DISCUSSION PAPER SERIES

IZA DP No. 7591

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

Todd Kumler
Eric Verhoogen
Judith Frías

August 2013

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

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

Todd Kumler
Columbia University

Eric Verhoogen
Columbia University
and IZA

Judith Frías
IMSS

Discussion Paper No. 7591
August 2013

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 7591
August 2013

ABSTRACT

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

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. We develop a simple partial-equilibrium model of endogenous compliance by heterogeneous firms to guide the empirical investigation. We then compare 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 investigate the extent of wage under-reporting and how it responded to an important change in the social security system. We document that under-reporting by formal firms is extensive, and that compliance is better in larger firms. Using a difference-in-differences strategy based on the 1997 Mexican pension reform, which effectively tied pension benefits more closely to reported wages for younger workers than for older workers, we show that the reform led to a relative decline in under-reporting for younger workers. Within metro area/sector/firm size cells, the decline in under-reporting was greater in cells initially employing a younger workforce on average. The empirical patterns are consistent with our theoretical model and suggest that giving employees incentives and information to improve the accuracy of employer reports can be an effective way to improve payroll-tax compliance.

JEL Classification: O17, H26, H55

Keywords: tax compliance, state capacity, Mexico, heterogeneous firms, pension reform

Corresponding author:

Eric Verhoogen
School of International and Public Affairs
Department of Economics
Columbia University
420 W. 118th St., Room 1022, MC 3308
New York, NY 10027
USA
E-mail: eric.verhoogen@columbia.edu

* 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, Roger Gordon, Santiago Levy, Suresh Naidu, Ben Olken, Kiki Pop-Eleches, Bernard Salanié, Monica Singhal and seminar participants at UC Berkeley, Washington University, Carnegie-Mellon University, the University of Houston, Warwick University, the CEPR-AMID Development Economics Symposium in Milan, IGC Growth Week 2012, Duke University, the University of Maryland, the 23rd BREAD conference, the World Bank, PUCRio, the NBER Public Economics Spring 2013 meeting, and MIT for helpful conversations. Verhoogen gratefully acknowledges 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*. This paper was previously circulated under the title, "Enlisting Workers in Monitoring Firms: Payroll Tax Compliance in Mexico."

1 Introduction

A growing body of research suggests that lack of state capacity — in particular, difficulty in raising taxes to fund public goods — is an important constraint on the growth of developing countries (Burgess and Stern, 1993; Besley and Persson, 2013). Developing countries generally tend to have low ratios of tax revenues to GDP and large informal sectors. Mexico, the focus of our study, 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 has been estimated to make up 40 percent or more of total output (OECD, 2011b; IMF, 2010; Schneider and Enste, 2000). Given weak enforcement institutions and widespread evasion, the task of improving the fiscal capacity in developing countries is a difficult one, and there is acute interest among policy-makers 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; Busso, Fazio, and Levy, 2012).[1] Governments in a number of countries have implemented programs to reduce registration costs and induce firms to formalize. An active literature has investigated the effects of such programs and other inducements to formalize on the behavior of firms.[2]

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, Knudsen, Kreiner, Pedersen, and Saez (2011) find little evasion when incomes are reported by employers or other third parties.[3] The view that third-party reporting is effective in ensuring compliance is widespread among

[1] Other notable theoretical work on firms' decisions about whether to formalize includes Rauch (1991), De Paula and Scheinkman (2011), and Galiani and Weinschelbaum (forthcoming).

[2] See e.g. Bruhn (2011), Kaplan, Piedra, and Seira (2011), Fajnzylber, Maloney, and Montes-Rojas (2011), Monteiro and Assunção (2012), McKenzie and Sakho (2010), de Mel, McKenzie, and Woodruff (2012).

[3] In 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).

practitioners and government agencies in developed countries (see e.g. Plumley (2004) and 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; Slemrod, 2007). 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 compare 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 (the *Encuesta Nacional de Empleo Urbano* (ENEU)) — to draw inferences about the extent of wage under-reporting and how it has responded to changes in the social security system. In particular, we focus on a major pension reform that introduced a system of personal retirement accounts, passed by the Mexican Congress in December 1995 and implemented on July 1, 1997. As discussed in more detail below, prior to the reform the social security benefits of most workers were largely insensitive to the wages reported by firms on their behalf. The reform tied individual pensions more closely to firms' wage reports and made it easier for employees to observe those reports. Workers already in the traditional 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, their expected pension was higher 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. We use this differential impact by age as the basis for a difference-in-differences estimation strategy.

To guide our empirical investigation, we develop a simple partial-equilibrium model of endogenous compliance by heterogeneous firms. Firms are assumed to be monopolistically competitive and to differ in productivity (in a setting similar to Melitz (2003) but without international trade). The cost of evasion is assumed to be increasing both in the unreported part of the wage per worker and in firm output, for reasons that may include the difficulty of maintaining collusion in large firms, as argued in a recent paper by Kleven, Kreiner, and Saez (2009), or simply the greater visibility of larger firms to auditors. One consequence of this assumption is that evasion is declining in firm size in cross-section. The model further predicts that a social-security reform like the one that occurred in Mexico, which raises the sensitivity of social-security benefits to the

reported wage, will lead to a within-firm decline in evasion for affected workers. The model carries ambiguous implications for the incidence of the reform on observable wage variables, but we will show in the context of the model that the predictions for evasion do not depend on the incidence.

A key limitation of our study is that, although evasion decisions are taken at the level of individual firms, the household labor-force survey does not contain firm identifiers and hence we are not able to construct measures of evasion at the firm level. Instead, we construct measures of evasion at the level of cells defined by different combinations of metropolitan areas, broad sectors, firm-size categories and age groups, depending on the specification. Within each cell, we compare the distributions of wage measures that in the absence of evasion should coincide — the post-tax wage from the administrative records of the social security agency and the take-home wage reported on the household survey — and measure evasion either as the difference in log median or log mean wage or as the excess mass in the administrative-records distribution at the left tail. We focus on male workers, for reasons discussed below.

Using these measures, we first document a set of cross-sectional relationships prior to the reform. Evasion is substantial, as one might expect given the weak relationship between wage reports and social security benefits for most workers. We also show that evasion is declining in firm size, consistent with our model and the previous theoretical work by Kleven, Kreiner, and Saez (2009), although evasion remains non-trivially positive even among quite large firms.

We then turn to the difference-in-differences strategy using the pension reform, conducting the analysis at two levels of aggregation. First, we focus on market-level outcomes by age group, where the market is defined either as the whole economy or as a local labor market (metropolitan area). We show that evasion rates by age group followed similar trends prior to the reform, but that the oldest age group, ages 55-65, saw a significant relative increase following the passage of the reform, a pattern that is robust across our measures of evasion. Second, we consider changes in evasion within cells defined by metropolitan area, sector and firm-size category. We show that cells with an initially greater share of younger workers saw a relative decrease in evasion following the reform. This pattern is again robust across evasion measures and does not appear to be driven entirely by sorting of workers of different age groups across cells. With the caveat that the results could in part reflect sorting of workers across individual firms within cells, the results suggest that the closer tying of social security benefits to reported wages for younger workers led to a relative improvement in compliance in firms that relied more heavily on such workers.

Overall, the results support the view that tying benefits to reported wages and making it easier

for workers to observe firms' reports can be an effective way to improve payroll-tax compliance. It seems likely that the change in incentives was the more important aspect of the reform: if incentives had not changed, and benefits had remained largely insensitive to wage reports, it is not clear why information alone would have led to changes in compliance. But the "experiment" we consider combined both elements, and the effects we estimate should be interpreted as combined effects of incentives and information.[4]

As will become clear in the theoretical discussion, we would expect reducing the payroll tax to have similar effects on compliance as tying benefits more closely to reported wages. But the two types of interventions have potentially very different effects on overall tax revenues, with the latter expected to generate a greater increase in the tax take. If the policy goal is to increase fiscal capacity, this is a potentially important advantage.

This paper is related to a number of different literatures. There is 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 in cross-section. Tonin (2011) investigates the effect of a minimum wage increase on reporting patterns in Hungary, using the relationship between reported household income and expenditures to draw inferences about the extent of evasion.[5] An independent paper by Bérgolo and Cruces (2012) examines an Uruguayan policy to extend health benefits to children of covered workers and finds an average increase in evasion for workers with children (as measured by workers' self-reports on a household survey), an effect which appears to be driven primarily by selection of workers from informal into formal employment (with under-reported wages).[6] Another independent paper by Mao, Zhang, and Zhao (2013) investigates the effect of average payroll tax rates at the city level in China on participation in the social security system and on the share of income subject to taxation.[7] This paper appears to be the first to analyze how tying benefits more closely to reported wages can contribute to improved compliance.[8] Relative to this literature, the paper is also distinctive in its explicit

[4] Our argument should not be interpreted as advocating a system of personal accounts *per se*; one could imagine a change in pension benefits under the traditional pay-as-you-go system that would have had similar effects.

[5] Also 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.

[6] A possible advantage of our approach relative to Bérgolo and Cruces (2012) is that it does not rely on workers' self-reports, which may not be entirely accurate (Slemrod, 2007).

[7] Advantages of our approach relative to Mao, Zhang, and Zhao (2013) are that we can construct payroll-tax and pension-benefit rates directly for different types of workers, without having to rely on average rates paid in a city, and that we have direct measures of evasion, as opposed the ratio of taxable salary to total salary (their measure), which may reflect legal subsidies and allowances rather than evasion.

[8] Bailey and Turner (2001) suggest verbally that tying pension benefits to contributions would have the effect of reducing evasion.

consideration of the endogenous compliance decisions of heterogeneous firms.

This paper is also related to an active recent literature on the role of firms in tax systems (Kopczuk and Slemrod, 2006; Gordon and Li, 2009; Dharmapala, Slemrod, and Wilson, 2011). An important strand of this literature has emphasized that, in the presence of differences in the costs of evasion, which parties remit taxes is not irrelevant for the incidence of those taxes (Slemrod, 2008; Kopczuk, Marion, Muehlegger, and Slemrod, 2013). Our study focuses not on changes in the point of taxation — firms remit all payroll taxes in Mexico, both before and after the pension reform — but rather on a change in the incentives of workers to ensure accurate reporting by their employers, and we see our study as complementary to this strand of work. Our argument can be thought of as a labor-market analogue of the claim that value-added taxes (VATs) have attractive enforcement properties relative to retail sales taxes in part because they give each party in a supply-chain transaction greater incentive to ensure that the other reports accurately (Kopczuk and Slemrod, 2006; Keen and Lockwood, 2010; Pomeranz, 2013). This paper is related more broadly 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).

This paper is part of a small but growing literature using administrative records from developing countries to document various aspects of taxpayer behavior (Pomeranz, 2013; Kleven and Waseem, 2013; Best, Brockmeyer, Kleven, Spinnewijn, and Waseem, 2013). It is in the spirit of a growing empirical literature in development economics examining how corruption and other forms of illegal behavior respond to economic incentives, recently surveyed by Olken and Pande (2012). Papers using the general strategy of comparing information from more than one data source to infer illicit behavior (in other contexts) include Pissarides and Weber (1989), Fisman and Wei (2004), Olken (2006), Gorodnichenko, Martinez-Vazquez, and Peter (2009), Marion and Muehlegger (2008), Hurst, Li, and Pugsley (2011), Braguinsky, Mityakov, and Liscovich (2010), and Niehaus and Sukhtankar (forthcoming).

2 Institutions: 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 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 in-

surance, 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.[9]

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. This reform replaced the entire PAYGO pension system with a system of personal retirement accounts (PRA). More extensive discussions of the pension reform are provided in Grandolini and Cerda (1998), Sales-Sarrapy, Solis-Soberon, and Villagomez-Amezcua (1996), and Aguila (2011).

Because of data constraints, discussed in more detail in Section 4 below, we focus on the years 1988-2003. 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 often inadequate, especially when analyzing the data separately by metropolitan area (or metropolitan area, firm size and sector), as explained below. We present the main tables and figures for women in Appendix D (online). To preview the results, the cross-sectional patterns are robust for women, but the 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

[9]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.

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, 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. Figures 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 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 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, independent of the reported wage. 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]

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

[11] There 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.

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

[13] 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.

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.

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 was above 25 percent in a number of subsequent years (1990-1991 and 1995-1996). (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] The congress subsequently 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,

[14] In 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.

[15] To 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.

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 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).[17]

Strictly speaking, pension values were insensitive to final wages only for infra-marginal workers whose *true* final wage corresponded to the minimum pension. If wages were under-reported to IMSS, as we argue below, then the graphs in Panel B of Figure 3 likely overstate the fraction of workers whose pensions were insensitive to under-reporting. To address this, in Panel C of Figure 3 we plot similar graphs using final average wage percentiles calculated from the ENEU household data (described in Section 4 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.

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"

[16] In 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).

[17] In addition, because pensions were calculated only on the basis of the last five years of employment, any worker who was certain that he or she would work for more than five years in covered employment could also be certain that the current reported wage would not affect the pension benefit. In unreported results, we have investigated whether we see an increase in reported wages five years before retirement, as one might expect if workers were being sophisticated in adjusting strategically to the five-year rule, but we do not find a significant change.

equal to 5.5 percent of the current minimum wage in Mexico City.[18] 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.[19] The 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.[20] 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). 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 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

[18] Employees also had the option to contribute to a voluntary retirement savings account. Few did so initially; in 2003, taxes on such voluntary contributions were reduced. Lara-Ibarra (2011) examines this reform.

[19] The 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.

[20] A worker who receives the minimum pension must choose this option.

interest rates and less rapid declines in the real minimum wage would be less favorable to the PRAs. Details of the simulation are in Appendix A.3 (online).

