Enlisting Employees in Improving Payroll-Tax Compliance: Evidence from Mexico

Todd Kumler
Eric Verhoogen
Judith Frías

Jan. 2020

Online Appendix

A Additional Institutional Background

In this section, we provide additional details about the Mexican tax system, IMSS contribution rates and non-pension benefits, the 1992 pension reform, and our pension simulation.

A.1 Other Dimensions of Tax System in Mexico

One reason that firms in developed countries engage in relatively little under-reporting of wages may be that it does little to reduce their overall tax burden. If corporate or personal income taxes are as high as payroll taxes and difficult to evade, then lower payroll taxes due to under-reporting will be offset by higher taxes on corporate or personal income. In Mexico, the corporate income tax is generally higher than the payroll tax on paper: it went from 39 to 34 percent over the 1988-2003 period.[1] But corporate tax evasion and avoidance are rife. For instance, the OECD in 1992 found that, in part due to various loopholes, 70 percent of corporate tax declarations reported no taxable income (OECD, 1992). By all accounts, tax evasion remains high (OECD, 2011). In addition, the social security agency and the Mexican tax authority first signed an agreement to share data in June 2002; thus for almost all of the period under study, there was no chance that information reported to the social security agency would affect the corporate tax burden. It appears, in other words, that evaded payroll taxes were not offset by increases in other taxes.

Also, it does not appear that individual income taxes provided a strong disincentive to most workers to have their wages reported accurately. Mexico provides extensive tax credits for low-wage workers, originally instituted to offset the regressive effects of VATs, with the consequence that many workers legally pay no income tax, or even receive funds from the tax authority (i.e. face a negative income tax.) In 1997, for instance, individuals making less than 3.2 times the minimum wage in Mexico City faced a zero or negative tax rate (OECD, 1999, p. 80).

A.2 Contribution Rates

During our study period, the total employer contribution varied between 18 percent and 22 percent of the wage over the wage range in which almost all workers fall. There were modest changes over time, illustrated in Appendix Figure A1.[2] Worker contributions, which varied between 2 percent and 5 percent over the relevant range, are illustrated in Appendix Figure A2. In the post-reform (PRA) regime, over the 1997-2003 period employers were required to contribute 5.15 percent of each employee's wage, and employees 1.125 percent; the government contributed 0.225 percent, as well as a "social quota" equal to 5.5 percent of the current minimum wage in Mexico City. Appendix Tables A1 and A2 provide full descriptions of the contribution schedules. In the post-reform regime, employees also have the option to contribute to a voluntary retirement savings account. See Lara-Ibarra (2011) for an analysis of the effects of a change in the tax rate on these contributions.

A.3 Non-Pension Benefits

In addition to the pension and health-care benefits mentioned in the main text, IMSS offers working mothers and widowed or divorced working fathers free child care during workdays for

[1] Source: OECD Tax Database, www.oecd.org/ctp/taxdatabase.

[2] The figure also illustrates that the maximum taxable income rose from 10 times to 25 times the minimum wage in Mexico City over the period. (There are three minimum wage zones in Mexico, corresponding to higher-, medium- and lower-wage municipalities, respectively.) But the topcodes apply to no more than 5 percent of wage-earners in any year and will play little role in our analysis.

children under four years old. As mentioned in the main text, the value of these benefits remained roughly constant over time.[3]

As an additional social-security benefit for workers, employers contribute five percent of a worker's wage, up to 10 times the applicable minimum wage, to the *Instituto del Fondo Nacional de la Vivienda de los Trabajadores* (INFONAVIT).[4] Employees can apply for housing loans through INFONAVIT. If approved, a worker can use the accumulated funds as a down payment on a house purchase; loan payments are subsequently deducted from employees' paychecks.

Prior to 1992, accumulated contributions in INFONAVIT not used toward housing were provided to the worker at the time of retirement. However, INFONAVIT only provided nominal accumulated contributions. Given the high inflation rates, the real value of nominal contributions was typically quite small. The 1992 IMSS reform sought to correct this problem by requiring that INFONAVIT provide workers with any unused contributions plus interest based on the operational surplus of the agency. In practice, however, INFONAVIT continued to suffer from several problems, including high rates of delinquency of loans, which limited the agency's ability to pay interest and resulted in a negative real rate of return (Grandolini and Cerda, 1998).

Following the 1997 reform, the INFONAVIT contributions are collected in a personal account managed by the same AFORE that manages an individual's retirement account (see the account statement in Appendix Figure A3). Workers who choose the PRA system at retirement receive the unused accumulated balances in their INFONAVIT account upon retirement. Transition workers that choose the PAYGO pension at retirement only receive the account balances that were accumulated between 1992 and 1997, thus forfeiting balances accumulated after 1997.

These changes in the housing account potentially made a substantial difference in the overall value of social security benefits. It is also worth noting that they tended to reinforce the differential change in incentives emphasized in the main text: the change in the administration of the housing accounts effectively made benefits more sensitive to reported wages, but only for those workers who in the end choose the PRA pension.

A.4 Fines

The social security law provides for fines if establishments are caught evading taxes. The fines ranged from 70-100 percent of the amount of evasion over the 1995-2001 period and have ranged from 40-100 percent, with most exactly at 40 percent, since 2001. Employees are not rewarded if firms are fined. In addition, the law requires employers with 300 or more employees to submit an audit by a certified public accountant to IMSS (since 1993) as well as to the Mexican tax authority (since 1991).

A.5 1992 Pension Reform

In an effort to restore financial stability to the IMSS system, the Mexican congress enacted a first attempt at pension reform in May 1992. This reform created a system of personal retirement accounts called the *Sistema de Ahorro para el Retiro* (SAR) to operate alongside the established PAYGO pension system. Employers contributed two percent of each worker's wage, depositing the contributions into a commercial bank of their own choosing. The commercial bank then transmitted the funds to the Mexican central bank. The central bank guaranteed a minimum

[3]There has been a secular decline in the number of IMSS hospital and clinic beds per covered individual, but there was no trend break in 1997 (IMSS, 2011, ch. 11). Below we will find no pre-trend in under-reporting prior to 1997.

[4]Note that the INFONAVIT contribution is the component of contributions that uses the local minimum wage, as opposed to the minimum wage in Mexico City, to calculate the topcode.

two percent real return, and workers were supposed to receive a lump-sum payment at retirement equal to the accumulated balances. Several problems plagued the implementation of the SAR system, however. The scheme suffered from poor regulatory oversight and management. Commercial banks received low fees for administering the collection of employer contributions, weakening incentives for these banks to provide efficient record-keeping or enforce the mandatory contributions. In addition, workers were often unaware of the balances in their accounts or even of which bank held their accounts. This led to the creation of multiple accounts, especially for workers who changed employers; by 1997, over half of the 20 million accounts were duplicates (Grandolini and Cerda, 1998). As a result of these difficulties, the reform was widely considered a failure, with workers and employers viewing the reform as simply another payroll tax (Grandolini and Cerda, 1998; Aguila, 2011).

A.6 Pension Simulation

As described in the main text, we conduct a pension simulation, based on the simulation in Aguila (2011), in order to compare pension wealth at retirement under the PAYGO and the PRA systems. Here we provide additional details of the simulation.

For the PAYGO system, the monthly pension is calculated based on the benefit schedule in Appendix Table A3, which was in effect from Jan. 1, 1991 to June 30, 1997.[5] To calculate the final average wage (the average nominal wage in the five years before retirement), we assume a constant real wage and an annual inflation rate of 13.65 percent (the average annual inflation rate from 1988-2003, excluding the high-inflation years of 1988, 1989, 1995, and 1996). Under these assumptions, the final average wage is approximately 79 percent of the wage at retirement. Following Aguila (2011), we discount the value of the monthly pension benefits to the present assuming a discount rate of one percent.

Under the PRA regime, we calculate the total value of contributions over a worker's career, assuming a constant real wage.[6] We then determine the total value of accumulated wealth at retirement assuming a 8.59 percent annual return (the average return from 1998-2002, and the higher of two rates of return considered by Aguila (2011)).[7] We next calculate the schedule of monthly annuity payments equivalent in expected value to total pension wealth at retirement, assuming a life expectancy of 93 years for men — the life expectancy assumption used by IMSS, which includes an adjustment for expected survivor benefits paid out to widows or dependent children. As with the PAYGO pension, we discount the monthly pension assuming a one percent discount rate.

There are a number of thorny issues in inflation indexing in the simulation. Beginning in 1997, the minimum PRA pension was indexed to the Mexican Consumer Price Index (CPI) and the minimum PAYGO pension was indexed to the minimum wage in Mexico City. The real value of the minimum wage declined steadily over the 1997-2001 period, as it had been doing for many

[5] Under the personal-account system, individuals have three options upon retirement. One is to receive programmed withdrawals from the individual's AFORE, where the withdrawal amount is calculated based on the account balance as well as the age and life expectancy of the individual and dependents. (A worker who receives the minimum pension must choose this option.) A second option is to purchase an annuity from a private insurance company that guarantees a fixed monthly pension. A third option, available to workers with a personal-account balance exceeding 130 percent of the cost of an annuity providing a monthly payment equal to the minimum pension, is to take a lump-sum payment upon retirement.

[6] For the PRA regime, we include housing account contributions, which assumes that workers do not use housing accounts for loans. Note that we do not include the housing account under the PAYGO system, as workers opting to take the PAYGO pension at retirement only receive housing account contributions accumulated between 1992 and 1997.

[7] These real interest rates are net of management fees charged by AFOREs.

years (see Appendix Table A4). The values of the minimum pensions under the two regimes thus diverged. In 2001, the Mexican congress passed a new law specifying that the CPI be used to index the PAYGO pension as well as the PRA pension and adjusting the PAYGO minimum pension to be equal to the PRA minimum pension. It is unclear what participants believed at the time of the reform about how the indexing would be carried out. Following Aguila (2011), we assume that participants expected a decline of 6.4 percent per year. At the same time, we assume that participants expected the minimum pension under the PAYGO regime to be the same as the minimum pension under the PRA regime, as the congress in fact subsequently legislated in 2001.

It is important to note that we have imposed a relatively optimistic assumption about future interest rates and a relatively pessimistic assumption about the decline in the real value of the minimum wage. Both tend to make the PRA pension appear relatively more attractive. We emphasize that the simulation is mainly to illustrate that younger workers are more likely to expect to opt for the PRA pension; we are not arguing that the pension reform made workers better off overall. Indeed, it seems clear from the simulation that workers with low wages who would have qualified for the minimum pension under the PAYGO system are generally worse off under the new regime.

