bar{y} \\ &+ (T-R)\left[\left(\rho^{M\prime}(R) - \rho^{DB\prime}(R)\right)\bar{y}\right] \\ &+ (T-R)\left[\tfrac{\rho^M(R)-\rho^{DB}(R)}{L}\left(y(R) - y(R-L)\right)\right] \\ &+ \tau^{SSC}\gamma\left(y(R) + r\int_0^R y(t)e^{r(R-t)}dt\right) \end{aligned}$$

So the difference in replacement rates across systems ($\rho^M(R) < \rho^{DB}(R)$) pushes against delaying in the mixed system, but the DC term (which is always positive) pushes strongly toward delaying retirement. The net effect depends on parameters, but it is likely that the higher accrual from the DC system will dominate (as is the case in our simulations below). This formula also illustrates how the DB system can reward high late-career earnings, although the compounding effect from capitalization is likely to be strong under most numerical simulations.

As for intensive margin responses, in the PAYG-DB system only earnings in the final years affect pensions, encouraging lower earnings earlier in life. The mixed system links all contributions to benefits, incentivizing higher earnings early-on. We can further pin down

the effect of such a reform on earnings incentives by looking at a continuous version of the effect on the net return to work from French et al. (2022). In particular, the marginal net return to producing an additional peso of earnings ($MNRE$) is given by:

$$MNRE^S(t) = \left(1 - \tau^{PIT} - \tau^{SSC}\right) + E_t\left(\frac{\partial PDV^S}{\partial y(t)}\right)$$

The worker gets $1-\tau^{PIT}-\tau^{SSC}$ post-tax pesos plus the effect that an additional peso has on the expected present discounted value of their future pension (PDV^S), which depends on the system. Assuming again a zero discount rate, the difference in the marginal net return to earnings at time t between the systems is given by:

$$MNRE^M(t) - MNRE^{DB}(t) = \gamma\tau^{SSC}e^{r(R-t)} + \mathbb{1}\{t \in (R-L, R)\}\left(\frac{T-R}{L}\left(\rho^M(R) - \rho^{DB}(R)\right)\right)$$

Given a retirement age, the net return to earnings in the mixed system early on is much higher than in the PAYG-DB system because of the share of contributions that are invested until retirement. This difference decreases as the worker draws closer to the window for benefit calculation because those savings will capitalize for fewer years. Once the worker enters the window for benefit calculation, the mixed system has a lower net return to earnings because the PAYG-DB system has a higher statutory replacement rate. In addition, for higher retirement ages, the mixed system offers a higher return to work because savings grow for a longer period of time.

Note that the difference in the $MNRE$ between the two systems is not simply that in the mixed system more years are used for the calculation. The implicit connection between contributions and subsequent pension benefits is different. The mixed system pays a *compounded* defined contribution payoff on every early peso contributed. Even if the PAYG-DB system took all periods into account, this compounding effect would still dominate incentives early on. It is straightforward to use our formula for the difference in the $MNRE$ across systems for a *hypothetical* (HDB) case in which the DB system takes into account all years:

$$MNRE^M(t) - MNRE^{HDB}(t) = \gamma\tau^{SSC}e^{r(R-t)} + \left(\frac{T-R}{R}\left(\rho^M(R) - \rho^{HDB}(R)\right)\right)$$

The first term that reflects compounding of contributions is still strictly positive and larger for early periods, while the second term is negative (because $\rho^M(R) - \rho^{HDB}$) but constant. In addition, the longer the window for benefit calculation, the less each individual period

of earnings matters for benefit calculation, which reduces the incentive to report earnings related to subsequent pension benefits.

In reality, the degree to which the present discounted value of pension benefits responds to an increase in earnings depends on parameters such as interest rates, annuitization factors, and statutory replacement rates. We expand our simple conceptual discussion with more realistic numerical simulations. We assume that the parameters of the institutional setting are the same as they were at the time of the reform and that workers expect them to remain that way. We assume a yearly return on pension savings of 6.46% (the 1996-2016 geometric average), the statutory replacement rates and taxes on wages from the relevant legislation, and the annuitization factors from actuarial tables. For simplicity, we assume constant earnings trajectories for workers based on administrative records.[10] For simulations for the mixed system, we separately simulate the default option and the Article 8 option, and then we calculate a weighted average assuming that 75% of workers in the mixed system choose the article 8 option (as documented in CESS, 2021). Finally, we assume that future flows are discounted at an annual risk-free rate of 2.5% and a maximum lifespan of 90 years.

Panel (a) of figure 1 shows that simulated replacement rates are indeed significantly more responsive to changes in the retirement age for workers in the mixed system than for workers in the PAYG-DB system. Panel (b) shows the simulated percentage increase in the net return to earnings from a the mixed system relative to the PAYG-system at each working age given two different retirement ages (60 and 64). With the mixed system representing about a 20% increase in the marginal net return to earnings in the early years that gradually fades and becomes about a 10% decrease in the net return to earnings once the worker enters the window for DB pension calculation.

4. Data

Social security records Our main data source is administrative records from Uruguay's Social Security Agency (*Banco de Previsión Social*), based on matched employer-employee payroll tax filings from 1997 to 2013. These records cover the universe of formal workers with positive reported earnings for at least one month and include monthly data on gross earnings, hours and days worked, firm identifiers, sector of employment, and whether the worker is a firm owner.[11] They also provide exact dates of birth, which serve as our running variable for treatment assignment. We supplement this with a second dataset—a random sample

[10] We take the earnings of workers born before the cutoff, so they were left in the PAYG-DB system by default.

[11] Ownership is indicated by whether a person is listed as an owner, partner, director, or administrator.

of 80,000 individuals—containing birth dates and assigned pension systems. Although we cannot merge this sample with the labor histories due to different identifiers, we use it to estimate a first stage. From the labor histories, we construct key employment variables. Employment is a binary indicator equal to one if the worker has positive earnings in a given period and zero otherwise (including unemployment, inactivity, retirement, or informal work). Labor earnings include regular wages, 13th salary (paid in July and December), and severance payments. The final dataset is a worker-level panel from 1997 to 2013.

Panel A of Table 2 presents summary statistics for workers born between 1955 and 1957—covering cohorts just before, during, and after the reform. Employment is defined as having positive earnings; labor earnings are reported in current pesos and winsorized at the 1st and 99th percentiles. We also report monthly averages for hours and days worked. Public sector employment and high-informality sector indicators are binary variables based on administrative and household survey data, respectively.

Census data Our second main data source is the 2011 Population and Household Census conducted by Uruguay's National Institute of Statistics (*Instituto Nacional de Estadística*, INE). This nationwide in-person survey, carried out between September and December 2011, collected standard socio-demographic information such as age, gender, education, ethnicity, health problems, and family structure (INE, 2012). It also includes limited labor market data, namely whether individuals are currently employed or retired, but lacks information on earnings or employment formality. Birth dates are reported at the monthly level, which we use as the running variable in our analysis. The final census dataset is a cross-section of individuals surveyed in 2011.

Panel B of Table 2 presents summary statistics for individuals born between 1955 and 1957. Among them, 67.6% report being employed, 16.2% retired, and 5.7% report poor health (defined as moderate or greater difficulty in vision, hearing, mobility, or cognition). We also summarize characteristics such as marital status, college completion, gender, parenthood, and a socioeconomic status index. The index is constructed via principal component analysis using indicators such as homeownership, higher education, appliance ownership, and internet access (see appendix F).

5. Econometric strategy

The cohort-based nature of the pension reform provides a natural setting for a Regression Discontinuity Design (RDD), comparing individuals born just before and after April 1st, 1956—the cutoff for assignment to the new mixed system. Our baseline specification

estimates:

$$Y_i = \alpha + \beta \mathbb{1}\{DOB_i > c\} + f(DOB_i) + \varepsilon_i \tag{1}$$

where Y_i represents any of our outcomes of interest (employment, earnings) for individual i, DOB_i is the individual's date of birth (at the daily level), c is the cutoff date of birth (April 1st 1956), and $f(DOB_i)$ is a polynomial based on the date of birth. Given that individuals born after the cutoff date of birth were assigned the new mixed system, the coefficient β measures the Intention-to-Treat (ITT) effect of being switched from the unfunded DB government system to the mixed system that includes a capitalization element. In our baseline specifications, we pool years to reduce the number of figures but results are similar for individual years.[12]

For our baseline specification, we estimate a local linear polynomial for workers born within 60 days of the cutoff date of birth, and we allow the polynomial to vary along both sides of the discontinuity. We also estimate a quadratic polynomial based on optimal bandwidths from Calonico et al. (2014). Given that the running variable is discrete, traditional polynomial-based RDD methods may violate smoothness assumptions (Kolesár and Rothe, 2018). Therefore, we also use the Local Randomization approach (Cattaneo et al., 2019), treating individuals within a narrow 20-day window around the cutoff as randomly assigned and estimate a simple difference in means. When pooling across years, we include year fixed effects and cluster standard errors at the individual level. For census data–where birth dates are monthly–we use a one-month window, comparing those born in March to April.[13]

Figure 2 shows the first stage using a separate random sample with pension system data. We find a sharp discontinuity at the cutoff, with an RD estimate of 0.841, indicating near-perfect compliance. Among those defaulted to the DB system, about 20% are in the mixed system—mostly due to voluntary opt-ins or lack of prior formal employment. While we cannot merge this subsample with the labor histories, the strong first stage confirms that the cohort-based assignment had substantial impact, supporting the validity of our ITT estimates.

6. Main results

Employment responses We analyze employment responses by estimating equation 1 using employment as the dependent variable and estimating separate RD coefficients for differ-

[12] Yearly RD plots for the main results are available upon request.

[13] April 1st births technically remained in the old system, so our census estimates may understate the ITT effect.

ent year groups from 1997 to 2013. Figure 3 shows RD plots using a 60-day window around the cutoff, with each panel corresponding to a group of years. As shown in panel (f), there is no significant effect in earlier years, but starting in 2012—when affected individuals are nearing retirement age—those in the new system are about 4 percentage points more likely to be employed. This suggests that the introduction of capitalization encouraged workers to remain formally employed longer.

We complement these findings using the 2011 Census, which captures both formal and informal employment and directly asks whether individuals are retired. Since census birth dates are at the monthly level, we compare individuals born in April and March 1956. Figure 4, panel (a), shows that workers in the new system are 2.5 percentage points more likely to be employed, while panel (b) shows they are 2.2 percentage points less likely to be retired. This aligns with our conceptual discussion from Section 3, as capitalization creates incentives to delay retirement by linking benefits to accumulated savings and retirement duration.

We also explore heterogeneity using two census variables that are predictive of early retirement: poor health status and socioeconomic status (SES). Poor health is defined as reporting at least moderate difficulty in eyesight, hearing, movement, or cognition. As shown in panels (c) and (d), workers with poor health are 10 points more likely to be employed and 10 points less likely to be retired if in the new system. Panels (e) and (f) split results by SES, with smaller but still notable differences across groups. These patterns suggest that capitalization particularly affects groups more prone to early retirement—those with lower SES or mild health problems.

Reported health problems in the census typically do not qualify for retirement via permanent disability, which requires being completely unable to work.[14] Thus, the observed effects likely reflect regular early retirement behavior. Because early retirement reduces pension wealth more sharply in a capitalization system, these findings indicate that marginal workers—especially lower SES or poor-health individuals—respond to these financial incentives by postponing retirement.[15]

Earnings responses We examine earnings responses by estimating equation 1 using the log of total labor earnings as the dependent variable, estimating separate RD coefficients for

[14] In disability retirements, DB benefits are tax-funded in both systems. In the mixed system, the DC portion is covered by mandatory insurance or can be withdrawn as a lump sum.

[15] One concern is that some individuals born before the cutoff are in the mixed system due to voluntary opt-in or lack of prior formal employment. If these individuals are systematically different—e.g., more likely to be high SES or have poor health—this could bias heterogeneity estimates. However, such bias is limited: our first stage (Figure 2) shows only 18% of pre-cutoff individuals are in the mixed system. In the worst-case scenario—if all of them were high SES—this could bias interaction coefficients by at most 36% (18/50). The same logic applies to poor health-related heterogeneity.

different year groups from 1997 to 2013. Figure 5 displays RD plots using a 60-day window around the cutoff. As shown in panel (f), workers in the new mixed system earned significantly more in the years immediately after the reform—about 20% higher salaries—though this effect diminishes over time and becomes insignificant by 2012–2013.

This pattern aligns with the conceptual discussion in Section 3: in the mixed system, early earnings are linked to subsequent pensions via retirement savings, providing incentives to report higher earnings early. In contrast, workers in the unfunded DB system only see pension benefits tied to the last 10 years of earnings, making earlier contributions potentially seem more like a tax. These aspects manifest in a much higher net-return-to-earnings in the mixed system relative to the PAYG-DB system.

To assess whether this earnings increase reflects real labor supply or reduced underreporting, we examine complementary evidence. First, we analyze days and hours worked as proxies for real effort.[16] Panels (a) and (b) of Figure 6 show no significant differences across systems, suggesting the higher earnings are not driven by higher labor supply. Second, we compare private and public sector workers. Since underreporting is rare in the public sector, any earnings effect there would likely reflect real labor responses. Panel (c) shows no significant earnings differences in the public sector, further supporting the underreporting explanation. Third, we explore heterogeneity by sector-level informality. Using household survey data, we construct an informality index combining (i) the share of workers not contributing to pensions and (ii) the share who admit to underreporting income.[17] Panel (d) shows that earnings gains are concentrated in high-informality sectors, again consistent with improved reporting rather than increased work effort. Fourth, we compare employees to firm owners and the self-employed—groups known to be more likely to underreport earnings. Panel (e) shows that earnings increases are substantially larger among owners and the self-employed, who have greater flexibility to underreport. Finally, panel (f) restricts the sample to dependent employees and shows the effect on earnings split by sector-level informality, indicating that reported earnings increase for employees as well.

Another possibility for trying to assess the degree to which tax evasion drives the earnings responses is to analyze discrepancies between income and consumption (e.g. Pissarides and Weber, 1989; Lyssiotou et al., 2004; Gorodnichenko et al., 2009; Hurst et al., 2014; Engström and Hagen, 2017). Although we do not have administrative records of consumption and expenditure household surveys do not have the date of birth and would probably have insufficient observations for our regression discontinuity analysis, we can replicate our regression discontinuity analysis using census data on several measures of consumption, from our SES

[16] We show how these measures of real labor supply correlate with earnings both across and within workers in Appendix G.

[17] See appendix table D.2 for details.

index and individual elements of consumption (e.g. owning a home, a car, a phone, number of TVs, and the sort). The results of this exercise can be found in Figure 7. All of the RD coefficients for these measures are small and statistically indistinguishable from zero. Thus, despite having had careers with substantially higher reported earnings for 15 years, we do not detect any differences in measures of consumption across the two groups, further supporting that the differences in reported earnings are driven by tax evasion.

Finally, we consider why the effect fades over time. One explanation is that, as discussed in the conceptual discussion, workers in the DB system start entering the 10-year window that determines pension benefits in their fifties, increasing their incentive to report higher earnings to the point that the marginal net return to earnings is lower in the mixed system. This has been documented for self-employed workers by Dean et al. (2022). Additionally, wage-setting became less flexible after the reintroduction of collective bargaining in 2005 (Mazzuchi, 2009), and income underreporting steadily declined over the period due to economic growth, income tax reforms, and enforcement efforts (Bergolo et al., 2021). Lastly, as shown earlier, low-SES and poor-health workers are more likely to remain employed under the new system, introducing selection that may also explain the fading earnings gap.

