

Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico*

Todd Kumler[†]

Eric Verhoogen[‡]

Judith Frías[§]

June 2012

Abstract

Non-compliance of firms with tax regulations is a major constraint on state capacity in developing countries. We focus on an arguably under-appreciated dimension of non-compliance: under-reporting of wages by formal firms to evade payroll taxes. Comparing wage distributions for similar sets of workers in the administrative records of the Mexican social security agency and a household labor-force survey, we document extensive under-reporting of wages. We further argue that the 1997 Mexican pension reform had a differential effect by age on the incentives of workers to ensure that their wages were reported accurately. Using a difference-in-differences strategy, we present evidence that the increase in the incentive for workers to ensure accurate reports led to a significant decline in under-reporting. The results suggest that enlisting workers in monitoring their employers is an effective way to increase payroll tax compliance.

* We would like to thank (without implicating) Emma Aguila for sharing information on tax rates and benefit formulas, Matthew Clifford for research assistance, and David Card, David S. Kaplan, Wojciech Kopczuk, Sebastian Galiani, Kiki Pop-Eleches, and seminar participants at UC Berkeley, Washington University, Carnegie-Mellon University, the University of Houston, Warwick University, and the CEPR-AMID Development Economics Symposium in Milan for helpful conversations. We gratefully acknowledge the support of the National Science Foundation (SES-0721068). The views expressed in this paper are those of the authors, and do not necessarily reflect the views of the *Instituto Mexicano del Seguro Social*.

[†] Columbia University, email: tjk2110@columbia.edu.

[‡] Columbia University, email: eric.verhoogen@columbia.edu. Corresponding author.

[§] Instituto Mexicano del Seguro Social (IMSS). email: judith.frias@imss.gob.mx.

1 Introduction

A growing body of research suggests that lack of state capacity — in particular, difficulty in raising taxes to fund the provision of public goods — is a major constraint on the growth of developing countries (Burgess and Stern, 1993; Besley and Persson, 2009, 2011). Developing countries tend to have low ratios of tax revenues to GDP and large informal sectors. Mexico is no exception: it has the lowest tax revenue share of GDP in the OECD, between 15 and 20 percent during the period we study, and the informal sector is estimated to make up 40 percent or more of total output (OECD, 2011b; IMF, 2010; Schneider and Enste, 2000). The Mexican social security agency, the *Instituto Mexicano del Seguro Social* (IMSS), is in principle supposed to cover all private-sector employees; in fact it covers slightly more than half.¹ Given weak enforcement institutions and widespread evasion, the task of improving the fiscal capacity of developing-country governments is a difficult one, and there is acute policy interest in potential remedies.

A key element of the general weakness of fiscal capacity is non-compliance of firms with tax regulations. A large literature has focused on one dimension of non-compliance: the failure of firms to register with tax authorities. Researchers have argued that this form of non-compliance generates a variety of market distortions, including limits on informal firms' employment growth and access to formal credit markets (Gordon and Li, 2009; La Porta and Shleifer, 2008; Levy, 2008).² Governments in a number of countries have implemented programs to reduce registration costs and induce firms to formalize.³

In this paper, we focus on a different dimension of non-compliance by firms, less appreciated but arguably no less important: the under-reporting of wages by registered firms to evade payroll taxes. This form of non-compliance has received surprisingly little empirical attention. One reason may be that it has been shown not to be a significant issue in developed countries. For instance, using audits of individual tax returns in Denmark, Kleven et al (2011) find little evasion when

¹In 2000, the ratio of remunerated employees (*trabajadores asalariados*) covered by IMSS to remunerated private sector employees in the Mexican population census was 0.528. Sources: IMSS (2009) for numerator and INEGI (2001), using information from 2000 population census, for denominator. We note that IMSS uses different series to calculate coverage in its published statistics; see details in Appendix B.1 (online).

²Notable theoretical work on firms' decisions about whether to formalize includes Rauch (1991), De Paula and Scheinkman (2011), and Galiani and Weinschelbaum (forthcoming).

³See for instance Bruhn (2011) and Kaplan, Piedra, and Seira (forthcoming) on the SARE program in Mexico and Fajnzylber, Maloney, and Montes-Rojas (2011) and Monteiro and Assunção (forthcoming) on the SIMPLES program in Brazil. A notable recent study by de Mel, McKenzie, and Woodruff (2012) randomized cash payments to induce small Sri Lankan firms to register, finding strong effects on registration but generally modest effects of registration on other firm-level outcomes. The paper finds significant effects on mean profitability, which are driven by a few fast-growing firms, and on the level of trust in the state expressed by entrepreneurs. In related work, McKenzie and Sakho (2010) find empirically that registering leads firms to be more profitable on average, but also find significant heterogeneity in the effects.

incomes are reported by employers or other third parties.⁴ The view that third-party reporting is effective in ensuring compliance is widespread among practitioners and government agencies in developed countries (see e.g. OECD (2006)). Another reason for the limited attention to wage under-reporting may be that it is difficult to study. It has been rare for researchers to have micro-level information on firms' wage reports, and rarer still to have access to an alternative source of wage information at a sufficiently disaggregated level to permit inferences about the extent of non-compliance (Slemrod and Yitzhaki, 2002, Sec. 4.1.1). As a consequence, it has not been clear to what extent the accuracy of third-party reporting carries over to developing-country settings.

In this paper, we draw on two independent sources of individual-level wage information from Mexico — firms' wage reports to the Mexican social security agency and workers' responses to a household labor-force survey — to draw inferences about the extent of wage under-reporting and how it responds to incentives inherent in the social security system. We make two main points. First, comparing wage distributions from each source for similar sets of workers, we show graphically that under-reporting of wages is substantial, and the pattern is consistent with what one would expect given the schedules of social security taxes and benefits. Second, using a major pension reform in 1997 as a source of exogenous variation, we show that under-reporting responds to changes in economic incentives, and in particular to the ability and incentives of employees to ensure the accuracy of their employers' reports. On July 1, 1997, the Mexican government switched from a pay-as-you-go pension system to a system of personalized savings accounts. Prior to the reform, the pensions of a significant subset of workers were largely insensitive to the wages reported by firms, for reasons we discuss below. The reform tied individual pensions more closely to firms' wage reports and made it easier for employees to observe those reports. The change affected different age groups differently. Workers already in the pay-as-you-go system prior to July 1, 1997 retained the right to choose, at the time of retirement, the pension that they would have received under the pre-reform regime. Because older workers had little time to accumulate sufficient balances in their personal accounts, they could be confident, even at the time of the reform, that they would be better off under the old regime. Younger workers had a greater expectation of being better off under the new regime and hence had stronger incentives to ensure accurate reporting. Using a difference-in-differences strategy, we show that under-reporting

⁴In another example, Saez (2010) finds significant bunching around the first kink point of the Earned Income Tax Credit, suggesting misreporting, only among the self-employed. The Internal Revenue Service has documented that compliance is higher for income groups with greater third-party reporting in the U.S. (Internal Revenue Service, 1996, 2006).

declined relatively more for younger workers.

Our results are broadly consistent with the theoretical model of firms as fiscal intermediaries of Kleven, Kreiner, and Saez (2009). In that model, firms are cooperatives of workers who may collude in under-reporting wages to the government.⁵ There are two reasons why collusion may be more difficult to sustain in larger firms, in cases where workers are unable to commit not to “blow the whistle” on evasion. First, workers may be subject to random shocks (e.g. may become disgruntled); since it only requires one worker to signal to the authorities that a firm should be audited, collusion is more difficult to sustain in larger firms. Second, the government may offer a reward for whistle-blowing that increases in the amount of taxes evaded, in which case workers in larger firms have a greater incentive to expose evasion and firms optimally evade less. It would be straightforward to tailor the model to our setting. Although there is no explicit reward for whistle-blowing in Mexico, the shift to personal retirement accounts can be thought of as an increase in the incentive for workers to ensure accurate reporting. This increase in the reward for monitoring would be expected to increase compliance by firms employing workers affected by the change.⁶

This paper appears to be the first empirical study of the extent to which under-reporting of wages by employers responds to changes in the incentives and ability of employees to monitor them. Our findings support the idea that the design of social-insurance systems should take into account the incentives of employees to ensure accurate reporting. This argument should not be interpreted as advocating a system of personal accounts *per se*; one could imagine a change in pension benefits under the pay-as-you-go system that would have had similar effects. The key point is that giving employees greater incentives to monitor their employers and making it easier for them to do so appear to be effective ways to improve payroll tax compliance.

The argument of this paper is in the spirit of recent work on the attractive enforcement properties of value-added taxes (VATs) relative to retail sales taxes (Kopczuk and Slemrod, 2006; Keen and Lockwood, 2010; Pomeranz, 2011). VATs are thought to reduce administrative costs of enforcement in part because each party in a supply-chain transaction has an incentive to ensure that the other reports accurately. Our argument that incentivizing employees to monitor reports of their employers can improve payroll-tax compliance is analogous.⁷

⁵The possibility of collusion between employees and employers in under-reporting wages had earlier been formalized by Yaniv (1992).

⁶In the presence of disgruntlement shocks, such a model would continue to predict greater compliance in larger firms, consistent with patterns we document below.

⁷Bailey and Turner (2001) suggest that tying pension benefits to contributions in this way would have the effect

This paper is related to a small literature on misreporting of social security contributions. Nyland, Smyth, and Zhu (2006) relate the outcomes of social-security tax audits in Shanghai to firm characteristics. Tonin (2011) investigates the effect of an imposition of a minimum wage on reporting patterns in Hungary, using the relationship between household income and expenditures to draw inferences about the extent of evasion.⁸ These papers do not focus on workers' incentives to monitor their firms.⁹

This paper is related more broadly to an active recent literature on the role of firms in tax systems, discussed in more detail in Kopczuk and Slemrod (2006), Slemrod (2008), and Gordon and Li (2009), and more broadly still to a voluminous literature on tax evasion and avoidance, reviewed by Andreoni, Erard, and Feinstein (1998), Slemrod and Yitzhaki (2002), and Saez, Slemrod, and Giertz (2012).

The next section describes the Mexican social security system and the 1997 pension reform. Section 3 describes the datasets. Section 4 presents cross-sectional comparisons of the wage distributions in the different datasets. Section 5 presents the difference-in-differences analysis of the effect of the pension reform. Section 6 concludes.

2 The Mexican Social Security System

Because our empirical strategy relies crucially on incentives in the Mexican social insurance system, this section describes the system and the 1997 pension reform in some detail. The *Instituto Mexicano del Seguro Social* (IMSS), the Mexican social security agency, is the primary source of social insurance for private-sector workers in Mexico. It administers pension benefits, disability insurance, work injury compensation, childcare centers, and a large number of clinics and hospitals, which are the primary source of health care for the formal, private-sector Mexican workforce. Public-sector workers and workers for PEMEX, the state-owned oil company, are covered by separate systems. In 2003, the government created an alternative system called *Seguro Popular*, which provides basic health coverage for all individuals and is not tied to formal employment. In this paper, we focus on the IMSS system and sectors with minimal government employment.

of reducing evasion.

⁸Also using Hungarian data, Elek, Köllő, Reizer, and Szabó (forthcoming) develop a method to identify workers whose wages are likely to be under-reported, using the responses of wages to a change in the level of reported wages at which audits may be triggered.

⁹Papers that have used the general strategy of comparing information from more than one data source in order to infer evasion (in other contexts) include Fisman and Wei (2004), Gorodnichenko, Martinez-Vazquez, and Peter (2009), Marion and Muehlegger (2008), Hurst, Li, and Pugsley (2011), and Braguinsky, Mityakov, and Liscovich (2010) among others.

Beginning with its creation in 1944, IMSS operated as a pay-as-you-go (PAYGO) scheme financed by payroll taxes. By the late 1980s, however, rising health care costs and an increase in the number of pensioners relative to the working-age population led to projected shortfalls in the IMSS financial accounts. Because of concerns about the financial viability of the system, the Mexican congress enacted a first attempt at pension reform in 1992. That reform created personal retirement accounts to exist alongside the PAYGO system. The personal retirement accounts were plagued by administrative problems and did not resolve the underlying financial imbalance in the PAYGO program, however. In December 1995, the congress enacted a new, more comprehensive pension reform, to take effect on July 1, 1997. As mentioned above, this reform replaced the entire PAYGO pension system with a system of personal retirement accounts (PRA). Because of data constraints, discussed in more detail in Section 3 below, we focus on the years 1988-2003. More extensive discussions of the pension reform are provided in Grandolini and Cerda (1998), Sales-Sarrapy, Solis-Soberon, and Villagomez-Amezcuca (1996), and Aguila (2011).

In describing the characteristics of the social security system and in the empirical work below, we will focus primarily on male workers. The incentives and empirical patterns for women are complicated by the facts that women’s labor force participation changed relatively rapidly over the study period and that many women receive IMSS benefits through their spouses, which provides an incentive to remain in the informal sector. In addition, because of relatively low labor force participation by older women, sample sizes in the ENEU household survey are inadequate, especially when analyzing the data separately by metropolitan area, as explained below. We present the main tables and figures for women in Appendix C (online). To preview the results, our cross-sectional point — that there is substantial evasion — is robust for women, but our difference-in-differences results are not, possibly for the reasons just discussed.