In addition to simulating the pension wealth of workers who started working at age 25 and expect to continue working until age 60, but who are of different ages in 1997, as we present in Table 1, we also simulate the pension wealth of a worker who entered the system on June 30, 1997, the day before the new personal retirement accounts (PRAs) went into effect. The results are presented in Table A5. If the worker expects to contribute for at least 10 years (and thus be guaranteed the minimum pension), the basic message is the same as in the previous simulation: the PRA is more likely to dominate the PAYGO pension as one moves up and to the right in the table. A new worker earning below 200 pesos/day (approximately 80% of our sample) would need to contribute for at least 20 years, if not more, for the PRA pension to dominate the PAYGO pension.

B Theory Appendix

In this section, we develop in full the model that we have summarized in Section 3 of the main text.

B.1 Set-up

Consider a setting with a competitive labor market populated by homogeneous workers and a continuum of heterogeneous, monopolistically competitive firms.[8] Let τ_f be the payroll tax statutorily imposed on the firm and τ_w the payroll tax statutorily imposed on workers; both are remitted by the firm and will enter similarly in our model.[9] Define $\tau = \tau_f + \tau_w$ and assume $0 < \tau < 1$.

Let w_r be the pre-tax wage reported by a firm to the government, and w_u the unreported wage, the wage paid to workers "under the table." The total wage paid by the firm is $w_f = w_r + w_u$. The net take-home wage received by workers is $w_{net} = w_u + (1 - \tau)w_r$. We assume in the theory that w_r, w_u, and w_{net} are observable to workers.[10] Both the reported wage, w_r, and the net wage, w_{net}, correspond to quantities that are in principle observable to the econometrician: w_r to the wages reported by firms in the administrative records of the social security agency, and w_{net} to the take-home pay reported by workers in the ENEU household survey. As mentioned in the main text, we do not observe w_{net} at the firm level, and hence cannot measure the unreported wage, $w_u = w_{net} - (1 - \tau)w_r$, at the firm level, but we will be able to construct measures at the level of more aggregate cells.[11]

Future pension benefits depend on the reported wage, w_r. In the interests of simplicity, we impose an assumption of linearity on these benefits and let bw_r be the amortized per-period value of future pension benefits for each worker, where $b \geq 0$. We further assume that $b < \tau$, which arguably corresponds to the Mexican institutional setting. We refer to the wage inclusive of pension benefits received by each worker as the "effective" wage, w_e, where $w_e = w_{net} + bw_r = w_u + (1 - (\tau - b))w_r$.

We assume that there is no stigma or other cost to workers of firms' under-reporting of their wages. Under this assumption, the effective wage is the wage relevant for workers' labor-supply decisions. We assume that the aggregate labor-supply function has constant elasticity:

$$L^S_{agg} = Bw^{\rho}_e \tag{A1}$$

where $\rho > 0$ and $B > 0$. This labor-supply function can be derived from maximization of a quasi-linear utility function for an individual choosing how many hours to devote to leisure versus

[8] It would be possible to incorporate different types of workers in our framework, along the lines, for instance, of Rothstein (2010). As long as evasion is worker-type-specific within the firm (i.e. as long as there is no internal equity constraint preventing firms from evading more for some groups of workers than for others) conceptually this exercise would be straightforward. In our view, however, it would complicate the exposition with relatively little payoff in additional insight. We leave the development of a more general framework to future work.

[9] As pointed out by Slemrod (2008), in settings in which the costs of evasion by employers and employees differ, it is in general not irrelevant who remits the tax. But since in Mexico the statutory payroll taxes on both firms and workers are remitted by firms, in our setting there is no conceptually important distinction between them.

[10] In making the assumption that w_r, w_u, and w_{net} are all observable to the worker, we are following the main strand of the related theoretical literature (Yaniv, 1992; Kleven, Kreiner, and Saez, 2009), which presumes that employees collude in under-reporting. An alternative, plausible assumption would be that workers observe w_{net} costlessly but only observe w_r (and hence w_u) at a cost. The pension reform could then be modeled as reducing this cost, in addition to increasing the sensitivity of benefits to reported wages. Asymmetric information of this type would complicate the model considerably, and we leave the analysis of this case to future work.

[11] In the theory, we also abstract from minimum wages (of which there are three in Mexico, depending on region).

wage work.[12] Note that we are not explicitly considering the extensive margin of labor supply or the lump-sum (i.e. independent of reported wage) benefits of participation in the social-security system. Such lump-sum benefits would affect employees' utility levels and participation decisions, but would not affect the evasion behavior that is the primary object of our analysis. An alternative approach to deriving the labor-supply elasticity (A1) would be to model individuals as choosing whether to supply labor to the formal sector (i.e. registered firms) or the informal sector (i.e. unregistered firms), as for instance in Marrufo (2001) or Galiani and Weinschelbaum (2012). In this paper we focus on evasion within the formal sector, and leave the analysis of individuals' and firms' choices about whether to enter the formal sector to future work.

B.2 The Firm's Problem

We build on a standard model of heterogeneous firms under monopolistic competition, similar to Melitz (2003) but without international trade. We adopt a monopolistic-competition framework because of its tractability and its comparability to much of the recent heterogeneous-firm literature (e.g. Hsieh and Klenow (2009)), but we acknowledge that other ways of modeling heterogeneous firms (e.g. perfect competition with decreasing returns to scale) would yield similar conclusions. Assuming a Dixit-Stiglitz (1977) representative consumer, the demand for each differentiated variety, ω, can be written:

$$x(\omega) = Ap(\omega)^{-\sigma} \tag{A4}$$

where $x(\omega)$ is the quantity consumed; $p(\omega)$ is the price; and σ is a parameter capturing the elasticity of substitution between varieties.[13] A captures the general level of demand, which individual firms treat as exogenous; in this partial-equilibrium framework, we abstract from the determinants of the level of demand and treat A as a parameter. We make the standard assumption that $\sigma > 1$. Firms are assumed to be heterogeneous in a productivity parameter, φ, with density $g(\varphi)$, with positive support over $[\varphi^{min}, \varphi^{max}]$ and zero support elsewhere.[14] There is assumed to

[12] Suppose that for each worker,

$$U(c, L) = u(c) - \beta L^{\frac{\rho+1}{\rho}} \tag{A2}$$

where c is consumption, L is hours supplied, and we assume $\beta > 0$, $\rho > 0$. Let the consumption good be the numeraire. The hours constraint is $L \leq \overline{L}$. As is standard in the literature on labor supply, we think of individuals as having "total income" $w_e\overline{L}$ and consuming leisure hours at a price w_e. Hence the budget constraint is:

$$c + (\overline{L} - L)w_e = w_e\overline{L} \tag{A3}$$

We assume that the hours constraint is not binding. Individual optimization then yields (A1), where $B = N\left(\frac{\rho}{\beta(\rho+1)}\right)^{\rho}$ and N is the number of workers in the workforce.

[13] If the representative consumer has utility

$$U = \left[\int_{\omega\in\Omega} (x(\omega))^{\frac{\sigma-1}{\sigma}}\, d\omega\right]^{\frac{\sigma}{\sigma-1}} \tag{A5}$$

where Ω represents the set of all differentiated varieties available in the market, then optimization yields (A4) with $A = UP^{\sigma}$, where U is defined in (A5) and

$$P \equiv \left[\int_{\omega\in\Omega} (p(\omega))^{1-\sigma}\, d\omega\right]^{\frac{1}{1-\sigma}} \tag{A6}$$

[14] In Melitz (2003), firms do not know their productivity before paying a fixed cost to get a productivity draw. Here, as in Chaney (2008), we simply take the set of firms in business as given and abstract from entry and exit of

be no cost of differentiation, and each firm differentiates and produces a distinct variety; hence φ also indexes varieties. There is a fixed cost of operation to be paid in each period, f. Each firm's production function is simply $x = \varphi L$, where L is labor input; this can be rewritten $L = \frac{x}{\varphi}$.

We assume that the cost of evasion is increasing in the unreported wage, w_u, and the output of the firm. In particular, we assume that the cost of evasion can be expressed in the multiplicatively separable form $xc(w_u)$, where $c(0) = 0$, $c'(w_u) > 0$, and $c''(w_u) > 0$.[15] As mentioned in the main text, there are a number of possible justifications for the assumption that costs of evasion are increasing in output. One is simply that auditors are more likely to audit larger firms because their operations are more visible, as suggested by Besley and Persson (2013, p. 66) — a conjecture that appears anecdotally to be relevant in Mexico.[16] Another is the argument of Kleven, Kreiner, and Saez (2009) that collusion in under-reporting is more difficult to sustain in larger firms.

The labor market is competitive, and firms are price-takers of the effective wage, w_e. The firm chooses the unreported wage, w_u; together w_e and w_u pin down w_r.[17] From the definitions of the wage variables above, the total wage paid by the firm is then $w_f = \frac{w_e-(\tau-b)w_u}{1-(\tau-b)}$. The firm also chooses the output price, p. Given the price, output, x, and hence labor demand, L, are pinned down by the firm-specific demand curve, (A4). Per-period profit for each firm can be written:

$$\pi(w_u, p; \varphi, w_e) = \left\{ p - \frac{w_e - (\tau - b)w_u}{\varphi(1 - (\tau - b))} - c(w_u) \right\} x - f \tag{A7}$$

The firm's problem is to choose w_u and p to maximize π.

The first order condition for the choice of w_u is:

$$c'(w_u) = \frac{\tau - b}{\varphi(1 - (\tau - b))} \tag{A8}$$

The left-hand side is the marginal cost of evasion and the right-hand side is the marginal benefit in the form of reduced tax payments, both per unit of output. Note that the solution to this equation, call it $w_u^*(\varphi)$, depends neither on the output price, p, nor on the market-determined effective wage, w_e. Note also that, given our assumptions on the $c(\cdot)$ function, in general we have that $w_u^*(\varphi) > 0$; we do not expect perfect compliance, even for highly productive (hence large in equilibrium) firms.

To derive an expression for labor demand, note first that the first order condition for price yields:

$$p^*(w_e, \varphi) = \left(\frac{\sigma}{\sigma - 1} \right) \left\{ \frac{w_e - (\tau - b)w_u^*(\varphi)}{\varphi(1 - (\tau - b))} + c(w_u^*(\varphi)) \right\} \tag{A9}$$

The term in brackets is simply marginal cost and, as usual with Dixit-Stiglitz demand, price is a fixed multiplicative mark-up over marginal cost. Given the optimal choices $w_u^*(\varphi)$ and $p^*(w_e, \varphi)$, the optimal output of the firm is given by:

$$x^*(w_e, \varphi) = Ap^*(w_e, \varphi)^{-\sigma} \tag{A10}$$

firms. It would be straightforward to add the initial investment decision, but would not add substantively to our analysis.