Robustness checks We conduct several robustness checks to validate our empirical strategy. A common concern with Regression Discontinuity Designs is manipulation of the running variable. In our case, this would imply individuals altering their date of birth to avoid assignment to the new mixed system. However, dates of birth come from official social security records based on birth certificates, making manipulation implausible. Panel (a) of Appendix Figure E.1 shows no visual evidence of bunching around the cutoff, and panel (b) confirms this using the local polynomial manipulation test of Cattaneo et al. (2020), yielding a p-value of 0.55.

Another concern is whether the April 1, 1956 cutoff may have introduced systematic differences in unobserved characteristics—e.g., if individuals born just after the cutoff were inherently different due to seasonality effects (Buckles and Hungerman, 2013). However, the reform's cohort-based design was not related to observable or unobservable traits (Forteza and Rossi, 2018). While pre-reform labor data is unavailable, we conduct placebo RD tests for cohorts born in 1955 and 1957, finding no significant effects in employment or earnings (Figure E.7). Additionally, Table D.1 confirms balance in observable characteristics across the cutoff using census data, supporting the validity of our design.

We also test robustness to alternative bandwidths around the cutoff. Figure E.8 shows that our RD estimates are robust to multiple alternative bandwidths both for employment rates and earnings. Given that our running variable is discrete, smoothness assumptions for

RD designs may not hold (Kolesár and Rothe, 2018). We show that our results are robust to using the local randomization approach (Cattaneo et al., 2016), treating observations in a narrow window around the cutoff as randomly assigned. To minimize seasonality bias (e.g., variation in educational attainment by birth month; see Appendix Figure E.11), we focus on a narrow window.[18] Figure E.9 shows that our baseline results are very similar to this approach, and closely match results from a standard continuity-based RD using a quadratic polynomial with optimal bandwidth from Cattaneo et al. (2016).

A further concern is that only workers in the mixed system face a contribution ceiling, which could mechanically raise earnings if contributions become voluntary above a threshold. This is unlikely to drive results for three reasons: (i) the ceiling lies at the 98th percentile of the wage distribution, (ii) we top-code earnings at the 99th percentile, and (iii) contributions above the ceiling for some workers are deducted and then workers can decide whether to get them reimbursed. As an additional check, we re-estimate earnings effects excluding all workers earning above the ceiling—an overly conservative restriction—and find similar results (Appendix Figure E.10).

Back-of-the-envelope calculations We conclude with some simple back-of-the-envelope calculations to approximate the elasticity of earnings with respect to the net return to earnings when young. This can serve as an analogue to the elasticity of reported earnings with respect to the net-of-tax rate of studies of the effect of tax rate changes on reported earnings. Taking our simulations from section 3, the reform increased the marginal net return to earnings by about 20 to 25 percent relative to the PAYG-DB system, depending on the expected retirement age. Taking our estimated earnings responses of an increase of about 15 to 20 percent during the early years of the reform, this yields an elasticity of reported earnings relative to the net return to earnings of around 0.72 to 0.9.

7. Conclusion and discussion

This paper examines the effects of pension privatization on worker behavior, leveraging a 1996 reform in Uruguay that replaced a pay-as-you-go defined benefit system with a mixed system that introduced a funded defined contribution component. We exploit a cohort-based discontinuity for identification using a Regression Discontinuity Design, and high-quality administrative and census data to track long-run outcomes.

We first analyze behavioral responses to the reform. Workers in the mixed system show

[18] Our baseline window is of 11 days around the cutoff, yielding 2,000 observations. Results are robust to narrower and wider windows as well.

similar employment rates early in life but are significantly more likely to remain formally employed closer to retirement—especially those with lower socioeconomic status or mild health problems. Earnings rise sharply after the reform among workers in the new system, particularly in the early years, but this gap fades with age. Evidence suggests these earnings increases reflect reduced underreporting rather than higher labor supply.

Given the global relevance of pension privatization, our findings carry important policy implications. Workers respond to incentives decades before retirement, suggesting that adding a DC component can reduce income underreporting and delay retirement—both desirable outcomes for fiscal sustainability. However, the stronger effects among disadvantaged groups raise concerns about regressivity. The complexity of pension privatization precludes from perfectly isolating the individual components that affect labor supply beyond the change in benefit formulas. Our findings suggest that workers indeed respond to incentives in the switch to a DC formula from a privatization, although the contribution of elements such as understanding of the system, risk aversion, trust in financial markets, and related factors is unclear. These present several productive avenues for future research.

References

Attanasio, O. P. and Brugiavini, A.: 2003, Social security and households' saving, *the Quarterly Journal of economics* **118**(3), 1075–1119.

Attanasio, O. P. and Rohwedder, S.: 2003, Pension wealth and household saving: Evidence from pension reforms in the united kingdom, *American Economic Review* **93**(5), 1499–1521.

Auerbach, A. J. and Kotlikoff, L. J.: 1985, The efficiency gains from social security benefit-tax linkage, *NBER Working Paper* (w1645).

Bachas, P. and Soto, M.: 2021, Corporate taxation under weak enforcement, *American Economic Journal: Economic Policy* **13**(4), 36–71.

Bergolo, M., Burdin, G., De Rosa, M., Giaccobasso, M. and Leites, M.: 2021, Digging into the channels of bunching: Evidence from the uruguayan income tax, *The Economic Journal* **131**(639), 2726–2762.

Bergolo, M. and Cruces, G.: 2014, Work and tax evasion incentive effects of social insurance programs: Evidence from an employment-based benefit extension, *Journal of Public Economics* **117**, 211–228.

Bjørneby, M., Alstadsæter, A. and Telle, K.: 2021, Limits to third-party reporting: Evidence from a randomized field experiment in norway, *Journal of Public Economics* **203**, 104512.

Brockmeyer, A. and Sáenz Somarriba, M.: 2025, Electronic payment technology and tax compliance: Evidence from uruguay's financial inclusion reform, *American Economic Journal: Economic Policy* **17**(1), 242–272.

Brown, K. M.: 2013, The link between pensions and retirement timing: Lessons from california teachers, *Journal of Public Economics* **98**, 1–14.

Buckles, K. S. and Hungerman, D. M.: 2013, Season of birth and later outcomes: Old questions, new answers, *Review of Economics and Statistics* **95**(3), 711–724.

Calonico, S., Cattaneo, M. D. and Titiunik, R.: 2014, Robust nonparametric confidence intervals for regression-discontinuity designs, *Econometrica* **82**(6), 2295–2326.

Cattaneo, M. D., Idrobo, N. and Titiunik, R.: 2019, *A practical introduction to regression discontinuity designs: Foundations*, Cambridge University Press.

Cattaneo, M. D., Jansson, M. and Ma, X.: 2018, Manipulation testing based on density discontinuity, *The Stata Journal* **18**(1), 234–261.

Cattaneo, M. D., Jansson, M. and Ma, X.: 2020, Simple local polynomial density estimators, *Journal of the American Statistical Association* **115**(531), 1449–1455.

Cattaneo, M. D., Titiunik, R. and Vazquez-Bare, G.: 2016, Inference in regression discontinuity designs under local randomization, *The Stata Journal* **16**(2), 331–367.

CESS: 2021, Diagnóstico del sistema previsional uruguayo, *Comisión de Expertos en Seguridad Social* .

Coile, C. and Gruber, J.: 2007, Future social security entitlements and the retirement decision, *The review of Economics and Statistics* **89**(2), 234–246.

Dean, A., Fleitas, S. and Zerpa, M.: 2022, Dynamic Incentives in Retirement Earnings-Replacement Benefits, *The Review of Economics and Statistics* pp. 1–45.
URL: *https://doi.org/10.1162/rest_a_01193*

Edwards, S. and Edwards, A. C.: 2002, Social security privatization reform and labor markets: The case of chile, *Economic Development and Cultural Change* **50**(3), 465–489.

Engström, P. and Hagen, J.: 2017, Income underreporting among the self-employed: A permanent income approach, *European Economic Review* **92**, 92–109.

Feinmann, J., Lauletta, M. and Hsu, R.: 2024, Employer-employee collusion and payments under the table: Evidence from brazil, *University of California, Berkeley* .

Feldstein, M. S.: 1995, Would privatizing social security raise economic welfare?, *NBER Working paper* (w5281).

Fetter, D. K. and Lockwood, L. M.: 2018, Government old-age support and labor supply: Evidence from the old age assistance program, *American Economic Review* **108**(8), 2174–2211.

Filmer, D. and Pritchett, L. H.: 2001, Estimating wealth effects without expenditure data—or tears: an application to educational enrollments in states of india, *Demography* **38**(1), 115–132.

Forteza, A. and Rossi, I.: 2018, Ganadores y perdedores en las primeras generaciones luego de una reforma estructural de la seguridad social: el caso de uruguay, *Apuntes* **45**(82), 99–119.

French, E., Lindner, A., O'Dea, C. and Zawisza, T.: 2022, Labor supply and the pension contribution-benefit link, *Technical report*, Working paper.

Gavoille, N. and Zasova, A.: 2023, What we pay in the shadows: Labor tax evasion, minimum wage hike and employment, *Journal of Public Economics* **228**, 105027.

Gelber, A. M., Isen, A. and Song, J.: 2016, The effect of pension income on elderly earnings: Evidence from social security and full population data, *NBER Working Paper* .

Goda, G. S., Levy, M., Manchester, C. F., Sojourner, A. and Tasoff, J.: 2019, Predicting retirement savings using survey measures of exponential-growth bias and present bias, *Economic Inquiry* **57**(3), 1636–1658.

Gorodnichenko, Y., Martinez-Vazquez, J. and Sabirianova Peter, K.: 2009, Myth and reality of flat tax reform: Micro estimates of tax evasion response and welfare effects in russia, *Journal of Political economy* **117**(3), 504–554.

Hernæs, E., Markussen, S., Piggott, J. and Røed, K.: 2016, Pension reform and labor supply, *Journal of Public Economics* **142**, 39–55.

Hosseini, R. and Shourideh, A.: 2019, Retirement financing: An optimal reform approach, *Econometrica* **87**(4), 1205–1265.

Hurst, E., Li, G. and Pugsley, B.: 2014, Are household surveys like tax forms? evidence from income underreporting of the self-employed, *Review of economics and statistics* **96**(1), 19–33.

INE: 2012, Resumen operativo censos 2011, *Instituto Nacional de Estadística* .

Kolesár, M. and Rothe, C.: 2018, Inference in regression discontinuity designs with a discrete running variable, *American Economic Review* **108**(8), 2277–2304.
URL: *https://www.aeaweb.org/articles?id=10.1257/aer.20160945*

Kotlikoff, L. J.: 1996, Privatization of social security: how it works and why it matters, *Tax policy and the economy* **10**, 1–32.

Kumler, T., Verhoogen, E. and Frías, J.: 2020, Enlisting employees in improving payroll tax compliance: Evidence from mexico, *Review of Economics and Statistics* **102**(5), 881–896.

Lachowska, M. and Myck, M.: 2018, The effect of public pension wealth on saving and expenditure, *American Economic Journal: Economic Policy* **10**(3), 284–308.

Leijten, F. R., de Wind, A., van den Heuvel, S. G., Ybema, J. F., van der Beek, A. J., Robroek, S. J. and Burdorf, A.: 2015, The influence of chronic health problems and work-related factors on loss of paid employment among older workers, *J Epidemiol Community Health* **69**(11), 1058–1065.

Liebman, J. B. and Luttmer, E. F.: 2012, The perception of social security incentives for labor supply and retirement: The median voter knows more than you'd think, *Tax Policy and the Economy* **26**(1), 1–42.

Liebman, J. B. and Luttmer, E. F.: 2015, Would people behave differently if they better understood social security? evidence from a field experiment, *American Economic Journal: Economic Policy* **7**(1), 275–299.

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

Lindbeck, A. and Persson, M.: 2003, The gains from pension reform, *Journal of Economic Literature* **41**(1), 74–112.

Londoño-Vélez, J. and Ávila-Mahecha, J.: 2021, Enforcing wealth taxes in the developing world: Quasi-experimental evidence from colombia, *American Economic Review: Insights* **3**(2), 131–48.

Lyssiotou, P., Pashardes, P. and Stengos, T.: 2004, Estimates of the black economy based on consumer demand approaches, *The Economic Journal* **114**(497), 622–640.

Manoli, D. and Weber, A.: 2016, Nonparametric evidence on the effects of financial incentives on retirement decisions, *American Economic Journal: Economic Policy* **8**(4), 160–82.

Mazzuchi, G.: 2009, Labour relations in uruguay: 2005–08, *Industrial and EmploymentRelations Department, Working Paper* **6**.

McKenzie, D. J.: 2005, Measuring inequality with asset indicators, *Journal of population economics* **18**(2), 229–260.

Naritomi, J.: 2019, Consumers as tax auditors, *American Economic Review* **109**(9), 3031–72.

Nishiyama, S. and Smetters, K.: 2007, Does social security privatization produce efficiency gains?, *The Quarterly Journal of Economics* **122**(4), 1677–1719.

Orenstein, M. A.: 2013, Pension privatization: Evolution of a paradigm, *Governance* **26**(2), 259–281.

Pissarides, C. A. and Weber, G.: 1989, An expenditure-based estimate of britain's black economy, *Journal of public economics* **39**(1), 17–32.

Pomeranz, D.: 2015, No taxation without information: Deterrence and self-enforcement in the value added tax, *American Economic Review* **105**(8), 2539–69.

Saez, E.: 2010, Do taxpayers bunch at kink points?, *American economic Journal: economic policy* **2**(3), 180–212.

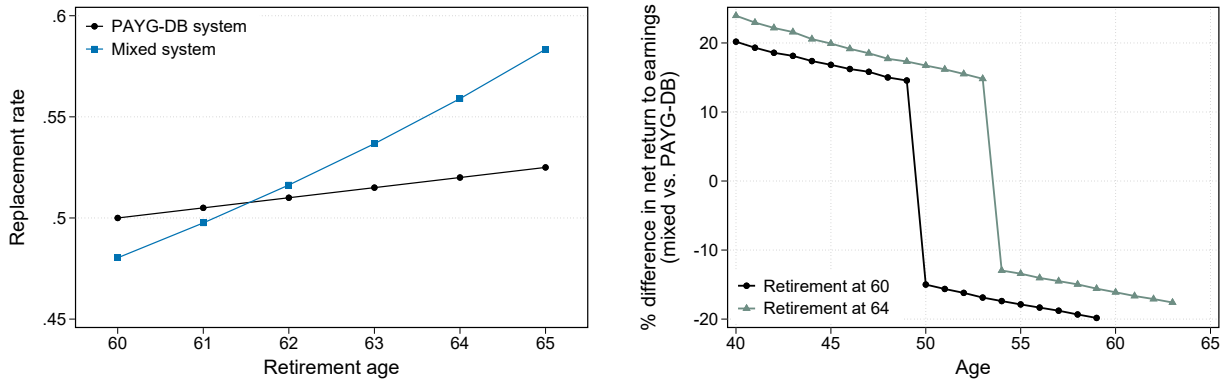
Stewart, F. and Yermo, J.: 2009, Pensions in africa.