2.1 Contribution Rates

IMSS requires contributions from both employers and employees based on reported wages; these are supplemented by government contributions. Figure 1 presents the contribution schedule for employers as a function of the reported real daily wages of each employee, for selected years. The schedule reflects a complicated set of formulas determining contributions to the various components of the IMSS system, principally health care, pension, and child care.¹⁰ The figure illustrates that the most significant changes in the schedule are for the highest-wage workers,

¹⁰Full details are presented in Appendix Tables A1 and A2 (online).

earning above 500 pesos per day, due to changes in the maximum taxable income over the period, from 10 times to 25 times the minimum wage in Mexico City.¹¹ The topcodes apply to no more than 5 percent of wage-earners in any year and will play little role in our analysis. The total employer contribution varied between 18 percent and 22 percent of the wage over the range in which almost all workers fall. There was an increase in the employer contribution from 1990 to 1993, and then the reform in 1997 introduced a kink in the schedule, which raised contributions disproportionately on the lowest-wage workers. Figure 2 displays worker contributions, which vary between 2 percent and 5 percent over the relevant range and declined with the 1997 reform. Overall, while there were changes in the contribution schedules, these were relatively modest over the relevant wage range. Looking ahead to the empirical strategy we implement in Section 5, we also note that the changes in contributions were the same for all age groups and their effects will be differenced out in our difference-in-differences procedure.

2.2 Non-pension Benefits

Any worker on whose behalf contributions are made to the system (regardless of the wage reported) is entitled to free health care at IMSS hospitals and clinics, for himself or herself, as well as for members of his or her immediate family. In addition, working mothers and widowed or divorced working fathers covered by IMSS in their jobs are entitled to free child care during workdays for children ages seven weeks to four years old.¹² It is difficult to estimate workers' valuations of these non-pension benefits. Conveniently for our empirical strategy, however, the health care and child care benefits did not change with the 1997 pension reform. Under the assumption that employees' valuations of the constant set of benefits did not change differentially by age group over the study period, the valuations will be differenced out in our difference-in-differences procedure.¹³

2.3 Pension Benefits

The pension system is the component of the social security system that experienced the largest change over our study period. Here we describe the pre-reform and post-reform regimes separately.

¹¹There are three minimum wage zones in Mexico, corresponding to higher-, medium- and lower-wage municipalities, respectively. The minimum wage in Mexico City is typically used for indexing purposes, and where we refer to the minimum wage (without specifying zone) we are referring to the minimum wage in Mexico City.

¹²IMSS also provides an individual savings account for housing expenditures, which in some cases can be used to contribute to an individual pensions. See Appendix A.2 (online) for details.

¹³There 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.

2.3.1 Pre-reform (pay-as-you-go) system

Under the pre-reform regime, workers became vested in the system after 10 years of contributions, and were then entitled to receive at least the minimum pension. Pensions were calculated on the basis of the final average wage, defined as the average nominal wage in the five years preceding retirement. Panel A of Figure 3 illustrates the expected daily pension as a function of the final average wage for workers with 10, 20 and 30 years of contributions in selected years. The schedules combine a minimum pension guarantee with a benefit proportional to an individual's wage. At first glance, the pension values illustrated in Panel A do not appear to be insensitive to the reported final average wage, but it is important to note that in the years leading up to the reform inflation had severely eroded the real value of wages and pensions, such that a large majority of workers had final average wages in the region in which the minimum was binding. Inflation exceeded 50 percent in every year in the volatile 1982-1988 period, and exceeded 100 percent in 1987 and 1988; it remained above 15 percent in a number of subsequent years (1989-1992 and 1995-1999); see Appendix Table A4 (online). In response to public pressure, the Mexican congress in 1989 increased the minimum pension to 70 percent of the minimum wage and indexed it to the minimum wage going forward, without raising the value of pensions greater than the minimum.¹⁴ Over time, the congress also raised the value of the minimum pension relative to the minimum wage, until it reached 100 percent of the minimum wage in Mexico City in 1995.

As a consequence of the erosion of the real value of pensions above the minimum and the legislative interventions to raise the minimum, the fraction of workers who expected to receive the minimum pension remained high throughout the pre-reform period. Panel B of Figure 3 plots the real value of the pension for male workers with 10, 20 or 30 years of contributions against the final average wage percentile of 60-65 year old men in the IMSS data, for selected years.¹⁵ In 1990, approximately 80 percent of male retirees with 10 years of contributions received the minimum pension. The corresponding numbers for male workers with 20 or 30 years of contributions were 70 percent and 60 percent respectively. In 1997, just prior to the implementation of the pension reform, nearly all workers with 10 years of contributions, roughly 50 percent of those with 20

¹⁴In 1991, benefits were indexed to the minimum wage, which slowed the erosion of the values of pensions above the minimum. That is, if a worker's final average wage was twice the minimum wage in 1991, the pension payment in 1992 was calculated on the basis of twice the minimum wage. The real minimum wage declined steadily over the period (see Appendix Table A4 (online)) so the slowing of the erosion of pensions as a result of this change was modest.

¹⁵To calculate the final average wage percentile, we calculate the nominal wage at each percentile of the IMSS wage distribution for 60-65 year old men in each of preceding five years, then take the average for each percentile.

years, and 40 percent of those with 30 years could expect to receive the minimum pension.¹⁶ Unfortunately, the data to which we have access do not contain total years of contributions by each individual worker, and hence we are not able to calculate the precise number of workers receiving the minimum pension. But analysts with access to this information report that approximately 80 percent of retirees were receiving the minimum pension prior to the reform (Grandolini and Cerda, 1998).

Strictly speaking, pension values were insensitive to under-reporting only for those workers whose *true* wage corresponded to the minimum pension. If wages were under-reported to IMSS, as we argue below, then the graphs in Panel B likely overstate the fraction of workers whose pensions were insensitive to under-reporting. To address this, in Panel C we plot similar graphs using final average wage percentiles calculated from the ENEU household data (described in Section 3 below), which should not be subject to under-reporting. We see that somewhat smaller fractions of workers with 10, 20 and 30 years of contributions would have received the minimum pension. But the key point is that the graph for 1997 resembles quite closely the corresponding graph in Panel B: essentially all workers with 10 years of contributions would have received the minimum pension, as well as more than 40 percent of workers with 20 years and more than 20 percent of workers with 30 years. The similarity between the graphs for 1997 in Panels B and C suggest that the number of workers who would have received the minimum pension under accurate reporting was not far below the 80 percent figure reported above.

2.3.2 Post-reform (personal retirement accounts) system

Under the personal retirement account (PRA) system, employees, employers and the government are required to make contributions to workers' personal retirement accounts in each period. 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. Each worker is required to choose an investment institution, known as an *Administrador de Fondos de Ahorro para el Retiro* (AFORE) [Retirement Savings Fund Administrator], to manage his or her account.¹⁷ The

¹⁶In addition, there was a penalty for retirement before age 65 of 5 percent per year (i.e. a worker who retired at age 60 would have his or her pension reduced by 25 percent), but this penalty was not allowed to reduce the pension below the minimum. This reduced the disincentive to retire early to workers with pensions near the minimum (Aguila, 2011).

¹⁷The AFORE management fees are in many cases substantial, and it is not clear that workers choose AFORES optimally. Duarte and Hastings (2010) investigate the role of behavioral issues in employees' choices of AFORES.

AFOREs are regulated by a government agency, the *Comisión Nacional del Sistema de Ahorro para el Retiro* (CONSAR) [National Retirement Savings Commission]. The reform also specified a minimum pension equal to the minimum wage on July 1, 1997, with further increases in the minimum pension indexed to the Consumer Price Index. Eligibility for the minimum pension was raised from 10 years of contributions to 25 years of contributions. The standard retirement age remained 65.

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.¹⁸ 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.

The establishment of the new pension regime created two categories of workers: "transition" workers who first registered with IMSS before July 1, 1997, and new workers who first registered after July 1, 1997. At retirement, transition workers are given a choice between receiving pension benefits under the PAYGO scheme or the PRA scheme. The PAYGO pension is calculated as if workers' post-reform contributions were under the old regime. If a transition worker opts for the PAYGO pension, IMSS appropriates the balance of his or her personal retirement account. The only option for new workers is the PRA.

To illustrate the impact of the reform on pension wealth, we conduct a simulation of pension wealth under the two regimes, based on a similar simulation by Aguila (2011). Table 1 displays the real present value of pension wealth under the two schemes for male workers who entered the IMSS system on June 30, 1997, and hence retained the right to opt for the old-regime pension. In carrying out the simulation, we choose a relatively optimistic annual return on the personal accounts: 8.59 percent, the average return from 1998-2002, as in the more optimistic of the two scenarios considered by Aguila (2011). We also assume that, for pensions above the minimum, participants expected the real value of the minimum wage to decline, as it had done for more than a decade (see Appendix Table A4). Assumptions of lower interest rates and less rapid declines in the real minimum wage are less favorable to the PRAs. Details of the simulation are in Appendix A.3 (online).

¹⁸A worker who receives the minimum pension must choose this option.

There are three main points to notice about the simulation. First, workers with fewer than 10 years of contributions are better off under the new regime, since they receive no pension under the old regime but a small pension under the new regime.¹⁹ Second, for workers with 10 or more years of contributions, the relative attractiveness of the new-regime pension is increasing as one moves to the right and up in the table: the PRA pension is relatively more attractive for workers with higher wages and more years of contributions. Third, conditional on having wages below 200 pesos per day (which applies to 80 percent of our samples) and qualifying for the minimum PAYGO pension, the PRA pension only dominates the PAYGO pension for workers who expect to contribute to the personal account for 20 or more years. For workers at the minimum wage, the PRA pension never dominates, and for workers earning 100 pesos per day the PRA pension only begins to dominate with more than 25 years of contributions. To save space, we do not report simulation results for transition workers with more years of contributions before 1997, but the basic points are the same. In particular, for workers with at least 10 years of contributions, in the lower four-fifths of the wage distribution, the PRA is expected to be preferable to the PAYGO pension only for those with 20+ years of contributions to the personal account.

We do not attempt to infer from the simulation exact crossing points at which the PRA becomes preferable to the PAYGO pension; any such calculation would be sensitive to assumptions about the path of interest and inflation rates, and it is not clear that workers are sophisticated in calculating the precise values of pensions under the different systems. The basic message of the simulation, which we believe was understood by participants at the time of the reform, is that for most workers, conditional on qualifying for the minimum pension under the old regime, the personal accounts could be expected to be relatively more attractive only for workers with a significant number of years of contributions after 1997.

Another aspect of the pension reform, which appears to be important in practice, is that the law requires AFOREs to send an account statement to each holder of a personal retirement account every four months. A redacted example of such an account statement appears as Figure 4. The account statement reports previous balances (*saldo anterior*), new contributions (*aportaciones*), withdrawals (*retiros*), interest earned (*rendimientos*), AFORE commissions charged (*comisiones*), and final balances (*saldo final*) for the pension account as well as for two additional accounts (a voluntary savings account and a housing savings account). The bottom section reports 3-year returns and commissions for each AFORE, as well as the average 5-year net return (at left). It

¹⁹Many women with weak labor-force attachment may fall into this category, which may explain part of the difference in results between women and men mentioned above and discussed in more detail in Appendix C (online).

appears that these account statements made it significantly easier for workers to discover how much employers were contributing on their behalf. This mechanism would not be expected to reduce evasion if employers and employees were colluding in under-reporting wages, but it may have reduced evasion in cases in which workers were unaware that their employers were under-reporting their wages.

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. In practical terms, however, IMSS has not had the resources to do extensive auditing of employers. Neither before nor after the reform was there a reward to employees for revealing evasion by their employers, beyond ensuring accurate reporting of their own wages.

As will be seen below, one aspect of IMSS reporting requirements does appear to be strictly enforced. By law, firms in Mexico are required to pay the relevant minimum wage and a holiday bonus called an *aguinaldo*, worth two weeks of salary — approximately 4.5 percent of annual earnings. In order to avoid fines, establishments are required to report wages of at least the corresponding minimum wage plus 4.5 percent throughout the year. Prior to 1991, there are a scattered few reports of wages below this level; beginning in 1991, IMSS stepped up enforcement of this rule and such wages have no longer been observed.

2.4 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, corporate and personal income taxes are generally higher than payroll taxes. The corporate income tax rate went from 39 to 34 percent over the 1988-2003 period.²⁰ But tax evasion and avoidance are rife in Mexico. 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, 2011a). 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

²⁰Source: OECD Tax Database, www.oecd.org/ctp/taxdatabase.

corporate tax burden. It appears, in other words, that over the study period payroll taxes evaded by under-reporting wages could be pocketed by firms, and were not offset by increases in other taxes.²¹

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 non-positive tax rate (OECD, 1999, p. 80).