[15] The assumption that the marginal cost of evasion incurs per unit of output is increasing is in the spirit of Slemrod (2001), who makes a similar assumption in the context of evasion by individuals.

[16] Also, as noted in Appendix A.4, employers with 300 or more employees have to submit an audit from a certified public accountant; this may be another reason why evasion is more costly for larger firms.

[17] We abstract from minimum wages, of which there are three in Mexico, depending on region. We consider the role of the minimum wages in the empirics below.

and the firm's labor demand is:

$$L^D(w_e, \varphi) = \frac{x^*(w_e, \varphi)}{\varphi} = \frac{Ap^*(w_e, \varphi)^{-\sigma}}{\varphi} \tag{A11}$$

Note that the firm's labor demand is decreasing in the effective wage, w_e, since price is increasing in w_e. Aggregate labor demand is the integral of firm-level labor demand (A11) over firms active in the market:

$$L^D_{agg}(w_e) = \int_{\varphi^{min}}^{\varphi^{max}} L^D(w_e, \varphi) g(\varphi) d\varphi \tag{A12}$$

Since we are assuming that the set of firms in the market stays fixed, and since each firm's labor demand is declining in the effective wage, w_e, we know that aggregate labor demand is also declining in w_e.

The equilibrium wage is the value of w_e that clears the labor market, i.e. that sets

$$L^S_{agg}(w_e) = L^D_{agg}(w_e) \tag{A13}$$

where aggregate labor supply, $L^S_{agg}(w_e)$, is given by (A1).

B.3 Evasion vs. Firm Size in Cross-Section

We now consider how the extent of evasion, as measured by the unreported wage, w_u, varies with firm size in cross-section, for a given effective wage w_e. Differentiating both sides of (A8) holding w_e fixed and rearranging, we have:

$$\frac{dw^*_u}{d\varphi} = -\frac{\tau - b}{\varphi^2 c''(w_u)(1 - (\tau - b))} < 0 \tag{A14}$$

That is, evasion is decreasing in firm productivity.[18]

Firm output is unambiguously increasing in productivity. To see this, first note that price is decreasing in productivity; differentiating both sides of (A9) (again, holding w_e fixed) and using (A8), we have:

$$\frac{dp^*}{d\varphi} = -\left(\frac{\sigma}{\sigma - 1}\right)\left\{\frac{w_e - (\tau - b)w^*_u(\varphi)}{\varphi^2(1 - (\tau - b))}\right\} < 0 \tag{A15}$$

Prices are lower in higher-φ firms for the standard reason that labor costs are lower per unit of output and price is a fixed multiplicative mark-up over costs. Then from (A10):

$$\frac{dx^*}{d\varphi} = -\sigma A(p^*)^{-\sigma-1}\left(\frac{dp^*}{d\varphi}\right) > 0 \tag{A16}$$

Together (A14) and (A16) imply an unambiguously negative relationship between firm output and evasion.

[18] Given that the effective wage, w_e, is constant across firms, this immediately implies that the reported wage is increasing in productivity: using the fact that $w_r = (w_e - w_u)/(1 - (\tau - b))$, we have:

$$\frac{dw^*_r}{d\varphi} = \frac{\tau - b}{\varphi^2 c''(w_u)(1 - (\tau - b))^2} > 0$$

The relationship between employment and productivity, and hence between employment and evasion, is more subtle. Differentiating (A11):

$$\frac{dL^D}{d\varphi} = -\frac{x^*(w_e, \varphi)}{\varphi^2} + \frac{1}{\varphi}\left(\frac{dx^*}{d\varphi}\right) \tag{A17}$$

Higher productivity leads firms to have greater output, which increases employment (the second term). But it also reduces the amount of labor required to produce a given level of output (the first term). Using (A9) and (A16)-(A17), it can be shown that labor demand will be increasing in productivity if and only if

$$\frac{c(w_u^*(\varphi))}{\sigma - 1} < \frac{1}{\varphi}\left\{\frac{w_e - (\tau - b)w_u^*(\varphi)}{1 - (\tau - b)}\right\} \tag{A18}$$

The term in brackets on the right-hand side is the total wage paid by the firm, w_f, and hence the right-hand side is labor cost per unit of output; $c(w_u^*(\varphi))$ is the cost of evasion per unit of output at the optimum. The condition thus requires that the equilibrium cost of evasion not be too large relative to labor costs. If enforcement were perfect, firms would set $w_u^*(\varphi) = 0$, condition (A18) would clearly be satisfied, and employment would be unambiguously increasing in productivity, as in Melitz (2003). But here the fact that the equilibrium cost of evasion per unit of output is positive dampens the responsiveness of output to productivity (since it raises prices and $(p^*)^{-\sigma-1}$ enters the expression for $\frac{dx^*}{d\varphi}$ in (A16)) and reduces the magnitude of the second term in (A17) relative to the first. In this context, it is theoretically possible that employment is declining in productivity. At the same time, previous work in Mexican data has found a positive correlation between employment and productivity (see Verhoogen (2008, Table A1)) and the positive correlation between size and productivity is robust across countries and datasets (see e.g. Foster, Haltiwanger, and Syverson (2008) for the U.S.). It seems clear that the empirically relevant case is the one in which (A18) holds. We will focus on this case hereafter. In this case, the extent of evasion is declining in employment, as it is in output.

B.4 Response of Evasion to Pension Reform

We now consider the response of evasion to the pension reform, which we model as an increase in the parameter b relating the reported wage to the amortized per-period pension benefits. Here we allow the equilibrium effective wage, w_e, to vary endogenously in response to the policy change. Differentiating (A8) with respect to b and rearranging:

$$\frac{dw_u^*}{db} = -\frac{1}{(1 - (\tau - b))^2 \varphi c''(w_u^*(\varphi))} < 0 \tag{A19}$$

The unreported wage unambiguously decreases within a given firm.[19]

It is worth emphasizing that the response of $w_u^*(\varphi)$ to the policy change here does not depend

[19]Given the discussion of cross-sectional patterns in Section B.3, it is interesting to consider the heterogeneous effects of the pension reform by firm size. The sign of this cross-partial depends on the third derivative of the cost-of-evasion function. Formally,

$$\frac{d}{d\varphi}\left(\frac{dw_u^*}{db}\right) = \frac{\varphi c'''(w_u^*(\varphi))\frac{dw_u^*}{d\varphi} + c''(w_u^*(\varphi))}{\left[(1 - (\tau - b))\varphi c''(w_u^*(\varphi))\right]^2} \tag{A20}$$

where $\frac{dw_u^*}{d\varphi}$ is given by (A14). Without imposing further structure on the cost of evasion, we do not have a clear prediction for a differential response by firm size.

on the market-determined effective wage, w_e, or the incidence of the policy change on that wage. In this sense, the model suggests that it is not unreasonable to examine the effect of the policy change on evasion separately from the question of incidence, which is how we proceed in the empirical analysis.

B.5 Incidence of Pension Reform on Wages

Now consider the effect of the reform on the effective wage, w_e. Differentiating both sides of the labor market equilibrium condition (A13), and using the fact that the integrand and its derivatives are continuous to pass the derivative through the integral in (A12),[20] we have:

$$\frac{dw_e}{db} = \frac{\int_{\varphi^{min}}^{\varphi^{max}} \left[w_r^*(w_e,\varphi)\right] \frac{(p^*)^{-\sigma-1}}{\varphi^2} g(\varphi)d\varphi}{\frac{1-\tau+b}{\sigma A}\left(\frac{\sigma-1}{\sigma}\right)\rho B w_e^{\rho-1} + \int_{\varphi^{min}}^{\varphi^{max}} \frac{(p^*)^{-\sigma-1}}{\varphi^2} g(\varphi)d\varphi} \tag{A21}$$

where $w_r^*(w_e,\varphi)$ is the optimal reported wage corresponding to $w_u^*(\varphi)$ (i.e. $w_r^*(w_e,\varphi) = (w_e - w_u^*(\varphi))/(1-\tau+b)$.)

The incidence depends in part on the labor-supply elasticity, ρ. In the limit as the labor supply elasticity becomes infinite (and hence B also becomes infinite), we have $\lim_{\rho\to\infty} \frac{dw_e}{db} = 0$. For finite positive values of ρ, we have $\frac{dw_e}{db} > 0$. That is, for a finite labor-supply elasticity, some of the increase in pension benefits will redound to workers. The upper bound on the incidence on workers is obtained as labor supply becomes perfectly inelastic:

$$\lim_{\rho\to 0} \frac{dw_e}{db} = \int_{\varphi^{min}}^{\varphi^{max}} \mu(\varphi)\left[w_r^*(w_e,\varphi)\right] g(\varphi)d\varphi \equiv \overline{w}_r^*(w_e) \tag{A22}$$

where

$$\mu(\varphi) = \frac{\left(\frac{(p^*)^{-\sigma-1}}{\varphi^2}\right)}{\int_{\varphi^{min}}^{\varphi^{max}} \left(\frac{(p^*)^{-\sigma-1}}{\varphi^2}\right) g(\varphi)d\varphi} \tag{A23}$$

That is, for a given set of active firms, the upper bound on $\frac{dw_e}{db}$ is equal to a weighted average of the equilibrium reported wages across firms, with weights $\mu(\varphi)$.

Now consider the effect of the pension reform on the (firm-specific) reported wage and net wage. Formally, using (A19) and (A21), it can be shown that:

$$\frac{dw_r^*}{db} = \frac{1}{\varphi c''(w_u^*(\varphi))(1-\tau+b)^2} + \frac{1}{1-\tau+b}\left\{\frac{dw_e}{db} - w_r^*(w_e,\varphi)\right\} \tag{A24}$$

$$\frac{dw_{net}^*}{db} = -\frac{b}{\varphi c''(w_u^*(\varphi))(1-\tau+b)} + \frac{1-\tau}{1-\tau+b}\left\{\frac{dw_e}{db} - w_r^*(w_e,\varphi)\right\} \tag{A25}$$

In general, without imposing additional assumptions on the distribution of firm productivities and/or other parameters, the signs of the responses are ambiguous.

There is a special case in which the response of the net wage can be signed, namely when firms are homogeneous. In the limit as $\rho \to 0$, which yields the upper bound on $\frac{dw_{net}^*}{db}$,

$$\lim_{\rho\to 0} \frac{dw_{net}^*}{db} = -\frac{b}{\varphi c''(w_u^*(\varphi))(1-\tau+b)} + \frac{1-\tau}{1-\tau+b}\left\{\overline{w}_r^*(w_e) - w_r^*(w_e,\varphi)\right\} \tag{A26}$$

[20] See e.g. Bartle (1976, Theorem 31.7).

where $\overline{w}_r^*(w_e)$ is from (A22). When firms are homogeneous, the term in brackets on the right-hand side is zero and we have:

$$\frac{dw_{net}^*}{db} < -\frac{b}{\varphi c''(w_u^*(\varphi))(1-\tau+b)} < 0 \tag{A27}$$

In this case, we have a clear prediction that the net wage will fall in response to the pension reform.