One way to see the differences in incentives by age in the system is to compare pension wealth for workers of different ages in 1997. Table 1 displays the real present value of pension wealth by wage level for male workers of different ages in 1997, all of whom began working at age 25 and expect to continue working until age 60, assuming real wages are constant over their lifetimes. Numbers in italics (and in blue where color is available) indicate that the PRA pension is more valuable than the PAYGO pension. The message of the simulation is clear: the PRA pension is expected to be more valuable only for younger workers who expect to contribute to the personal account for 25 or more years, and among these workers the PRA pension is relatively more attractive for higher-wage workers.[21]

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 may have been 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 the voluntary savings account (see footnote 18) and the housing savings account (see footnote 12). The bottom section reports 3-year returns and commissions for each AFORE, as well as the average 5-year net return (at

[21] Another way to see the effect of the reform is to consider the values of the pensions for different numbers of years of expected contributions, for a worker who entered then system on June 30, 1997, as presented in Appendix Table A5. Note that 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. But conditional on a worker having at least 10 years of contributions, we again see that the attractiveness of the PRA pension is increasing in the number of years of contributions and the wage. The median wage for male workers is just above 100 pesos/day, and for a worker at this level the PRA only becomes more attractive if he expects to contribute for more than 25 years.

left). It 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.[22] 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 on paper: the corporate income tax rate went from 39 to 34 percent over the 1988-2003 period.[23] 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

[22] One form of auditing is written into the social security law: employers with 300 or more employees are required to submit an audit by a certified public accountant to IMSS (since 1993) as well as to the Mexican tax authority (since 1991). This may be another reason why evasion is more difficult for larger firms, as we assume below.

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

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.[24]

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).

3 Theoretical Framework

To organize our thinking about the empirical analysis, this section develops a simple model of endogenous payroll-tax compliance by heterogeneous firms, building on earlier theoretical work by Yaniv (1992), Kopczuk and Slemrod (2006), Kleven, Kreiner, and Saez (2009), Besley and Persson (2013) and others. The model is special in a number of ways: we assume a particular market structure (which is standard but nonetheless restrictive), we impose a particular structure on the cost-of-evasion function faced by firms, and so on. Our goal is not to be completely general, and the reduced-form empirical analysis that follows does not require the restrictive theoretical assumptions to be literally correct. Rather, our aim is to spell out in a precise way why it makes sense to look at the relationships we examine empirically.

3.1 Basic Set-up

Consider a setting with a competitive labor market populated by homogeneous workers and a continuum of heterogeneous, monopolistically competitive firms.[25] Let τ_f be the payroll tax statutorily imposed on the firm and τ_w the payroll tax statutorily imposed on workers; both are

[24] Madzharova (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.

[25] 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 significantly with relatively little payoff in additional insight. We leave the development of a more general framework to future work.

remitted by the firm and will enter similarly in our model.[26] 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.[27] Both the reported wage, w_r, and the net wage, w_{net}, correspond to quantities that are also 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 above, 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.[28]

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 corresponds to the Mexican institutional setting, where the tax payment includes contributions for health care as well as pension benefits. 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:

$$L^S_{agg} = Bw_e^{\rho} \tag{1}$$

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

[26] As pointed out by Slemrod (2008), in settings in which the costs of evasion by employers and employers 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.

[27] 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.

[28] In the theory, we also abstract from the minimum wages in Mexico (of which there are three, depending on region); we will consider them in the empirical section below.

wage work.[29] 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 analysis below. An alternative approach to deriving the labor-supply elasticity (1) 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 (forthcoming). 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.

3.2 The Firm's Problem

We build on a standard model of heterogeneous firms under monopolistic competition, similar to Melitz (2003) but without international trade. Assuming a Dixit-Stiglitz (1977) representative consumer, the demand for each differentiated variety, ω, can be written:

$$x(\omega) = Ap(\omega)^{-\sigma} \tag{4}$$

where $x(\omega)$ is the quantity consumed; $p(\omega)$ is the price; and σ is a parameter capturing the elasticity of substitution between varieties.[30] A captures the general level of demand, which individual firms treat as exogenous; in this partial-equilibrium framework, we abstract from the determi-

[29] Suppose that for each worker,

$$U(c, L) = u(c) - \beta L^{\frac{\rho+1}{\rho}} \tag{2}$$

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:

$$c + (\overline{L} - L)w_e = w_e\overline{L} \tag{3}$$

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

[30] If the representative consumer has utility

$$U = \left[\int_{\omega\in\Omega} (x(\omega))^{\frac{\sigma-1}{\sigma}}\, d\omega\right]^{\frac{\sigma}{\sigma-1}} \tag{5}$$

where Ω represents the set of all differentiated varieties available in the market, then optimization yields (4) with $A = UP^{\sigma}$, where U is defined in (5) and

$$P \equiv \left[\int_{\omega\in\Omega} (p(\omega))^{1-\sigma}\, d\omega\right]^{\frac{1}{1-\sigma}} \tag{6}$$

nants 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.[31] There is assumed to 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 $c''(w_u) > 0$.[32] 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.[33] 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.[34] (From the definitions of the wage variables above, we have that $w_r = (w_e - w_u)/(1 - (\tau - b))$. 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, will are pinned down by the firm-specific demand curve, (4). Per-period profit for each firm can be written:

$$\pi(w_u, p; \varphi, w_e) = \left\{ p - \frac{w_e - (\tau - b)w_u}{\varphi(1 - (\tau - b))} - c(w_u) \right\} x - f \tag{7}$$

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:

$$c'(w_u) = \frac{\tau - b}{\varphi(1 - (\tau - b))} \tag{8}$$

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

[32] 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.

[33] Also, as noted in footnote 22, 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.

[34] 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.

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:

$$p^*(w_e, \varphi) = \left(\frac{\sigma}{\sigma - 1}\right) \left\{ \frac{w_e - (\tau - b) w_u^*(\varphi)}{\varphi(1 - (\tau - b))} + c(w_u^*(\varphi)) \right\} \tag{9}$$

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:

$$x^*(w_e, \varphi) = A p^*(w_e, \varphi)^{-\sigma} \tag{10}$$

and the firm's labor demand is:

$$L^D(w_e, \varphi) = \frac{x^*(w_e, \varphi)}{\varphi} = \frac{A p^*(w_e, \varphi)^{-\sigma}}{\varphi} \tag{11}$$

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 (11) over firms active in the market:

$$L_{agg}^D(w_e) = \int_{\varphi^{min}}^{\varphi^{max}} L^D(w_e, \varphi) g(\varphi) d\varphi \tag{12}$$

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

$$L_{agg}^S(w_e) = L_{agg}^D(w_e) \tag{13}$$

where aggregate labor supply, $L^S_{agg}(w_e)$, is given by (1).

3.3 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 (8) holding w_e fixed and rearranging, we have:

$$\frac{dw_u^*}{d\varphi} = -\frac{\tau - b}{\varphi^2 c''(w_u)(1-(\tau-b))} < 0 \tag{14}$$

That is, evasion is decreasing in firm productivity.[35]

Firm output is unambiguously increasing in productivity. To see this, first note that price is decreasing in productivity; differentiating both sides of (9) (again, holding w_e fixed) and using (8), we have:

$$\frac{dp^*}{d\varphi} = -\left(\frac{\sigma}{\sigma-1}\right)\left\{\frac{w_e - (\tau-b)w_u^*(\varphi)}{\varphi^2(1-(\tau-b))}\right\} < 0 \tag{15}$$

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 (10):

$$\frac{dx^*}{d\varphi} = -\sigma A (p^*)^{-\sigma-1}\left(\frac{dp^*}{d\varphi}\right) > 0 \tag{16}$$

Together (14) and (16) imply an unambiguously negative relationship between firm output and evasion.

The relationship between employment and productivity, and hence between employment and evasion, is more subtle. Differentiating (11):

$$\frac{dL^D}{d\varphi} = -\frac{x^*(w_e,\varphi)}{\varphi^2} + \frac{1}{\varphi}\left(\frac{dx^*}{d\varphi}\right) \tag{17}$$

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

[35] 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:

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

first term). Using (9) and (16)-(17), it can be shown that labor demand will be increasing in productivity if and only if

$$\frac{c(w_u^*(\varphi))}{\sigma - 1} < \frac{1}{\varphi}\left\{\frac{w_e - (\tau - b)w_u^*(\varphi)}{1 - (\tau - b)}\right\} \tag{18}$$

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 (18) 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 (16)) and reduces the magnitude of the second term in (17) 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 (18) holds. We will focus on this case hereafter. In this case, the extent of evasion is declining in employment, as it is in output.

3.4 Response 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 (8) with respect to b and rearranging:

$$\frac{dw_u^*}{db} = -\frac{1}{(1 - (\tau - b))^2 \varphi c''(w_u^*(\varphi))} < 0 \tag{19}$$

The unreported wage unambiguously decreases within a given firm. Note that $\frac{dw_u^*}{db}$ is firm-specific, but since the effect is negative for all values of φ, the aggregate effect will be negative for a given

set of firms over time.[36]

It is worth emphasizing that the response of $w_u^*(\varphi)$ to the policy change here does not depend 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 will proceed in the empirical analysis.

Because it is not our primary focus, we relegate the formal discussion of the incidence of the reform to the appendix. Briefly, there we show that for a finite labor-supply elasticity $\frac{dw_e}{db} > 0$: some of the incidence of the increase in benefits will be on workers. However, in general it is not possible to sign the effects of the reform on the firm-specific reported wage, $w_r(\varphi)$, or the firm-specific take-home wage, $w_{net}(\varphi)$. Intuitively, there are several potentially offsetting effects. 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, a higher b raises the equilibrium effective wage. It also induces induces the firm to report a greater share of the total wage payment (i.e. raise w_r relative to w_u). For the special case of homogeneous firms, it is possible to show that $\frac{dw_{net}}{db} < 0$; the take-home wage falls with the increase in benefits. But a priori it is not clear which of the effects will dominate. Appendix B makes these points in greater detail.

To sum up, this simple partial-equilibrium model carries a number of testable implications. First, and most obviously, we expect the unreported wage to be positive. Second, we expect a negative relationship between evasion and firm size as proxied by output or — under a plausible additional assumption — by employment. Third, we expect the pension reform to have a negative effect on evasion within a given firm or set of firms.

It is worth noting that, in this model, a decrease in the payroll-tax rate, τ, would be expected to have the same effect on evasion as an increase of the same size in the parameter relating reported wages to pension benefits, b. (Note that b always appears as part of the quantity $\tau - b$.) But the two types of interventions should not be expected to have the same effects on government

[36] Given the discussion of cross-sectional patterns in Section 3.3, 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,

$$\frac{d}{d\varphi}\left(\frac{dw_u^*}{db}\right) = \frac{\varphi c'''(w_u^*(\varphi))\frac{dw_u^*}{d\varphi} + c''(w_u^*(\varphi))}{\left[(1-(\tau-b))\varphi c''(w_u^*(\varphi))\right]^2} \tag{20}$$

where $\frac{dw_u^*}{d\varphi}$ is given by (14). Without imposing further structure on the cost of evasion, we do not have a clear prediction for a differential response by firm size.

revenues. To the extent that the policy goal is to raise revenue, increasing the sensitivity of benefits to reported wages is likely to be preferable.

4 Data

The establishments' wage reports are drawn from 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.[37] 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 for wage reporting (which occurred in 1991) in all years. We drop establishments with a single insured worker, since these are likely to be self-employed workers.[38] In the interests of comparability with the ENEU household 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 the only available formal-sector social insurance program: manufacturing, construction, and retail/hotel/restaurants. Other sectors contain a substantial share of public employees, who are covered by a separate system.[39] We focus primarily on men, for the reasons discussed in Section 2 above. (Results for women are reported in Appendix D (online).) We refer to the sample selected following these criteria as our IMSS baseline sample. Further details on sample selection and data processing in Appendix C (online).

The household data we use are from the *Encuesta Nacional de Empleo Urbano* (ENEU) [Na-

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

[38] Including these single-worker establishments has no effect on the results reported below.

[39] We focus on manufacturing, construction, and retail/hotel/restaurants in part so that we can be confident that respondents to the household survey are not mistaking coverage under the public-sector system for IMSS coverage.

tional 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 on 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, after social security taxes have been 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 C (online).

Although the ENEU survey does not contain a firm identifier, it does ask respondents about the size of the firm at which a he or she works. We use this information to generate a firm-size indicator taking on values 1-10, 11-50, 51-100, 101-250, or 250+ employees.[40] In addition, we drop workers with reported daily wages below 30 pesos (in 2002 constant pesos, approximately US$3, which is approximately 50 percent of the lowest legal 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.[41]

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 sample and various ENEU samples for 1990 and 2000, for a set of variables that are common between the sources: daily (post-tax) wage, age, and share in large establishments (with more than 100

[40]The survey allows for eight responses, 1, 2-5, 6-10, 11-15, 16-50, 51-100, 101-250 and 250+. To ensure that sample sizes within cells are sufficiently large, we create the five categories listed above.

[41]There 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 apparent benefit, since the IMSS data are available only until 2005) we focus on the period before the redesign.

employees). In constructing the daily post-tax wage in the IMSS data, we subtract the required social security contributions (discussed in Section 2 above) from the reported pre-tax IMSS wage. 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). 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. Nor does it appear that there was a significant large inflow to (or outflow from) formal employment in response to the pension reform.

As an additional check on the comparability of the baseline IMSS and ENEU samples, Table 3 reports the composition of the samples across two dimensions that will be crucial in our analysis, age and firm size. In order to ensure that we have sufficient sample size in the ENEU to calculate the evasion measures below, we group individuals into five age categories (ages 16-25, 26-35, 36-45, 46-55, 56-65). Maintaining sufficient sample size is also the motivation for focusing on the five firm-size categories mentioned above (1-10, 11-50, 51-100, 100-250, and 250+ employees.) 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 Figure 5 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. Given the issues in reconciling the samples discussed above (permanent workers in IMSS vs. full-time in ENEU, place-of-work in IMSS vs. place-of-residence in ENEU), it is perhaps not surprising that formal

Kolmogorov-Smirnov or Wilcoxon tests reject the null hypothesis that the samples are drawn from the same underlying distribution. Caution is thus warranted in interpreting our results below. We nonetheless feel that the samples appear to be sufficiently similar that it is not unreasonable to use the wage discrepancy between them as a measure of evasion. We also note again that any differences between the samples that are constant over time will be differenced out in our difference-in-differences procedure below.