3 Data

The source for establishments' wage reports are IMSS administrative records. All private Mexican employers are in principle legally obligated to report wages for their employees, and pay social-security taxes on the basis of the reports. The IMSS dataset contains the full set of wage reports for employees in registered, private-sector establishments over the period 1985-2005.²² The dataset contains a limited set of variables: age, sex, daily wage (including benefits), state and year of the individual's first registration with IMSS, an employer-specific identifier, and industry and location of the employer. Wages are reported in spells (with a begin and end date for each wage level) and in theory we could construct a day-by-day wage history for each individual. To keep the dataset manageable, we extract wages for a single day, June 30, in each year. Prior to 1997, records for temporary workers were not collected in digital form. To ensure comparability before and after 1997, we focus on workers identified in the IMSS data as permanent, defined as having a written contract of indefinite duration.

We select ages 16-65. To maintain consistency across years, we impose the lowest real value of the IMSS topcode (which occurred in 1991) in all years. We drop establishments with a single insured worker, since these are likely to be self-employed workers.²³ In the interests of comparability with the ENEU data, we include only the metropolitan areas included in the ENEU samples (described below). We also focus on sectors for which we are confident that IMSS is

²¹Madzharova (2011) provides a formalization of the idea that lower corporate taxes give firms greater incentives to under-report wages. Empirically, however, the evidence she presents suggests that changes in corporate tax rates do not have a large effect on wage reports in Bulgaria.

²²The data have been used in several previous papers, including Castellanos, Garcia-Verdu, and Kaplan (2004), and Frías, Kaplan, and Verhoogen (2009).

²³Including these single-worker establishments has no effect on the results reported below.

the only available formal-sector social insurance program: manufacturing, construction, and retail/hotel/restaurants. Other sectors contain a larger share of public employees, and it is not clear that we would be able to treat workers not covered by IMSS as belonging to the informal sector. We refer to the sample selected following these criteria as our IMSS baseline sample. Further details are in Appendix B.1 (online).

The household data we use are from the *Encuesta Nacional de Empleo Urbano* (ENEU) [National Urban Employment Survey], a household survey modeled on the Current Population Survey (CPS) in the United States, collected by the *Instituto Nacional de Estadísticas y Geografía* (INEGI), the Mexican statistical agency. The original ENEU sample, beginning in 1987, focused on the 16 largest Mexican metropolitan areas; although the coverage expanded over time, to maximize the number of pre-reform years we focus on the original 16 areas. As in the IMSS data, we include male workers ages 16-65, focus the second quarter of each year, exclude self-employed workers, impose the 1991 IMSS topcode in all years, and include only manufacturing, construction, and retail/hotels/restaurants. All calculations below use the sampling weights provided by INEGI.

A very useful feature of the ENEU for our purposes is that it asks respondents whether they receive IMSS coverage as an employment benefit. Beginning in the third quarter of 1994, the ENEU also asked respondents whether they had a written contract of indefinite duration, the legal definition of a permanent employee used by IMSS. Hourly wages are calculated as monthly wages divided by 4.3 times hours worked in the previous week, and daily wages as 8 times hourly wages. The ENEU wage measures are based on respondents' reports of take-home pay, and do not include respondents' social security contributions or other taxes paid. They also exclude bonuses paid less frequently than monthly, and hence exclude the yearly *aguinaldo* bonus. The differences between the IMSS and the ENEU wage measures are discussed further in Appendix B (online). We drop workers with reported daily wages below 30 pesos (in 2002 constant pesos, approximately US\$3, which is approximately 50 percent of the minimum wage.) In principle, both the IMSS and the ENEU data are available over the 1987-2005 period, but in the interest of consistency over time in the ENEU we focus on the years 1988-2003.²⁴

Our goal in the preparation of the datasets is to construct samples in the IMSS and ENEU data that are as similar as possible. Table 2 presents summary statistics for the IMSS baseline

²⁴There appear to be a number of data inconsistencies in the ENEU in 1987, the first year of the survey. The ENEU sampling scheme was redesigned in the third quarter of 2003; to avoid introducing inconsistencies (with little benefit, since the IMSS data are available only until 2005) we focus on the period before the redesign.

sample and various ENEU samples for 1990 and 2000, for a set of variables that are common between the sources: daily wage, age, and share in large establishments (with more than 100 employees). Column 2 contains the “full” ENEU sample, containing all non-self-employed men satisfying the age and sector criteria. Comparing columns 3 and 4, we see that ENEU workers with IMSS coverage tend to be higher-wage and more likely to work in large establishments than workers without IMSS coverage.

Column 5 contains the sample that in principle should be the best match for the IMSS baseline sample: ENEU workers who report receiving IMSS coverage and having a written contract of indefinite duration — that is, who satisfy the definition of “permanent” used by IMSS. The average wage for this ENEU sample is greater than for the IMSS baseline sample, consistent with our argument below that there is under-reporting of wages in the IMSS data. Because the contract-type variable is available only beginning in 1994, however, we have prohibitively few years of pre-reform data for this sample. Instead, we will focus hereafter on the Column 6 sample, ENEU workers who report receiving IMSS coverage and working full-time (i.e. at least 35 hours in the previous week), which can be defined consistently over the entire period. We refer to the Column 6 sample as our ENEU baseline sample.

The ENEU baseline sample is not an ideal comparison group, for several reasons. Some temporary workers may work full-time, and some permanent workers may work part-time. Comparing Columns 5 and 6 for the year 2000, we see that average wages are significantly lower in the Column 6 sample; this is attributable to the facts that temporary full-time workers earn relatively low wages and that permanent part-time workers earn relatively high wages on average. It may also be that firms interpret “permanent” to mean something different from the legal definition (i.e. written contract of indefinite duration) when reporting wages. In addition, patterns of non-response may differ between the IMSS and ENEU samples. It is well known, for instance, that richer households tend to be less likely to respond to income questions in household surveys (Groves and Couper, 1998; Korinek, Mistiaen, and Ravallion, 2006). The weighted employment totals from the ENEU data in Columns 5 and 6 are below the IMSS totals in Column 1; this may in part reflect such non-response. These potential discrepancies recommend caution in interpreting cross-sectional differences between the IMSS and ENEU baseline samples. It is worth emphasizing, however, that our difference-in-difference strategy will focus on changes over time in the discrepancies between the samples, and any time-invariant sources of discrepancy will be differenced out.

As a further comparison, Figure 5 plots employment totals over the 1988-2003 period for the same samples as in Table 2. Perhaps surprisingly, we see that over most of the period the number of workers in the IMSS sample is slightly *greater* than the numbers in any of the ENEU samples. There are several potential explanations. The difference may reflect non-response by households in the ENEU (perhaps varying systematically with income, as mentioned above). It may be that some respondents are unaware that they receive IMSS coverage from their employer, or believe that they are covered by the public-sector social security agency (known by the acronym ISSSTE) when in fact they are covered by IMSS. It may also be that individuals live outside of the boundaries of the metropolitan area in which they work, and hence are included in our IMSS sample but not our ENEU sample. For our purposes, however, the most important lesson of the figure is that there does not appear to have been a large change over time in the extent of the employment discrepancy between the IMSS and ENEU samples in response to the pension reform.

4 Cross-Sectional Comparisons of Wage Distributions

In this section, we consider cross-sectional differences in wage distributions between the IMSS and ENEU baseline samples prior to the 1997 pension reform. As discussed in Section 2 above, in the pre-reform period the key issue for workers was simply whether or not they were in the IMSS system; for almost all workers, the benefits that they enjoyed were insensitive to the wages that employers reported on their behalf.

Figure 6 plots simple histograms of daily wages in the IMSS baseline sample (gray bars) and the ENEU baseline sample (bars with black borders and no fill color) in 1990, using bins that are 5 pesos wide. The three vertical lines between 50 and 70 pesos (approximately US\$5-US\$7/day) represent the the three minimum wages in Mexico, with the rightmost corresponding to the minimum wage in Mexico City. Figure 7 plots similar histograms using the same samples but using only observations below 200 pesos (approximately US\$20), with bins 2 pesos wide. The pattern is striking: there is clear evidence of stacking in the IMSS sample slightly above the three minimum wages. These stacks correspond to 104.5 percent of the minimum wages in each zone — the minimum reports to IMSS that did not incur penalties. It is also evident that the IMSS distribution lies largely to the left of the ENEU distribution.²⁵ The stacking and shift to the left

²⁵The exception to this generalization is at the far right tail. In Figure 6, we see that there is relatively more weight at the topcode in the IMSS sample; there is also slightly more weight at high wage values just below the topcode. This appears to reflect non-response by high-income households in the ENEU — a common pattern in household surveys, as mentioned above.

of the distribution is precisely what one would have expected, given the incentives inherent in the social security system.

A key prediction of the theoretical framework of Kleven et al. (2009) is that collusion in under-reporting wages is more difficult to sustain in larger firms. It may also be the case simply that the intensity of IMSS monitoring and auditing is greater in larger firms. In either case, we would expect the difference between the IMSS and ENEU wage distributions to be smaller in larger firms. Figure 8 presents figures similar to Figure 7 (focused on daily wages below 200 pesos), separately for five firm sizes. Caution is warranted in interpreting these figures, since observed establishment size in the IMSS data is itself affected by firms' compliance decisions. Subject to this caveat, it appears that there is less stacking on the minimum allowable wage reports at larger firm sizes, suggesting greater compliance. Even in establishments with 250 workers or more, however, there is evidence of stacking at the minimum allowable wage report, suggesting some under-reporting even in quite large firms.

Another insight from the taxation literature is that the larger the number of reports the tax authorities receive from firms, the more difficult it is for firms to evade taxes. To investigate this prediction, we consider the wage distributions for two subsets of plants that appear in plant-level datasets collected by INEGI, the Mexican statistical agency.²⁶ The plant-level datasets do not contain individual-level wage information on the full distribution of wages, but it is nevertheless instructive to consider firms' wage reports in plants that appear in the plant-level datasets. Figure 9 plots the IMSS wage distribution for workers in manufacturing establishments that also appear in the main Mexican longitudinal plant panel, the *Encuesta Industrial Anual* (EIA) [Annual Industrial Survey], which excludes assembly-for-export *maquiladora* plants.²⁷ There is little evidence of stacking at the minimum allowable wage values, suggesting relatively little under-reporting of wages. The EIA sample consists mainly of plants with more than 100 employees, and we saw above that larger plants are less likely to under-report. But the EIA plants display less stacking even than plants in the 100-250 and >250 employees categories in Figure 8.

For the sake of completeness, Figure 10 plots the IMSS wage distribution for workers in assembly-for-export *maquiladora* plants, on which monthly statistics are reported in the *Estadísticas Mensuales de la Industria Maquiladora de Exportación* (EMIME) [Monthly Statistics

²⁶INEGI does not share plant-level information collected in INEGI surveys with the Mexican tax authorities. It is not clear that plants are aware of this, however.

²⁷The links between establishments in the IMSS data and the EIA were constructed and first exploited in Frías, Kaplan, and Verhoogen (2009); see that paper for details of the linking.

on Export Maquiladora Industry]. Maquiladoras captured in the EMIME generally tend to have lower wages than the non-*maquiladoras* captured in the EIA. Although there is significant stacking at the minimum allowable wage values, it is difficult to determine whether this reflects under-reporting or simply the fact that many maquiladoras pay wages at or near the minimum.

5 Effect of 1997 Pension Reform on Under-reporting

We now consider how the differences in wage distributions between the IMSS and ENEU samples varied over time, and in particular whether the 1997 pension reform had the expected differential effect by age. We begin with a set of figures illustrating the key patterns. Figure 11 plots non-parametric estimates of the wage distribution for men by age group, for three years, 1990, 1997 (the quarter before the reform), and 2003. The non-parametric densities, solid red for the IMSS baseline sample and dashed blue for the ENEU baseline sample, are analogous to the histograms for low wage levels in Figure 7; non-parametric densities are chosen for visual clarity. Each column of graphs corresponds to an age group (indicated in the x-axis titles) and each row to a year. There is an evident reduction in the extent of stacking for the youngest workers, but little apparent change in stacking for the highest age group.

To illustrate the differential reduction in the discrepancy between wage distributions in another way, Figure 12 plots the difference in log median wages between the IMSS and ENEU samples by age group over the 1988-2003 period. We refer to the difference in log median wages between the IMSS and ENEU samples as the “wage gap.” One noticeable fact is that the gap is initially larger for the youngest age group, 16-25. But the key point is that the gap for the oldest age group, 56-65, clearly increases relative to the gaps for the other groups. The other group that appears to see a relative increase, albeit smaller, is the second-oldest age group, 46-55.

A possible concern with Figure 12 is that the differential changes may reflect shocks to local labor markets which differ in their age composition. To remove the effects of local labor-market shocks, we calculate the wage gaps at the age group-year-metro area level, regress them on a full set of metro area-year indicators, and average the residuals at the age group-year level. Figure 13 plots these averages. Again we see a clear relative increase in the wage gap in the oldest group. It is worth noting that although there is some volatility in the estimates prior to 1997, the relative increase in wage gaps for the oldest group only becomes evident in 1998, the first observation after the July 1997 pension reform.