When firms are heterogeneous, however, the term in brackets on the right-hand side of (A26) may be sufficiently positive that the upper bound is positive and hence the sign of $\frac{dw_{net}^*}{db}$ remains ambiguous. Intuitively, the strength of the reduction in $w_{net}^*(\varphi)$ depends on the firm's own $w_r^*(w_e, \varphi)$, while the strength of the increase in the equilibrium effective wage, w_e, depends on the aggregate increase in labor demand, which is related to the weighted-average reported wage, $\overline{w}_r^*$. For lower-productivity firms with relatively low w_r^*, the latter may dominate the former.[21]

For the reported wage, there is an additional reason for ambiguity in the sign of the incidence effect. On one hand, a higher b means that the government offers a greater benefit for a given w_r, which allows the firm to reduce its wage payment conditional on a given effective wage, w_e. On the other hand, the reform induces the firm to report a greater share of the total wage payment (i.e. raise w_r relative to w_u). The $\frac{dw_r^*}{db}$ term cannot be signed, even under the assumption of homogeneous firms.

Summarizing this subsection, we have no clear predictions for the incidence of the pension reform on the observable wage variables, w_{net} and w_r. At the same time, we emphasize again that the predictions for the response of evasion to the reform do not depend on the incidence.

[21] In principle, it is also possible to sign an average effect on $w_{net}^*(w_e, \varphi)$. Using the $\mu(\varphi)$ weights from (A23), define:

$$\overline{\frac{dw_{net}^*}{db}} = \int_{\varphi^{min}}^{\varphi^{max}} \mu(\varphi) \left[\frac{dw_{net}^*}{db}\right] g(\varphi) d\varphi \tag{A28}$$

Then the upper bound is:

$$\lim_{\rho \to 0} \overline{\frac{dw_{net}^*}{db}} = -\int_{\varphi^{min}}^{\varphi^{max}} \mu(\varphi) \left[\frac{b}{\varphi c''(w_u^*(\varphi))(1-\tau+b)}\right] g(\varphi) d\varphi < 0 \tag{A29}$$

and we can conclude that $\overline{\frac{dw_{net}^*}{db}} < 0$. The $\mu(\varphi)$ weights will not be observable, however, and it will not be feasible in our data to construct this average effect.

C Data Appendix

C.1 IMSS administrative data

As mentioned above, all private Mexican employers are legally required to report wages for their employees to the Mexican social security agency, *Instituto Mexicano del Seguro Social (IMSS)*. Not all employers comply; those that do not are commonly defined as being in the informal sector. The raw IMSS data can thus be considered a census of private, formal-sector, non-petroleum-sector employers and their workforces for 1985-2005.[22] (Public-sector workers and employees of the state-run oil company are covered by other insurance programs.) The IMSS data contain information on the daily wage of individuals. The wages are a measure of total compensation, called the *salario base de cotización*, which includes earnings and benefits, including payments made in cash, bonuses, commissions, room and board, overtime payments, and in-kind benefits.[23] The data are reported as a sequence of spells for each worker, with beginning and end dates. In principle it is possible to recover a wage for every individual for every day of every year. We extracted data for June 30 for each year. At the level of individuals, the data also contain information on age, sex, and state and year of the individual's first registration with IMSS. At the establishment level, the data contain information only on location and industry (using IMSS's own 4-digit industrial categories, of which there are 276.)

We impose the following criteria in constructing and cleaning the IMSS baseline sample. (1) In its internal records, IMSS classifies wage records according to different types, referred to as *modalidades*. We use only *modalidades* corresponding to permanent workers for which consistent, reliable wage figures are available.[24] (2) We require that an individual have a positive wage. (3) We treat workers in single-worker establishments as self-employed and exclude them.[25] (4) If two observations appear for the same individual, we select the standard *modalidad*[26] and the highest-wage observation within it. (5) We require that individuals be 16 years or older and 65 years or younger. (6) We drop observations with missing sex, industry, or location. (7) We drop workers outside of non-petroleum manufacturing, construction, retail, restaurants, and hotels, as explained in the main text. (8) We include workers employed by firms located in the original 16 metropolitan areas sampled in the ENEU (described below).

Over time, IMSS raised the maximum reported income in conjunction with the maximum taxable income. Initially, the maximum reported income was set at 10 times the minimum wage; this was increased to 18 times the minimum wage in 1993 and to 25 times the minimum wage in 1994. The lowest real value for the maximum reported income occurred in 1991; we impose a topcode equal this value to ensure comparability across years. Prior to 1991, IMSS allowed firms to report wages below the relevant minimum wage, and in other years a very small number of observations have wages below the minimum wage. We drop all observations with real daily wages

[22] Generally an employer identification number in the data refers to an establishment, rather than a firm. However, more than one establishment can be associated with a single identification number. We do not have access to information on which establishments belong to which firms, and throughout the paper we treat each employer with a separate identification number as a separate firm.

[23] In principle, the wage reports should include the *aguinaldo*, the holiday bonus that employers are required to pay each December, on a pro-rated basis throughout the year. IMSS has no means to enforce this requirement for wages above the minimum reportable wage, however, and it is possible that some establishments do not comply.

[24] In the internal classification system, we use *modalidades* 10, 13, and 17. This excludes rural casual laborers, self-employed individuals who are insured through IMSS, employees of rural agricultural cooperatives and credit unions, freelance workers, taxi drivers, miscellaneous public-sector workers insured through IMSS, and a number of smaller categories.

[25] Including establishments with single workers yields similar results to those reported in the main text.

[26] The order of priority of *modalidades* we impose is 10, 13, 17. See footnote 24.

below 30 pesos; observations with wages greater than 30 pesos but lower than the minimum wage are kept unchanged.

C.2 ENEU Household Data

In the *Encuesta Nacional de Empleo Urbano (ENEU)*, households are interviewed quarterly for five quarters, and then rotate out of the sample. The original ENEU sample focused on the 16 largest Mexican metropolitan areas.[27] Over time, the coverage of metropolitan areas expanded but we focus on the original 16 metropolitan areas in order to maintain a consistent sample for as many years as possible.[28]

Wages reported in the ENEU are based on the take-home (after-tax) pay reported by respondents, and include bonuses that workers receive on a regular basis (monthly or more frequently).[29] This measure differs from the wage measure in the IMSS administrative records in that it excludes workers' social security contributions and other taxes as well as the *aguinaldo*, the holiday bonus employers are required to pay in December of each year.

The hourly wage figures were constructed as follows. (1) We recovered monthly wages for the job worked last week (as converted from hourly, daily, weekly or bi-weekly basis by INEGI enumerators). For a small number of workers, wages were reported in wage categories relative to the minimum wage; we calculated monthly wages for these workers using the midpoint of the wage categories.[30] Individuals who reported not working in the previous week or who were missing wage information were dropped. (2) Monthly hours were calculated as 4.3 times hours worked in the previous week. Responses of "irregular hours, less than 35", "irregular hours, between 35 and 48" and "irregular hours, more than 48" were assigned values of 20, 42 and 60 hours per week, respectively. Workers with missing weekly hours were dropped. (3) The hourly wage was calculated as monthly wage divided by monthly hours. (4) The daily wage was calculated by multiplying the hourly wage by eight. (5) The wage was deflated to constant 2002 pesos using the main consumer price index (INPC) from *Banco de Mexico*, the Mexican central bank. (6) We

[27] The 16 metropolitan areas are: Mexico City, Guadalajara, Monterrey, Puebla, Leon, Torreon, San Luis Potosi, Merida, Chihuahua, Tampico, Orizaba, Veracruz, Ciudad Juarez, Tijuana, Matamoros and Nuevo Laredo. Mexico also had a nationally representative survey that covered rural areas, the *Encuesta Nacional de Empleo (ENE)*, but until it was combined with the ENEU in 2000 it was carried out at less regular intervals: 1991, 1993, 1995, 1996, 1997, 1998, and 1999. Beginning in 2000, the combined ENEU/ENE survey was referred to as the *Encuesta Nacional de Empleo Trimestral (ENET)* and following a redesign in 2004 it has been known as the *Encuesta Nacional de Ocupación y Empleo (ENOE)*.

[28] The 16 metropolitan areas gradually expanded to include additional municipalities over our sample period. We include the new municipalities (and include establishments in the IMSS data in the new municipalities.)

[29] The wording of the relevant question (7a on 1994 survey) is: "En el trabajo principal de la semana pasada ... ¿[c]uánto ganó o en cuánto calcula sus ingresos? (Asegúrese de que la cantidad sea lo que la persona recibe efectivamente.)" Translation: "In the job you worked last week ... how much did you earn or how much do you calculate your income to be? (Assure yourself that the quantity is what the person actually received.)" The phrase in parentheses is an instruction to enumerators. The manual for enumerators goes into greater detail (p. 225): "Para los trabajadores a sueldo, marcará la opción que corresponda al periodo de pago y la cantidad que debe anotar es el INGRESO NETO, lo que realmente recibió y pudo disponer la última vez que le pagaron en su trabajo, es decir, no debe anotar los descuentos hechos por impuestos, cuotas al seguro social, ... " Translation: "For salaried workers, mark the pay period and the quantity that should be noted is the NET INCOME, that which the respondent really received and could spend the last time he/she was paid, that is to say, you should not include taxes, contributions to social security ..."

[30] Prior to 1994, the categories were (for some wage w and the relevant minimum wage MW): $w < .25 * MW$; $.25 * MW \leq w < .5 * MW$; $.5 * MW < w \leq 1 * MW$; $1 * MW < w \leq 2 * MW$; $2 * MW < w \leq 3 * MW$; $3 * MW < w \leq 5 * MW$; $5 * MW < w \leq 10 * MW$; and $10 * MW < w$. In 1994, the survey combined the first three categories into a single "less than the minimum wage" category and added a category of greater than 20 times the minimum wage.

imposed a topcode equal to the real value of the IMSS topcode in 1991 and dropped workers with a real daily wage below 30 pesos (approximately US$3).