Van Rijn, R. M., Robroek, S. J., Brouwer, S. and Burdorf, A.: 2014, Influence of poor health on exit from paid employment: a systematic review, *Occupational and environmental medicine* **71**(4), 295–301.

Figures

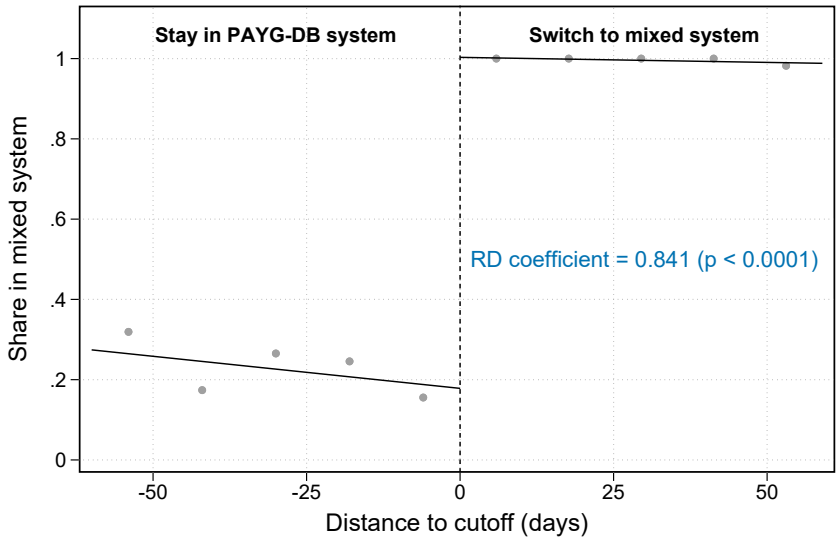
Figure 1: Reform incentives simulations

(a) Replacement rate comparison across systems **(b)** Change in net marginal return to earnings



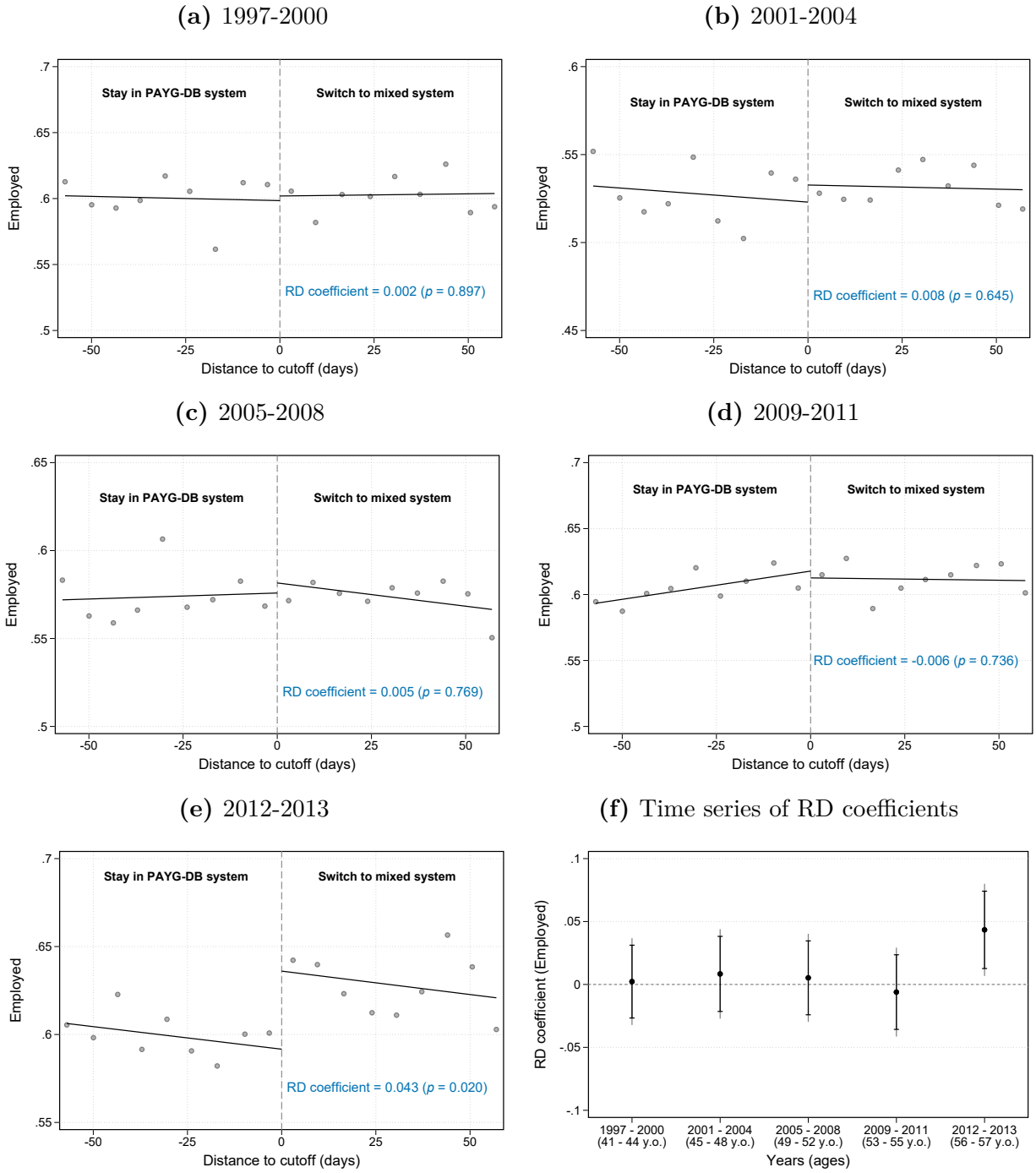
Notes: Panel (a) shows average replacement rates by retirement age from numerical simulations. Panel (b) shows the percentage difference in the simulated marginal net return to earnings of the mixed system relative to the PAYG-DB system by age. The black line corresponds to a worker that retires at 60 and the green line corresponds to a worker that retires at 64.

Figure 2: First stage



Notes: Panel (a) shows the empirical first stage: the share in the new mixed pension system by equal-sized bins around the cutoff birth date. Those born before the cutoff remained in the PAYG-DB system; those born at or after switched to the mixed system. The dependent variable is an indicator for being in the mixed system. p is the p-value for no difference across groups.

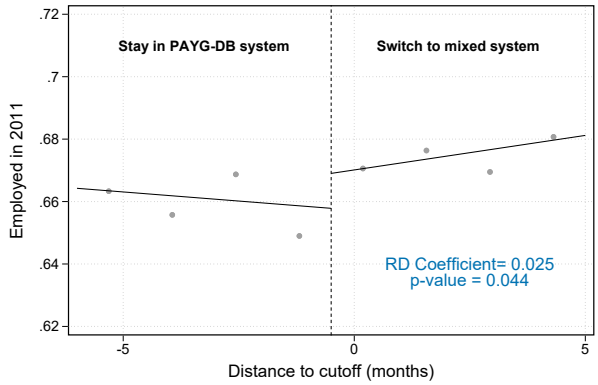
Figure 3: Effect of the reform on employment rates - SSA data



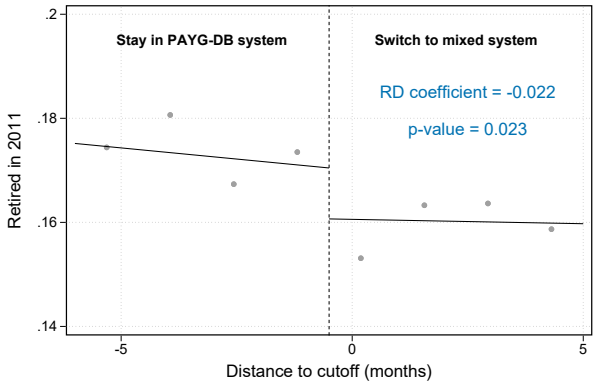
Notes: This figure shows individual RD plots and a time series plot of RD coefficients. Panels (a) through (e) plot the share of individuals who are employed by equal-sized bins of distance to the cutoff date of birth using social security data. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is an indicator of whether the worker was employed. p denotes the p-value of the null hypothesis of no difference in the outcome variable across the two groups, calculated using worker-level cluster-robust inference. The dependent variable is residualized from year fixed effects and evaluated at the mean. Panel (f) shows the time series of RD coefficients. The numbers underneath the years indicate the ages of workers in those years. Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference.

Figure 4: Effect of the reform on employment and retirement - census data (inc. heterogeneity)

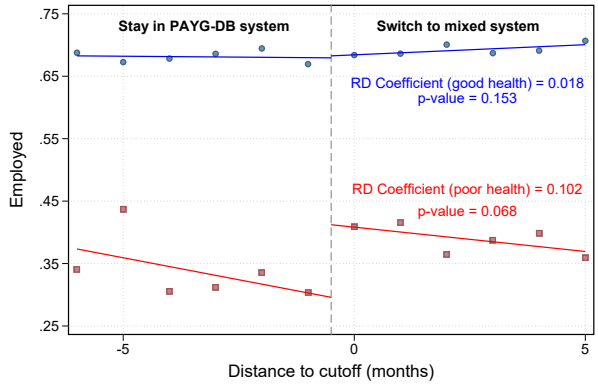
(a) Employment



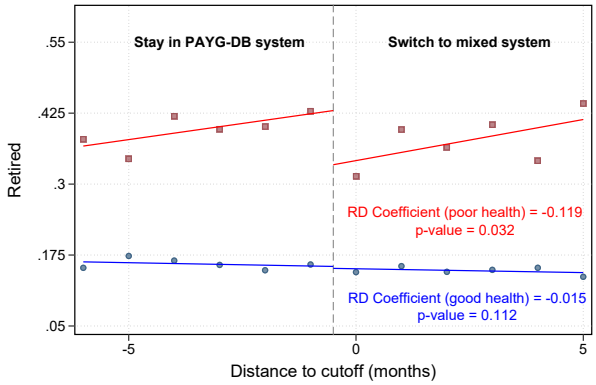
(b) Retirement



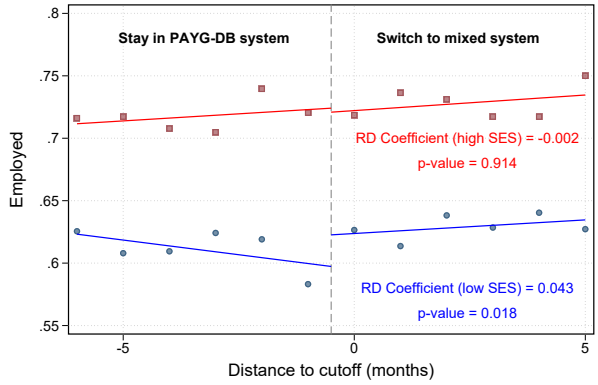
(c) Employment (by poor health)



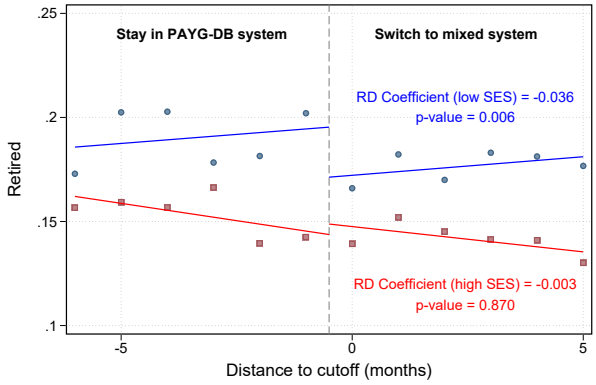
(d) Retirement (by poor health)



(e) Employment (by SES)

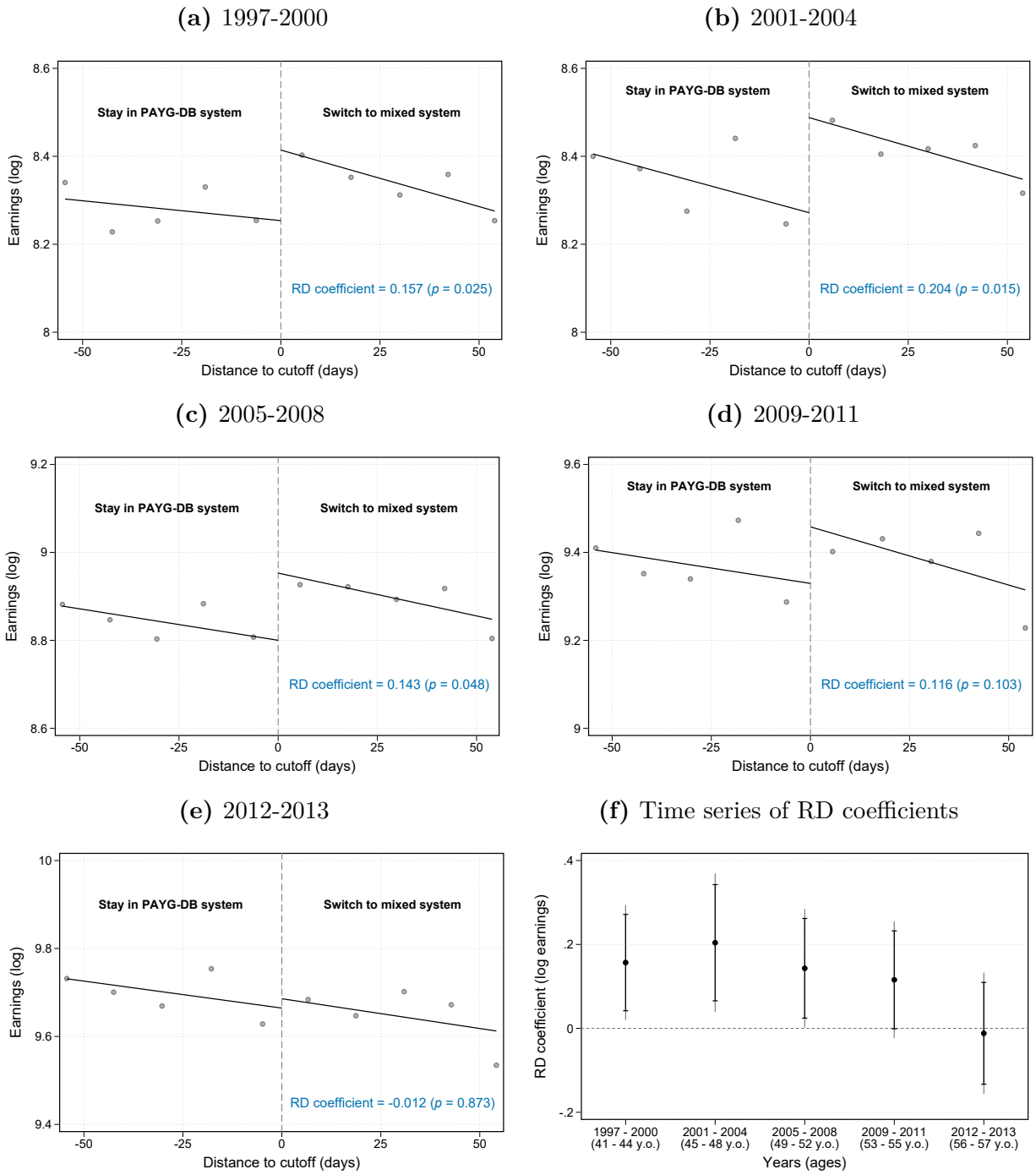


(f) Retirement (by SES)



Notes: This figure plots the share of workers that are employed by bins of distance to the cutoff date of birth using census data. Panel (a) reports the effects for the probability of being employed. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Panel (b) reports effects for the probability of being retired. Panels (c) and (e) shows the RD-plot heterogeneity for the probability of being employed and panels (d) and (f) for the probability of being retired. In panels (c) and (d) the color blue corresponds to individuals with no health problems and the color red corresponds to individuals with some health problem. In panels (d) and (f) blue corresponds to individuals with a below-median socioeconomic status index and red corresponds to individuals with an above median socioeconomic status index. RD coefficients are estimated by calculating average differences for individuals born in April and individuals born in March.