The difference in log medians ignores much of the information contained in the shapes of the IMSS and ENEU wage distributions. An alternative measure of the divergence between two distributions is the Kullback-Liebler divergence (Kullback and Leibler, 1951), which has been used by DiNardo, Fortin, and Lemieux (1996) among others. The Kullback-Liebler divergence between two densities, f_1 and f_2 , is given by:

$$D_{12} = \int_0^{\infty} [f_1(w) - f_2(w)] \frac{f_1(w)}{f_2(w)} dw \quad (1)$$

While in principle this measure could be computed using the histograms displayed in Figures 6-8 (and a discrete version of the definition of D_{12}), the measure is only defined if $f_2(w)$ is positive, and in the histograms this cannot be guaranteed. For this reason, we first estimate non-parametric densities at the age group-year level and then apply (1). Figure 14 plots the Kullback-Liebler divergence measure over the 1988-2003 period, separately by age group. Once again a relative increase in divergence for the oldest wage group is evident.

To put standard errors on the estimates, we turn to simple regressions. To keep the empirical model parsimonious, motivated by the pension simulation in Table 1 and by the patterns in Figures 12- 14 above, we model the differential effect by age as captured by an interaction of an 56-65 age-group indicator and a full set of year indicators. Table 3 reports the regressions. The columns differ in the sets of dummy variables included (year effects only, year and metro area effects, metro area-year effects in Columns 1-3, respectively) but the coefficient estimates hardly change. Figure 15 plots the coefficients and 95-percent intervals for the column-3 specification. Although there is some volatility in the estimates, the coefficients are relatively flat in the pre-reform period and there is a clear increase beginning in 1997. Coefficients are significant at the 95-percent level in 1999 and 2001-2003, and at the 90-percent level in 2000. These results support the hypothesis that the reform reduced under-reporting to a greater extent for younger workers.

Table 4 presents analogous regressions for employment gaps (log differences between employment in the ENEU and IMSS samples). Consistent with our earlier impressions from Figure 5, there does not appear to be a systematic response of discrepancies in employment reports to the pension reform.

An interesting question that remains open is to what extent workers are aware of under-reporting by their employers. The theoretical framework of Kleven, Kreiner, and Saez (2009), mentioned in the introduction, assumes that employees and employers must agree to under-report

in order to do so without triggering an audit. But our empirical results can also potentially be reconciled with an extension of the model in which workers do not actively collude but face a cost of obtaining information about the reported wage, which they may or may not pay. The current empirical setting is arguably not well-suited to teasing apart these two interpretations, in part because the change in incentives (in the form of the increased dependence of pensions on reported wages) and the change in information (in the form of the regular AFORE account statements) occurred simultaneously. A potential way forward would be to investigate the consequences of the reform for wage bargaining between employers and employees, for instance by estimating differential changes in the formal-sector vs. informal-sector wage difference for observably similar workers. We do not have exogenous variation in whether workers are in the formal or informal sector, however, and any such estimates would be subject to concerns about endogenous selection.²⁸ We feel that a definitive answer to the question of how much workers know about their employers' under-reporting must await a setting in which incentives vary separately from costs of information, and there is exogenous variation in workers' attachment to the formal or informal sector. It is worth emphasizing, however, that our key empirical conclusion does not depend on the answer to this question: regardless of whether workers actively collude in under-reporting, it appears that the differential increase in the incentive to ensure accurate reports, together with the decrease in the cost of observing employers' reports, generated a differential improvement in compliance.

6 Conclusion

Improving compliance of firms with tax regulations is a first-order policy issue in many developing countries. Much of the debate has focused on how to induce firms to register with tax authorities in the first place — what we might call the *extensive* margin of non-compliance. In this paper, we have shown that under-reporting of wages among firms that are already registered — non-compliance on an *intensive* margin — is also substantial and responds to the incentives and ability of workers to monitor their employers. These results suggest that providing such incentives should be a consideration in the design of social-insurance systems.

Our results also point to two potentially interesting questions for further research. First, what

²⁸In the same empirical setting, Marrufo (2001) argues that the difference in wages between formal and informal workers in part reflects endogenous selection and is not a sufficient basis for calculating the incidence of social security taxes and benefits.

are the implications of the form of non-compliance we have highlighted for measurement of the incidence of social-insurance costs and benefits? There is a large literature on incidence, including early influential papers by Summers (1989) and Gruber (1994) and continuing more recently with Kugler and Kugler (2009), and Cruces, Galiani, and Kidyba (2010), but relatively little attention has been paid to estimating incidence in the presence of endogenous non-compliance. Second, and relatedly, does greater pressure to report accurately (i.e. which increases compliance on the intensive margin) induce more firms to remain informal (i.e. reduce compliance on the extensive margin)? If the costs of increased intensive-margin compliance are not entirely borne by workers in the form of lower wages, we might well expect such an effect. More research is needed to weigh this potential cost against the benefits of reduced payroll-tax evasion by registered firms.

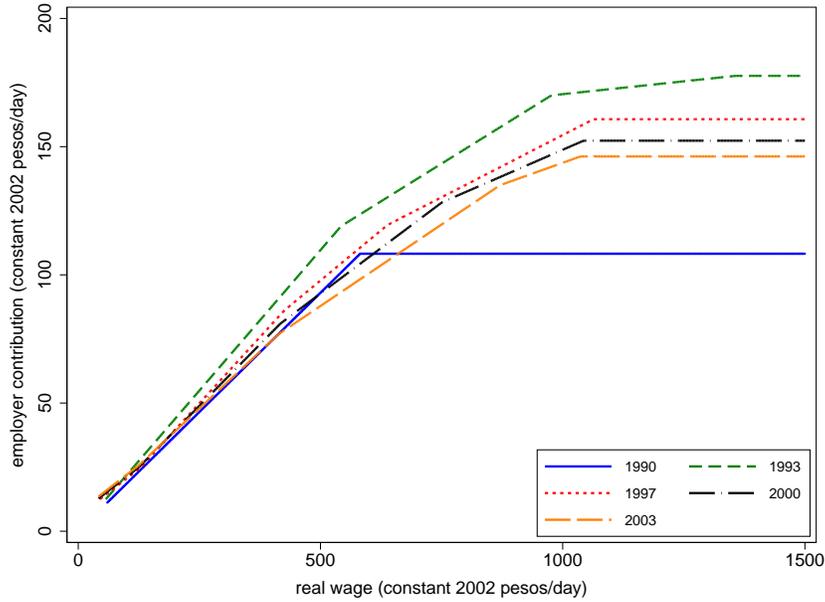
References

- AGUILA, E. (2011): “Personal Retirement Accounts and Saving,” *American Economic Journal: Economic Policy*, 3(4), 1–24.
- ANDREONI, J., B. ERARD, AND J. FEINSTEIN (1998): “Tax Compliance,” *Journal of Economic Literature*, 36(2), 818–860.
- BAILEY, C., AND J. TURNER (2001): “Strategies to Reduce Contribution Evasion in Social Security Financing,” *World Development*, 29(2), 385–393.
- BESLEY, T., AND T. PERSSON (2009): “The Origins of State Capacity: Property Rights, Taxation, and Politics,” *American Economic Review*, 99(4), 1218 – 1244.
- (2011): *Pillars of Prosperity: The Political Economy of Development Clusters*. Princeton University Press.
- BRAGUINSKY, S., S. MITYAKOV, AND A. LISCOVICH (2010): “Direct Estimation of Hidden Earnings: Evidence from Administrative Data,” Unpub. paper, Carnegie-Mellon University.
- BRUHN, M. (2011): “License to Sell: The Effect of Business Registration Reform on Entrepreneurial Activity in Mexico,” *Review of Economics and Statistics*, 93(1), 382–386.
- BURGESS, R., AND N. STERN (1993): “Taxation and Development,” *Journal of Economic Literature*, 31(2), pp. 762–830.
- CASTELLANOS, S., R. GARCIA-VERDU, AND D. KAPLAN (2004): “Wage Rigidities in Mexico: Evidence from Social Security Records,” *Journal of Development Economics*, 75(2), 507–33.
- CRUCES, G., S. GALIANI, AND S. KIDYBA (2010): “Payroll Taxes, Wages and Employment: Identification Through Policy Changes,” *Labour Economics*, 17(4), 743 – 749.
- DE MEL, S., D. J. MCKENZIE, AND C. WOODRUFF (2012): “The Demand for, and Consequences of, Formalization among Informal Firms in Sri Lanka,” NBER working paper no. 18019.
- DE PAULA, A., AND J. SCHEINKMAN (2011): “The Informal Sector: An Equilibrium Model and Some Empirical Evidence from Brazil,” *Review of Income and Wealth*, 57, S8–S26.
- DI NARDO, J., N. M. FORTIN, AND T. LEMIEUX (1996): “Labor Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach,” *Econometrica*, 64(5), 1001–1044.
- DUARTE, F., AND J. HASTINGS (2010): “Fettered Consumers and Sophisticated Firms: Evidence from Mexico’s Privatized Social Security Market,” Unpub. paper, March.
- ELEK, P., J. KÖLLÖ, B. REIZER, AND P. A. SZABÓ (forthcoming): “Detecting Wage Under-reporting Using a Double Hurdle Model,” Forthcoming, *Research in Labor Economics*.
- FAJNZYLBER, P., W. F. MALONEY, AND G. V. MONTES-ROJAS (2011): “Does formality improve micro-

- firm performance? Evidence from the Brazilian SIMPLES program,” *Journal of Development Economics*, 94(2), 262 – 276.
- FISMAN, R., AND S.-J. WEI (2004): “Tax Rates and Tax Evasion: Evidence from Missing Imports in China,” *Journal of Political Economy*, 112(2), 471–500.
- FRÍAS, J. A., D. S. KAPLAN, AND E. A. VERHOOGEN (2009): “Exports and Wage Premia: Evidence from Mexican Employer-Employee Data,” Unpub. paper, Columbia University.
- GALIANI, S., AND F. WEINSCHELBAUM (forthcoming): “Modeling Informality Formally: Households and Firms,” Forthcoming, *Economic Inquiry*.
- GORDON, R., AND W. LI (2009): “Tax Structures in Developing Countries: Many Puzzles and a Possible Explanation,” *Journal of Public Economics*, 93(7-8), 855 – 866.
- GORODNICHENKO, Y., J. MARTINEZ-VAZQUEZ, AND K. S. PETER (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.
- GRANDOLINI, G., AND L. CERDA (1998): “The 1997 Pension Reform in Mexico,” World Bank Policy Research Working Paper no. 1933.
- GROVES, R. M., AND M. P. COUPER (1998): *Nonresponse in Household Interview Surveys*. John Wiley & Sons, New York.
- GRUBER, J. (1994): “The Incidence of Mandated Maternity Benefits,” *The American Economic Review*, 84(3), pp. 622–641.
- HURST, E., G. LI, AND B. PUGSLEY (2011): “Are Household Surveys Like Tax Forms? Evidence from Income Underreporting of the Self-Employed,” Unpub. paper, University of Chicago.
- IMF (2010): *Government Finance Statistics Yearbook*. International Monetary Fund Statistics Department.
- IMSS (2009): “Memoria Estadística,” Mexico City, Mexico.
- (2011): “Informe al Ejecutivo Federal y al Congreso de la Unión sobre la Situación Financiera y los Riesgos del Instituto Mexicano del Seguro Social, 2010-2011,” Mexico City, Mexico.
- INEGI (2001): “Indicadores Sociodemográficos de México, 1930-2000,” Aguascalientes, Mexico.
- INTERNAL REVENUE SERVICE (1996): “Federal Tax Compliance Research: Individual Income Tax Gap Estimates for 1985, 1988, and 1992,” IRS Publication 1415 (Rev. 4-96), Washington DC.
- (2006): “Updated Estimates of the Tax Year 2001 Individual Income Tax Underreporting Gap. Overview,” Washington DC.
- KAPLAN, D. S., E. PIEDRA, AND E. SEIRA (forthcoming): “Entry regulation and Business Start-ups: Evidence from Mexico,” Forthcoming, *Journal of Public Economics*.
- KEEN, M., AND B. LOCKWOOD (2010): “The Value Added Tax: Its Causes and Consequences,” *Journal of Development Economics*, 92(2), 138 – 151.
- KLEVEN, H., M. KNUDSEN, C. T. KREINER, S. PEDERSEN, AND E. SAEZ (2011): “Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark,” *Econometrica*, 79, 651–692.
- 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.
- KOPCZUK, W., AND J. SLEMRUD (2006): “Putting Firms into Optimal Tax Theory,” *American Economic Review Papers & Proceedings*, 96(2), 131–134.
- KORINEK, A., J. A. MISTIAEN, AND M. RAVALLION (2006): “Survey Nonresponse and the Distribution of Income,” *Journal of Economic Inequality*, 4(1), 33 – 55.
- KUGLER, A., AND M. KUGLER (2009): “Labor Market Effects of Payroll Taxes in Developing Countries: Evidence from Colombia,” *Economic Development and Cultural Change*, 57(2), 335 – 358.
- KULLBACK, S., AND R. LEIBLER (1951): “On Information and Sufficiency,” *Annals of Mathematical Statistics*, 22, 79–86.
- LA PORTA, R., AND A. SHLEIFER (2008): “The Unofficial Economy and Economic Development,” *Brookings Papers on Economic Activity*, 2008, pp. 275–352.
- LEVY, S. (2008): *Good Intentions, Bad Outcomes: Social Policy, Informality and Economic Growth in Mexico*. Brookings Institution Press, Washington D.C.
- MADZHAROVA, B. (2011): “The Effect of Low Corporate Tax Rate on Payroll Tax Evasion,” CERGE-EI Working Paper No. 448.