The ENEU asks individuals about their main job as well as any secondary employment. We focus solely on the main job reported in the ENEU. The survey also asks number of people working in the firm/establishment of that job in the previous week; we use this information to construct the firm size variable.[31] We follow sample selection criteria similar to those we impose on the IMSS. Our baseline ENEU sample includes men ages 16-65 working in non-petroleum manufacturing, construction, retail, restaurants, and hotels. We also drop self-employed workers, owners, and unpaid workers (retaining workers who receive a fixed wage, work for a commission, or work in a cooperative). The sample thus constructed is referred to as the "full ENEU sample" in Table 2. As explained in Section 4, in our ENEU baseline sample we focus on workers who report receiving IMSS through their main employment and who work full-time, defined as at least 35 hours in the last week.

C.3 Further Comparison of IMSS and ENEU Samples

To further explore the comparability of the IMSS and ENEU samples, Appendix Table A6 compares the distributions in each sample across two dimensions that will be important in our analysis, age and firm size. We group individuals into five age categories (ages 16-25, 26-35, 36-45, 46-55, 56-65). Comparing the rightmost columns for the two panels, which indicate the share of employment in each firm size category as a share of total employment, it appears that firm sizes in the ENEU are skewed slightly away from the smallest and toward the largest size category (although there is non-monotonicity at intermediate sizes.) This may be because respondents in the household survey do not distinguish between employees directly hired by their employer and sub-contracted employees, or simply that respondents systematically overestimate employment. It may also be that firms under-report employment to IMSS, although the patterns of employment differences in Table 2 and Appendix Figure A4 tend to cast doubt on this interpretation. The distributions of employment across age groups conditional on a particular firm-size category also reveal some differences. In general, in the ENEU it appears that employment in smaller firms is shifted a bit toward younger workers relative to the IMSS (with the opposite shift among larger firms). But the overall distributions across age categories (in the "all firm sizes" rows) appear to be fairly similar.

D Additional Results

D.1 Differential Effect of Pension Reform on Employment

Table A7 reports regressions similar to those in Table 4, but where the outcome variable is the difference in log employment — what we can call the "employment gap" — between the ENEU and IMSS baseline samples within an age group-metro area cell. There is no evidence of a differential change in the employment gap for older workers in response to the reform.[32] It appears that the

[31] The exact wording is "¿Cuántas persons, incluyendo el dueño, trabajaron en el negocio, empresa o establecimiento de [su trabajo principal] la samana pasada?" [How many people, including the owner, worked in the business, firm, or establishment of [your main job] in the previous week?] We elide the distinction between establishments and firms and take the response as indicators of firm size. Also, we note that the question does not distinguish between permanent and temporary employees.

[32] Note that the employment gap here captures discrepancies in the number of workers reported to the IMSS system and the number who report on the household survey that they receive IMSS coverage; it does not capture changes in the propensity of firms to offer formal (i.e. covered by IMSS) employment. As mentioned above, we

results for wage gaps reported in Table 4 are not explained by differential changes in employment between the IMSS and ENEU data.

References

Aguila, E. (2011): "Personal Retirement Accounts and Saving," *American Economic Journal: Economic Policy*, 3(4), 1–24.

Bartle, R. G. (1976): *The Elements of Real Analysis*. John Wiley & Sons, New York, 2nd edn.

Besley, T., and T. Persson (2013): "Taxation and Development," in *Handbook of Public Economics, vol. 5*, ed. by A. Auerbach, R. Chetty, M. Feldstein, and E. Saez. Elsevier.

Chaney, T. (2008): "Distorted Gravity: The Intensive and Extensive Margins of International Trade," *American Economic Review*, 98(4), 1707 – 1721.

Dixit, A., and J. E. Stiglitz (1977): "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 67, 297–308.

Foster, L., J. Haltiwanger, and C. Syverson (2008): "Reallocation, Firm Turnover and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, 98(1), 394–425.

Galiani, S., and F. Weinschelbaum (2012): "Modeling Informality Formally: Households and Firms," *Economic Inquiry*, 50(3), 821–838.

Grandolini, G., and L. Cerda (1998): "The 1997 Pension Reform in Mexico: Genesis and Design Features," World Bank Policy Research Working Paper 1933.

Hsieh, C.-T., and P. J. Klenow (2009): "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124(4), pp. 1403–1448.

IMSS (2011): "Informe al Ejecutivo Federal y al Congreso de la Unión sobre la Situatión Financiera y los Riesgos del Instituto Mexicano del Seguro Social, 2010-2011," Mexico City, Mexico.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009): "Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries," NBER working paper no. 15218.

Lara-Ibarra, G. (2011): "Essays on the Determinants of Pension Savings and Retirement Management Decisions," Ph.D. thesis, University of Maryland.

Marrufo, G. M. (2001): "The Incidence of Social Security Regulation: Evidence from the Reform in Mexico," Ph.D. thesis, University of Chicago.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

OECD (1992): *OECD Economic Surveys: Mexico 1992*. OECD Publishing.

——— (1999): *OECD Economic Surveys: Mexico 1999*. OECD Publishing.

——— (2011): *OECD Economic Surveys: Mexico 2011*. OECD Publishing.

Rothstein, J. (2010): "Is the EITC as Good as an NIT? Conditional Cash Transfers and Tax Incidence," *American Economic Journal: Economic Policy*, 2(1), 177–208.

Slemrod, J. (2001): "A General Model of the Behavioral Response to Taxation," *International Tax and Public Finance*, 8, 119–128.

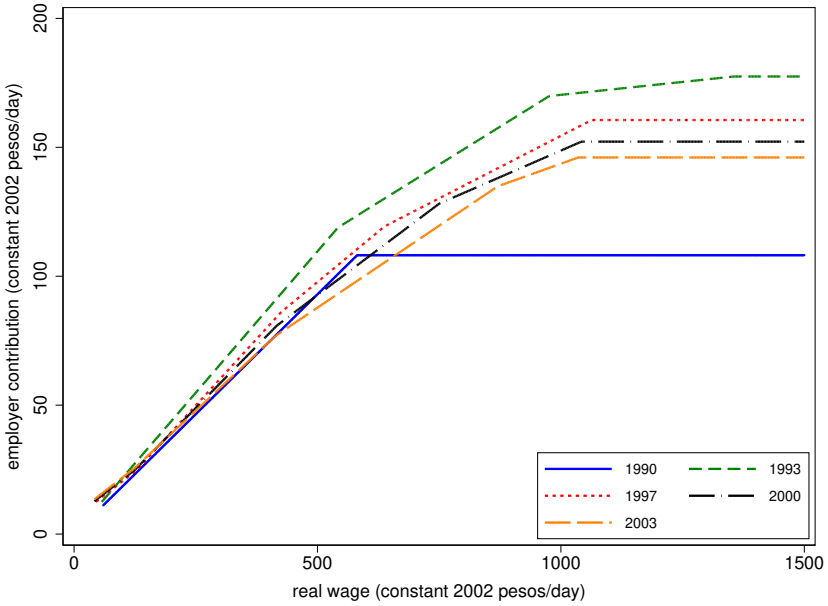
——— (2008): "Does It Matter Who Writes the Check to the Government? The Economics of Tax Remittance," *National Tax Journal*, 61(2), 164 – 165.

Verhoogen, E. (2008): "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 123(2), 489–530.

Yaniv, G. (1992): "Collaborated Employee-Employer Tax Evasion," *Public Finance*, 47(2), 312–321.

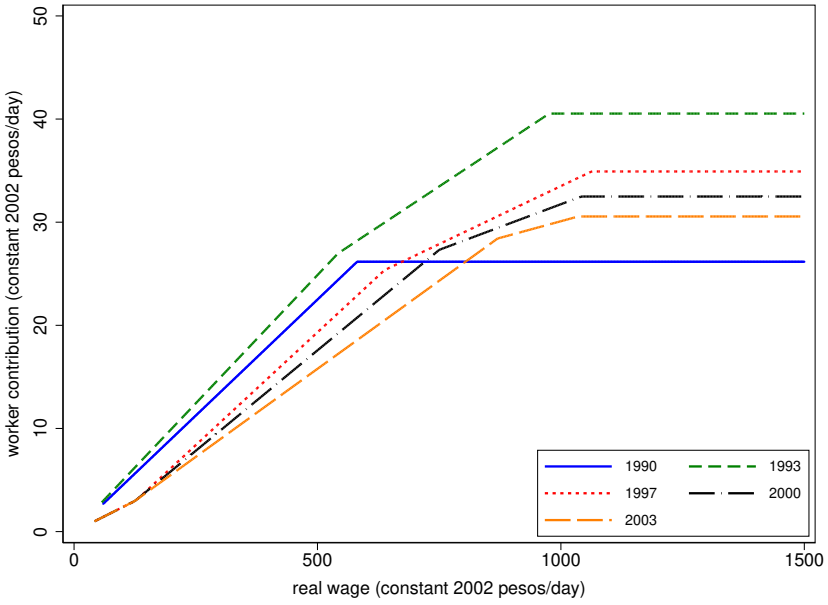
leave the investigation of the effects of the reform on the informal/formal employment margin for future work.