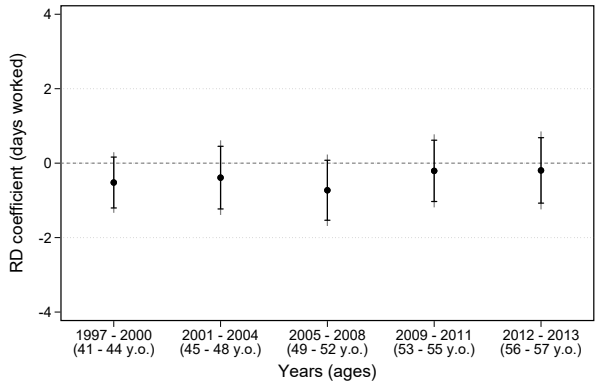
Figure 5: Effect of the reform on labor earnings - SSA data



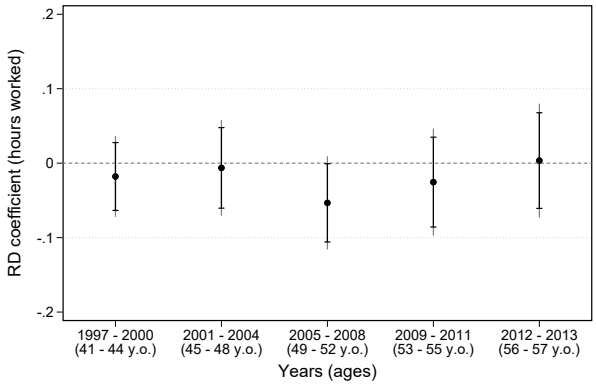
Notes: This figure shows individual RD plots and a time series plot of RD coefficients. Panels (a) through (e) plot the average log of labor earnings by equal-sized bins of distance to the cutoff date of birth using social security data. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is the natural logarithm of total reported labor earnings. p denotes the p-value of the null hypothesis of no difference in the outcome variable across the two groups, calculated using worker-level cluster-robust inference. The dependent variable is residualized from year fixed effects and evaluated at the mean. Panel (f) shows the time series of RD coefficients. The numbers underneath the years indicate the ages of workers in those years. Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference.

Figure 6: Time series plot RD coefficients (labor supply and earnings heterogeneity - SSA data)

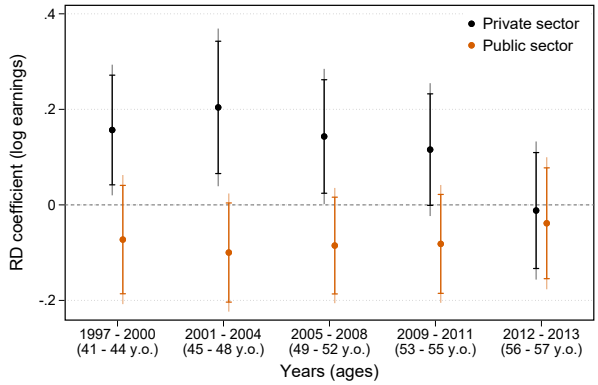
(a) Days worked



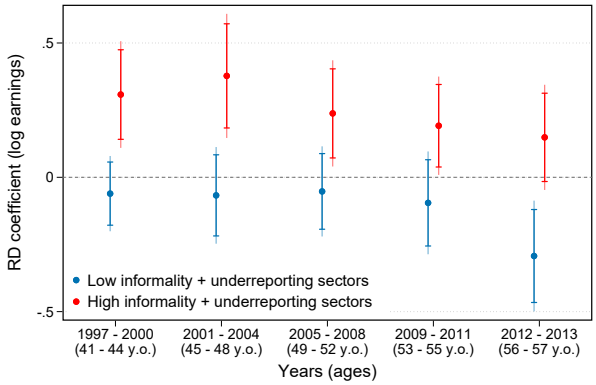
(b) Hours worked



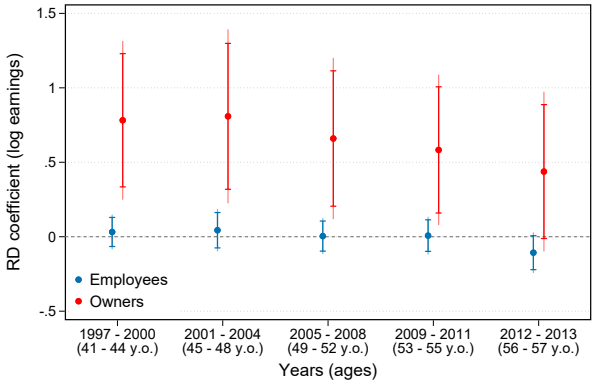
(c) Earnings - public vs private sector



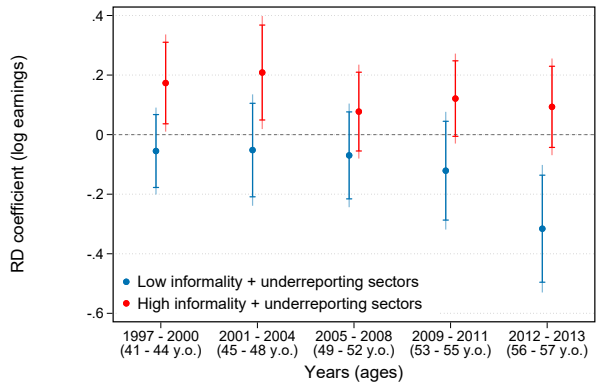
(d) Earnings - by sector informality



(e) Earnings - owner vs employees

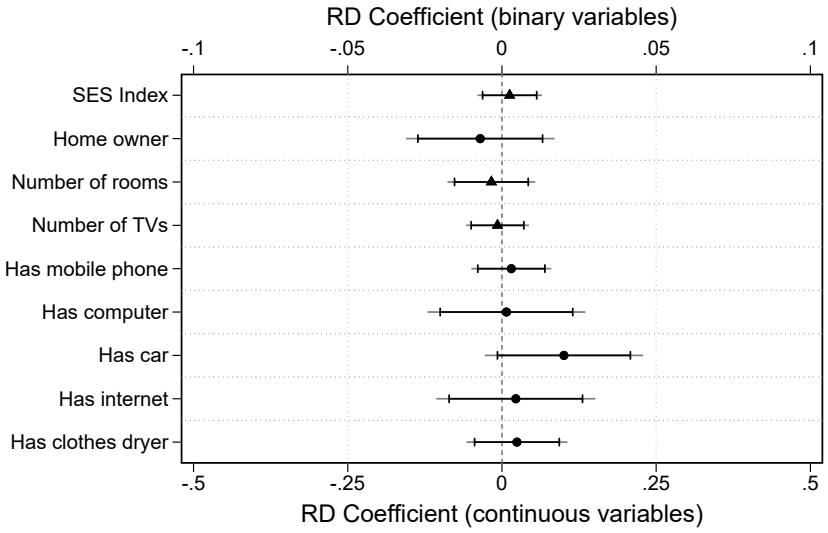


(f) Earnings - by sector informality (employees only)



Notes: This figure shows several time series plots for the RD coefficients from equation 1 using SSA data. Panel (a) shows the effect on days worked in the month. Panel (b) shows the effect on the natural logarithm of monthly hours worked. Panels (c) through (d) show effects on earnings heterogeneity. In panel (c) black coefficients correspond to private-sector workers and orange coefficients correspond to public-sector workers. In panel (d) red coefficients correspond to sectors with high levels of informality and income underreporting and blue coefficients correspond to sectors with low informality and income underreporting (see appendix table D.2). In panel (e) red coefficients correspond to firm owners and blue coefficients correspond to firm employees. In panel (f) we restrict the sample to dependent employees, with coefficients in red corresponding to sectors with high levels of informality and income underreporting and blue coefficients corresponding to sectors with low informality and income underreporting. The numbers underneath the years indicate the ages of workers in those years. Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference. The individual RD plots can be found in appendix figures E.2 through E.6.

Figure 7: RD Coefficients for measures of consumption



Notes: This figure shows RD coefficients for several measures of consumption. Coefficients for continuous variables (triangles) correspond to the bottom X axis and coefficients for binary variables (circles) correspond to the top X axis. SES index is our index of socioeconomic status. Home owner is a binary variable equal to 1 if the individual owns their home. Number of rooms is the number of rooms of the dwelling. Number of TVs is the number of television sets in the dwelling. Has mobile phone is a binary variable equal to 1 if the individual has a mobile phone. Has computer is a binary variable equal to 1 if the household has a computer. Has car is a binary variable equal to 1 if the household owns a car. Has internet is a binary variable equal to 1 if the household has internet connection. Has clothes dryer is a binary variable equal to 1 if the household has a clothes drying machine. Vertical bars represent 95% and 90% confidence intervals.

Tables

Table 1: Key Institutional Features

Feature	PAYG-DB only	Mixed system
System structure	Single pay-as-you-go, defined benefit (DB) scheme. Benefits financed entirely from current payroll taxes.	Two-pillar system: (i) public DB pension, and (ii) funded individual retirement account (defined contribution, DC) managed by a licensed pension fund.
Contribution rate	15% of gross wage to DB system.	15% of gross wage, split between DB and DC pillars depending on income thresholds and chosen option.
Choices within system	None.	Two elements of choice: (i) default contribution split or *Article 8* option; (ii) pension fund administrator for DC pillar.
Contribution base	Pension calculated using the average of last 10 years of indexed earnings.	DB component: average of last 10 years of indexed earnings; DC annuity: based on accumulated contributions made in all years.
Adjustments to benefits based on retirement age	Minor adjustments to statutory replacement rate (increases around 1-2 percent in pension benefits for an extra year of employment)	Minor adjustments to statutory replacement rate for DB part, major adjustments to annuity based on accumulated savings (increases of 6-10 percent in total pension for an extra year of employment)
Minimum normal retirement age	60 (with provisions for early retirement depending on sectors).	60 (with provisions for early retirement depending on sectors).

Notes: This table presents the main institutional features of each system. The full details of each pension system and the reform are discussed in section 2.

Table 2: Summary statistics

	Observations	Mean	Standard Deviation	Median
Panel A. Social Security data				
Employed	1552882	0.579	0.494	1.000
Total labor earnings	929,373	14858.658	19262.985	8728.500
Monthly hours worked	902,771	163.922	54.419	171.429
Days worked in the month	929,153	24.642	8.985	30.000
Public sector	893,059	0.288	0.453	0.000
Owner	922,403	0.121	0.327	0.000
High inf. sector	840,859	0.387	0.487	0.000
Panel B. Census data				
Employed	109,583	0.676	0.468	1.000
Retired	109,583	0.162	0.368	0.000
Poor health	109,575	0.057	0.231	0.000
SES Index	109,354	0.007	1.001	0.108
Married	109,584	0.674	0.469	1.000
College complete	109,828	0.224	0.417	0.000
Female	109,828	0.524	0.499	1.000
Has children	109,828	0.468	0.499	0.000

Notes: This table reports summary statistics from our main datasets. All individuals are born between 1955 and 1957. Panel A shows summary statistics from Social Security data. Employed is an indicator equal to 1 if the worker reported positive earnings in that period. Total labor earnings is the average of total monthly labor earnings. Monthly hours worked is the average monthly hours worked. Public sector is an indicator equal to 1 if the worker is employed in the public sector. Owner is an indicator equal to 1 if the worker is listed as a firm owner. High inf. sector is an indicator equal to 1 if the firm's sector is categorized as a high informality sector. Panel B shows summary statistics from census data. Employed is an indicator equal to 1 if the individual reported being employed. Retired is an indicator equal to 1 if the individual reported being retired. Poor health is an indicator equal to 1 if the individual reported experiencing at least some moderate difficulty related to eyesight, hearing, mobility, or cognitive ability. SES index is the socioeconomic status index. Married is an indicator of whether the individual is married. College complete is an indicator of whether the individual has completed college education. Has children is an indicator equal to 1 if the individual reported having at least one child.

Pension Privatization and Worker Behavior: Evidence from a Cohort-Based Reform

Appendix - For Online Publication

Maximiliano Lauletta[19] - Marcelo Bérgolo[20]

A. Additional pension system context

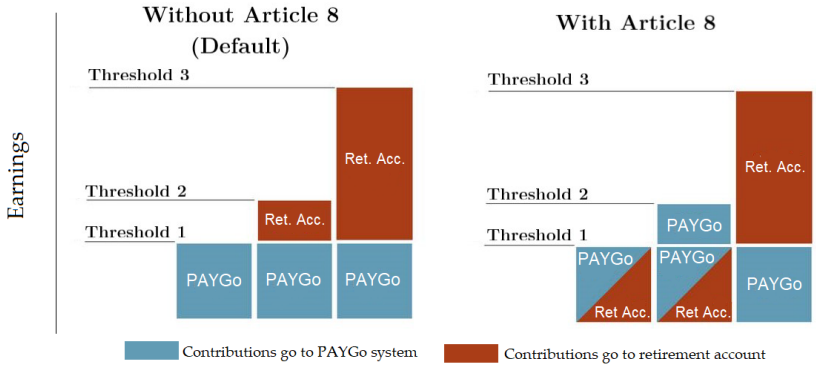
Figure A.1 shows the options on how to distribute contributions in the mixed system. The default option (without Article 8) is the option that workers are assigned by default. In it, contributions on earnings up until an earnings threshold (around the 70th percentile of the wage distribution) go entirely to the pay-as-you-go DB government system, while contributions on earnings above that threshold go entirely to the retirement account.

The alternative option (known as Article 8) allows workers whose earnings lie below the first threshold to contribute to their retirement account. For workers with earnings below threshold 1, their contributions are evenly divided between the unfunded DB system and the funded DC system. Workers whose earnings lie between thresholds 1 and 2 evenly divide contributions between the public and the private systems until threshold 1, while contributions on earnings above threshold 1 go entirely to the unfunded DB system. Finally, workers whose earnings exceed threshold 2 face the same contribution schedule as in the default option. No mandatory contributions are made on earnings above threshold 3, but workers can arrange with their employer to make those deductions and transfer them to their pension fund. The contribution rate is 15% of the pre-tax wage in all cases.

The Article 8 option implies a reduction in the government unfunded DB pension that workers will receive. This is implemented by reducing the "contributory salary" to which the replacement rate is applied by 25%. Note that this implies a subsidy for the Article 8 option: contributions to the public unfunded DB system fall by 50% but the pension received falls by 25%. This subsidy is phased-out such that the maximum government pension that a worker who chooses Article 8 can receive is the replacement rate applied to threshold 1. This is implemented by threshold 2 being set up such that a worker who chose Article 8 with earnings at threshold 2 makes the same government pension as a worker in the default option with earnings at threshold 1 and above.