5 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. We focus on the year 1990, in part because (in unreported results) we have been able to validate the ENEU sample against the population census from that year. To begin, Figure 6 plots simple histograms of raw data: *pre-tax* daily wages from the IMSS baseline sample (gray bars) and daily take-home wages the ENEU baseline sample (bars with black borders and no fill color), 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 clear: there is bunching in the IMSS sample slightly above the three minimum wages. These bunches 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.[42] The bunching and shift to the left of the distribution is precisely what one would have expected, given that, for most workers, social security benefits were insensitive to reported wages, as long as their firms made the minimum contributions on their behalf.[43]

A key empirical implication of our model, as well as of the previous theoretical work by Kleven et al. (2009), is that there is less evasion 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

[42]The 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.

[43]Note that we are showing the raw, pre-tax IMSS data here, to illustrate the bunching at the top-code and the minimum reportable values; when we use the post-tax IMSS wage below, the discrepancy in the distributions is even more evident.

interpreting these figures, since observed establishment size in the IMSS data may itself affected by firms' compliance decisions. Subject to this caveat, it appears that there is less bunching 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 bunching at the minimum allowable wage report, suggesting some under-reporting even in quite large firms.

To quantify the extent of non-compliance, we construct three measures of evasion. Recall from the theory that the unreported wage is the difference between the worker's net wage and the post-tax wage reported by the firm: $w_u = w_{net} - (1-\tau)w_r$. As noted above, the ENEU survey asks individuals their take-home wage, which corresponds to w_{net}, and the raw IMSS administrative records contain the reported wage, w_r. We know the social security tax scheduled in each year, discussed in Section 2, and hence can calculate the IMSS post-tax wage. The ENEU data do not contain firm identifiers, but we can construct an estimate of w_u at the level of cells defined by metropolitan area, sector, firm size categories and/or age groups.

At the cell level, our first measure of evasion is the log median ENEU take-home wage minus the log median IMSS post-tax wage. Our second measure is the log mean ENEU take-home wage minus the log mean IMSS post-tax wage. We refer to the first as the "wage gap (medians)" and to the second as the "wage gap (means)". Our preferred measure is the medians measure, because the mean is sensitive to the treatment of topcodes and to the non-response of richer households in the household survey mentioned above. Below we report estimates using both measures, however.

Our third measure of evasion is an estimate of the excess mass at the left tail of the IMSS wage distribution. Figure 9 illustrates the calculation. The dotted (blue) curve is a non-parametric estimate of the ENEU distribution, the same one that underlies the hollow-rectangle histogram in Figure 6. The solid (red) curve is a non-parametric estimate of the *post-tax* IMSS distribution. (Note that this differs from the distribution underlying the sold-gray histogram in Figure 6, which was the *pre-tax* IMSS distribution.[44]) In principle, the ENEU take-home wage and the IMSS post-tax wage are an "apples-to-apples" comparison; in the absence of evasion, they should coincide. We calculate the excess mass as the fraction of the IMSS sample minus the fraction of the ENEU sample to the left of some critical quantile of the ENEU distribution. In Figure 9 and our baseline estimates, we choose the 25^{th} percentile as the critical value, indicated by the vertical line; we also show that results are robust to other choices. Intuitively, our excess mass measure reflects the share of the sample that has to be moved from right to left across the vertical line in order to

[44]This is the reason why the spikes at the right tails of the red and blue distributions, which correspond to the IMSS and ENEU topcodes, no longer coincide.

transform the blue distribution into the red distribution.[45]

The level of aggregation is an important issue when constructing these evasion measures. Although sample size is not a severe constraint in the IMSS administrative records, the ENEU contains on the order of 10,000-14,000 raw observations on male full-time workers in each quarter in the country as a whole. When we divide these by age group, metropolitan area, firm size category and sector, cell sizes in the ENEU can become prohibitively small. We cannot avoid doing some aggregation. As discussed above, we focus on five age categories and five firm size categories. We also aggregate four-digit industries into three broad sectors: manufacturing, construction, and retail/services. In addition, when constructing the evasion measures, we pool all four quarters within a given year in the ENEU data.[46] In this section, we present cross-sectional statistics using the measures of evasion calculated at the metro area/sector/firm size category/age group level. Below we will conduct the analyses are higher levels of aggregation, as appropriate to the questions being investigated.

Table 4 reports simple cross-sectional regressions of our three evasion measures on age-group, firm-size and sector indicators in 1990. For each evasion measure, we report simple regressions on a set of age-group or firm-size indicators without controls (Columns 1-2, 4-5, and 7-8) and then a regression including sector indicators and metro-area indicators (Columns 3, 6, and 9). For age groups, there is clear evidence that evasion is higher on average for the youngest age group, ages 16-25 (the omitted category), perhaps not surprisingly, since the youngest workers tend to have lower labor-force attachment and higher mobility across jobs. This pattern is not significant using the wage gap (medians) measure without controls in Column 1, but it is significant once the firm-size and metro-area controls are included, and the pattern is robust for the other evasion measures. The differences in coefficients among the over-25 age groups are generally not significantly different from one another. For firm size, the general pattern that evasion is declining in firm size, consistent with the pattern in the raw histograms in Figure 8. There appears to be some non-monotonicity in the relationship between evasion and firm size for the intermediate

[45]This excess mass measure differs from other excess mass measures that have been used in the literature, for instance by Saez (2010), Chetty, Friedman, Olsen, and Pistaferri (2011), and Kleven and Waseem (2013), in two ways. First, in our setting we do not have to construct a counterfactual distribution, requiring assumptions about how individuals in the region of bunching are distributed in the counterfactual; here effectively we observe the counterfactual distribution in the ENEU household data. Second, it is common to scale the excess mass by the density of the counterfactual distribution in the region of bunching; in our setting at the left tail the density of the ENEU wage distribution is near zero, and dividing by this density might introduce significant errors, so we do not re-scale.

[46]Because the survey follows households for five quarters, observations in different quarters are not independent, but our econometric tests do not require an independence assumption.

size categories (51-100 and 101-250 employees), but it appears robust that evasion is lower in 11-50 employee firms than in 1-10 employee firms (the omitted category), and lower still in 250+ employee firms. The estimates are largely unaffected by controlling for age group, metro area, and sector, which suggests that the pattern we observed in the raw data in Figure 8 is not due to differing age or metro area composition in different firm size categories. Finally, evasion follows a consistent pattern across broad sectors, with construction displaying the greatest extent of evasion, followed by manufacturing, followed by retail/services.

Before moving on to the difference-in-difference analysis, we briefly examine the wage distributions for two sets of manufacturing establishments that can be linked to plant-level datasets collected by INEGI, the Mexican statistical agency. An important insight from the taxation literature is that the larger the number of reports the tax authorities receive on firms' tax liabilities, the more difficult it is for firms to evade (Kopczuk and Slemrod, 2006).[47] The plant-level datasets from INEGI do not contain individual-level wage information on the full distribution of wages, but it is nevertheless instructive to consider the reports of covered plants to the social security agency. Figure 10 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.[48] There is little evidence of bunching 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 bunching even than plants in the 100-250 and >250 employees categories in Figure 8. For the sake of completeness, Figure 11 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 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 bunching 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.

[47] In fact, the INEGI 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, and hence being required to report to INEGI may lead firm to report more accurately to the tax authorities.

[48] The 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.

6 Effect of Pension Reform on Compliance

We now consider how evasion varied over time in response to the pension reform, conducting the analysis at two levels of aggregation. In subsection 6.1, we examine the aggregate market-level response of evasion for different age groups, where "market" may refer to the entire economy or to a local labor market. In subsection 6.2, we consider changes within cells defined by metropolitan area, sector and firm-size categories, and relate the initial age composition to subsequent within-cell changes in evasion.

6.1 Differential Responses by Age Group

A simple set of graphs illustrates the intuition of our first approach, analyzing the response of the reform at the level of age groups. Figure 12 plots non-parametric estimates of the male wage distributions, similar to those in Figure 9 but for wages levels below 200 pesos/day, by age group, for three years, 1990, 1997, and 2003.[49] Each column of graphs corresponds to an age group (indicated in the x-axis titles) and each row to a year. The key empirical pattern is illustrated by the contrast between the first column, corresponding to ages 16-25, and the last column, corresponding to ages 56-65. There is a clear decline in bunching and shift to the right of the IMSS distribution for the youngest age group. For the oldest age group, there is little evident decline in bunching or shift to the right in the IMSS distribution.

The same basic message, taking into account more years, is conveyed by Figure 13, which plots the wage gap (medians) measure, calculated at the age-group level, over the 1988-2003 period. One point to notice is that, as we saw in Table 4, evasion is highest for the youngest age group. But the more important point is that the gap for the oldest age group, 56-65, increased relative to the gaps for the other groups. To the extent that there is another group that appears to have seen a relative increase, it is the second-oldest age group, 46-55.

A possible concern with Figures 12 and 13 is that the differential changes by age group 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-metro area-year level, regress them on a full set of metro area-year indicators, and average the residuals at the age group-year level. Figure 14 plots these averages. The relative increase for the oldest age group is again evident. (We discuss the timing of the trend break below.)

[49] We present non-parametric density estimates, rather than histograms, because they are easier to interpret visually. Histograms would convey a qualitatively similar message.

Table 5 reports regressions that capture the pattern illustrated in Figure 14. Motivated by the figure, we look for a divergence in the time-path of evasion for the oldest age group relative to the younger groups. We interact an indicator for the oldest age group with year effects for each year over the 1988-2003 period, omitting the interaction with 1995, the year the reform was passed. We control flexibly for metro area-year interactions, and include either age group fixed effects (Columns 1, 3, 5) or age group-metro area fixed effects (Columns 2, 4, 6). Across all three measures of evasion, we see little evidence of a differential pre-trend, but robust evidence of a relative increase in evasion for the oldest age group following the passage of the reform. To facilitate interpretation, Figures 15 and 16 plot the coefficients and a 95% confidence interval for the estimates from Columns 1 and 3, using the two wage gap measures. The differential response is evident.

The main qualitative difference across specifications is in the timing of the relative increase. For the wage gap (medians) and excess mass measures, the sustained relative increase appears to begin in 1996, the year after the passage of the reform but before the implementation of the reform in July 1997.[50] For the wage gap (means) measure, the relative increase appears to begin in 1998, the first full year after the implementation of the reform.[51] It is possible that forward-looking individuals responded to the announcement of the policy change — which was widely discussed in the Mexican media — by pressuring their employers to reduce under-reporting, differentially by age group. But because the results on timing are not robust, we do not pursue this possibility further. We interpret the results of Table 5 simply as suggesting that there was a relative increase in evasion for the oldest age group following the reform, which showed up robustly by 1998 or 1999.

For our third measure of evasion, the excess mass measure, there is some arbitrariness in the choice of critical value. It turns out, however, that the results are relatively insensitive to the choice of critical value. Table 6 reports regressions similar to Column 5 of Table 5, for different choices for the critical value in the excess mass calculation. (The third column replicates Column 5 of Table 5.) The results are qualitatively similar across columns.

The results for our evasion measures do not seem to be driven by discrepancies in the reporting of employment in the two data sources. Table 7 reports regressions similar to those in Table 5,

[50] For the wage gap (medians) measure, there is also a spike in 1994 relative to 1995, consistent with Figure 14; the volatility, due in part to sampling error in the ENEU survey, warrants caution in interpreting the results on timing.

[51] The interaction term for 1998 is not significant in Column 3 with age group effects, but becomes significant once age group-metro area effects are included in Column 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, rather than one of the evasion measures. There is no evidence of a differential change in the employment gap for older workers in response to the reform.[52]

To conclude this sub-section, we briefly consider the incidence of the reform on market-level net wages, using the ENEU household survey alone. Table 8 reports specifications similar to those in Table 5, where the outcome is the log net wage reported on the ENEU survey, and we are able to control flexibly for individual characteristics. Column 1 includes just age-group and metro area-year effects, then moving across the columns we add sets of dummy variables for schooling categories, marital status, occupation, industry and firm size. The consistent message is that there is no systematic differential effect of the reform on net wages of the oldest age group. While this non-result may appear surprising, it is consistent with recent evidence from other countries on the effects of tax reforms on take-home wages. For instance, Saez, Matsaganis, and Tsakloglou (2012) find little evidence of differences in take-home pay for workers of different cohorts facing different payroll-tax rates. The authors speculate that institutional factors such as fairness norms may constrain firms' ability to offer different wages to otherwise similar workers facing different effective payroll tax rates. Such factors may also be present in Mexico and may similarly prevent firms from offering different take-home wages to different age groups. Also, as discussed briefly in Section 3.4 and in more detail in Appendix B, in our theoretical framework the presence of firm heterogeneity alone is sufficient to render the sign of the relationship between the pension reform and the net wage ambiguous, even without such institutional factors. In short, based both on prior empirical work and on our model, it is not clear that we should have expected to find a differential effect of the pension reform on take-home wages.

6.2 Differential Responses within Metro-Area/Sector/Firm-Size Cells

Our theoretical framework predicts that a pension reform will lead to a reduction in evasion with a given firm or set of firms for affected workers — affected in the sense that their pensions become more sensitive to the reported wage, as we argue was true for younger workers in Mexico. A natural extension of this idea is that a reform will have a relatively greater effect within sets of

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

firms initially employing younger workers. Because of the data constraints discussed above, it is not possible to examine within-firm responses directly. Here instead we conduct what we consider to be a reasonable second-best analysis of changes in evasion within cells defined by metropolitan area, sector, and firm-size categories. We relate indicators of initial age composition to within-cell changes in evasion over time.