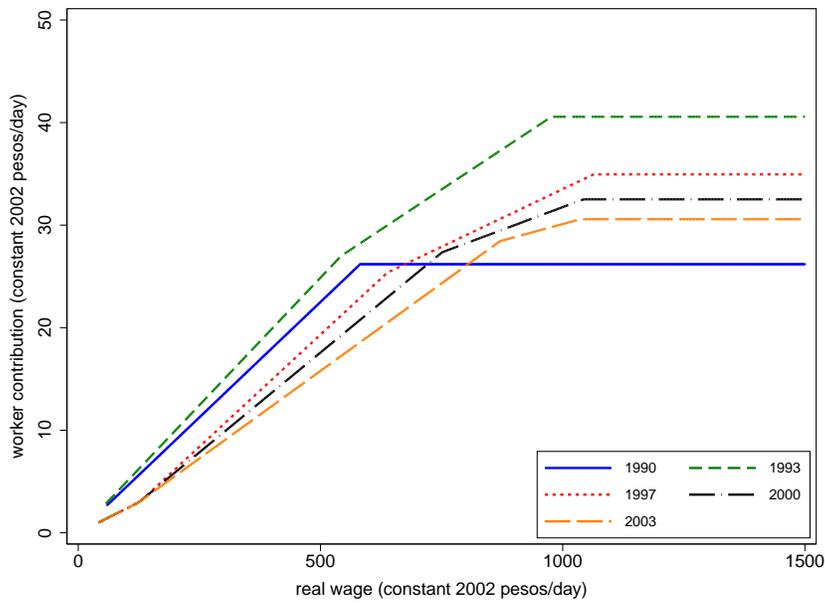
- MARION, J., AND E. MUEHLEGGGER (2008): “Measuring Illegal Activity and the Effects of Regulatory Innovation: Tax Evasion and the Dyeing of Untaxed Diesel,” *Journal of Political Economy*, 116(4), 633 – 666.
- MARRUFO, G. M. (2001): “The incidence of Social Security regulation: Evidence from the Reform in Mexico,” PhD Dissertaton, University of Chicago.
- MCKENZIE, D., AND Y. S. SAKHO (2010): “Does It Pay Firms to Register for Taxes? The Impact of Formality on Firm Profitability.,” *Journal of Development Economics*, 91(1), 15 – 24.
- MONTEIRO, J. C., AND J. J. ASSUNÇÃO (forthcoming): “Coming Out of the Shadows? Estimating the Impact of Bureaucracy Simplification and Tax Cut on Formality in Brazilian Microenterprises,” Forthcoming, *Journal of Development Economics*.
- NYLAND, C., R. SMYTH, AND C. ZHU (2006): “What Determines the Extent to Which Employers Will Comply with their Social Security Obligations? Evidence from Chinese Firm-level Data,” *Social Policy & Administration*, 40(2), 196–214.
- OECD (1992): *OECD Economic Surveys: Mexico 1992*. OECD Publishing.
- (1999): *OECD Economic Surveys: Mexico 1999*. OECD Publishing.
- (2006): *Tax Administration in OECD and Selection Non-OECD Countries: Comparative Information Series*. OECD, Prepared by the Forum on Tax Administration, Centre for Tax Policy and Administration.
- (2011a): *OECD Economic Surveys: Mexico 2011*. OECD Publishing.
- (2011b): *Revenue Statistics 1965-2010*. OECD, Paris.
- POMERANZ, D. (2011): “No Taxation without Information: Deterrence and Self-Enforcement in the Value-Added Tax,” Unpub. paper, Harvard Business School.
- RAUCH, J. E. (1991): “Modelling the Informal Sector Formally,” *Journal of Development Economics*, 35(1), 33 – 47.
- SAEZ, E. (2010): “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2(3), 180–212.
- SAEZ, E., J. B. SLEMROD, AND S. H. GIERTZ (2012): “The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review,” *Journal of Economic Literature*, 50(1), 3–50.
- SALES-SARRAPY, C., F. SOLIS-SOBERON, AND A. VILLAGOMEZ-AMEZCUA (1996): “Pension System Reform: The Mexican Case,” NBER Working Paper No. 5780.
- SCHNEIDER, F., AND D. H. ENSTE (2000): “Shadow Economies: Size, Causes, and Consequences,” *Journal of Economic Literature*, 38(1), 77–114.
- SLEMROD, J. (2008): “Does It Matter Who Writes the Check to the Government? The Economics of Tax Remittance,” *National Tax Journal*, 61(2), 164 – 165.
- SLEMROD, J., AND S. YITZHAKI (2002): “Tax Avoidance, Evasion and Administration,” in *Handbook of Public Economics*, vol. 3, ed. by A. J. Auerbach, and M. Feldstein. Elsevier, Amsterdam.
- SUMMERS, L. H. (1989): “Some Simple Economics of Mandated Benefits,” *The American Economic Review*, 79(2), pp. 177–183.
- TONIN, M. (2011): “Minimum Wage and Tax Evasion: Theory and Evidence,” *Journal of Public Economics*, 95(11-12), 1635 – 1651.
- YANIV, G. (1992): “Collaborated Employee-Employer Tax Evasion,” *Public Finance*, 47(2), 312–321.

Figure 1. Employer contributions



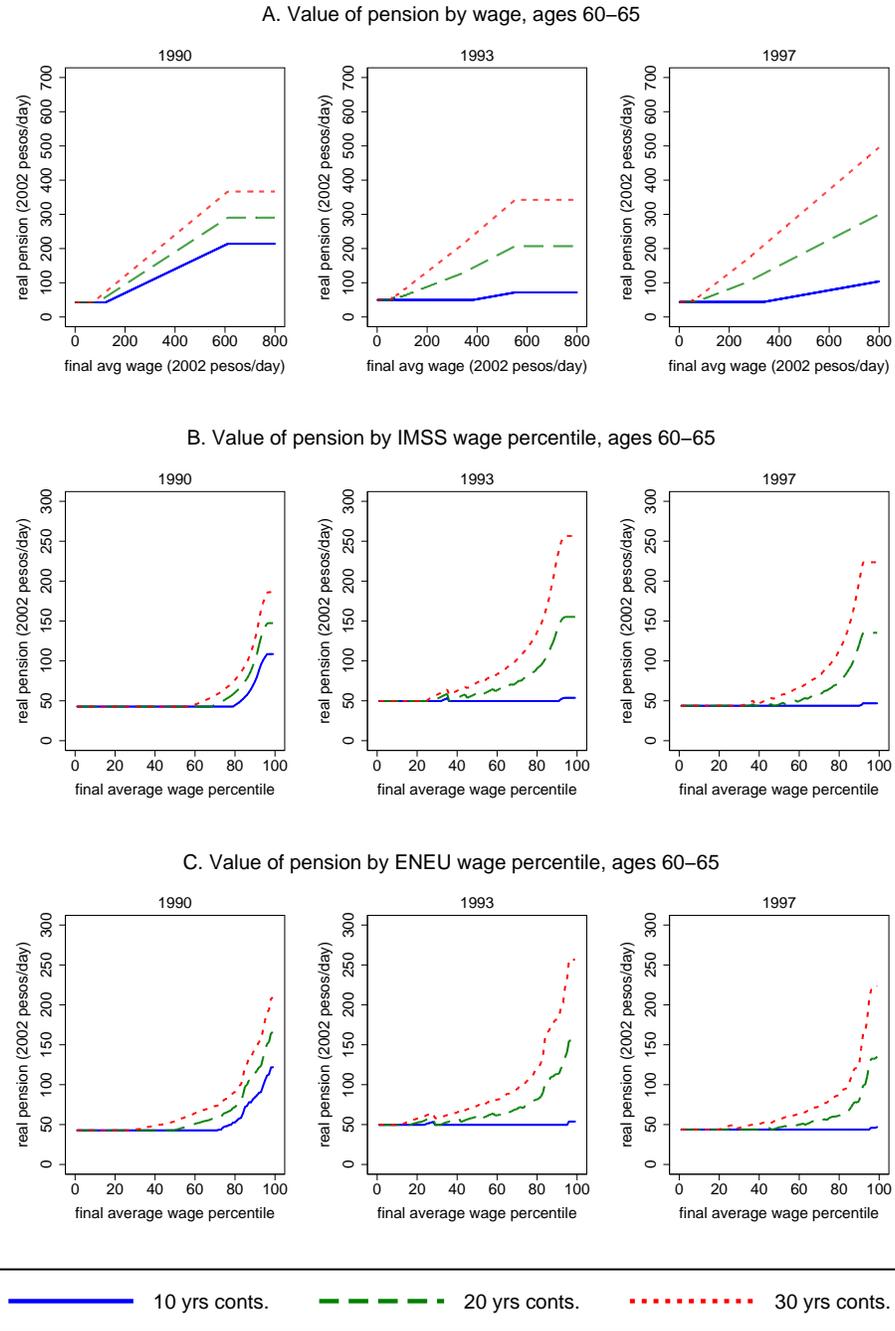
Notes: Variation in IMSS employer contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 2. Worker contributions



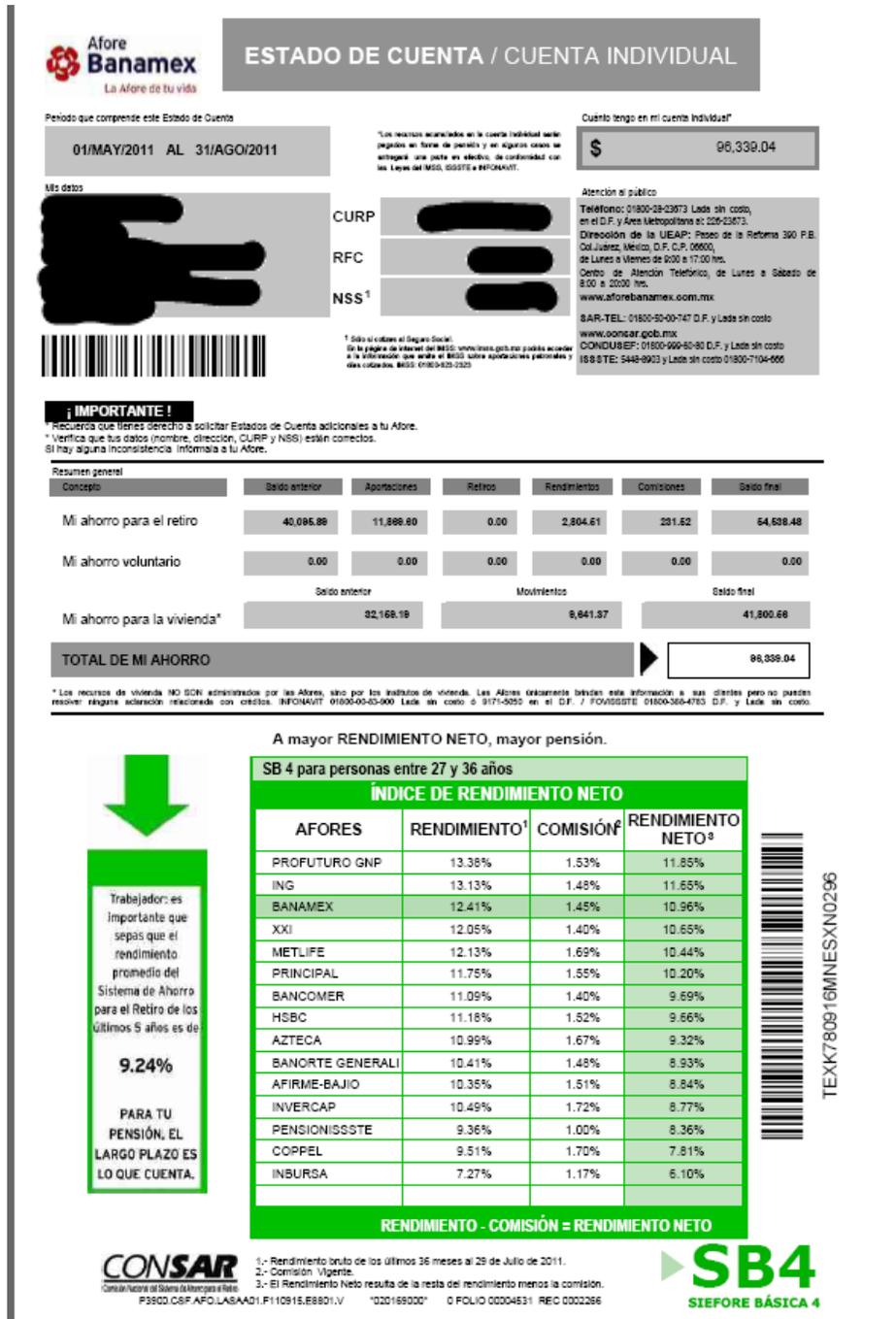
Notes: Variation in IMSS worker contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 3. Pension values, selected years, men