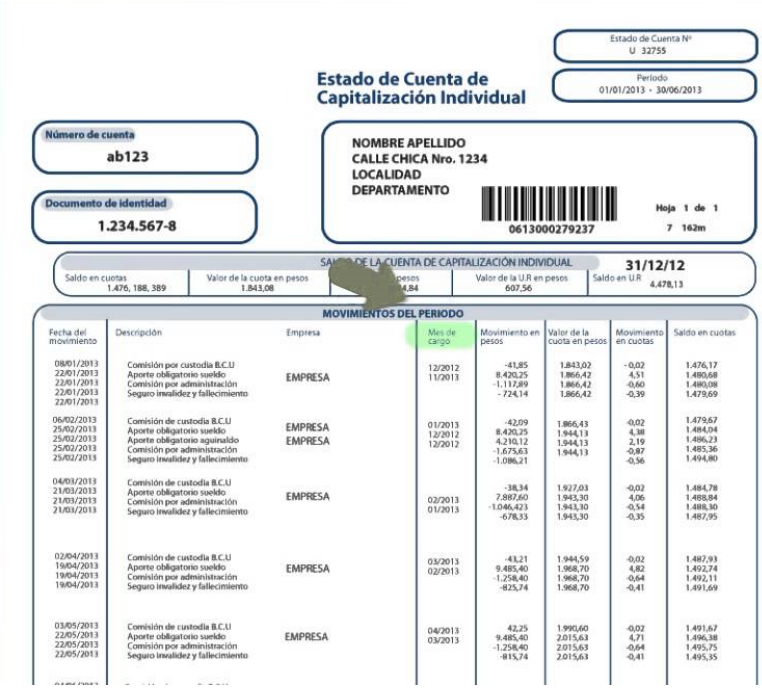
[19] Federal Reserve Board. Corresponding author. E-mail: maximiliano.lauletta@gmail.com.
[20] IECON-UDELAR & IZA. E-mail: marcelo.bergolo@fcea.edu.uy.

Figure A.1: Options in two-pillar system



Notes: This figure shows the options on how to distribute social security contributions in the two-pillar system. Contributions on earnings indicated in blue go entirely to the unfunded DB government system. Contributions on earnings indicated in red go entirely to the worker's retirement account. The default option (without Article 8) is the option that workers are assigned by default. The alternative option (with Article 8) has to be actively chosen by the workers.

Figure A.2: Example of account summary

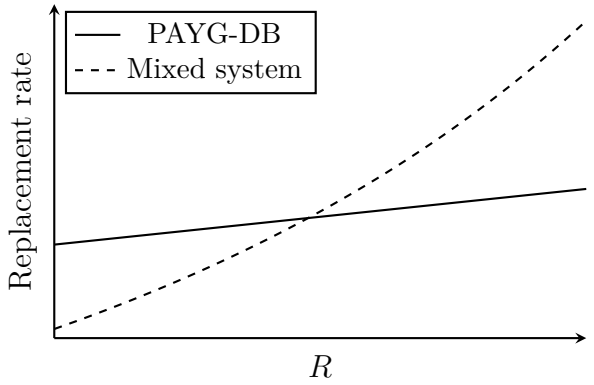


Notes: This figure shows an example of the retirement account summary that workers in the mixed system receive periodically. The document indicates the type of activity and the date in which it occurred. The types of account activities displayed are: (i) the mandatory contributions for a given month, (ii) the commission charged by the pension fund, (iii) a commission charged by the Central Bank (who regulates the pension funds), and (iv) a fee for disability and death insurance pension funds have to purchase for all workers.

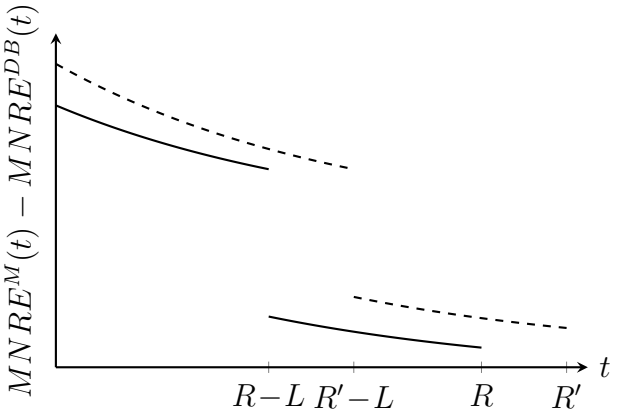
B. Conceptual framework figures

Figure B.3: Conceptual framework figures

(a) Replacement-rate comparison (conceptual)



(b) Gap in net return-to-earnings (conceptual)



Notes: Panel (a) shows replacement rates by retirement age from the conceptual framework. The solid line represents the PAYG-DB system and the dashed line represents the mixed system. Panel (b) plots the difference in marginal net return to earnings between the mixed and PAYG-DB systems by age for two retirement ages (R and R', with $R' > R$).

C. Simulations detail

In this section, we describe our simulations in more detail. We closely follow the simulations established by Forteza and Rossi (2018) specifically for the Uruguayan case. For simplicity, we assume constant earnings trajectories, but the intuitions hold for more complex cases as well. We apply our simulations to the earnings distribution of workers born as far back as March 20, 1956. Specifically, we follow the following process:

- Workers are assumed to be age 40 with 15 years of prior contributions.
- They continue contributing uninterrupted until their chosen retirement age.
- DB pensions are computed with the system's statutory replacement-rate formulas, applying any caps or floors as required.
- DC pensions are calculated with a two-step process
  1. Fund accumulation:
     - Contributions grow at the historical average nominal return of 6.46%.
     - 20% of contributions is deducted as management fees.
     - Two scenarios are simulated:
       * The worker stays in the default investment option.
       * The worker opts for the Article 8 alternative.
  2. Annuity conversion:
     - At retirement, each scenario's balance is converted to an annuity using the Central Bank's minimum payout per 1,000 pesos, which differs by retirement age.
- Mixed-system pension:
  - The overall mixed system pension equals a weighted average of the two DC scenarios:
    * 75% weight on the Article 8 option (chosen by about 75% of workers — CESS, 2021).
    * 25% weight on the default option.
- Present discounted value (PDV) of retirement benefits:
  - Workers are assumed to live to age 90.
  - Future pension payments are discounted at a real interest rate of 2.5% to obtain the PDV.

To simulate the marginal net return to earnings, we simulate the effect of increasing the wage in each year by 1 peso in the first month of the year. For the effect of this in the pension, we undertake the previous simulation process for the increased earnings and then calculate the difference in the present value of the pension benefits. We then calculate the present value of this increase at the time of the wage increase using the 2.5% interest rate. For the effect of taxes on the marginal net return to earnings, we take the tax rates relevant at the time and discount that to the additional peso earned (3% for healthcare, 6% for the personal income tax, and 15% for social security contributions). We do this for each worker and then calculate the average percentage increase in the marginal net return to earnings in the mixed system relative to the PAYG-DB system.

D. Additional tables

Table D.1: Balance - Census data

Variable	(1) Unfunded DB system	(2) Mixed system	(3) Difference
Married	0.692	0.706	0.014
	(0.462)	(0.456)	(0.012)
College complete	0.238	0.238	-0.000
	(0.426)	(0.426)	(0.011)
Has children	0.481	0.474	-0.007
	(0.500)	(0.499)	(0.013)
Poor health	0.056	0.049	-0.007
	(0.230)	(0.216)	(0.006)
SES Index	0.012	0.025	0.012
	(1.014)	(1.014)	(0.027)
Female	0.540	0.528	-0.012
	(0.499)	(0.499)	(0.013)
Observations	3,004	2,810	5,814

Notes: This table shows the balance on demographic characteristics across the originally treated and control groups using census data. Married is a dummy variable equal to 1 if the respondent is married. College complete is a dummy variable equal to 1 if the respondent has completed some college education. Has children is a dummy variable equal to 1 if the respondent reports having any children. Poor health is an indicator equal to 1 if the individual reported experiencing at least some moderate difficulty related to eyesight, hearing, mobility, or cognitive ability. SES index is the socioeconomic status index. Female is an indicator for equal to 1 if the individual reported being female. * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level.

Table D.2: Informality by sector

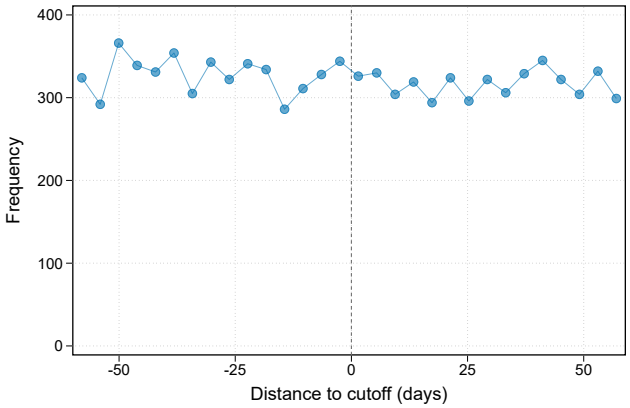
	(1) Proportion informal	(2) Proportion underreports	(3) Informality index
Panel A. Low informality sectors			
Education	0.0975	0.0430	-1.629
Financial services	0.0605	0.0591	-1.453
Social and Health services	0.137	0.0492	-1.369
Professional services	0.224	0.0557	-0.935
Water and sewage	0.0296	0.105	-0.678
Information and communication	0.189	0.0790	-0.613
Real Estate	0.233	0.0807	-0.426
Arts and entertainment	0.372	0.0571	-0.385
Mining	0.343	0.0672	-0.293
Electricity and gas	0.0303	0.126	-0.270
Panel B. High informality sectors			
Agriculture	0.312	0.124	0.686
Commerce	0.504	0.0961	0.831
Administrative support services	0.417	0.113	0.845
Hotels and Restaurants	0.416	0.135	1.259
Construction	0.607	0.106	1.395
Other services	0.678	0.120	1.909
Home services	0.635	0.199	3.265

Notes: This table reports measures of informality for each sector constructed using household surveys. Column 1 reports the proportion of workers that report being informal. Column 2 reports the proportion of formal workers that admit to underreporting their income for their contributions. Column 3 reports an index constructed as the first component of a principal component analysis of the proportion of informal workers and the proportion of workers who underreport their earnings by sector. The sample corresponds to surveys conducted in the year 2006 (the first year in which the underreporting question was included in the questionnaire).

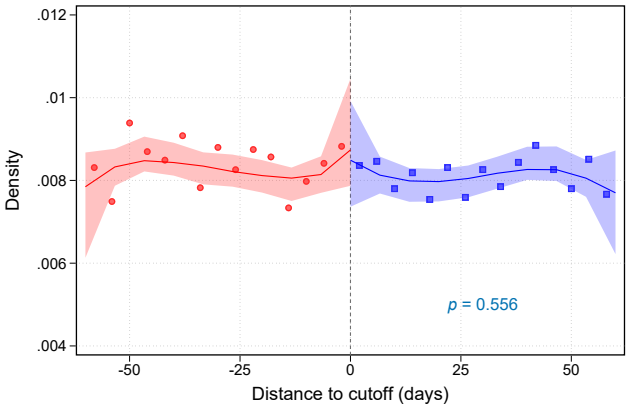
E. Additional figures

Figure E.1: Density around the cutoffs and manipulation test

(a) Density around the cutoff



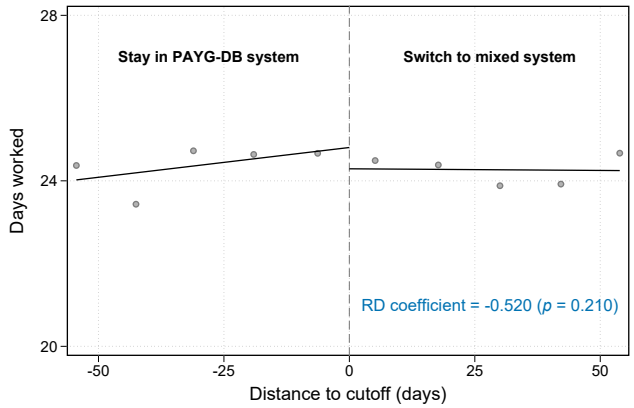
(b) Manipulation test



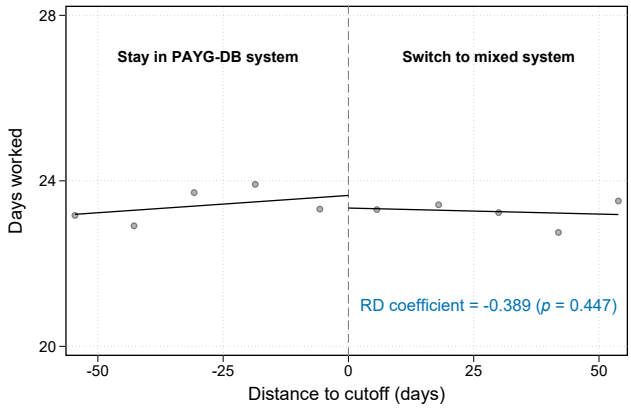
Notes: This figure shows the density of observations around the cutoff and a manipulation test for the running variable. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Panel (a) shows a frequency histogram of the number of observations in 30 equally-spaced bins and Panel (b) shows a manipulation testing plot and p-value for manipulation of the running variable based on local polynomials from (Cattaneo et al., 2020) using the `rddensity` routine from (Cattaneo et al., 2018).