Table 9 presents our baseline specifications for this section. We calculate the average age within metropolitan area/sector/firm-size cells from the IMSS data in 1988, and interact this initial average age with year indicators, omitting the year of passage of the reform, 1995. We include a full set of dummy variables for firm-size/sector/metro area cells and drop the initial year, 1988, from the estimation sample. We also include metro area-year dummies and, in some specifications, sector-year dummies. We report results for the three measures of evasion, wage gap (medians), wage gap (means), and excess mass (below 25th percentile). We see that cells with a higher initial age saw a significant relative increase in evasion following the passage of the reform, consistent with the predictions of our theoretical model. Although there are a few significant coefficients in the early years, there do not appear to be important differences pre-trends in the years leading up to the passage of the reform. As in the market-level results discussed in subsection 6.1, there are differences in the timing of the beginning of the relative increase across evasion measures, with the relative increase becoming significant in 1996 for the wage gap (medians) measure and not until 1999 for the wage gap (means) measure.

To investigate the robustness of these patterns, we repeat the exercise using the initial share of workers in the 55-65 age group as our measure of initial age composition. Table 10 reports the results. The patterns are consistent with those in Table 9: a higher initial share of workers in the oldest age group predicts a relative increase in evasion within cells following the passage of the reform.

It does not appear that the within-cell patterns can be explained entirely by sorting of workers of different ages across cells. Table 11 presents regressions analogous to those in Tables 9-10, but with the average age or the share in the oldest age group as the dependent variable. (The initial age measure in Columns 1-2 is the mean age in 1988; in Columns 3-4 it is the share in the oldest age group in 1988.) There appears to be some serially correlated measurement error in the age measures, explaining the positive coefficients in the first row or two. But the overall message of the table is that there does not appear to have been a systematic change in age composition following the pension reform, differentially by initial age composition, that could explain the within-cell

patterns in Tables 9-10.

Although we cannot definitively rule out the possibility that the results in Tables 9-10 reflect sorting of workers across firms within cells, or changes in age composition that are not captured by the age measures in Table 11, we interpret the within-cell patterns as consistent with the hypothesis that the pension reform led to a relatively greater increase in compliance for firms initially employing younger workers.

7 Conclusion

Improving firms' compliance 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 incentives and the availability of information in the social security system.

The results suggest that providing incentives to workers to ensure accurate reporting, as well as information about firms' reports, should be a consideration in the design of social-insurance systems. Conceptually, our theoretical model suggests that an increase in such incentives and a reduction of payroll taxes should have equivalent effects on evasion, other things equal. But the effects on government revenues are decidedly non-equivalent. If the policy goal is to increase the fiscal capacity of the state, it appears that the tying benefits more closely to wage reports may be the preferable option.

A number of interesting questions remain open. One is to what extent workers are aware of under-reporting by their employers and, relatedly, to what extent the effects of the pension reform we observe are due to the change in incentives versus the change in information. It seems unlikely that the results we observe could be due solely to the increased availability of information post-reform; if benefits had remained truly insensitive to wage reports, then workers would have had no reason to act on the information once they received it.[53] But it remains possible that there was an important interaction between the change in incentives and the change in information. Separating the two effects definitively will require a research design in which incentives vary separately from costs of information.

[53]Relatedly, a reduction of the cost of acquiring information about firms' reports would not be expected to have an effect, as long as the cost remains positive, as workers would not be willing to pay even an ε cost to acquire it.

Another important open question is whether increased pressure on firms to report accurately (which increases compliance on the intensive margin) induces more firms to remain informal (that is, reduces compliance on the extensive margin). Because of the nature of the IMSS data, we are not able to observe firms as they move from formality to informality or vice-versa. But clearly a full accounting of the costs and benefits of policies to increase intensive-margin compliance will have to take such a response into effect.

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.

Bérgolo, M., and G. Cruces (2012): "Work and Tax Evasion Incentive Effects of Social Insurance Programs: Evidence from an Employment Based Benefit Extension," Unpub. paper, CEDLAS, Universidad Nacional de la Plata.

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.

Best, M. C., A. Brockmeyer, H. J. Kleven, J. Spinnewijn, and M. Waseem (2013): "Production vs Revenue Efficiency with Limited Tax Capacity: Theory and Evidence from Pakistan," Unpub. apaper, London School of Economics.

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.

Busso, M., M. V. Fazio, and S. Levy (2012): "(In)Formal and (Un)Productive: The Productivity Costs of Excessive Informality in Mexico," Inter-American Development Bank Working Paper no. 341.

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.

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

Chetty, R., J. N. Friedman, T. Olsen, and L. Pistaferri (2011): "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 126(2), pp. 749–804.

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.

Dharmapala, D., J. Slemrod, and J. D. Wilson (2011): "Tax policy and the missing middle: Optimal tax remittance with firm-level administrative costs," *Journal of Public Economics*, 95(910), 1036 – 1047, ¡ce:title¿Special Issue: The Role of Firms in Tax Systems¡/ce:title¿.

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

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. Köllő, 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.

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.

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.

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 (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.

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 (2011): "Entry regulation and Business Start-ups: Evidence from Mexico," *Journal of Public Economics*, 95, 1501–1515.

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.

Kleven, H. J., and M. Waseem (2013): "Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan," *The Quarterly Journal of Economics*, 128(2), 669–723.

Kopczuk, W., J. Marion, E. Muehlegger, and J. Slemrod (2013): "Do the Laws of Tax Incidence Hold? Point of Collection and the Pass-through of State Diesel Taxes," Unpub. paper, Columbia University.

Kopczuk, W., and J. Slemrod (2006): "Putting Firms into Optimal Tax Theory," *American Economic Review Papers & Proceedings*, 96(2), 131–134.

La Porta, R., and A. Shleifer (2008): "The Unofficial Economy and Economic Development," *Brookings Papers on Economic Activity*, 2008, pp. 275–352.

Lara-Ibarra, G. (2011): "Essays on the Determinants of Pension Savings and Retirement Management

Decisions," Ph.D. thesis, University of Maryland.

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.

Mao, J., L. Zhang, and J. Zhao (2013): "Social Security Taxation and Compliance: The Chinese Evidence," Unpub. paper, Tsinghua University.

Marion, J., and E. Muehlegger (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," Ph.D. thesis, 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.

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

Monteiro, J. C., and J. J. Assunção (2012): "Coming Out of the Shadows? Estimating the Impact of Bureaucracy Simplification and Tax Cut on Formality in Brazilian Microenterprises," *Journal of Development Economics*, 99, 105–115.

Niehaus, P., and S. Sukhtankar (forthcoming): "Corruption Dynamics: the Golden Goose Effect," *American Economic Journal: Economic Policy*.

Nyland, C., R. Smyth, and C. Zhu (2006): "What Determines the Extent to Which Employers Will Comply with their Social Security Oblications? 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.

Olken, B. A. (2006): "Corruption and the Costs of Redistribution: Micro Evidence from Indonesia.," *Journal of Public Economics*, 90(4-5), 853 – 870.

Olken, B. A., and R. Pande (2012): "Corruption in Developing Countries.," *Annual Review of Economics*, 4(1), 479 – 509.

Pissarides, C. A., and G. Weber (1989): "An Expenditure-Based Estimate of Britain's Black Economy," *Journal of Public Economics*, 39(1), 17 – 32.

Plumley, A. (2004): *Overview of the Federal Tax Gap*. Washington DC: U.S. Department of the Treasury, Internal Revenue Service, NHQ Office of Research.

Pomeranz, D. (2013): "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.

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.

Saez, E. (2010): "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2(3), 180–212.

Saez, E., M. Matsaganis, and P. Tsakloglou (2012): "Earnings Determination and Taxes: Evidence from a Cohort-Based Payroll Tax Reform in Greece.," *Quarterly Journal of Economics*, 127(1), 493 – 533.

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

——— (2007): "Cheating Ourselves: The Economics of Tax Evasion," *The Journal of Economic Perspectives*, 21(1), 25–48.

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

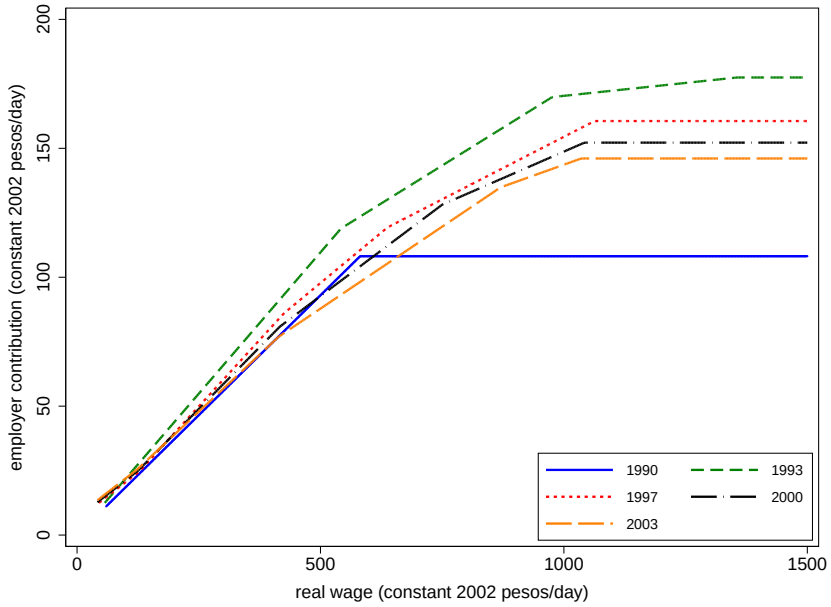
Slemrod, J., and S. Yitzhaki (2002): "Tax Avoidance, Evasion and Administration," in *Handbook of Public Economics, vol. 3*, ed. by A. J. Auerbach, and M. Feldstein. Elsevier, Amsterdam.

Tonin, M. (2011): "Minimum Wage and Tax Evasion: Theory and Evidence," *Journal of Public Economics*, 95(11-12), 1635 – 1651.

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

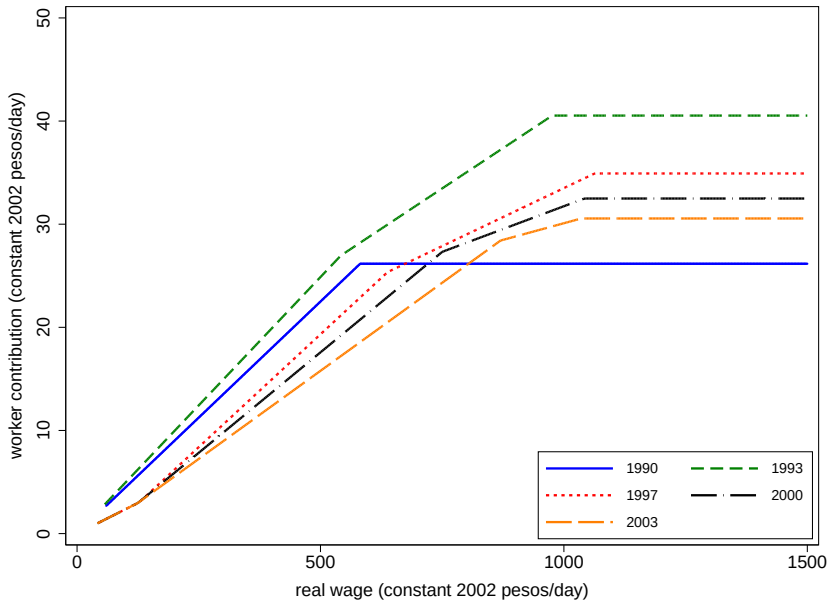
Yaniv, G. (1992): "Collaborated Employee-Employer Tax Evasion," *Public Finance*, 47(2), 312–321.

Figure 1. Employer contributions



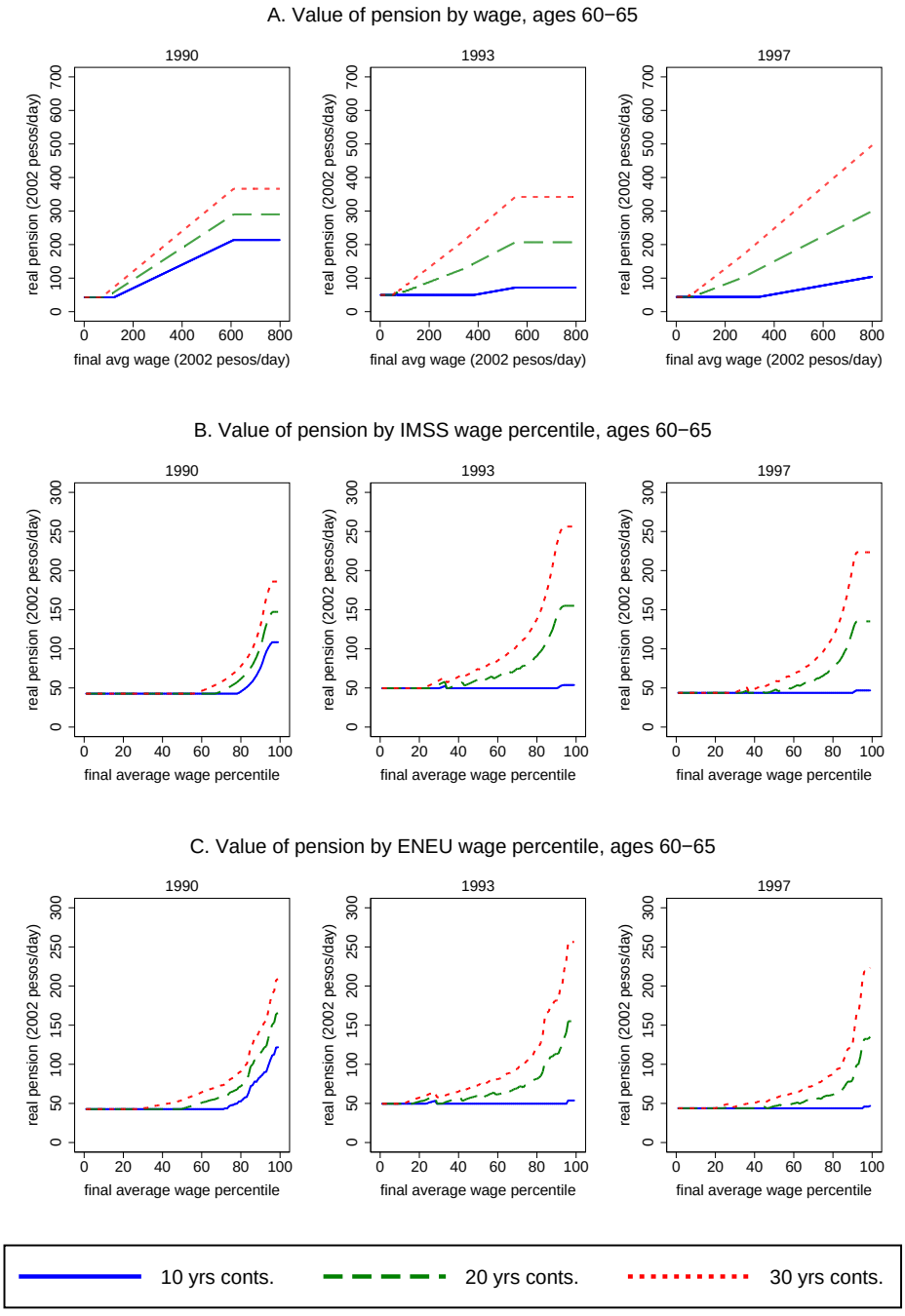
Notes: Variation in IMSS employer contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 2. Employee contributions