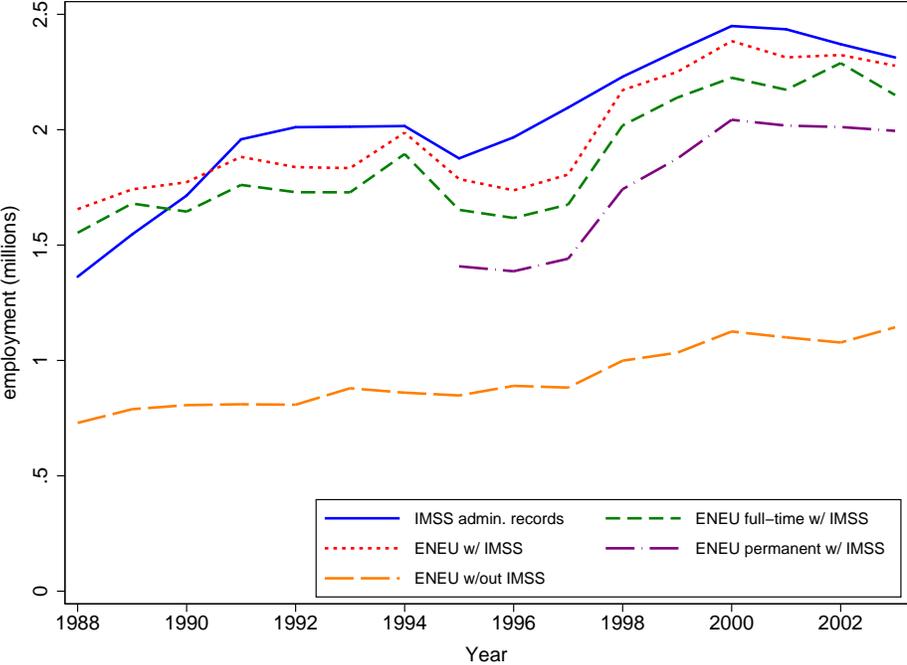
Notes: Final average wage (2002 pesos/day) is average nominal daily wage over five years prior to retirement, deflated to constant 2002 pesos. Figure indicates pension values for individuals with 10, 20 and 30 years of contributions to IMSS. In Panel B, we calculate the nominal wage at each quantile of the IMSS wage distribution for 60-65 year old men in each year and take the average for that quantile over the preceding five years. Panel C is constructed similarly using wage distributions from the ENEU baseline samples. See Section 3 for details of samples and Section 2.3 for details on pension benefits. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 4. Estado de Cuenta



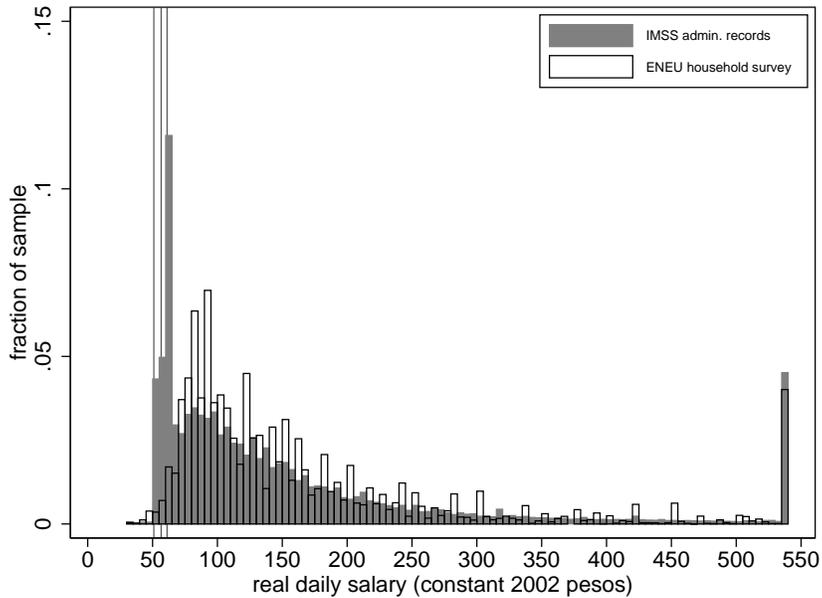
Notes: The box at top right (“Cuánto tengo en mi cuenta individual”) reports total balance. The first row of boxes in the middle section (“Mi ahorro para el retiro”) pertains to the retirement pension and reports previous balance (“Saldo anterior”), new contributions (“Aportaciones”), withdrawals (“Retiros”), 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 5. Employment, IMSS admin. records vs. ENEU household data, men



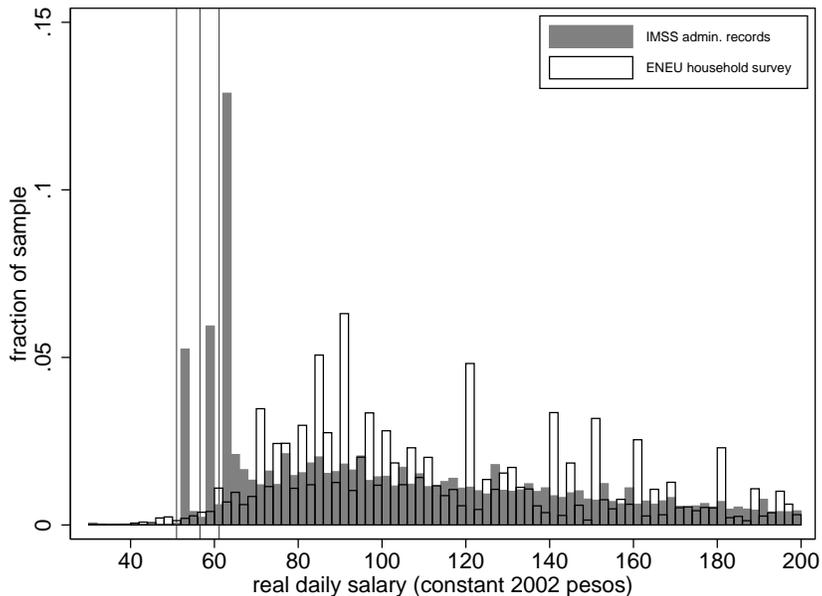
Notes: Samples are the same as those in Columns 1 and 3-6 of Table 2; refer to that table for details. ENEU totals are calculated using sampling weights. See Section 3 and Appendix B (online) for details of sample selection.

Figure 6. Wage histograms, men, 1990



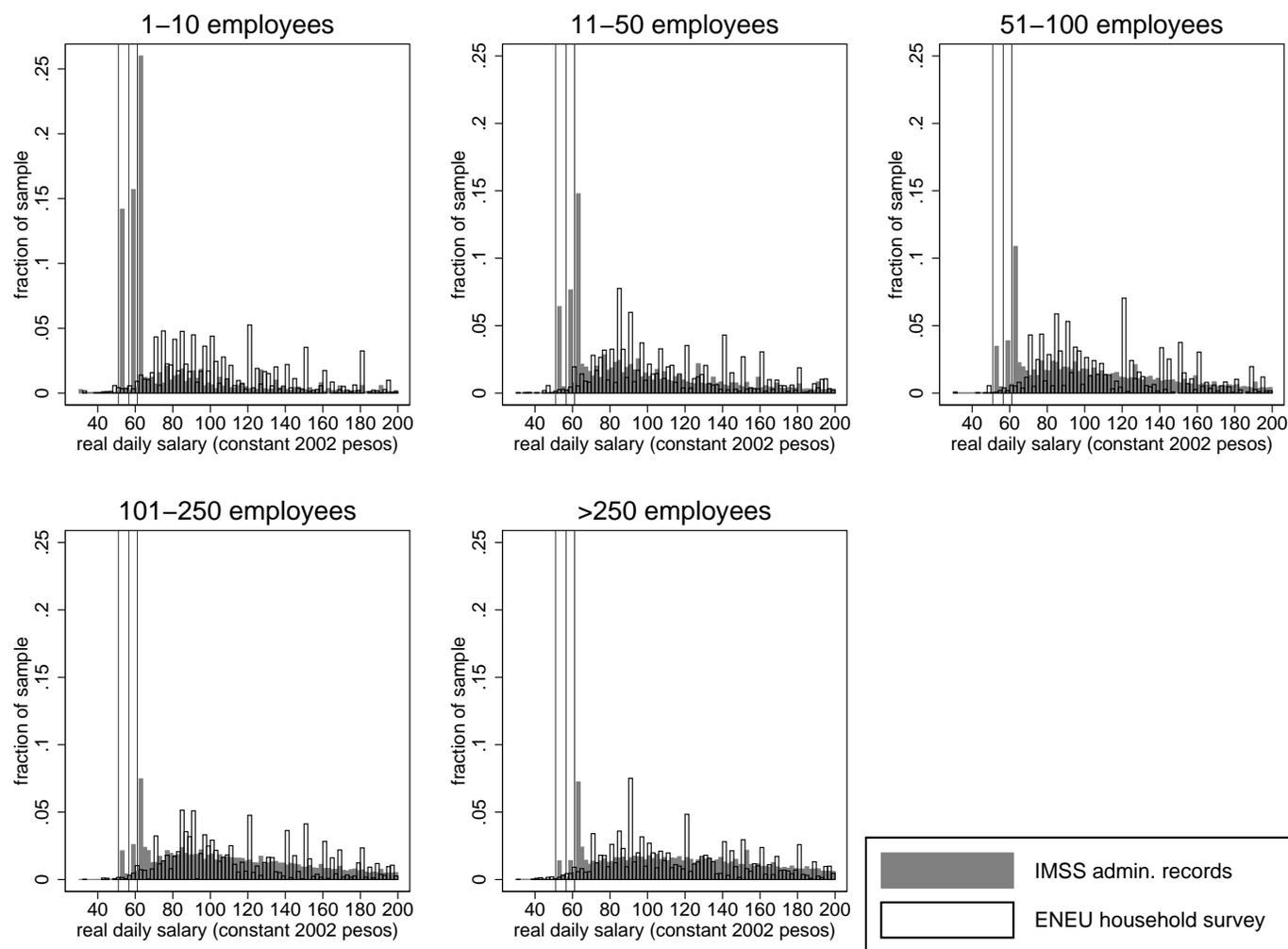
Notes: Samples are “baseline” samples of men: IMSS sample is permanent male workers ages 16-65 in 16 cities in original ENEU sample; ENEU sample includes full-time male workers ages 16-65 in 16 cities in original ENEU sample who report receiving IMSS coverage as an employment benefit. Data are from second quarter. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 5 pesos wide. The rightmost bin captures all individuals with reported wages at or above the minimum IMSS topcode over the study period (from 1991). Average 2002 exchange rate: 9.66 pesos/dollar. See Section 3 and Appendix B (online) for details of sample selection.

Figure 7. Wage histograms, men, 1990, low wage levels



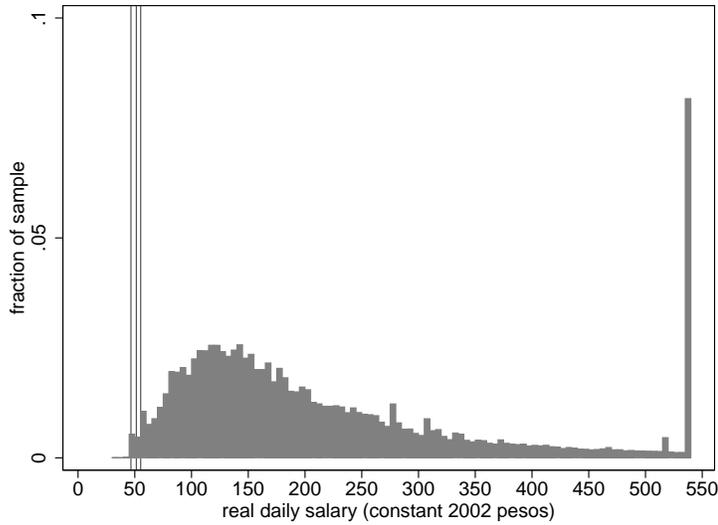
Notes: Histogram is similar to Figure 6 but only includes workers with wages less than 200 pesos/day (approx. \$20/day) in constant 2002 pesos. Bins are 2 pesos wide.