Figure A1. Employer contributions



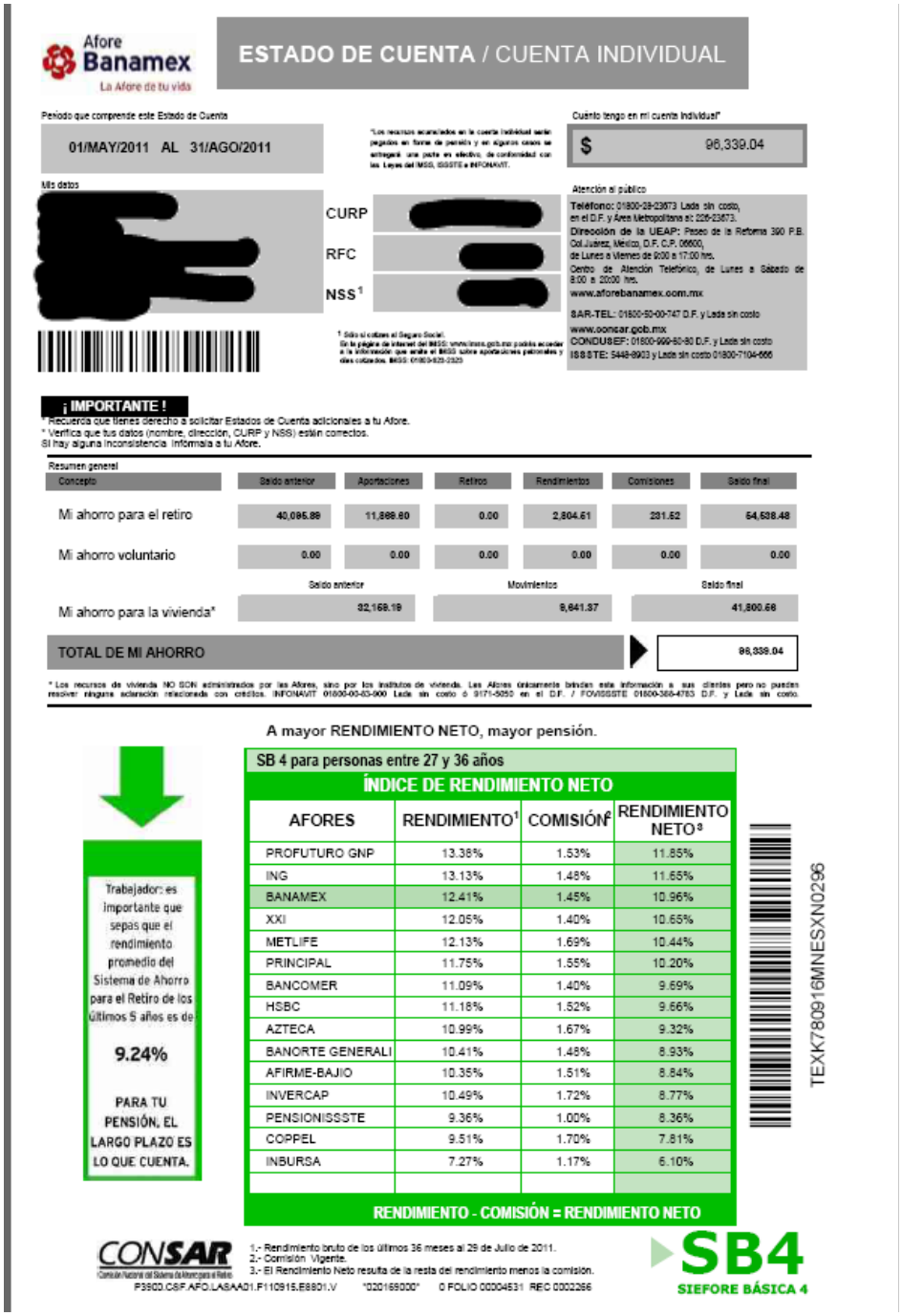
Notes: Variation in IMSS employer contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure A2. Employee contributions



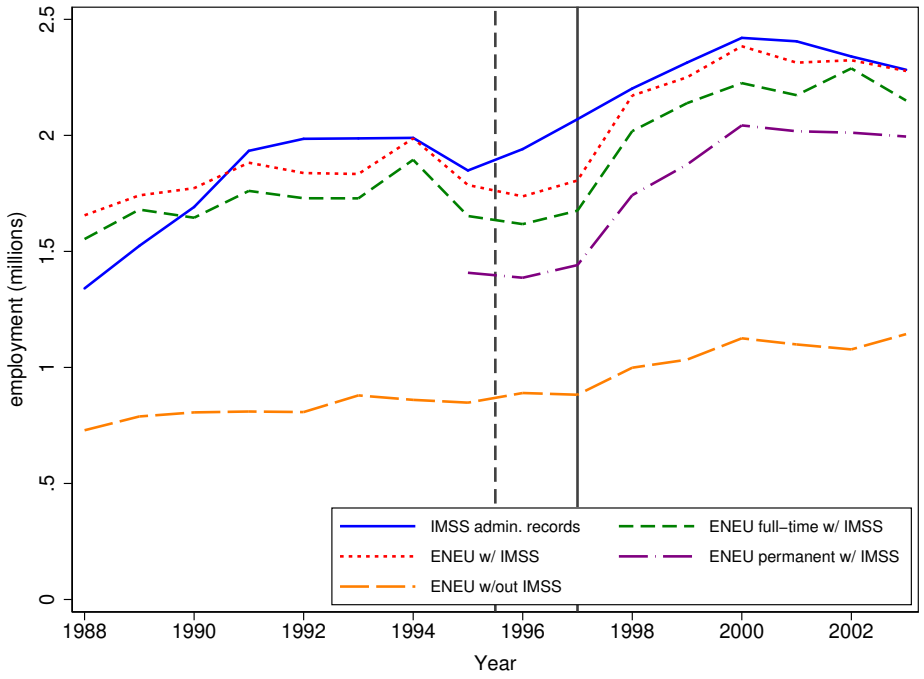
Notes: Variation in IMSS worker contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure A3. Account statement



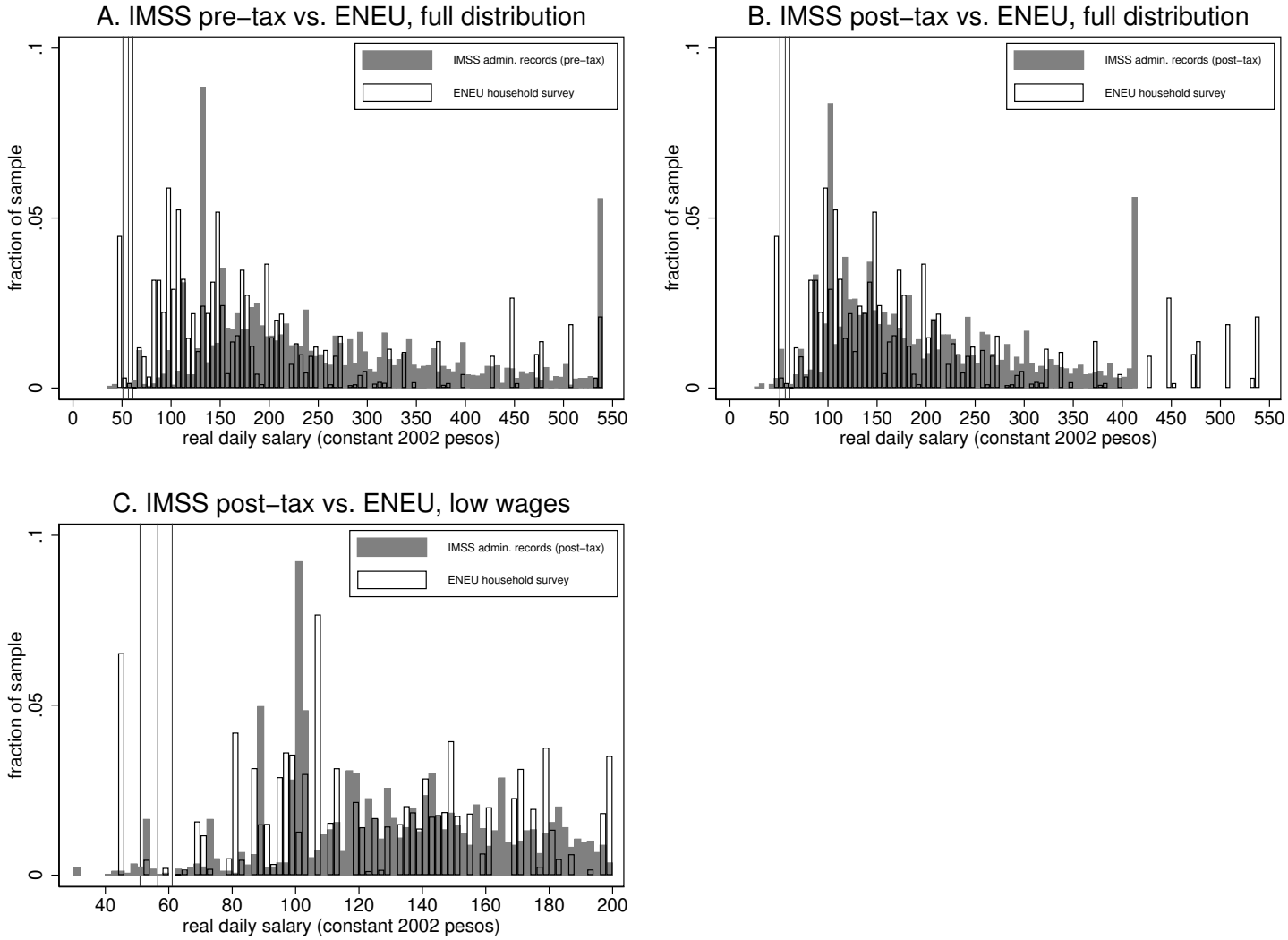
Notes: The box at top right ("Cuánto tengo en mi cuenta individual") reports total balance. The first row of boxes in the middle section ("Mi ahorro para el retiro") pertains to the retirement pension and reports previous balance ("Saldo anterior"), new contributions ("Aportaciones"), withdrawals ("Retiros"), interest earned ("Rendimientos"), AFORE commission charged ("Comisiones"), and final balance ("Saldo final"). The second and third rows in the middle section report balances in the individual's voluntary savings account and housing account. The bottom section reports 3-year returns and commissions for each AFORE, as well as the average 5-year net return (at left).

Figure A4. Employment, IMSS admin. records vs. ENEU household data, men



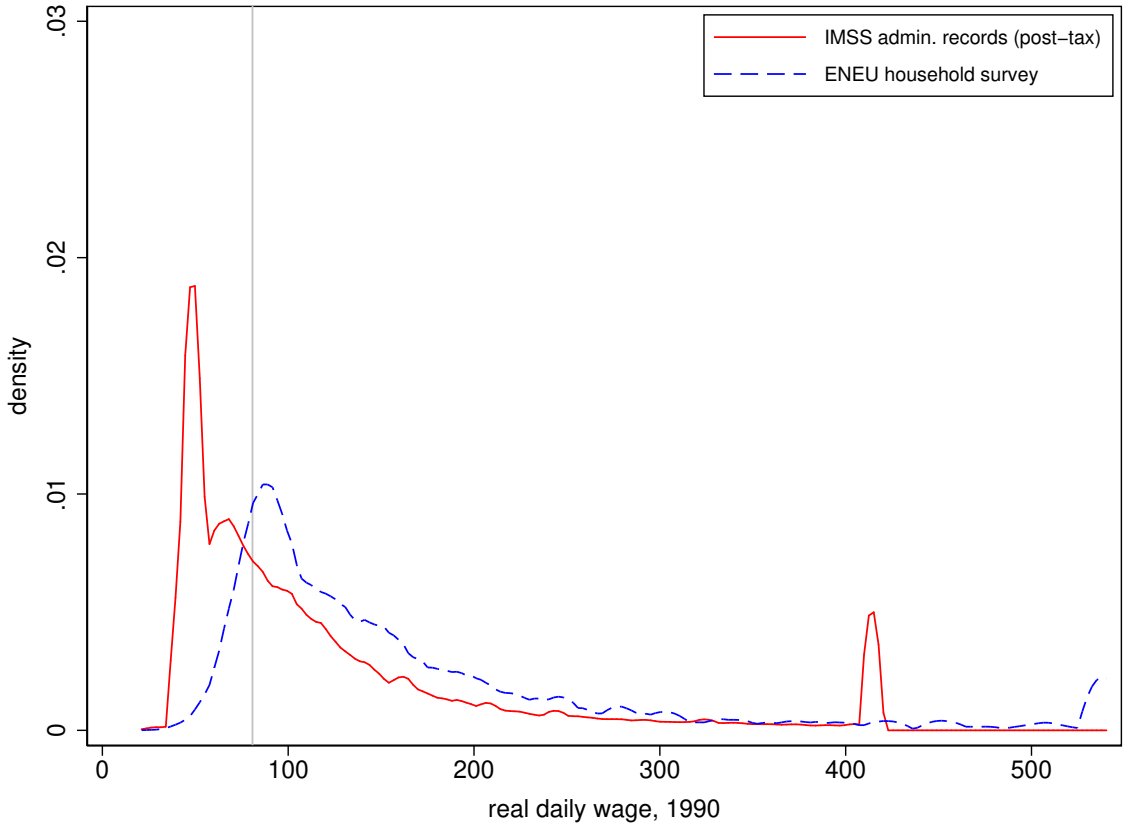
Notes: Samples are the same as those in Columns 1 and 3-6 of Table 2; refer to that table for details. ENEU totals are calculated using sampling weights. The dashed vertical line indicates the date the pension reform was passed by Congress (Dec. 21, 1995); the solid vertical line indicates the date the reform took effect (July 1, 1997). Observations correspond to the second quarter of each year. See Section 4 and Appendix C for details of sample selection.