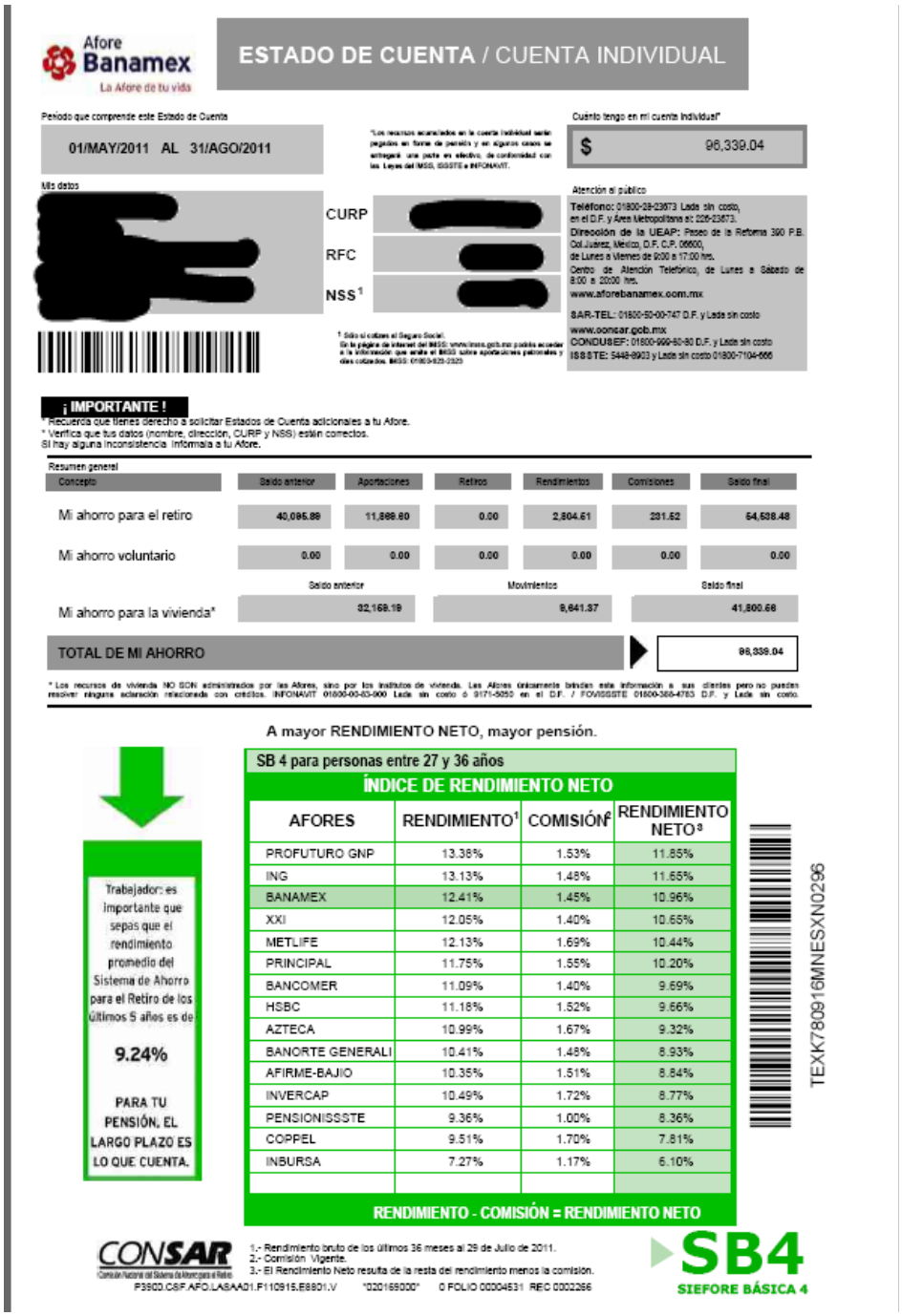
Notes: Variation in IMSS worker contribution rates at levels above 500 pesos/day are primarily due to changes in topcodes, which varied from 10 to 25 times the minimum wage in Mexico City over the period. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 3. Pension values, selected years, men



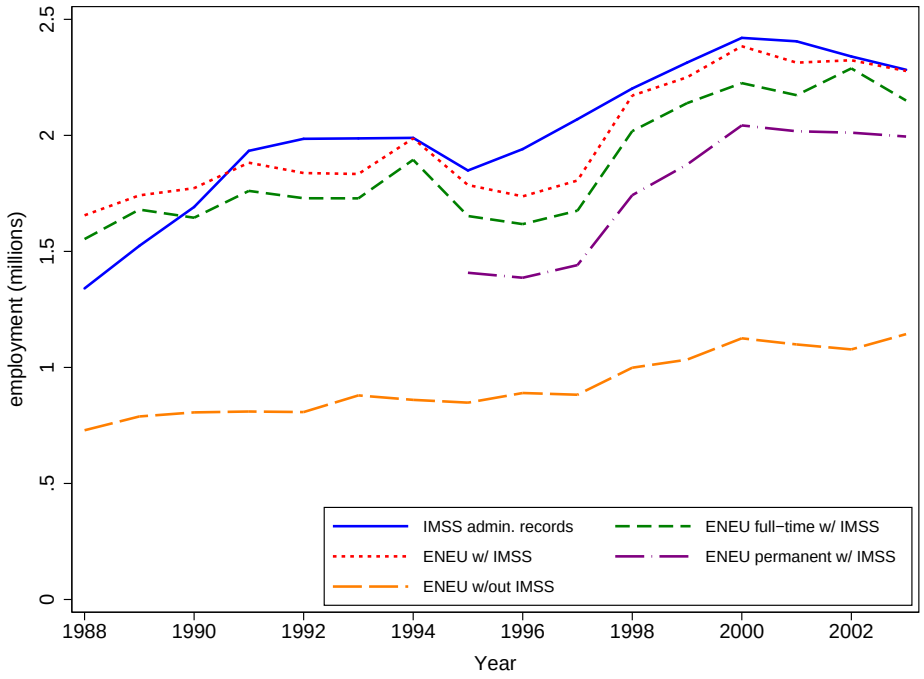
Notes: Final average wage (2002 pesos/day) is average nominal daily wage over five years prior to retirement, deflated to constant 2002 pesos. Figure indicates pension values for individuals with 10, 20 and 30 years of contributions to IMSS. In Panel B, we calculate the nominal wage at each quantile 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 4 for details of samples and Section 2.3 for details on pension benefits. Average 2002 exchange rate: 9.66 pesos/dollar.

Figure 4. Account statement



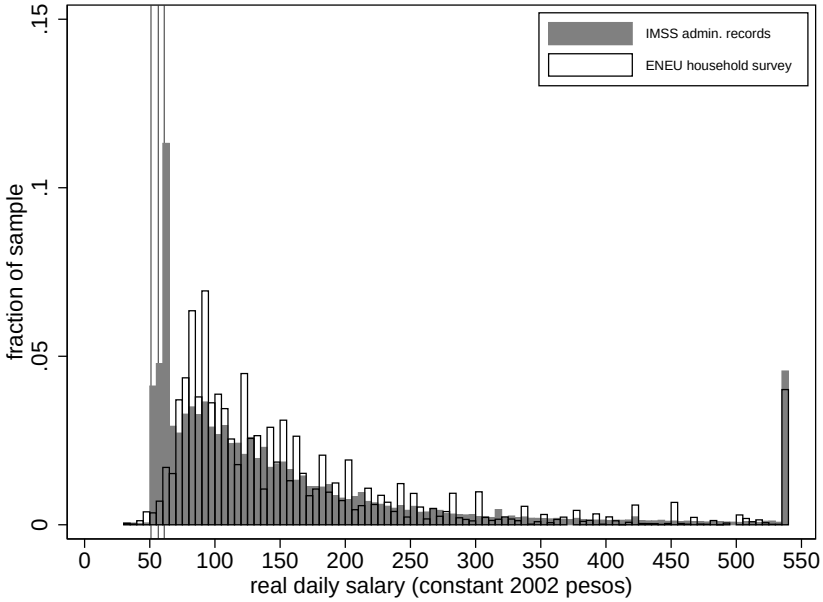
Notes: The box at top right ("Cuánto tengo en mi cuenta individual") reports total balance. The first row of boxes in the middle section ("Mi ahorro para el retiro") pertains to the retirement pension and reports previous balance ("Saldo anterior"), new contributions ("Aportaciones"), withdrawals ("Retiros"),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



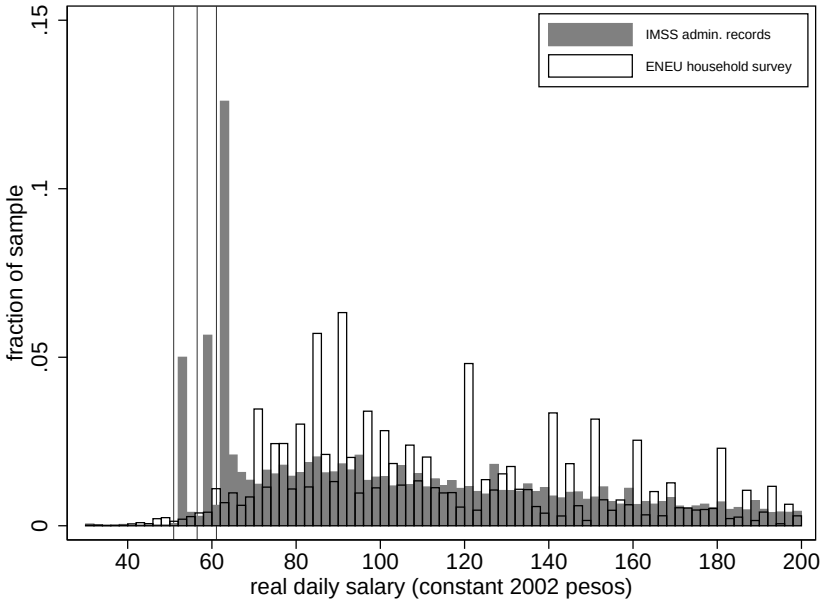
Notes: Samples are the same as those in Columns 1 and 3-6 of Table 2; refer to that table for details. ENEU totals are calculated using sampling weights. See Section 4 and Appendix C (online) for details of sample selection.