Figure E.2: Effect of the reform on days worked - SSA data

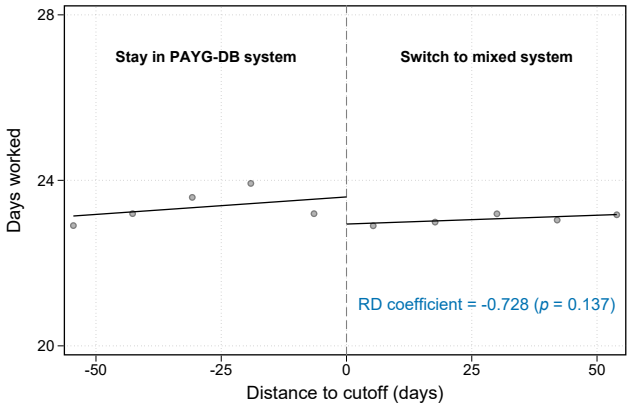
(a) 1997-2000



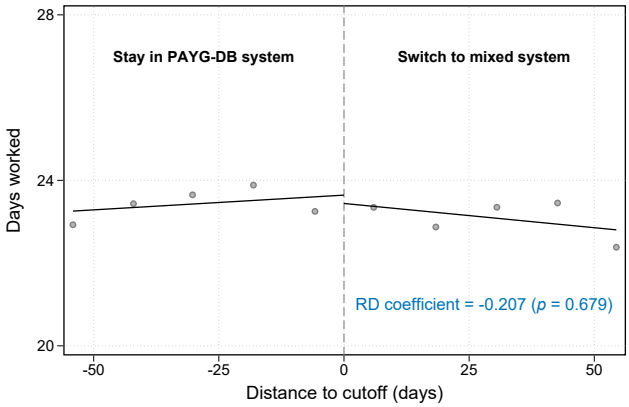
(b) 2001-2004



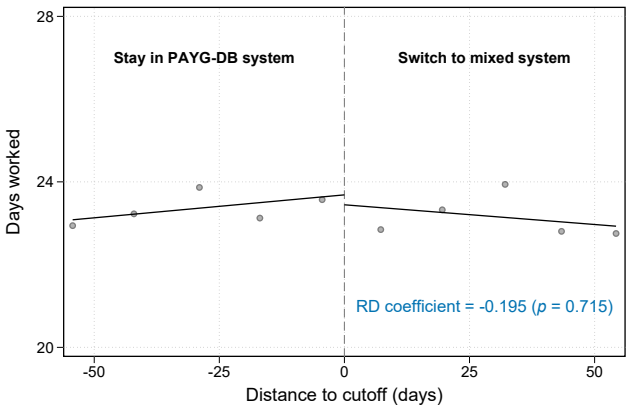
(c) 2005-2008



(d) 2009-2011



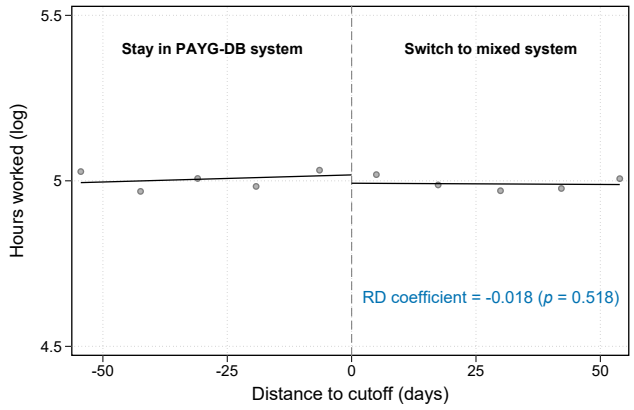
(e) 2012-2013



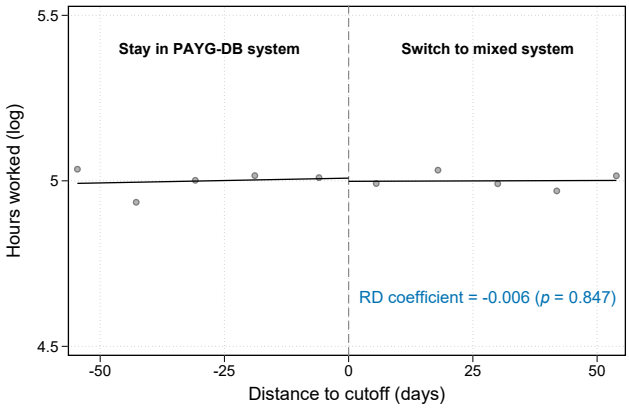
Notes: This figure shows RD-plots for the average days worked from equation 1. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is number of days worked. p-values are calculated using clustered standard errors at the worker level. The dependent variable is residualized from year fixed effects and evaluated at the mean.

Figure E.3: Effect of the reform on hours worked - SSA data

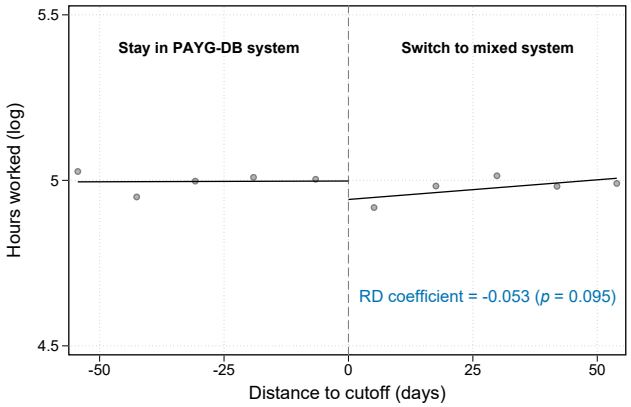
(a) 1997-2000



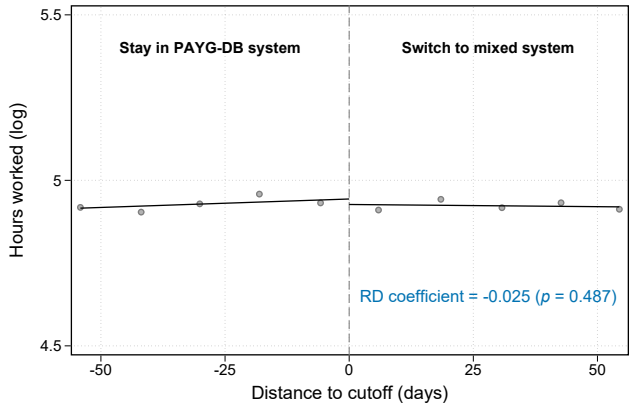
(b) 2001-2004



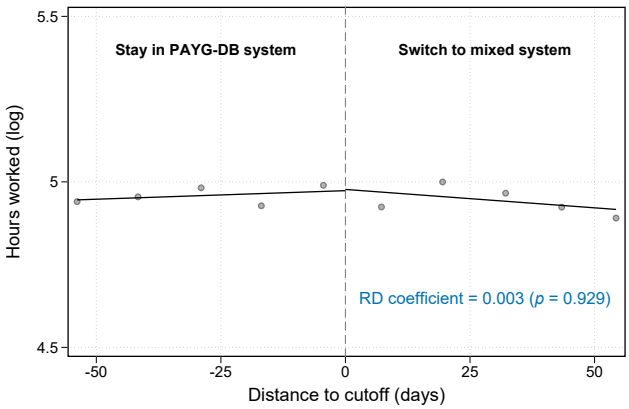
(c) 2005-2008



(d) 2009-2011



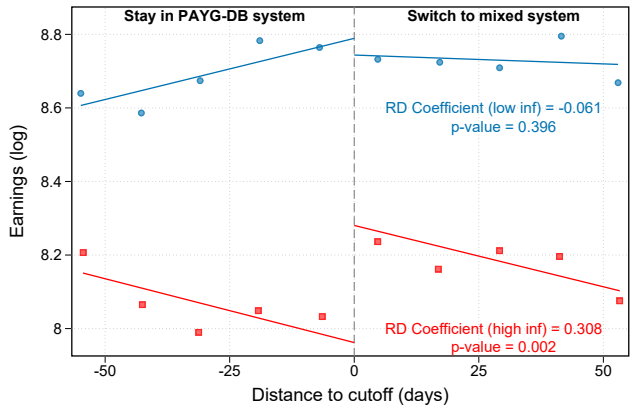
(e) 2012-2013



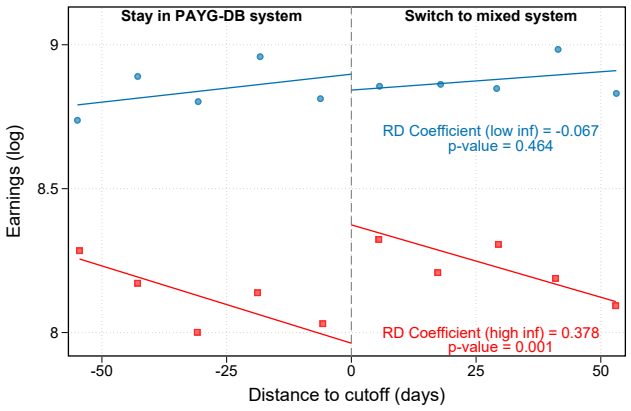
Notes: This figure shows RD-plots for the average monthly hours worked from equation 1. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is natural logarithm of the total monthly hours worked. p-values are calculated using clustered standard errors at the worker level. The dependent variable is residualized from year fixed effects and evaluated at the mean.

Figure E.4: Effect of the reform on earnings (heterogeneity by sector-level informality and evasion) - SSA data

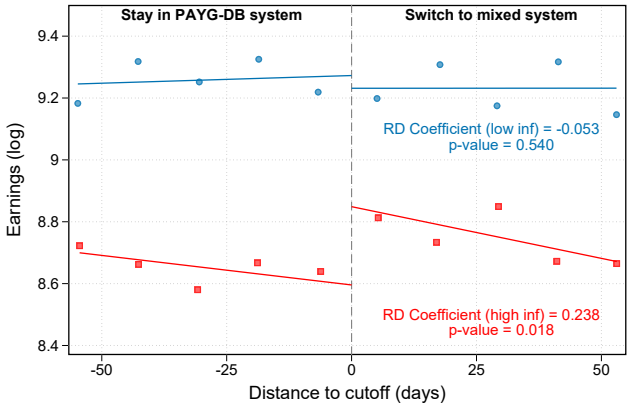
(a) 1997-2000



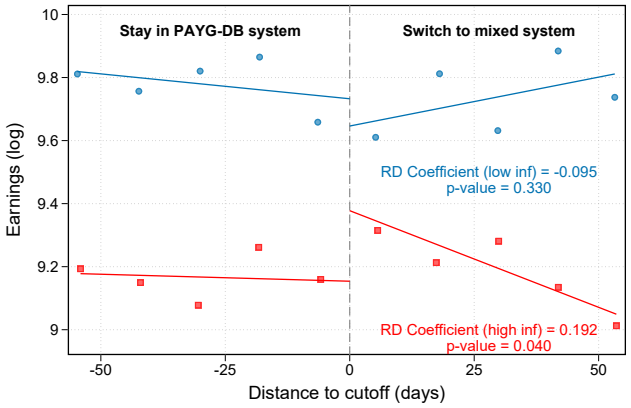
(b) 2001-2004



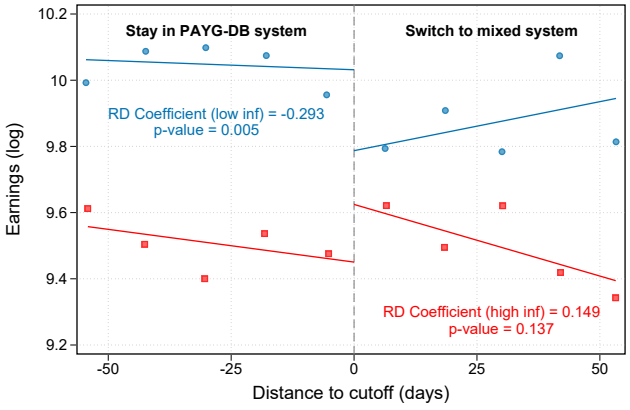
(c) 2005-2008



(d) 2009-2011



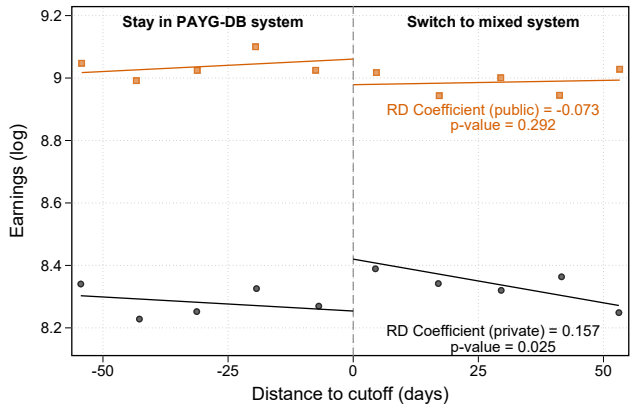
(e) 2012-2013



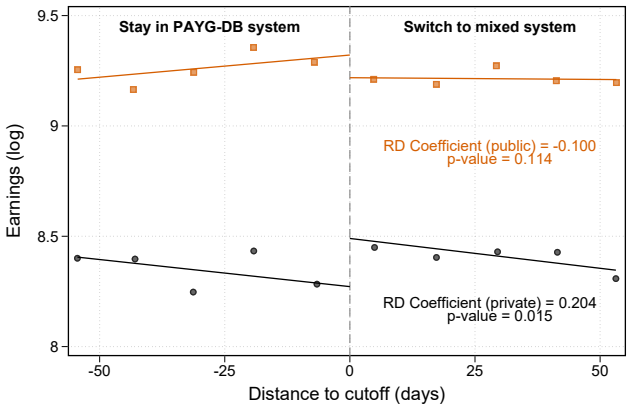
Notes: This figure shows RD-plots for the total labor earnings from equation 1. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is natural logarithm of total reported labor earnings. p-values are calculated using clustered standard errors at the worker level. Red indicates high informality and underreporting sectors and blue indicates low informality and underreporting sectors. The dependent variable is residualized from year fixed effects and evaluated at the mean.

Figure E.5: Effect of the reform on earnings (heterogeneity by public and private sector) - SSA data

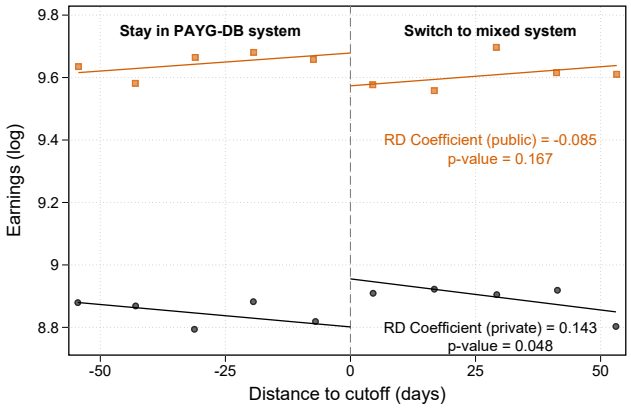
(a) 1997-2000



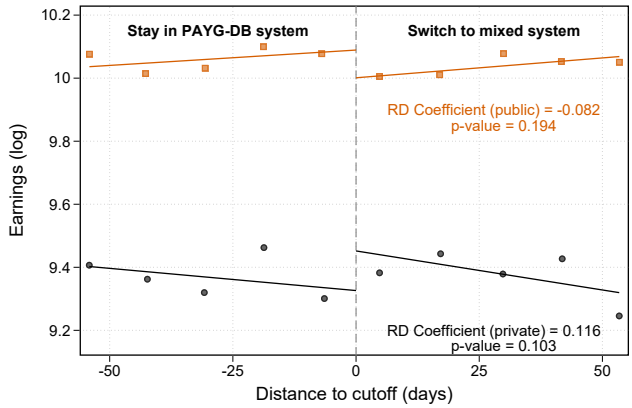
(b) 2001-2004



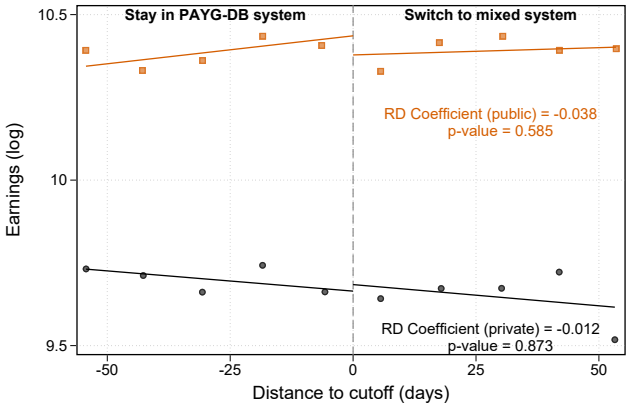
(c) 2005-2008



(d) 2009-2011



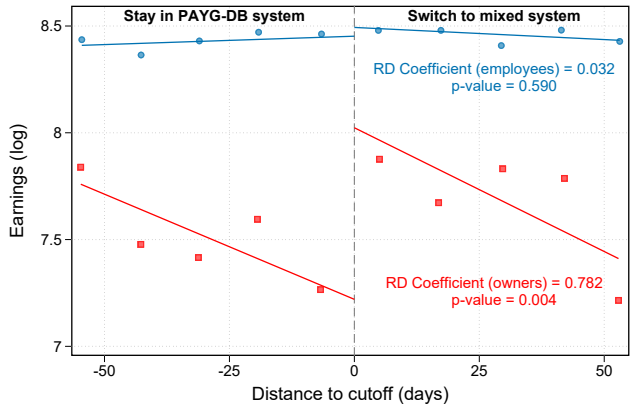
(e) 2012-2013



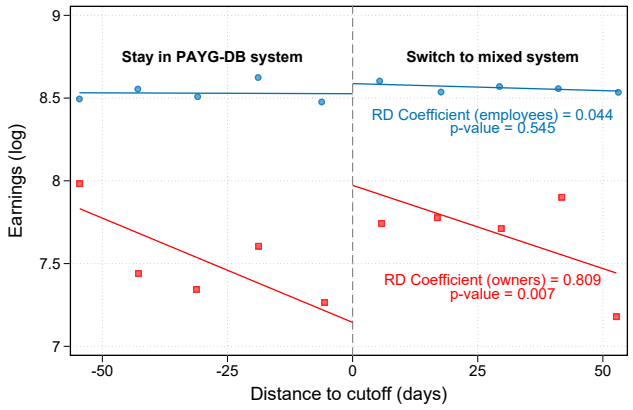
Notes: This figure shows RD-plots for the total labor earnings from equation 1. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is natural logarithm of total reported labor earnings. *p*-values are calculated using clustered standard errors at the worker level. Black indicates the private sector and orange indicates the public sector. The dependent variable is residualized from year fixed effects and evaluated at the mean.