Figure A5. Wage histograms, 1990, social security employees



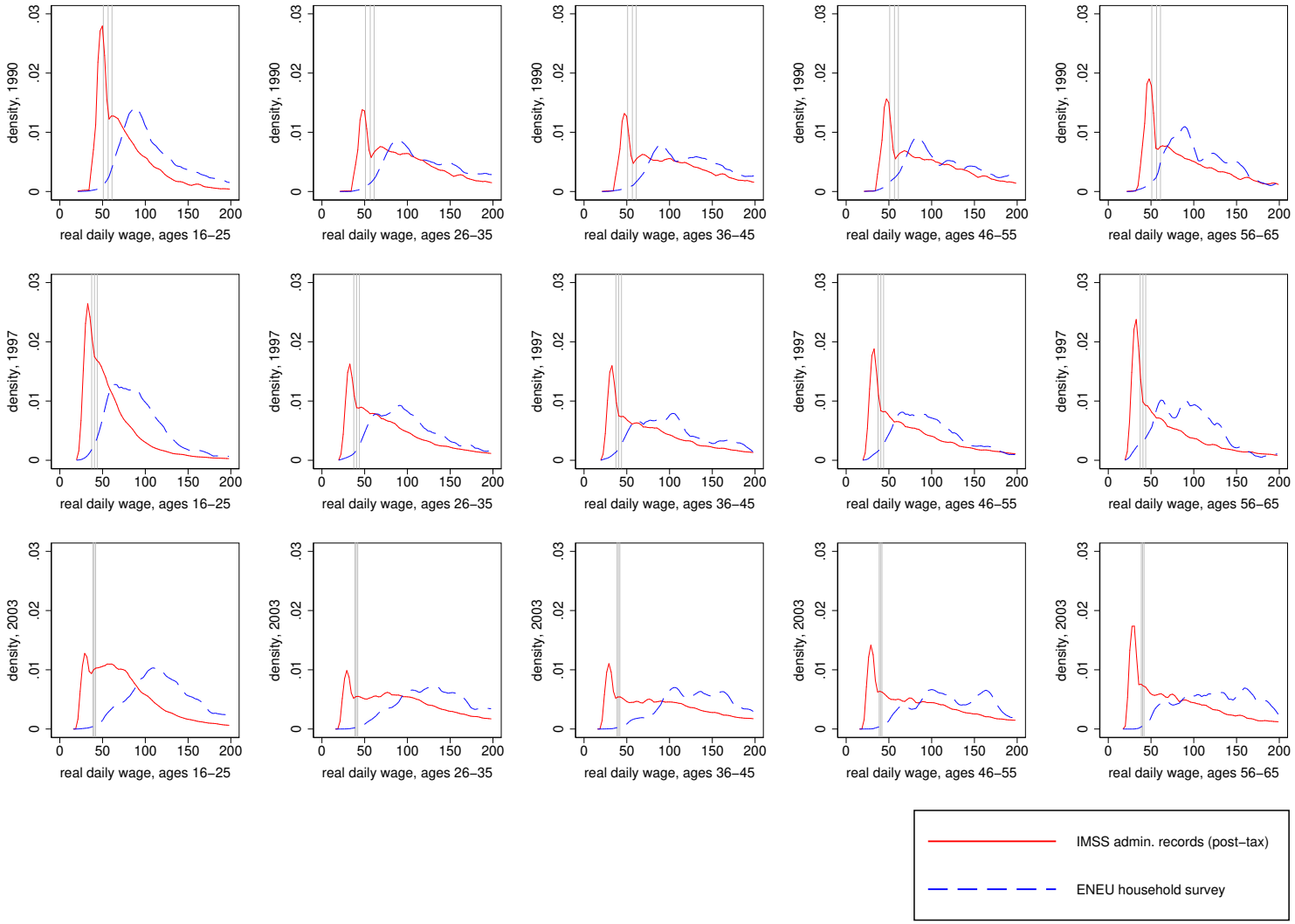
Notes: Histogram A is similar to Figure 2, comparing the IMSS *pre-tax* wage to the ENEU take-home wage, but for male employees in the "social security" sector. Because of small sample size, ENEU observations are pooled across quarters for 1990. Histogram B compares the IMSS *post-tax* wage to the ENEU take-home wage, and histogram C makes the same comparison for wages below 200 pesos/day. Bins in histograms A and B are 5 pesos wide, and in histogram C are 2 pesos wide. See notes to Figure 2, Section 4, and Appendix C for further details.

Figure A6. Excess mass calculation



Notes: The wage variables are the real daily take-home wage from ENEU and real daily *post-tax* reported wage from IMSS. Densities are estimated using 1990 data and an Epanechnikov kernel with bandwidth 3 pesos for IMSS data and 6 pesos for ENEU data (using Stata kdensity command). Wages are in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. (Densities are smoothed versions of histograms in Figure 2.) Vertical line is at 15th percentile of the ENEU wage distribution. Data are from second quarter of 1990. Excess mass for 15th percentile defined as (area under red, left of vertical line) - (area under blue, left of vertical line). Excess mass defined analogously for other percentiles. See Section 4 and Appendix C for further details of data processing.

Figure A7. Wage densities by age group, 1990, 1997, 2003



Notes: The wage variables are the real daily take-home wage from ENEU and real daily *post-tax* reported wage from IMSS. Densities are estimated and an Epanechnikov kernel with bandwidth 3 pesos for IMSS data and 6 pesos for ENEU data (using Stata kdensity command). Wages in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. Rows correspond to years 1990, 1997, 2003; columns to age groups 16-25, 26-35, 36-45, 46-55, 56-65. Samples are IMSS and ENEU baseline samples of men, only including workers with wages less than 200 pesos/day. Data in both samples are from second quarter. See Section 4 and Appendix C for further details.

Table A1. IMSS Contribution Rates, Pensions and Child Care

	General pension fund				Personal retirement account				Childcare			
	E	W	G	TC	E	W	G	TC	E	W	G	TC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Jan 01 1973 - Jun 28 1986	3.750	1.500	0.750	10	0.000	0.000	0.000	10	1.000	0.000	0.000	10
Jun 29 1986 - Jan 04 1989	4.200	1.500	0.300	10	0.000	0.000	0.000	10	1.000	0.000	0.000	10
Jan 05 1989 - Dec 31 1990	4.200	1.500	0.300	10	0.000	0.000	0.000	10	1.000	0.000	0.000	10
Jan 01 1991 - Dec 31 1991	4.900	1.750	0.350	10	0.000	0.000	0.000	10	1.000	0.000	0.000	10
Jan 01 1992 - Apr 30 1992	5.040	1.800	0.350	10	0.000	0.000	0.000	10	1.000	0.000	0.000	10
May 01 1992 - Dec 31 1992	5.040	1.800	0.360	10	2.000	0.000	0.000	25	1.000	0.000	0.000	10
Jan 01 1993 - Jul 20 1993	5.180	1.850	0.370	10	2.000	0.000	0.000	25	1.000	0.000	0.000	10
Jul 21 1993 - Dec 31 1993	5.180	1.850	0.370	10	2.000	0.000	0.000	25	1.000	0.000	0.000	18
Jan 01 1994 - Dec 31 1994	5.670	2.025	0.405	10	2.000	0.000	0.000	25	1.000	0.000	0.000	25
Jan 01 1995 - Dec 31 1995	5.810	2.075	0.415	10	2.000	0.000	0.000	25	1.000	0.000	0.000	25
Jan 01 1996 - Jun 30 1997	5.950	2.125	0.425	10	2.000	0.000	0.000	25	1.000	0.000	0.000	25
Jul 01 1997 - Jun 30 1998	2.800	1.000	0.200	15	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 1998 - Jun 30 1999	2.800	1.000	0.200	16	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 1999 - Jun 30 2000	2.800	1.000	0.200	17	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2000 - Jun 30 2001	2.800	1.000	0.200	18	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2001 - Jun 30 2002	2.800	1.000	0.200	19	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2002 - Jun 30 2003	2.800	1.000	0.200	20	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2003 - Jun 30 2004	2.800	1.000	0.200	21	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2004 - Jun 30 2005	2.800	1.000	0.200	22	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2005 - Jun 30 2006	2.800	1.000	0.200	23	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2006 - Jun 30 2007	2.800	1.000	0.200	24	5.150	1.125	0.225	25	1.000	0.000	0.000	25
Jul 01 2007 - .	2.800	1.000	0.200	25	5.150	1.125	0.225	25	1.000	0.000	0.000	25

Notes: Numbers represent contribution rates as a percentage of a worker's daily wage. E = employer contribution, W = worker contribution, G = government contribution, and TC = topcode, representing the maximum taxable income for the purposes of each calculation, as a multiple of the minimum wage in Mexico City. The "General pension fund" columns (columns 1-4) indicate contributions to the general IMSS pension fund; The "Personal retirement account" columns (columns 5-8) indicate contributions to individuals' personal retirement accounts. See Table A2 for details on health care contributions.