Figure 6. Wage histograms, men, 1990



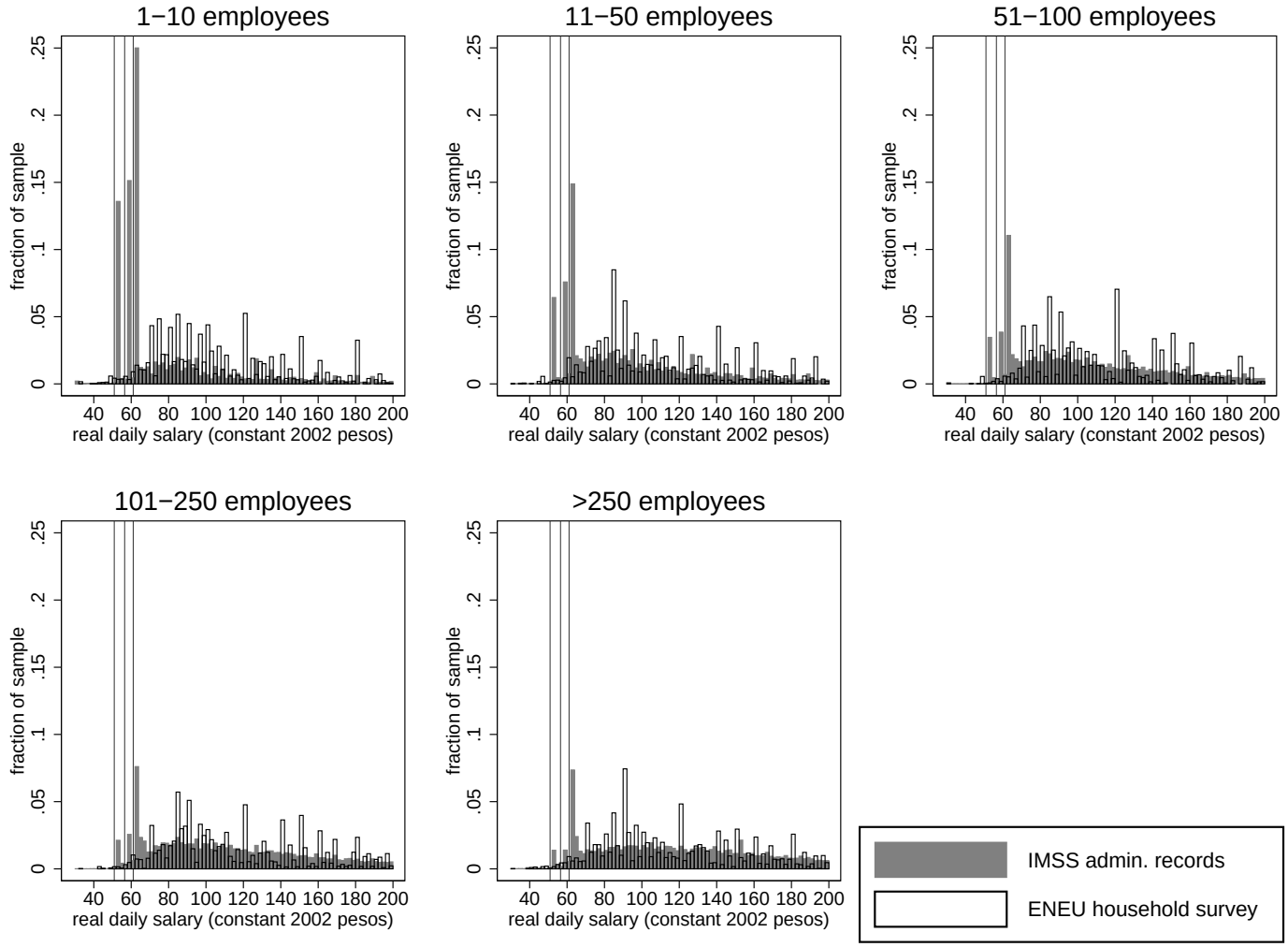
Notes: The ENEU wage is the real daily take-home wage reported to the ENEU household survey. The IMSS wage is the real daily *pre-tax* reported wage from the IMSS administrative records. Wages in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. Samples are IMSS and ENEU "baseline" samples of men. Data in both samples are from second quarter of 1990. 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). See Section 4 and Appendix C (online) for further details.

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



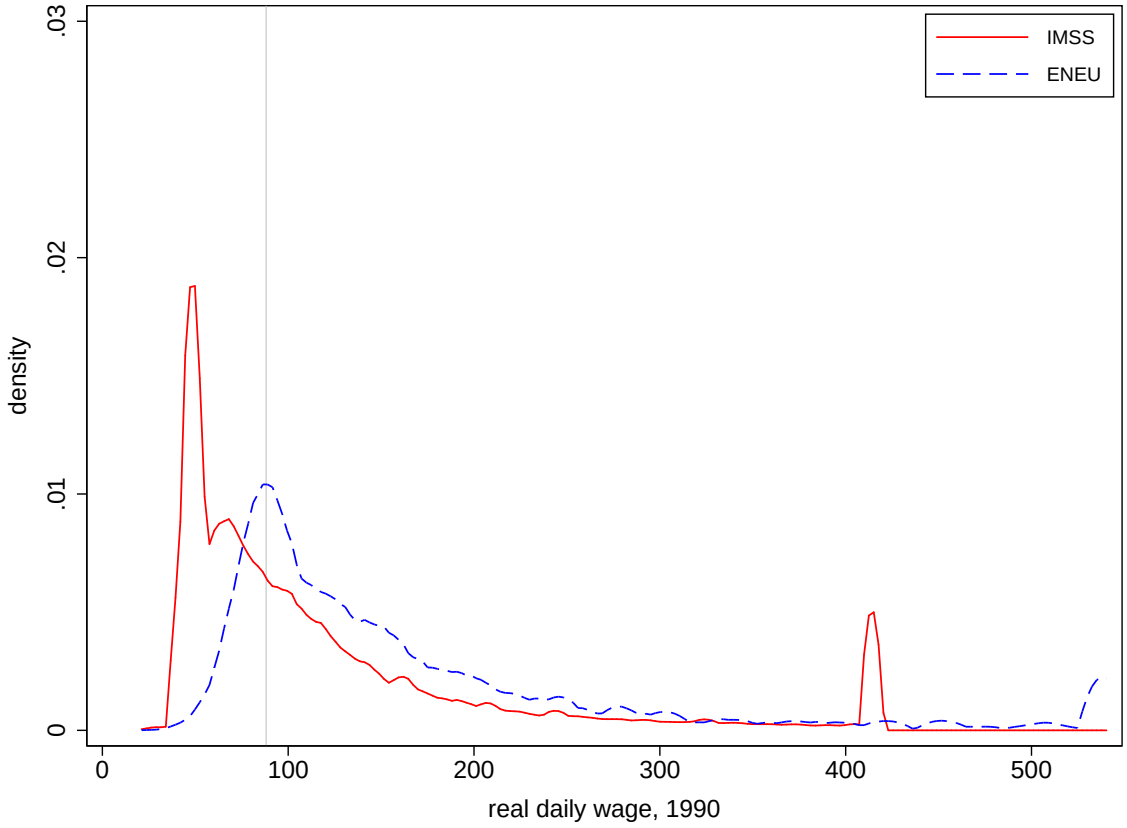
Notes: Histogram is similar to Figure 6 but only includes workers with wages less than 200 pesos/day (approx. $20/day) in constant 2002 pesos. Bins are 2 pesos wide.

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



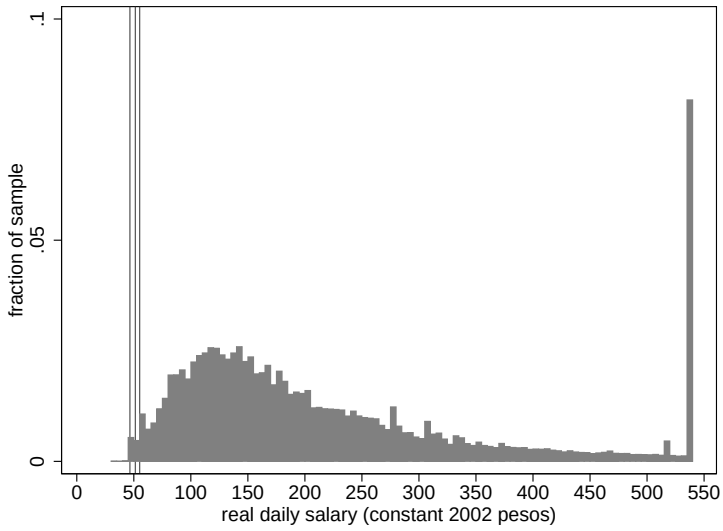
Notes: Histograms are similar to those in Figure 7. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 2 pesos wide. Average 2002 exchange rate: 9.66 pesos/dollar. See Section 4 and Appendix C (online) for further details.

Figure 9. Excess mass calculation



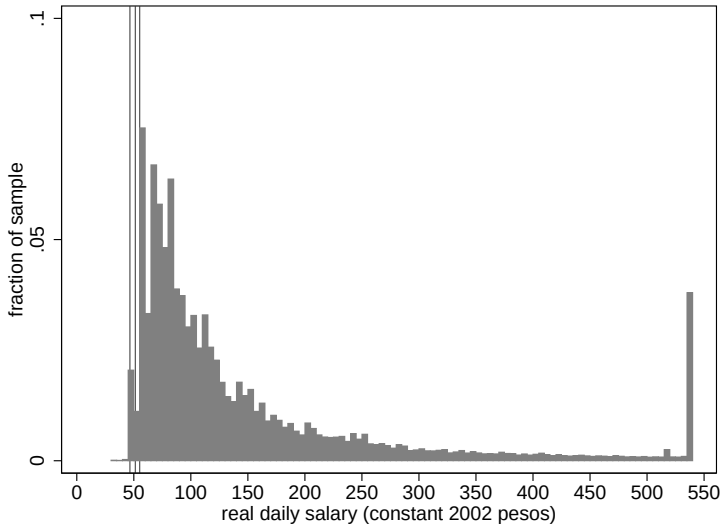
Notes: The wage variables are the real daily take-home wage from ENEU and real daily *post-tax* reported wage from IMSS. Densities are estimated using 1990 data and an Epanechnikov kernel with bandwidth 3 pesos for IMSS data and 6 pesos for ENEU data (using Stata kdensity command). Wages are in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. (Densities are smoothed versions of histograms in Figure 6.) Vertical line is at 25th percentile of the ENEU wage distribution. Data are from second quarter of 1990. Excess mass for 25th 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 (online) for further details of data processing.

Figure 10. Wage histograms, men, 1993, IMSS data, establishments linked to EIA



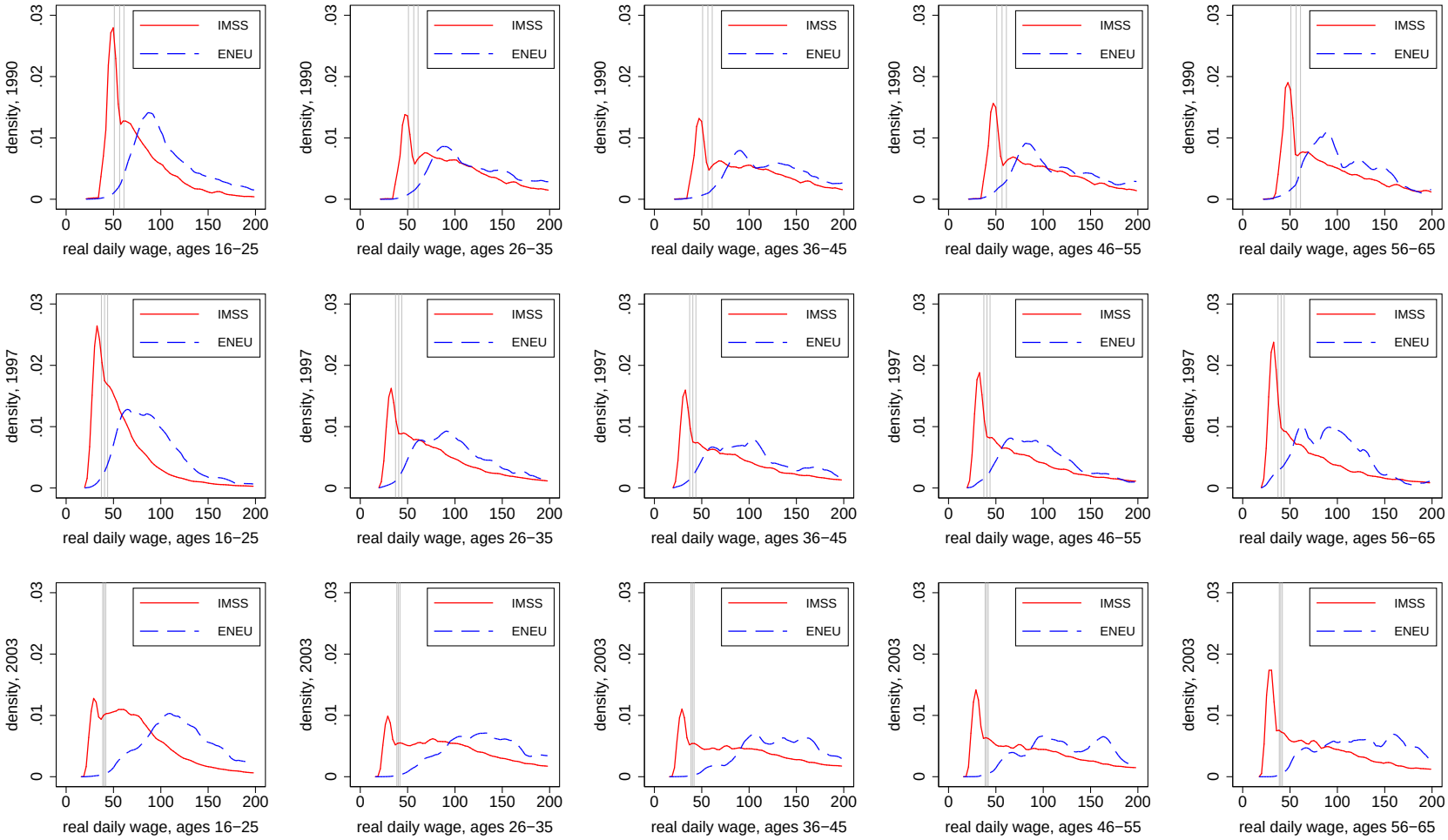
Notes: Sample is permanent male workers ages 16-65 in IMSS data in 2389 establishments that can be linked to a balanced 1993-2003 panel from the *Encuesta Industrial Anual (EIA)* [Annual Industrial Survey], which excludes assembly-for-export *maquiladora* plants. Data are from second quarter. Vertical lines indicate minimum wages in the three minimum-wage zones in Mexico (A, B, C). Bins are 5 pesos wide. See Section 4 and Appendix C (online) for further details of data processing.