Figure E.6: Effect of the reform on earnings (heterogeneity by ownership) - SSA data

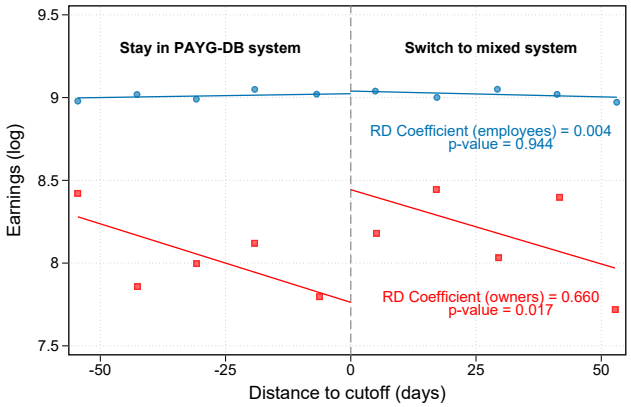
(a) 1997-2000



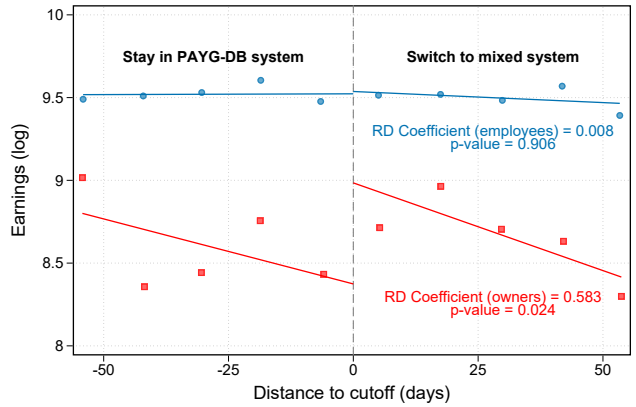
(b) 2001-2004



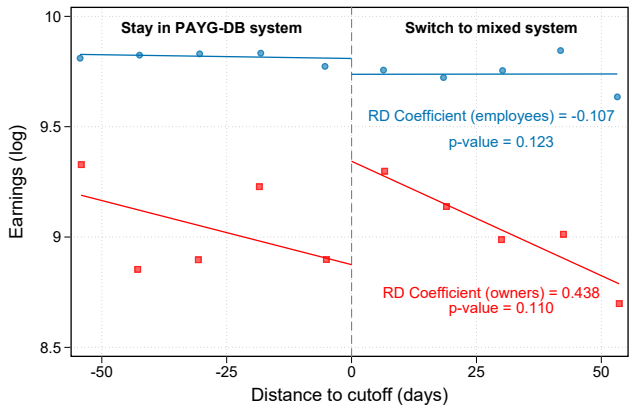
(c) 2005-2008



(d) 2009-2011

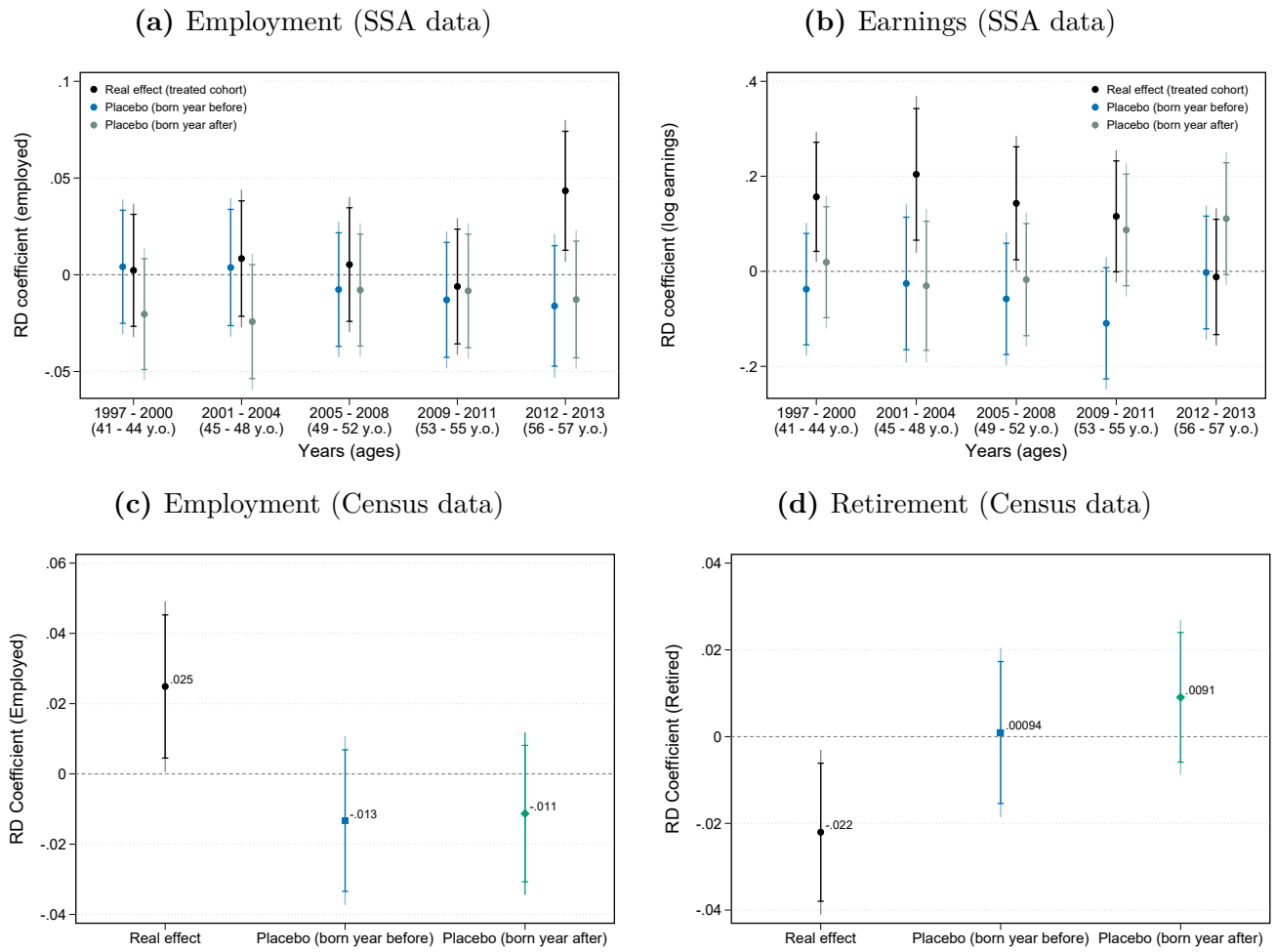


(e) 2012-2013



Notes: This figure shows RD-plots for the total labor earnings from equation 1. Individuals born before the cutoff were left by default in the pay-as-you-go system with defined benefits and individuals born at the cutoff or after were switched to the mixed system with retirement accounts. Each panel corresponds to a different group of years. In all panels, the dependent variable is natural logarithm of total reported labor earnings. p-values are calculated using clustered standard errors at the worker level. Red indicates owners and blue indicates employees. The dependent variable is residualized from year fixed effects and evaluated at the mean.

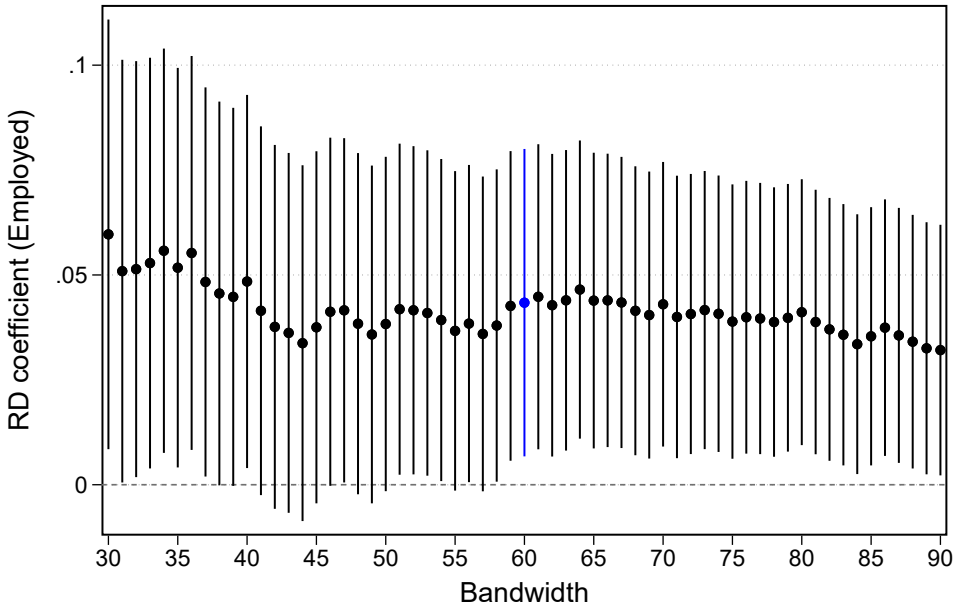
Figure E.7: RD coefficients - comparison with placebos (SSA and census data)



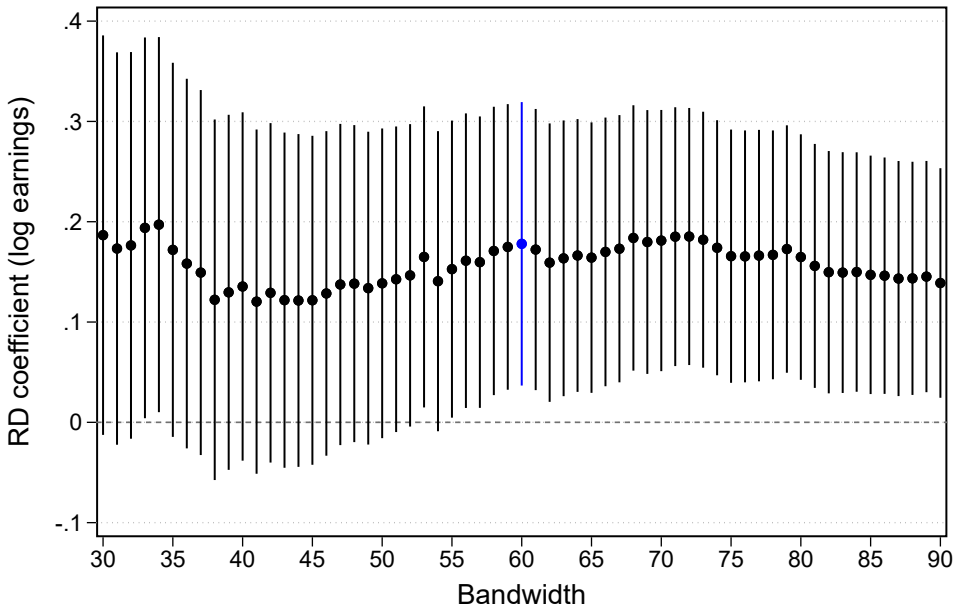
Notes: This figure shows a comparison of the main RD coefficients with placebos estimated using cohorts born in the year before and the year after the cohort affected by the reform. Panel (a) shows coefficients for the effect on the probability of being employed and panel (b) shows coefficients for the natural logarithm of labor earnings, both using SSA data. The numbers underneath the years indicate the ages of workers in those years. Panel (c) shows coefficients for the probability of being employed and panel (d) for the probability of being retired, both using census data. Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference. Black corresponds to estimates for the cohort affected by the reform. Green corresponds to estimates for the cohort born in the year after the cohort affected by the reform. Blue corresponds to estimates for the cohort born in the year before the cohort affected by the reform.

Figure E.8: Different bandwidths (SSA data)

(a) Employment rates



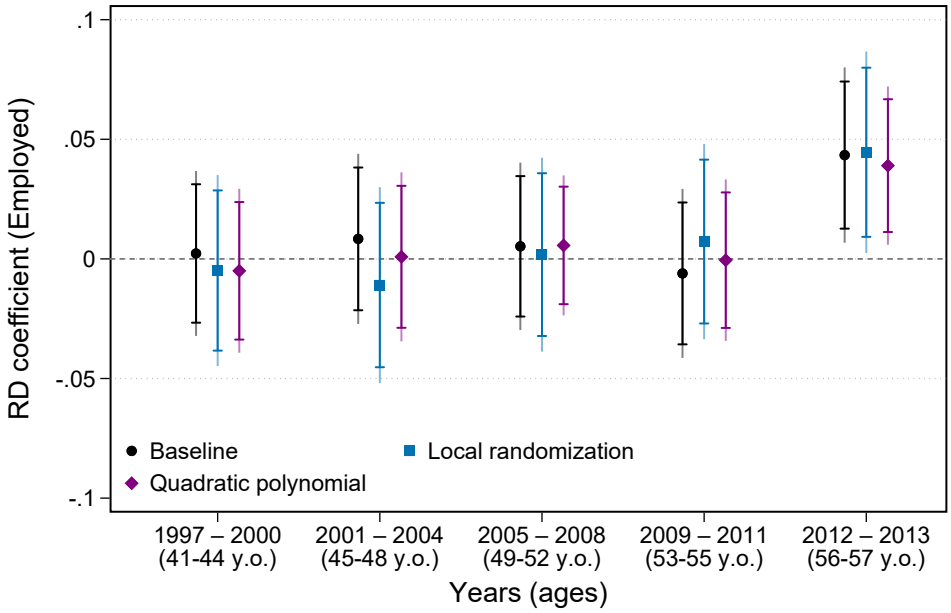
(b) Earnings



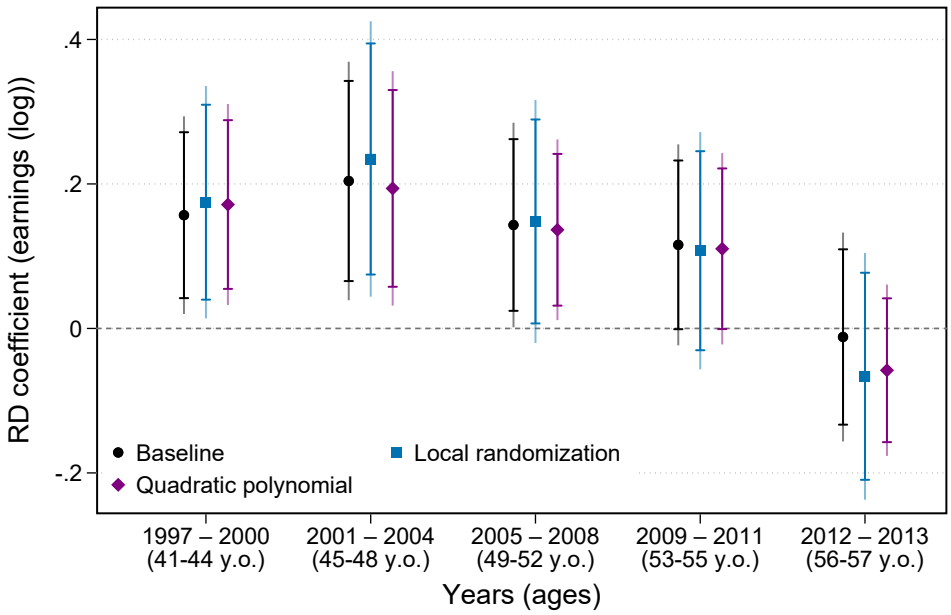
Notes: This figure shows RD estimates for different bandwidths around the cutoff. Panel (a) shows coefficients for the effect on the probability of being employed in 2012-2013 and panel (b) for the effect on the natural logarithm of monthly earnings for the years 1997-2004. Blue corresponds to our baseline estimates and black to all other bandwidths. Vertical bars represent 95% from worker-level cluster-robust inference.