Table A2. IMSS Contribution Rates, Health Care

	Variable component				Fixed component				Additional for wage>3xMW			
	E	W	G	TC	E	W	G	TC	E	W	G	TC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Jan 01 1973 - Jun 28 1986	5.625	2.250	1.125	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jun 29 1986 - Jan 04 1989	6.300	2.250	0.450	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jan 05 1989 - Dec 31 1990	8.400	3.000	0.600	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jan 01 1991 - Dec 31 1991	8.400	3.000	0.600	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jan 01 1992 - Apr 30 1992	8.400	3.000	0.600	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
May 01 1992 - Dec 31 1992	8.400	3.000	0.600	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jan 01 1993 - Jul 20 1993	8.400	3.000	0.625	10	0.000	0.000	0.000	10	0.000	0.000	0.000	10
Jul 21 1993 - Dec 31 1993	8.750	3.125	0.625	18	0.000	0.000	0.000	18	0.000	0.000	0.000	18
Jan 01 1994 - Dec 31 1994	8.750	3.125	0.625	25	0.000	0.000	0.000	25	0.000	0.000	0.000	25
Jan 01 1995 - Dec 31 1995	8.750	3.125	0.625	25	0.000	0.000	0.000	25	0.000	0.000	0.000	25
Jan 01 1996 - Jun 30 1997	8.750	3.125	0.625	25	0.000	0.000	0.000	25	0.000	0.000	0.000	25
Jul 01 1997 - Jun 30 1998	0.700	0.250	0.050	25	13.900	0.000	13.900	25	6.000	2.000	0.000	25
Jul 01 1998 - Jun 30 1999	0.700	0.250	0.050	25	14.550	0.000	13.900	25	5.510	1.840	0.000	25
Jul 01 1999 - Jun 30 2000	0.700	0.250	0.050	25	15.200	0.000	13.900	25	5.020	1.680	0.000	25
Jul 01 2000 - Jun 30 2001	0.700	0.250	0.050	25	15.850	0.000	13.900	25	4.530	1.520	0.000	25
Jul 01 2001 - Jun 30 2002	0.700	0.250	0.050	25	16.500	0.000	13.900	25	4.040	1.360	0.000	25
Jul 01 2002 - Jun 30 2003	0.700	0.250	0.050	25	17.150	0.000	13.900	25	3.550	1.200	0.000	25
Jul 01 2003 - Jun 30 2004	0.700	0.250	0.050	25	17.800	0.000	13.900	25	3.060	1.040	0.000	25
Jul 01 2004 - Jun 30 2005	0.700	0.250	0.050	25	18.450	0.000	13.900	25	2.570	0.880	0.000	25
Jul 01 2005 - Jun 30 2006	0.700	0.250	0.050	25	19.100	0.000	13.900	25	2.080	0.720	0.000	25
Jul 01 2006 - Jun 30 2007	0.700	0.250	0.050	25	19.750	0.000	13.900	25	1.590	0.560	0.000	25
Jul 01 2007 - .	0.700	0.250	0.050	25	20.400	0.000	13.900	25	1.100	0.400	0.000	25

Notes: E = employer contribution, W = worker contribution, G = government contribution, and TC = topcode, representing the maximum taxable income for the purposes of each calculation, as a multiple of the minimum wage in Mexico City. The "Variable component" columns (columns 1-4) indicate contribution rates as a percentage of a worker's daily wage. The "Fixed component" columns (columns 5-8) indicate fixed contribution rates for each worker, shown as a percentage of the minimum wage in Mexico City. The variable and fixed components (columns 1-8) must be paid for all workers. The "Additional for wage>3xMW" columns (columns 9-12) indicates additional contributions as a percentage of a worker's daily wage exceeding three times the minimum wage in Mexico City, applicable only for workers whose daily wage exceeds that value. See Table A1 for details on pension and child care contributions.

Table A3. IMSS Pension Benefit Schedule, 1/1/1991 to 6/30/1997

Wage as multiple of min wage	Base replacement rate (%)	Extra benefit for each additional year (%)
1 or less	80.00	0.563
1.01 to 1.25	77.11	0.814
1.26 to 1.5	58.18	1.178
1.51 to 1.75	49.23	1.430
1.76 to 2	42.67	1.615
2.01 to 2.25	37.65	1.736
2.26 to 2.5	33.68	1.868
2.51 to 2.75	30.48	1.958
2.76 to 3	27.83	2.033
3.01 to 3.25	25.60	2.096
3.26 to 3.5	23.70	2.149
3.51 to 3.75	22.07	2.195
3.76 to 4	20.65	2.235
4.01 to 4.25	19.39	2.271
4.26 to 4.5	18.29	2.302
4.51 to 4.75	17.30	2.330
4.76 to 5	16.41	2.355
5.01 to 5.25	15.61	2.377
5.26 to 5.5	14.88	2.398
5.51 to 5.75	14.22	2.416
5.76 to 6	13.62	2.433
6.01 or higher	13.00	2.450

Notes: Base replacement rate is based on average nominal wage in the five years preceding retirement. Extra benefit is for each year of contribution exceeding the minimum contribution length of ten years.

Table A4. Inflation rate and minimum wage over time

Year	Inflation rate	Real value of minimum wage
1983	112.50	118.01
1984	67.14	110.16
1985	53.43	109.99
1986	83.17	99.20
1987	126.73	77.54
1988	135.81	71.88
1989	17.58	66.02
1990	26.11	61.08
1991	23.10	58.57
1992	15.85	56.63
1993	9.87	55.18
1994	6.85	55.26
1995	37.72	48.09
1996	31.82	45.05
1997	20.35	43.81
1998	15.31	43.38
1999	17.39	42.15
2000	9.41	42.39
2001	6.57	42.34
2002	4.94	42.15
2003	4.27	41.86
2004	4.37	41.57
2005	4.33	41.22

Notes: Data from second quarter of each year. Real minimum wages are for daily wages in Mexico City (zone A), reported in constant 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar.

Table A5. Pension wealth simulation, male worker entering system on June 30, 1997

Years of Contributions	Plan	Real Daily Wage					
		43	100	200	300	500	1079
35	PRA	398.6	815.0	1626.2	2437.3	4059.7	8751.9
	PAYGO	398.6	398.6	603.8	890.2	1483.6	3200.1
30	PRA	398.6	523.4	1044.3	1565.3	2607.1	5620.5
	PAYGO	398.6	398.6	510.7	743.3	1238.9	2672.1
25	PRA	398.6	398.6	659.1	987.8	1645.3	3546.9
	PAYGO	398.6	398.6	406.9	579.5	965.8	2083.2
20	PRA	87.9	202.4	403.9	605.4	1008.4	2173.9
	PAYGO	398.6	398.6	398.6	449.6	749.3	1616.2
15	PRA	51.1	117.8	235.0	352.2	586.6	1264.7
	PAYGO	398.6	398.6	398.6	398.6	504.5	1088.2
10	PRA	26.8	61.7	123.1	184.5	307.4	662.6
	PAYGO	398.6	398.6	398.6	398.6	398.6	560.3
5	PRA	10.7	24.6	49.0	73.5	122.4	264.0
	PAYGO	0.0	0.0	0.0	0.0	0.0	0.0

Notes: Values represent real present discounted value of the future stream of pension benefits in thousands of 2002 pesos under the pre-reform pay-as-you-go (PAYGO) and personal retirement account (PRA) systems, for a male worker who enters the system on June 30, 1997. 43 pesos is real daily minimum wage (in Mexico City) in 1997, 1,079 pesos is the topcode we impose (corresponding to the lowest real value of IMSS topcode over study period.) See Sections 2.2 and 2.3 and Appendix A.6 for further details.

Table A6. Age composition by firm size category, 1990, men

	Age category (employment as % of row total)					
	16-25	26-35	36-45	46-55	56-65	employment as % of column total
A. IMSS						
1-10 employees	29.9	32.6	19.8	11.9	5.8	14.5
11-50 employees	33.6	32.2	18.7	10.6	4.9	22.6
51-100 employees	35.0	32.5	18.5	9.8	4.2	10.8
101-250 employees	36.3	33.3	17.8	9.0	3.5	14.7
$>$ 250 employees	37.7	34.8	17.5	7.6	2.5	37.5
all firm sizes	35.1	33.4	18.3	9.3	3.8	
B. ENEU						
1-10 employees	35.9	28.3	18.0	12.5	5.3	12.4
11-50 employees	33.5	33.3	18.4	10.3	4.5	21.0
51-100 employees	35.6	33.4	15.2	10.7	5.1	11.6
101-250 employees	30.2	31.2	21.5	12.4	4.7	10.5
$>$ 250 employees	34.0	33.4	21.5	8.5	2.7	44.5
all firm sizes	33.9	32.5	19.7	10.1	3.9	

Notes: Data are from IMSS and ENEU baseline samples. Percentages are calculated based on employment (using sampling weights, in the case of the ENEU) in each cell. Panel B drops observations in ENEU baseline sample that are missing the firm-size variable (which make up less than 1% of sample). Panel B uses ENEU sampling weights. For further details, see Section 4 and Appendix C.

Table A7. Differential effects of pension reform on employment gap

	dep. var.: log(empl., ENEU) - log(empl., IMSS)	
	(1)	(2)
1(age <= 45)*1988	0.104	0.098
	(0.067)	(0.065)
1(age <= 45)*1989	0.028	0.022
	(0.059)	(0.053)
1(age <= 45)*1990	-0.042	-0.048
	(0.053)	(0.053)
1(age <= 45)*1991	-0.088	-0.093
	(0.068)	(0.062)
1(age <= 45)*1992	-0.031	-0.036
	(0.056)	(0.051)
1(age <= 45)*1993	-0.018	-0.022
	(0.058)	(0.051)
1(age <= 45)*1994	0.051	0.048
	(0.080)	(0.082)
1(age <= 45)*1995	-0.024	-0.025
	(0.048)	(0.038)
1(age <= 45)*1996	-0.016	-0.016
	(0.078)	(0.071)
1(age <= 45)*1998	0.096	0.098
	(0.060)	(0.060)
1(age <= 45)*1999	-0.029	-0.026
	(0.055)	(0.048)
1(age <= 45)*2000	-0.014	-0.010
	(0.062)	(0.053)
1(age <= 45)*2001	-0.055	-0.051
	(0.063)	(0.061)
1(age <= 45)*2002	-0.022	-0.020
	(0.072)	(0.069)
1(age <= 45)*2003	0.045	0.046
	(0.050)	(0.046)
age group effects	Y	
age group-metro area effects	N	Y
metro-year effects	Y	Y
R-squared	0.61	0.71
N	1280	1280

Notes: Samples are IMSS and ENEU baseline samples of men, collapsed to metro area/age group/year level. Regressions are weighted by IMSS employment in each cell. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C for further details of data processing.