Figure 11. Wage histograms, men, 1993, establishments linked to EMIME



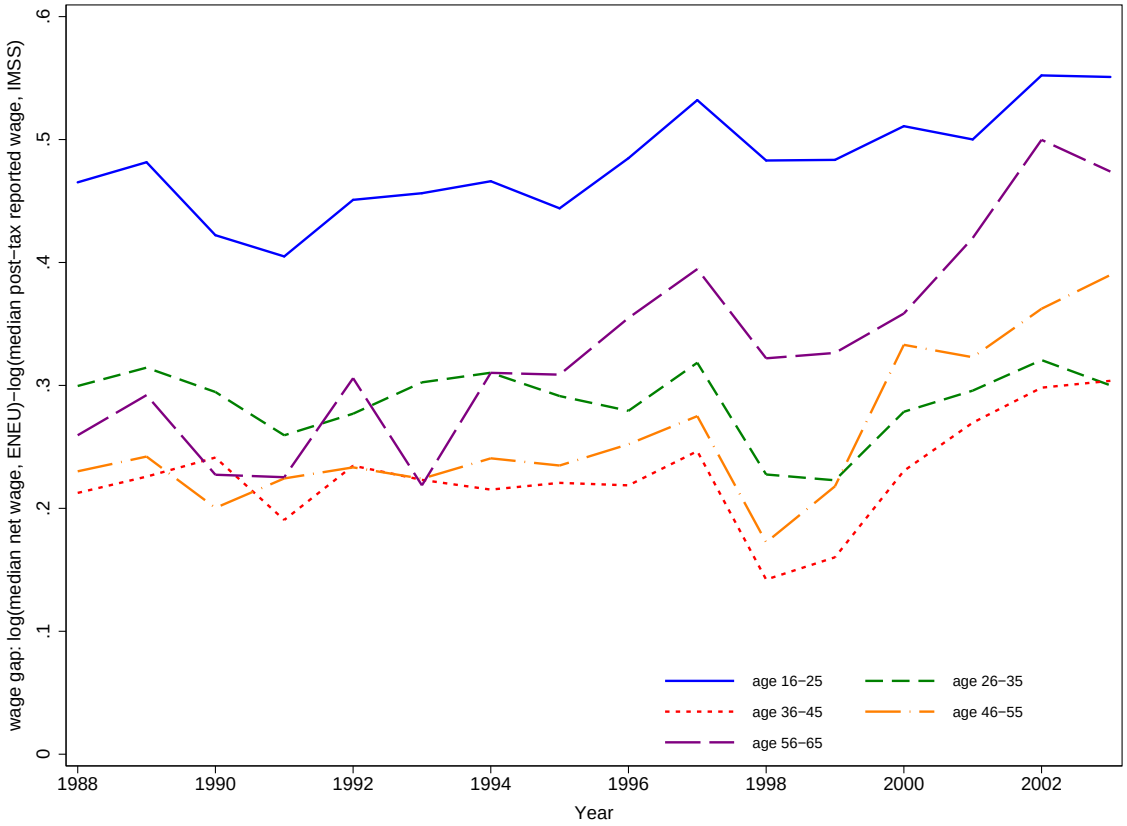
Notes: Sample is permanent male workers ages 16-65 in IMSS data in 520 establishments that can be linked to a balanced 1993-2003 panel from the *Estadísticas Mensuales de la Industria Maquiladora de Exportación (EMIME)* [Monthly Statistics on Maquiladora Export Industry], a dataset made up exclusively 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 4 and Appendix C (online) for further details of data processing.

Figure 12. Wage densities by age group, 1990, 1997, 2003, men



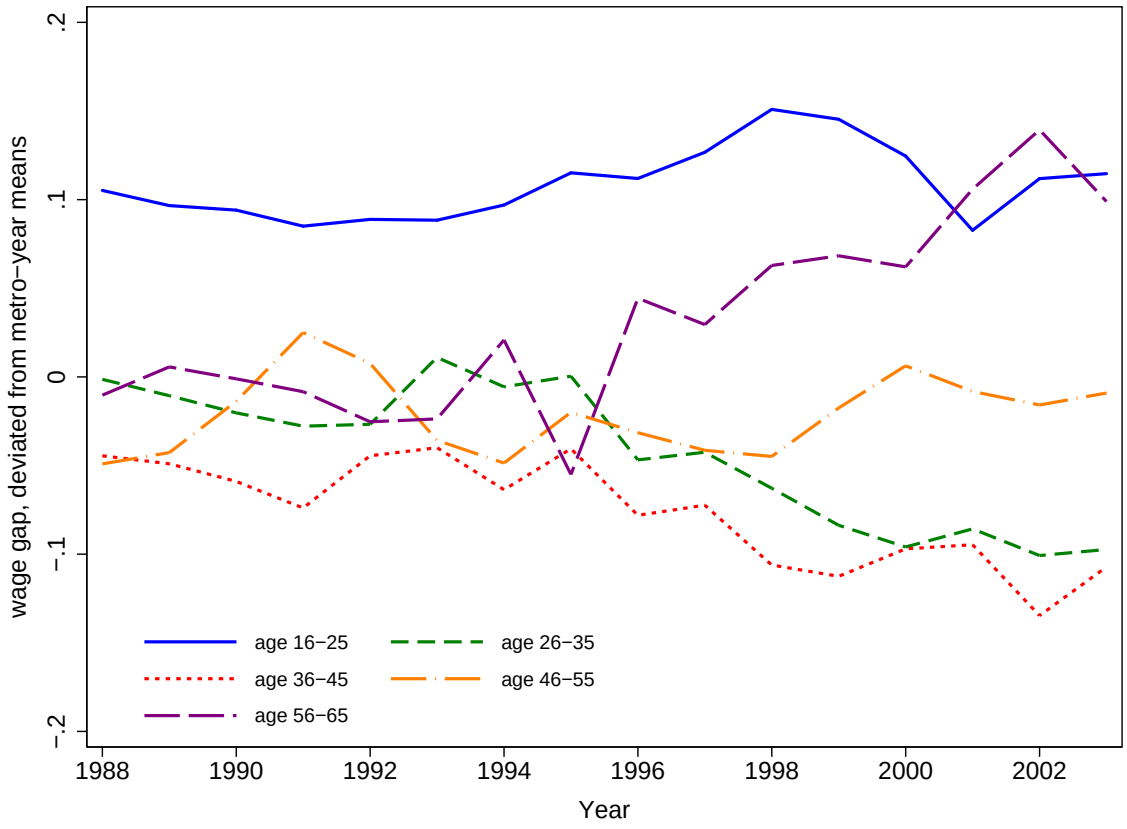
Notes: The wage variables are the real daily take-home wage from ENEU and real daily *post-tax* reported wage from IMSS. Densities are estimated using 1990 data and an Epanechnikov kernel with bandwidth 3 pesos for IMSS data and 6 pesos for ENEU data (using Stata kdensity command). Wages are in 2002 pesos. Average 2002 exchange rate: 9.66 pesos/dollar. Rows correspond to years 1990, 1997, 2003; columns to age groups 16-25, 26-35, 36-45, 46-55, 56-65. Samples are 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 (online) for further details.

Figure 13. Wage gaps (medians) by age group, men



Notes: Each wage gap is the difference between the log median net wage from the ENEU survey and the log median post-tax reported wage from the IMSS administrative records, using the ENEU and IMSS baseline samples. See Section 4 and Appendix C (online) for further details.

Figure 14. Wage gaps (medians) by age group, men, deviated from metro-year means

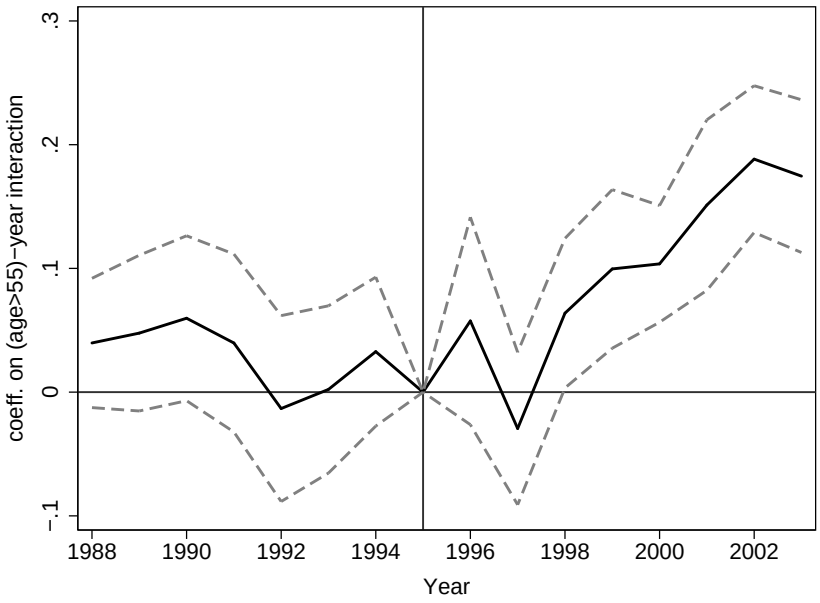


Notes: Each wage gap is the difference between the log median net wage from the ENEU survey and the log median post-tax reported wage from the IMSS administrative records, using the ENEU and IMSS baseline samples. To calculate deviated wage gaps, we calculate wage gaps separately by age group-year-metro area, regress them on a full set of metro area-year dummies, and average the residuals at the age-group level. See Section 4 and Appendix C (online) for details of sample selection.

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

Notes: Figure plots coefficients for 1(age>55)*year interaction term from Column 2 of Table 5. The dotted lines indicate the 95 percent confidence interval.

Figure 16. Differential effect of reform on wage gap (means), ages 55-65, men



Notes: Figure plots coefficients for 1(age>55)*year interaction term from Column 4 of Table 5. The dotted lines indicate the 95 percent confidence interval.

Table 1. Pension wealth simulation, by age in 1997, male worker with 35 years of expected contributions

Age in 1997	Years of Expected PRA Contributions	Plan	Real Daily Wage					
			43	100	200	300	500	1079
25	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	30	PRA	398.6	*523.4*	*1044.3*	*1565.3*	*2607.1*	*5620.5*
		PAYGO	398.6	*398.6*	*603.8*	*890.2*	*1483.6*	*3200.1*
35	25	PRA	398.6	398.6	*659.1*	*987.8*	*1645.3*	*3546.9*
		PAYGO	398.6	398.6	*603.8*	*890.2*	*1483.6*	*3200.1*
40	20	PRA	398.6	398.6	403.9	605.4	1008.4	2173.9
		PAYGO	398.6	398.6	603.8	890.2	1483.6	3200.1
45	15	PRA	398.6	398.6	398.6	398.6	586.6	1264.7
		PAYGO	398.6	398.6	603.8	890.2	1483.6	3200.1
50	10	PRA	398.6	398.6	398.6	398.6	398.6	662.6
		PAYGO	398.6	398.6	603.8	890.2	1483.6	3200.1
55	5	PRA	398.6	398.6	398.6	398.6	398.6	398.6
		PAYGO	398.6	398.6	603.8	890.2	1483.6	3200.1

Notes: Values are real present discounted value of the future stream of pension benefits in thousands of 2002 pesos, for a male worker who began contributing at age 25 and expects to continue until age 60. Numbers is italics (and blue where color is available) indicate that personal retirement account (PRA) has a higher expected payoff than the pre-reform pension (PAYGO). Average 2002 exchange rate: 9.66 pesos/dollar. 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 post-tax wage	121.02	163.88	172.98	143.88		166.73
	(0.07)	(1.58)	(1.94)	(2.62)		(1.85)
age	31.75	31.46	32.13	29.98		32.22
	(0.01)	(0.15)	(0.17)	(0.29)		(0.17)
fraction employed in ests >100 employees	0.52	0.43	0.55	0.18		0.55
	(0.00)	(0.01)	(0.01)	(0.01)		(0.01)
N (raw observations)	1691417	16169	11592	4577		10978
N (population, using weights)	1691417	2578847	1772523	806324		1645229
B. 2000						
real avg. daily post-tax wage	123.60	148.20	161.15	120.78	166.42	155.80
	(0.07)	(1.31)	(1.60)	(2.16)	(1.80)	(1.59)
age	32.70	32.22	32.82	30.94	33.22	32.88
	(0.01)	(0.14)	(0.16)	(0.28)	(0.17)	(0.16)
fraction employed in ests >100 employees	0.58	0.44	0.59	0.10	0.63	0.59
	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
N (raw observations)	2420307	19171	14063	5108	11918	13246
N (population, using weights)	2420307	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 reports statistics for IMSS baseline sample; Column 2 for full ENEU (household survey) sample (satisfying aforementioned criteria); Column 3 for employees in ENEU who report receiving IMSS benefit in current employment; Column 4 for employees in ENEU who report not receiving IMSS benefit; Column 5 for employees in ENEU who report receiving IMSS benefit and having a written contract of indefinite duration; and Column 6 for employees in ENEU who report receiving IMSS benefit and working at least 35 hours in previous week (the ENEU baseline sample). 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 4 and Appendix C (online).

Table 3. 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). For further details, see Section 4 and Appendix C (online).

Table 4. Cross-sectional patterns of evasion, 1990, men

	wage gap (medians)			wage gap (means)			exc. mass (25th percentile)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
age 26-35	-0.054*		-0.054**	-0.081***		-0.081***	-0.145***		-0.145***
	(0.029)		(0.021)	(0.024)		(0.019)	(0.015)		(0.013)
age 36-45	-0.072**		-0.073***	-0.149***		-0.150***	-0.167***		-0.168***
	(0.034)		(0.027)	(0.028)		(0.024)	(0.016)		(0.013)
age 46-55	-0.029		-0.026	-0.154***		-0.151***	-0.145***		-0.144***
	(0.035)		(0.031)	(0.031)		(0.027)	(0.017)		(0.014)
age 56-65	-0.026		-0.034	-0.165***		-0.172***	-0.108***		-0.112***
	(0.044)		(0.040)	(0.037)		(0.034)	(0.019)		(0.016)
11-50 employees		-0.332***	-0.333***		-0.173***	-0.173***		-0.129***	-0.128***
		(0.026)	(0.023)		(0.025)	(0.023)		(0.011)	(0.009)
51-100 employees		-0.480***	-0.478***		-0.281***	-0.281***		-0.218***	-0.214***
		(0.033)	(0.031)		(0.030)	(0.028)		(0.015)	(0.014)
101-250 employees		-0.393***	-0.374***		-0.242***	-0.233***		-0.214***	-0.203***
		(0.039)	(0.037)		(0.035)	(0.032)		(0.017)	(0.015)
$>$ 250 employees		-0.499***	-0.465***		-0.231***	-0.200***		-0.237***	-0.218***
		(0.035)	(0.034)		(0.030)	(0.029)		(0.017)	(0.016)
construction			0.128***			0.122***			0.064***
			(0.029)			(0.025)			(0.013)
retail/services			-0.073***			-0.108***			-0.045***
			(0.024)			(0.021)			(0.010)
constant	0.559***	0.854***	0.639***	0.501***	0.574***	0.505***	0.483***	0.524***	0.495***
	(0.017)	(0.018)	(0.047)	(0.016)	(0.018)	(0.039)	(0.009)	(0.006)	(0.019)
metro area effects	N	N	Y	N	N	Y	N	N	Y
R-squared	0.00	0.20	0.31	0.03	0.08	0.27	0.09	0.20	0.42
N	1068	1068	1068	1068	1068	1068	1068	1068	1068

Notes: Data are from IMSS and ENEU baseline samples, collapsed to metro area/age group/firm-size category/sector level for 1990. The omitted category for age is 16-25, for firm size is 1-10 employees, and for sector is manufacturing. The wage gap (medians) is log median real daily take-home wage from the ENEU minus log median real daily post-tax reported wage from IMSS, calculated. Wage gap (means) is analogous, using mean in place of median. Excess mass is calculated as described in Section 5 and Figure 9. In calculating evasion measures, we pool ENEU data across quarters within year. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details of data processing.