Figure E.9: Time series plot of RD coefficients - different specifications (SSA data)

(a) Employment rates

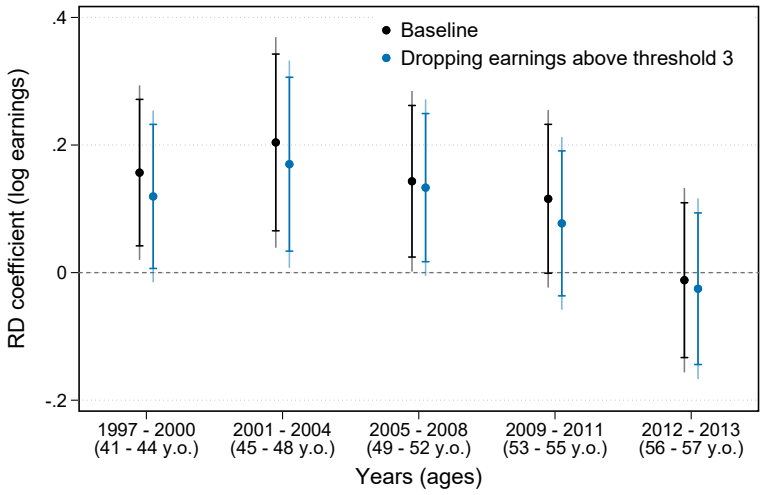


(b) Earnings



Notes: This figure shows a time series plot for the RD coefficients for different specifications. Panel (a) shows coefficients for the effect on the probability of being employed and panel (b) for the effect on the natural logarithm of monthly earnings. The numbers underneath the years indicate the ages of workers in those years. Black corresponds to estimates using our baseline bandwidth of 60 days around the cutoff. Blue corresponds to local randomization estimates calculated using a window of 11 days around the cutoff. Purple corresponds to estimates calculated fitting a quadratic polynomial with the continuity-based approach, using a triangular kernel, and an optimal bandwidth following Calonico et al. (2014). Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference.

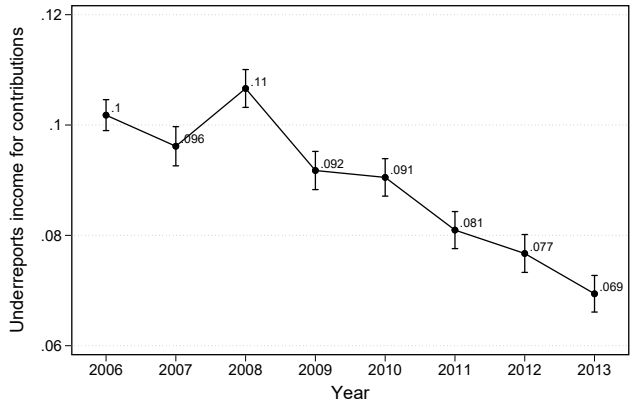
Figure E.10: Time series plot of RD coefficients - Earnings dropping earnings above ceiling (SSA data)



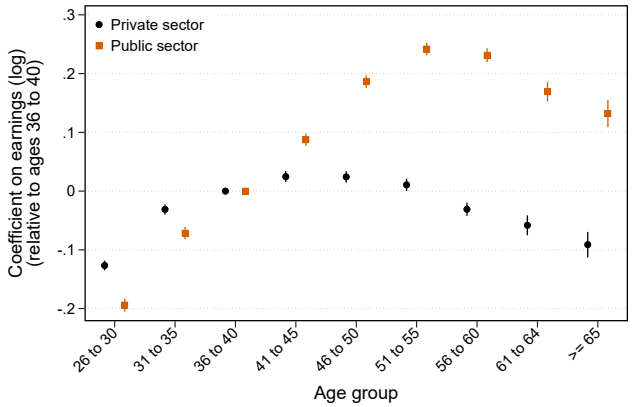
Notes: This figure shows a time series plot for the RD coefficients for each group of years. The dependent variable is the natural logarithm of labor earnings. The numbers underneath the years indicate the ages of workers in those years. Vertical bars represent 95% and 90% confidence intervals, both from worker-level cluster-robust inference. Black corresponds to estimates for the cohort using all workers. Blue corresponds to estimates using only workers with earnings below the ceiling.

Figure E.11: Additional survey and census data results

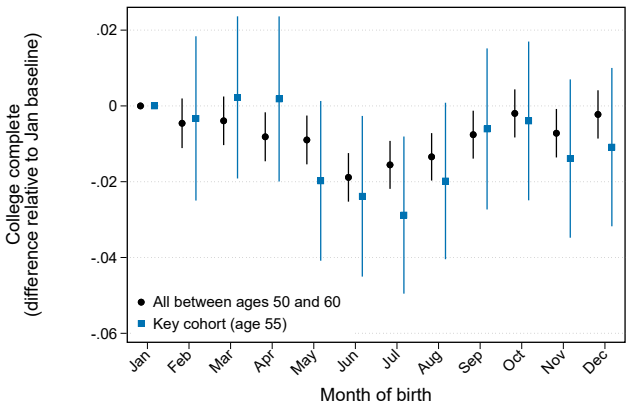
(a) Earnings underreporting over time (HH Surveys)



(b) Age-earnings profiles by public and private sector (HH Surveys)



(c) College completion rates by month of birth



Notes: This figure shows additional auxiliary results using survey and census data. Panel (a) shows the proportion of formal workers that admit to underreporting their labor earnings for their social security contributions in labor-market household surveys for each year. We define as underreporting workers who answer "no" to the following question "do you contribute to social security based on the totality of your labor earnings?". Panel (b) shows the estimated age-earnings profiles for workers in the private sector and in the public sector for workers of at least 26 years of age. Coefficients for the private sector are shown in black and for the public sector are shown in orange. The dependent variable is the natural logarithm of total labor earnings. Each point represents the OLS coefficient of each age group dummy variable, relative to the omitted category of 36 to 40 years old. Panel (c) shows the differences in college completion rates in the 2011 census according to the month of birth, relative to the baseline of a January date of birth. Black corresponds to coefficients estimated using all individuals between the ages of 50 and 60. Blue corresponds to estimates calculated using only workers who are 55 years of age at the time of the census (that means that they were born in 1956, the year of the cohort-based discontinuity). In panels (a) and (b) estimates are calculated using household surveys from 2006 to 2019 and include year fixed effects, in panel (c) estimates are calculated using 2011 census data. Vertical bars represent 95% confidence intervals.

F. Construction of the SES index

In this section we describe the procedure to construct the socioeconomic status (SES) index. We proceed in two steps, first we select several characteristics indicative of SES (such as whether the individual owns their dwelling, has completed college, and owns several durable goods). Then we compute the index as a weighted sum of these characteristics, the weights of which we obtain via Principal Component Analysis (PCA).

We have information on several characteristics indicative of socioeconomic status. We use an indicator of whether the individual has completed a college degree, an indicator of being a home owner, an indicator for having a clothes drying machine, the number of television sets owned, an indicator for owning a mobile phone, an indicator of owning a computer, the number of cars owned, and an indicator of having an internet connection.[1] Table F.1 presents summary statistics of the variables we use to construct the SES index.

Table F.1: Summary statistics for variables used to construct the SES index

	Observations	Mean	Standard Deviation	Median
College complete	109,828	0.224	0.417	0.000
Home owner	109,354	0.676	0.468	1.000
Has clothes dryer	109,354	0.109	0.312	0.000
Number of TVs	109,828	1.848	0.981	2.000
Has mobile phone	109,354	0.930	0.255	1.000
Has computer	109,828	0.556	0.497	1.000
Number of cars	109,828	0.542	0.677	0.000
Has internet	109,354	0.506	0.500	1.000

Notes: This table reports summary statistics for the variables used to construct the socioeconomic status index. College complete is a dummy variable equal to 1 if the individual has completed any college degree and zero otherwise. Home owner is a dummy variable equal to 1 if the individual owns their home and zero otherwise. Has clothes dryer is a dummy variable equal to 1 if the household owns a clothes drying machine. Number of TVs is the number of television sets owned in the household. Has mobile phone is a dummy variable equal to 1 if the individual owns a mobile phone and zero otherwise. Has computer is a dummy variable equal to 1 if the household owns at least one computer and zero otherwise. Number of cars is the total number of cars owned by the household. Has internet is a dummy variable equal to 1 if the household has an internet connection and zero otherwise.

Table F.2 reports the results for the PCA, where we report the main 3 components. Panel A shows the variable weights and Panel B shows the statistics associated to each component. The typical approach in the SES literature is to retain only the first component, based on the fact that it tends to provide a good estimation of the SES of the household (Filmer and Pritchett, 2001; McKenzie, 2005). The first component in our case positively correlates with all the variables, and

[1] The census data also contains other variables that are frequently used to infer socioeconomic status, such as having a bathroom or having electricity. However, these have little variation, since most households in the sample have access to such amenities. Thus, we exclude them for the derivation of the socioeconomic status index.

has an eigenvalue of almost 2.8 while explaining almost 35% of the variance. We normalize this first component to have mean zero and standard deviation of one, and use it as our SES index.

Table F.2: Principal component analysis for SES index

	Component 1	Component 2	Component 3
Panel A. Variable loadings			
College complete	.2957395	.0577124	-.3226414
Home owner	.1573287	.6922537	.5696147
Has clothes dryer	.2262674	.3437633	-.6546112
Number of TVs	.3858493	.060771	.0300353
Has mobile phone	.2058759	-.3836007	.3673276
Has computer	.5070501	-.2670452	.0579324
Number of cars	.3636602	.3293137	.0448591
Has internet	.5059405	-.2620643	.0415409
Panel B. Component statistics			
Eigenvalue	2.784	1.055	0.934
Proportion explained	0.348	0.132	0.117

Notes: This table reports the results from the principal component analysis. We keep the 3 main components. Panel A reports the variable weights for each component and Panel B reports the component statistics. College complete is a dummy variable equal to 1 if the individual has completed any college degree and zero otherwise. Home owner is a dummy variable equal to 1 if the individual owns their home and zero otherwise. Has clothes dryer is a dummy variable equal to 1 if the household owns a clothes drying machine. Number of TVs is the number of television sets owned in the household. Has mobile phone is a dummy variable equal to 1 if the individual owns a mobile phone and zero otherwise. Has computer is a dummy variable equal to 1 if the household owns at least one computer and zero otherwise. Number of cars is the total number of cars owned by the household. Has internet is a dummy variable equal to 1 if the household has an internet connection and zero otherwise. Eigenvalue is the eigenvalue associated to each component. Proportion of the variance explained is the proportion of the variance explained by each component.

G. Correlation of earnings with days and hours worked

In this section we assess the relationship of labor earnings with real measures of labor supply. Specifically, we correlate our measures of real labor supply (days and hours worked) with labor earnings. We estimate equations of the form:

$$Y_{it} = \alpha + \beta \text{LaborSupply}_{it} + u_{it} \tag{G.1}$$

Where Y_{it} represents the earnings of worker i at time t. LaborSupply_{it} is a measure of real labor supply (days worked in the month or the natural logarithm of hours worked). In different specifications we include time fixed effects and worker fixed effects. We use the full sample of individuals born between 1955 and 1957 and cluster standard errors at the worker level.

Table G.1 presents OLS estimates of equation G.1. Panel A includes the monthly days worked as the measure of labor supply. Panel B includes the natural logarithm of the total monthly hours worked. Panel C includes both days and hours worked. Column 1 includes no additional controls, column 2 includes year fixed effects, column 3 includes worker-fixed effects, and column 4 includes worker and year fixed effects.

Across specifications, both measures of labor supply positively correlate with labor earnings. An additional day worked is associated to an increase in labor earnings between 3 and 4 percent. Similarly, monthly hours worked also positively correlate with earnings: a 10 percent increase in monthly hours worked is associated to a 2.6 percent increase in earnings. Both correlations are robust to estimating the coefficients using within-person variation by including worker fixed effects (columns 3 and 4) or while including both measures of labor supply (in Panel C).

Table G.1: Regressions of earnings on hours and days worked

	Total labor earnings (log)			
	(1)	(2)	(3)	(4)
Panel A. Days worked				
Days worked in the month	0.0366***	0.0392***	0.0228***	0.0258***
	(0.000406)	(0.000400)	(0.000287)	(0.000235)
Year fixed effects		✓		✓
Worker fixed effects			✓	✓
Number of workers	121356	121356	108831	108831
Panel B. Hours worked				
Monthly hours worked (log)	0.315***	0.343***	0.266***	0.272***
	(0.00626)	(0.00602)	(0.00446)	(0.00335)
Year fixed effects		✓		✓
Worker fixed effects			✓	✓
Number of workers	120728	120728	107957	107957
Panel C. Days and hours worked				
Days worked in the month	0.0447***	0.0463***	0.0236***	0.0260***
	(0.000445)	(0.000434)	(0.000328)	(0.000271)
Monthly hours worked (log)	-0.00512	0.0120*	0.123***	0.114***
	(0.00645)	(0.00615)	(0.00470)	(0.00345)
Year fixed effects		✓		✓
Worker fixed effects			✓	✓
Number of workers	120728	120728	107957	107957

Notes: this table reports OLS estimates of equation G.1. In all specifications the dependent variable is the natural logarithm of total labor earnings. Standard errors are clustered at the worker level. Days worked in the month is the total number of days worked in the month. Monthly hours worked (log) is the natural logarithm of total monthly hours worked. Column 2 includes year fixed effects. Column 3 includes worker fixed effects. Column 4 includes year fixed effects and worker fixed effects. * Significant at the 10% level ** Significant at the 5% level *** Significant at the 1% level.