Figure 8. Wage histograms by firm size, men, 1990, low wage levels



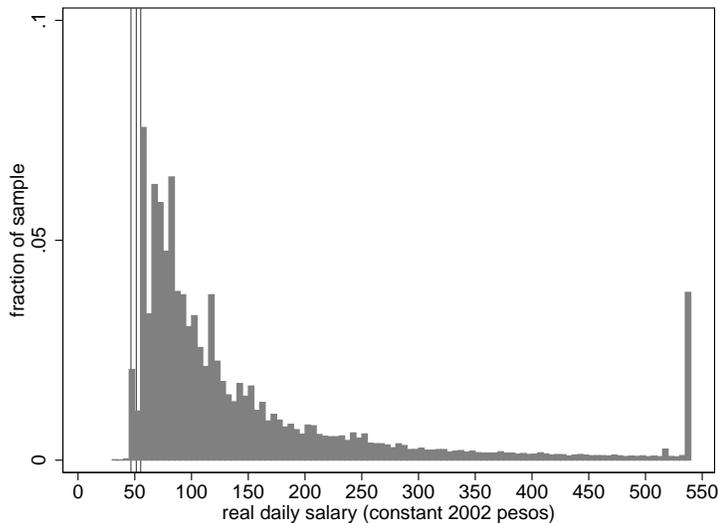
Notes: Samples are “baseline” samples of men, similar to those used in Figure 6 but only including workers with wages less than 200 pesos/day in constant 2002 pesos. ENEU sample only includes observations for which firm size variable is observed. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 2 pesos wide. Average 2002 exchange rate: 9.66 pesos/dollar. See Section 3 and Appendix B (online) for details of sample selection.

Figure 9. Wage histograms, men, 1993, establishments linked to EIA



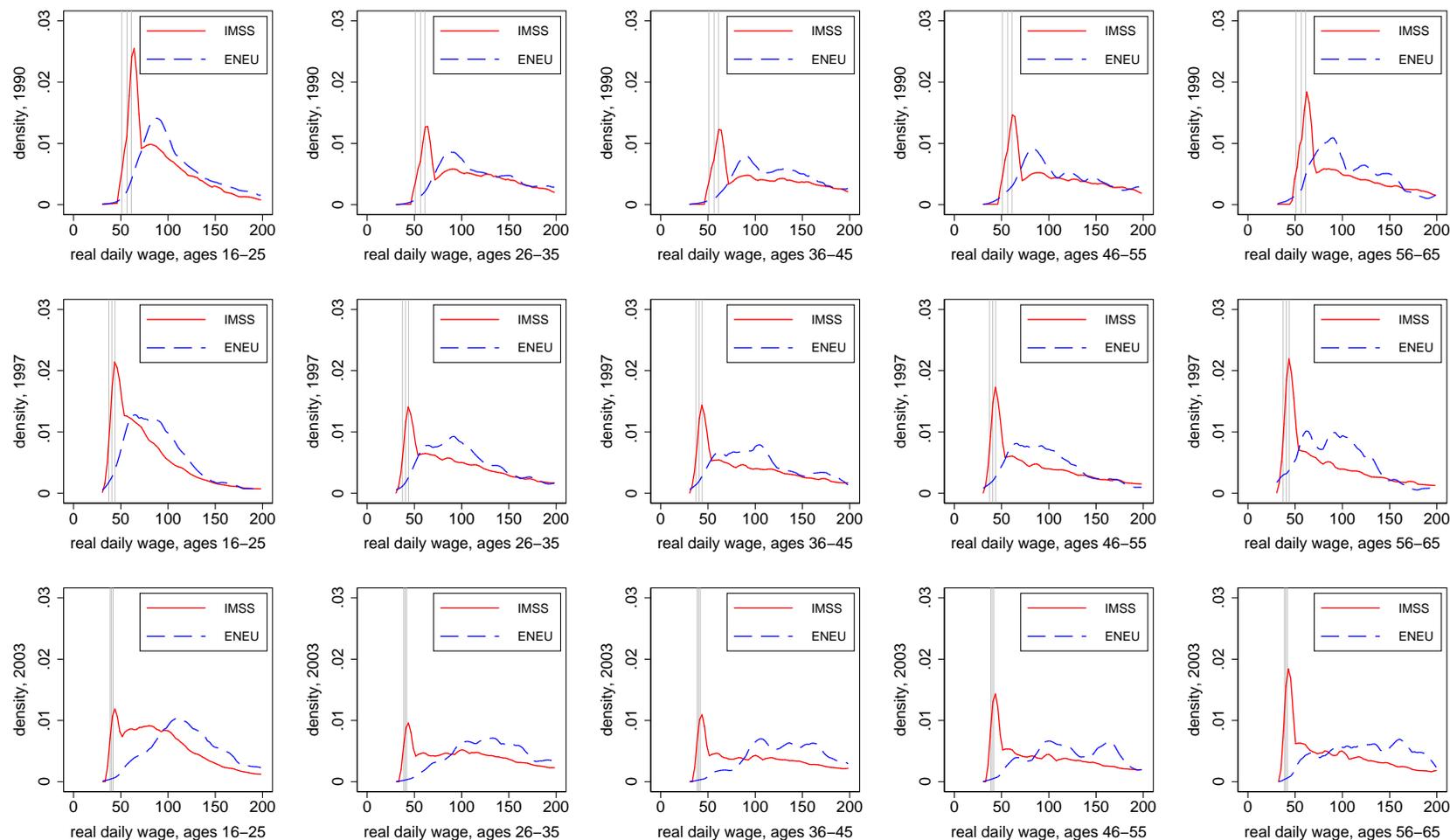
Notes: Sample is permanent male workers ages 16-65 in IMSS data in 2405 establishments that can be linked to a balanced 1993-2003 panel from the *Encuesta Industrial Anual (EIA)* [Annual Industrial Survey], which excludes assembly-for-export *maquiladora* plants. Data are from second quarter. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 5 pesos wide. See Section 3 and Appendix B (online) for details of sample selection.

Figure 10. Wage histograms, men, 1993, establishments linked to EMIME



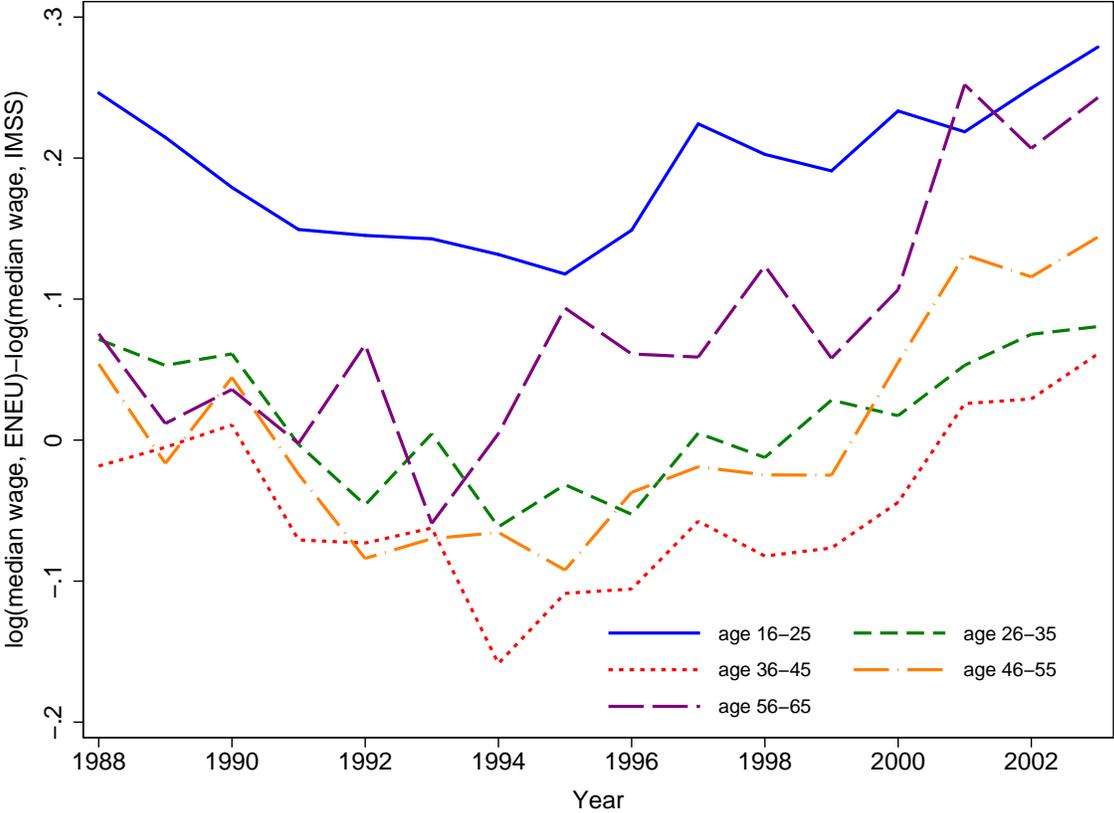
Notes: Sample is permanent male workers ages 16-65 in IMSS data in 506 establishments that can be linked to a balanced 1993-2003 panel from the *Estadísticas Mensuales de la Industria Maquiladora de Exportación (EMIME)* [Monthly Statistics on Maquiladora Export Industry], a dataset made up exclusively of assembly-for-export *maquiladora* plants. Data are from second quarter. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 5 pesos wide. See Section 3 and Appendix B (online) for details of sample selection.

Figure 11. Wage densities by age group, 1990, 1997, 2003, men



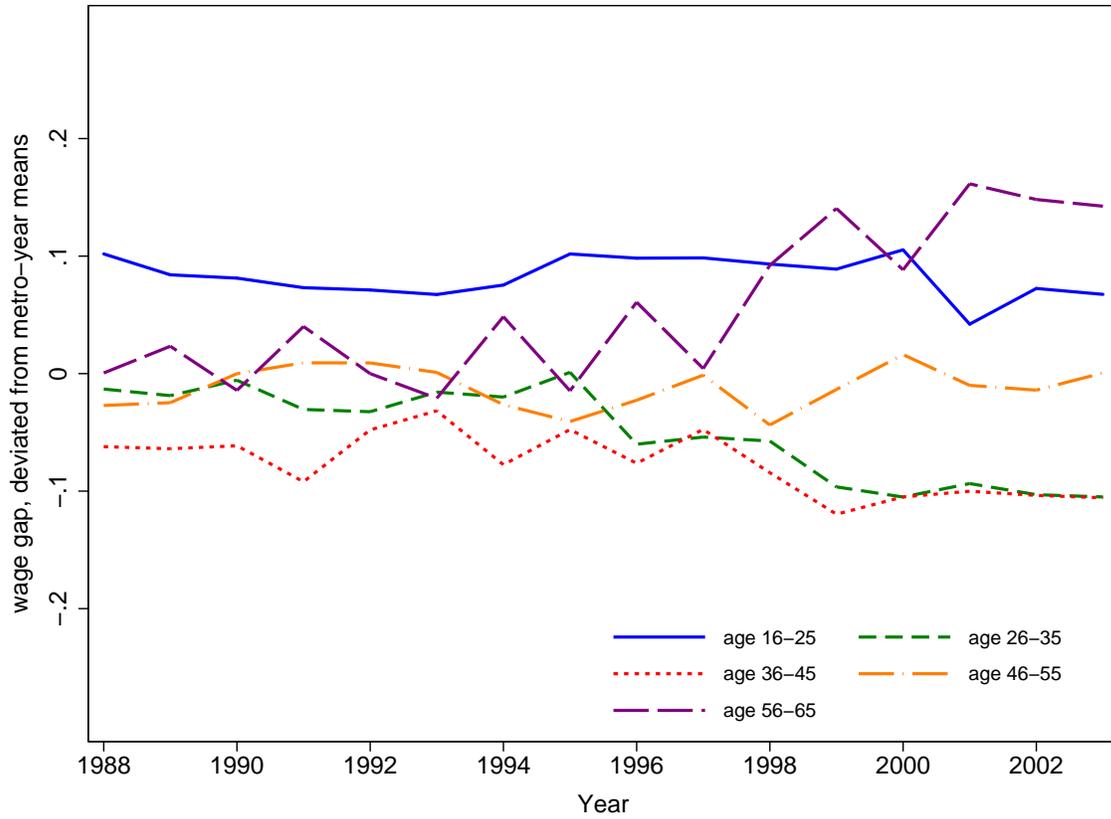
Notes: Densities are estimated using an Epanechnikov kernel and bandwidth 3 pesos for IMSS data and 6 pesos for ENEU data (using Stata `kdensity` command). Rows correspond to years 1990, 1997, 2003; columns to age groups 16-25, 26-35, 36-45, 46-55, 56-65. Wages (on x-axes) are real daily wages in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. Samples are baseline samples of men, only including workers with wages less than 200 pesos/day. See Section 3 and Appendix B (online) for details of sample selection.

Figure 12. Wage gaps by age group, men



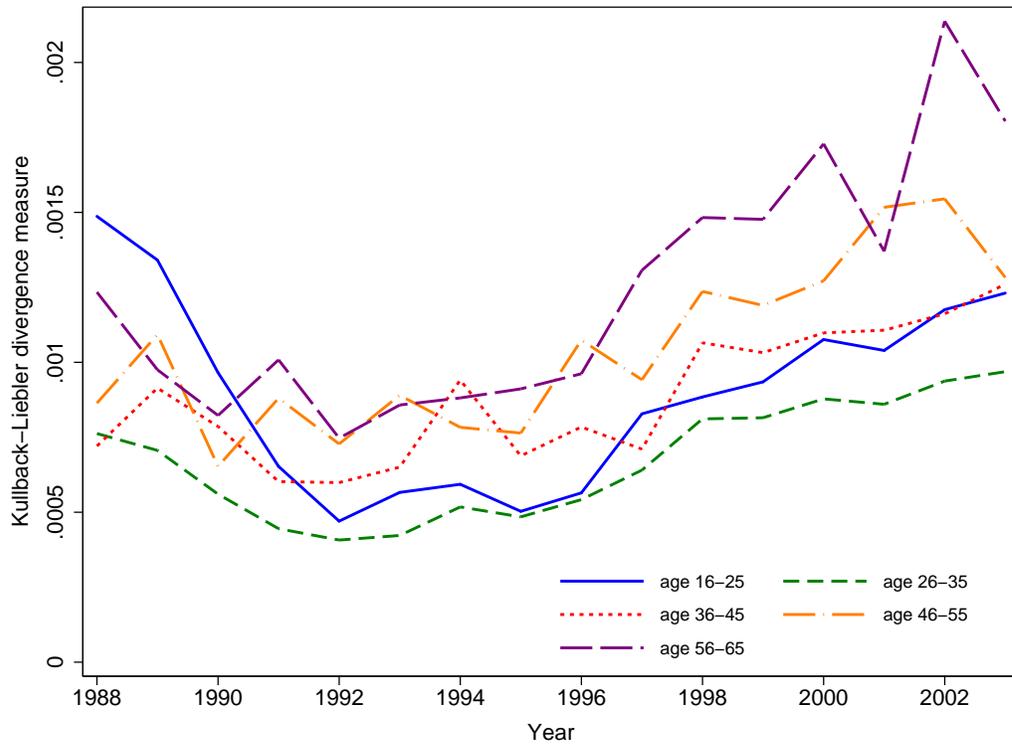
Notes: Each wage gap is the difference in log median real daily wages between ENEU and IMSS baseline samples, here calculated separately by age group. See Section 3 and Appendix B (online) for details of sample selection.

Figure 13. Wage gaps by age group, men, deviated from metro-year means



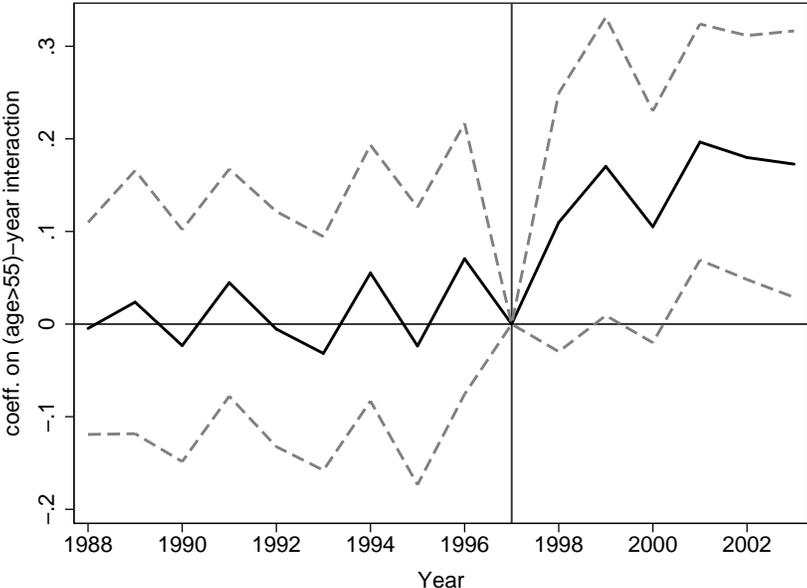
Notes: Each wage gap is the difference in log median real daily wages between ENEU and IMSS baseline samples. To calculate deviated wage gaps, we calculate wage gaps separately by age group-year-metro area, regress them on a full set of metro area-year dummies, and average the residuals at the age-group level. See Section 3 and Appendix B (online) for details of sample selection.

Figure 14. Kullback-Liebler divergence by age group, men



Notes: Samples are IMSS and ENEU baseline samples. For details on calculation of Kullback-Liebler divergence measure, see Section 5 of text. See Section 3 and Appendix B (online) for details of sample selection.

Figure 15. Differential effect of reform on wage gap, ages 55-65, men



Notes: Figure plots coefficients for $1(\text{age} > 55) \cdot \text{year}$ interaction term from Column 3 of Table 3. The dotted lines indicate the 95 percent confidence interval.

Table 1. IMSS pension wealth simulation, men

Years of Contributions	Plan	Real Daily Wage					
		43	100	200	300	500	1079
35	PRA	399.0	815.0	1626.2	2437.3	4059.7	8759.2
	PAYGO	399.0	399.0	603.8	890.2	1483.6	3202.7
30	PRA	399.0	523.4	1044.3	1565.3	2607.1	5625.1
	PAYGO	399.0	399.0	510.7	743.3	1238.9	2674.3
25	PRA	399.0	399.0	659.1	987.8	1645.3	3549.9
	PAYGO	399.0	399.0	406.9	579.5	965.8	2085.0
20	PRA	88.0	202.4	403.9	605.4	1008.4	2175.7
	PAYGO	399.0	399.0	399.0	449.6	749.3	1617.5
15	PRA	51.2	117.8	235.0	352.2	586.6	1265.7
	PAYGO	399.0	399.0	399.0	399.0	504.5	1089.1
10	PRA	26.8	61.7	123.1	184.5	307.4	663.2
	PAYGO	399.0	399.0	399.0	399.0	399.0	560.7
5	PRA	10.7	24.6	49.0	73.5	122.4	264.2
	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, 1079 pesos is the topcode we impose (corresponding to the lowest real value of IMSS topcode over study period.) See Section 2.3 and Appendix A.3 (online) for further details.

Table 2. Comparison of IMSS baseline sample and various ENEU samples, men

	IMSS baseline sample (1)	full ENEU sample (2)	ENEU w/ IMSS (3)	ENEU w/o IMSS (4)	ENEU permanent w/ IMSS (5)	ENEU full-time w/ IMSS (6)
A. 1990						
real avg daily wage	156.41 (0.09)	164.01 (1.58)	173.12 (1.94)	144.00 (2.62)		166.87 (1.85)
age	31.81 (0.01)	31.46 (0.15)	32.13 (0.17)	29.98 (0.29)		32.22 (0.17)
fraction employed in ests >100 employees	0.52 (0.00)	0.43 (0.01)	0.55 (0.01)	0.18 (0.01)		0.55 (0.01)
N (raw observations)	1714518	16169	11592	4577		10978
N (population, using weights)	1714518	2578847	1772523	806324		1645229
B. 2000						
real avg daily wage	160.28 (0.09)	148.32 (1.31)	161.28 (1.60)	120.88 (2.16)	166.56 (1.80)	155.93 (1.59)
age	32.77 (0.01)	32.22 (0.14)	32.82 (0.16)	30.94 (0.28)	33.22 (0.17)	32.88 (0.16)
fraction employed in ests >100 employees	0.58 (0.00)	0.44 (0.01)	0.59 (0.01)	0.10 (0.01)	0.63 (0.01)	0.59 (0.01)
N (raw observations)	2449442	19171	14063	5108	11918	13246
N (population, using weights)	2449442	3509828	2384267	1125561	2042988	2225318

Notes: All columns focus on wage-earning male workers ages 16-65 in manufacturing, construction, and retail/hotel/restaurant sectors in 16 metropolitan areas from the original ENEU sample. Column 1 includes IMSS baseline sample; Column 2 includes full ENEU (household survey) sample (satisfying aforementioned criteria); Column 3 includes employees in ENEU who report receiving IMSS benefit in current employment; Column 4 includes employees in ENEU who report not receiving IMSS benefit; Column 5 includes employees in ENEU who report receiving IMSS benefit and having a written contract of indefinite duration; and Column 6 includes employees in ENEU who report receiving IMSS benefit and working at least 35 hours in previous week. Standard errors of means in parentheses. In IMSS data, the fraction in establishments with >100 employees variable refers to permanent employees. In the ENEU survey, the establishment-size question asks the total number of employees (without specifying permanent vs. temporary.) For further details, see Section 3 and Appendix B (online).

Table 3. Differential effects on wage gap, men

	dep. var.: log(median wage, ENEU) - log(median wage, IMSS)		
	(1)	(2)	(3)
1(age > 55)*1988	-0.005 (0.088)	-0.005 (0.065)	-0.005 (0.058)
1(age > 55)*1989	0.024 (0.103)	0.024 (0.083)	0.024 (0.072)
1(age > 55)*1990	-0.023 (0.099)	-0.023 (0.071)	-0.023 (0.064)
1(age > 55)*1991	0.045 (0.106)	0.045 (0.070)	0.045 (0.063)
1(age > 55)*1992	-0.005 (0.090)	-0.005 (0.068)	-0.005 (0.065)
1(age > 55)*1993	-0.032 (0.100)	-0.032 (0.071)	-0.032 (0.064)
1(age > 55)*1994	0.055 (0.103)	0.055 (0.076)	0.055 (0.071)
1(age > 55)*1995	-0.024 (0.101)	-0.024 (0.080)	-0.024 (0.076)
1(age > 55)*1996	0.071 (0.102)	0.071 (0.077)	0.071 (0.075)
1(age > 55)*1998	0.110 (0.092)	0.110 (0.077)	0.110 (0.071)
1(age > 55)*1999	0.170 (0.120)	0.170* (0.092)	0.170** (0.082)
1(age > 55)*2000	0.105 (0.104)	0.105 (0.069)	0.105* (0.064)
1(age > 55)*2001	0.197** (0.098)	0.197*** (0.071)	0.197*** (0.065)
1(age > 55)*2002	0.180* (0.103)	0.180** (0.071)	0.180*** (0.067)
1(age > 55)*2003	0.173* (0.104)	0.173** (0.075)	0.173** (0.073)
metro area effects	N	Y	
year effects	Y	Y	
metro-year effects	N	N	Y
age category effects	Y	Y	Y
R-squared	0.14	0.67	0.77
N	1280	1280	1280

Notes: Samples are IMSS and ENEU baseline samples. The dependent variable (wage gap) is the difference in log median real daily wages between the ENEU and IMSS baseline samples, calculated separately by metro area-age group-year. *** 1%, ** 5%, * 10% level. See Section 3 and Appendix B (online) for further details of data processing

Table 4. Differential effects on employment gap, men

	dep. var.: $\log(\text{employment, ENEU}) - \log(\text{employment, IMSS})$		
	(1)	(2)	(3)
1(age > 55)*1988	-0.026 (0.105)	-0.026 (0.098)	-0.026 (0.100)
1(age > 55)*1989	0.048 (0.103)	0.048 (0.097)	0.048 (0.103)
1(age > 55)*1990	0.077 (0.096)	0.077 (0.088)	0.077 (0.097)
1(age > 55)*1991	0.109 (0.111)	0.109 (0.110)	0.109 (0.109)
1(age > 55)*1992	0.054 (0.101)	0.054 (0.096)	0.054 (0.100)
1(age > 55)*1993	0.098 (0.092)	0.098 (0.087)	0.098 (0.091)
1(age > 55)*1994	-0.224** (0.098)	-0.224** (0.095)	-0.224** (0.101)
1(age > 55)*1995	0.029 (0.112)	0.029 (0.105)	0.029 (0.107)
1(age > 55)*1996	0.005 (0.102)	0.005 (0.100)	0.005 (0.102)
1(age > 55)*1998	0.045 (0.106)	0.045 (0.099)	0.045 (0.104)
1(age > 55)*1999	0.031 (0.112)	0.031 (0.104)	0.031 (0.106)
1(age > 55)*2000	-0.006 (0.094)	-0.006 (0.093)	-0.006 (0.093)
1(age > 55)*2001	0.014 (0.110)	0.014 (0.109)	0.014 (0.105)
1(age > 55)*2002	0.091 (0.113)	0.091 (0.107)	0.091 (0.103)
1(age > 55)*2003	0.034 (0.094)	0.034 (0.093)	0.034 (0.091)
metro area effects	N	Y	
year effects	Y	Y	
metro-year effects	N	N	Y
age category effects	Y	Y	Y
R-squared	0.28	0.37	0.54
N	1280	1280	1280

Notes: Samples are IMSS and ENEU baseline samples, collapsed to metro area-age group-year level. *** 1%, ** 5%, * 10% level. See Section 3 and Appendix B (online) for further details of data processing