Table 5. Differential effects of pension reform on evasion, men

	wage gap (medians)		wage gap (means)		excess mass (25^{th} perc.)	
	(1)	(2)	(3)	(4)	(5)	(6)
1(age > 55)*1988	0.056	0.056	0.040	0.040	0.022	0.022
	(0.040)	(0.037)	(0.035)	(0.027)	(0.024)	(0.019)
1(age > 55)*1989	0.076*	0.076*	0.048	0.048	0.026	0.026
	(0.045)	(0.042)	(0.039)	(0.032)	(0.021)	(0.016)
1(age > 55)*1990	0.067	0.067*	0.060	0.060*	0.027	0.027
	(0.044)	(0.039)	(0.041)	(0.034)	(0.022)	(0.017)
1(age > 55)*1991	0.058	0.058	0.040	0.040	0.042**	0.042***
	(0.039)	(0.038)	(0.036)	(0.037)	(0.019)	(0.014)
1(age > 55)*1992	0.037	0.037	-0.013	-0.013	0.029	0.029*
	(0.042)	(0.043)	(0.042)	(0.038)	(0.021)	(0.016)
1(age > 55)*1993	0.039	0.039	0.002	0.002	0.015	0.015
	(0.040)	(0.040)	(0.036)	(0.034)	(0.018)	(0.015)
1(age > 55)*1994	0.095**	0.095**	0.033	0.033	0.002	0.002
	(0.045)	(0.045)	(0.035)	(0.031)	(0.019)	(0.016)
1(age > 55)*1996	0.124***	0.124***	0.058	0.058	0.053**	0.053***
	(0.048)	(0.040)	(0.048)	(0.043)	(0.021)	(0.018)
1(age > 55)*1997	0.106**	0.106**	-0.029	-0.029	0.037*	0.037**
	(0.052)	(0.045)	(0.039)	(0.031)	(0.022)	(0.017)
1(age > 55)*1998	0.147***	0.147***	0.064	0.064**	0.054***	0.054***
	(0.043)	(0.037)	(0.040)	(0.031)	(0.018)	(0.013)
1(age > 55)*1999	0.154***	0.154***	0.100***	0.100***	0.062***	0.062***
	(0.045)	(0.041)	(0.032)	(0.033)	(0.017)	(0.013)
1(age > 55)*2000	0.146***	0.146***	0.104***	0.104***	0.053***	0.053***
	(0.044)	(0.039)	(0.030)	(0.024)	(0.017)	(0.014)
1(age > 55)*2001	0.201***	0.201***	0.151***	0.151***	0.074***	0.074***
	(0.049)	(0.047)	(0.041)	(0.035)	(0.018)	(0.015)
1(age > 55)*2002	0.243***	0.243***	0.188***	0.188***	0.071***	0.071***
	(0.046)	(0.039)	(0.033)	(0.030)	(0.018)	(0.013)
1(age > 55)*2003	0.192***	0.192***	0.175***	0.175***	0.051***	0.051***
	(0.044)	(0.040)	(0.035)	(0.031)	(0.018)	(0.014)
age group effects	Y		Y		Y	
age group-metro area effects	N	Y	N	Y	N	Y
metro-year effects	Y	Y	Y	Y	Y	Y
R-squared	0.85	0.92	0.83	0.89	0.91	0.96
N	1280	1280	1280	1280	1280	1280

Notes: Data are from IMSS and ENEU baseline samples, collapsed to metro area/age group/year level. Wage gap (medians) is log median real daily net wage from ENEU minus log median post-tax daily wage from IMSS. Wage gap (means) is defined analogously, using means in place of medians. Excess mass is calculated as described in Section 5 and Figure 9. In calculating evasion measures, we pool ENEU data across quarters within year. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details of data processing.

Table 6. Robustness to different definitions of excess mass, men

	dep. var.: excess mass (below indicated ENEU percentile)							
	15^{th}	20^{th}	25^{th}	30^{th}	35^{th}	40^{th}	45^{th}	50^{th}
1(age > 55)*1988	-0.002	0.012	0.022	0.019	0.015	0.022	0.026	0.026
	(0.026)	(0.025)	(0.024)	(0.024)	(0.025)	(0.024)	(0.023)	(0.022)
1(age > 55)*1989	0.013	0.023	0.026	0.013	0.016	0.024	0.031	0.030
	(0.020)	(0.021)	(0.021)	(0.023)	(0.024)	(0.025)	(0.024)	(0.023)
1(age > 55)*1990	0.019	0.026	0.027	0.016	0.016	0.017	0.017	0.015
	(0.020)	(0.021)	(0.022)	(0.022)	(0.024)	(0.025)	(0.024)	(0.023)
1(age > 55)*1991	0.037**	0.044**	0.042**	0.033	0.032	0.034*	0.029	0.021
	(0.018)	(0.019)	(0.019)	(0.020)	(0.021)	(0.020)	(0.020)	(0.019)
1(age > 55)*1992	0.010	0.024	0.029	0.025	0.021	0.027	0.021	0.024
	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.021)	(0.020)	(0.019)
1(age > 55)*1993	0.003	0.011	0.015	-0.001	-0.005	0.007	0.016	0.017
	(0.020)	(0.020)	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)	(0.019)
1(age > 55)*1994	-0.002	-0.006	0.002	-0.001	-0.000	0.012	0.027	0.028
	(0.019)	(0.019)	(0.019)	(0.021)	(0.022)	(0.021)	(0.021)	(0.019)
1(age > 55)*1996	0.040**	0.045**	0.053**	0.050**	0.047*	0.046*	0.048*	0.050**
	(0.019)	(0.020)	(0.021)	(0.024)	(0.025)	(0.025)	(0.026)	(0.024)
1(age > 55)*1997	0.017	0.032	0.037*	0.032	0.030	0.034	0.033	0.031
	(0.021)	(0.022)	(0.022)	(0.022)	(0.024)	(0.023)	(0.024)	(0.023)
1(age > 55)*1998	0.040**	0.052***	0.054***	0.046**	0.045**	0.048**	0.047**	0.044**
	(0.019)	(0.018)	(0.018)	(0.018)	(0.019)	(0.020)	(0.020)	(0.020)
1(age > 55)*1999	0.052***	0.057***	0.062***	0.048**	0.048**	0.049**	0.046**	0.048**
	(0.019)	(0.018)	(0.017)	(0.019)	(0.020)	(0.021)	(0.021)	(0.019)
1(age > 55)*2000	0.051***	0.052***	0.053***	0.050***	0.047**	0.051***	0.051**	0.045**
	(0.017)	(0.017)	(0.017)	(0.018)	(0.019)	(0.020)	(0.020)	(0.019)
1(age > 55)*2001	0.061***	0.069***	0.074***	0.067***	0.064***	0.057***	0.057***	0.054***
	(0.018)	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.020)	(0.020)
1(age > 55)*2002	0.049**	0.058***	0.071***	0.062***	0.056***	0.065***	0.070***	0.067***
	(0.020)	(0.019)	(0.018)	(0.018)	(0.019)	(0.020)	(0.021)	(0.020)
1(age > 55)*2003	0.046**	0.047**	0.051***	0.044**	0.042**	0.046**	0.049**	0.050***
	(0.019)	(0.019)	(0.018)	(0.019)	(0.020)	(0.020)	(0.021)	(0.019)
metro-year effects	Y	Y	Y	Y	Y	Y	Y	Y
age group effects	Y	Y	Y	Y	Y	Y	Y	Y
R-squared	0.92	0.92	0.91	0.91	0.90	0.90	0.89	0.89
N	1280	1280	1280	1280	1280	1280	1280	1280

Notes: Data are from IMSS and ENEU baseline samples, collapsed to metro area/age group/year level. Excess mass is calculated as described in Section 5 and Figure 9. In calculating excess mass, we pool ENEU data across quarters within year. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details.

Table 7. Differential effects of pension reform on employment gap, men

	dep. var.: log(empl., ENEU) - log(empl., IMSS)	
	(1)	(2)
1(age > 55)*1988	-0.052	-0.052
	(0.099)	(0.086)
1(age > 55)*1989	0.023	0.023
	(0.103)	(0.082)
1(age > 55)*1990	0.048	0.048
	(0.096)	(0.087)
1(age > 55)*1991	0.083	0.083
	(0.108)	(0.093)
1(age > 55)*1992	0.026	0.026
	(0.099)	(0.078)
1(age > 55)*1993	0.073	0.073
	(0.091)	(0.070)
1(age > 55)*1994	-0.248**	-0.248***
	(0.100)	(0.077)
1(age > 55)*1996	-0.015	-0.015
	(0.101)	(0.088)
1(age > 55)*1997	-0.017	-0.017
	(0.108)	(0.093)
1(age > 55)*1998	0.025	0.025
	(0.103)	(0.088)
1(age > 55)*1999	0.010	0.010
	(0.105)	(0.092)
1(age > 55)*2000	-0.028	-0.028
	(0.093)	(0.079)
1(age > 55)*2001	-0.008	-0.008
	(0.104)	(0.094)
1(age > 55)*2002	0.070	0.070
	(0.102)	(0.084)
1(age > 55)*2003	0.016	0.016
	(0.090)	(0.075)
age group effects	Y	
age group-metro area effects	N	Y
metro-year effects	Y	Y
R-squared	0.55	0.68
N	1280	1280

Notes: Samples are IMSS and ENEU baseline samples, collapsed to metro area/age group/year level. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details of data processing.

Table 8. Differential effects of pension reform on ENEU take-home wage, men

	dep. var.: log daily net wage, ENEU				
	(1)	(2)	(3)	(4)	(5)
1(age > 55)*1988	-0.058	-0.072**	-0.045	-0.042	-0.038
	(0.037)	(0.032)	(0.028)	(0.028)	(0.028)
1(age > 55)*1989	-0.034	-0.036	-0.031	-0.034	-0.030
	(0.041)	(0.034)	(0.032)	(0.031)	(0.031)
1(age > 55)*1990	-0.045	-0.054	-0.049	-0.047	-0.043
	(0.041)	(0.034)	(0.031)	(0.030)	(0.030)
1(age > 55)*1991	-0.065*	-0.051	-0.030	-0.033	-0.030
	(0.039)	(0.034)	(0.030)	(0.030)	(0.030)
1(age > 55)*1992	-0.064	-0.063*	-0.044	-0.044	-0.037
	(0.043)	(0.034)	(0.032)	(0.031)	(0.032)
1(age > 55)*1993	-0.092**	-0.049	-0.053	-0.050	-0.048
	(0.043)	(0.036)	(0.033)	(0.033)	(0.033)
1(age > 55)*1994	0.004	-0.001	0.015	0.019	0.023
	(0.044)	(0.036)	(0.034)	(0.034)	(0.034)
1(age > 55)*1996	0.032	0.032	0.043	0.038	0.040
	(0.043)	(0.036)	(0.032)	(0.032)	(0.032)
1(age > 55)*1997	-0.024	-0.013	0.010	0.017	0.022
	(0.043)	(0.036)	(0.032)	(0.032)	(0.031)
1(age > 55)*1998	0.013	-0.004	0.001	0.001	0.003
	(0.045)	(0.037)	(0.032)	(0.032)	(0.032)
1(age > 55)*1999	-0.008	-0.057*	-0.042	-0.051*	-0.047
	(0.041)	(0.033)	(0.030)	(0.030)	(0.030)
1(age > 55)*2000	-0.043	-0.091***	-0.055*	-0.055*	-0.049
	(0.040)	(0.035)	(0.032)	(0.032)	(0.031)
1(age > 55)*2001	-0.006	-0.071**	-0.044	-0.044	-0.040
	(0.041)	(0.033)	(0.030)	(0.029)	(0.029)
1(age > 55)*2002	0.038	-0.044	-0.012	-0.009	-0.006
	(0.039)	(0.032)	(0.029)	(0.029)	(0.028)
1(age > 55)*2003	-0.009	-0.051	-0.029	-0.032	-0.031
	(0.044)	(0.038)	(0.034)	(0.034)	(0.033)
age group effects	Y	Y	Y	Y	Y
metro-year effects	Y	Y	Y	Y	Y
schooling effects	N	Y	Y	Y	Y
married indicator	N	Y	Y	Y	Y
occupation effects	N	N	Y	Y	Y
industry effects	N	N	N	Y	Y
firm-size effects	N	N	N	N	Y
R-squared	0.13	0.40	0.46	0.47	0.48
N	667566	667566	667566	667566	667566

Notes: Sample is ENEU baseline sample. Take-home wage is the post-payroll-tax net wage as reported on ENEU. Estimates use population sampling weights provided in ENEU dataset. Controls include sets of 9 schooling indicators, 22 occupation indicators, and/or 50 industry indicators, in addition to the sets of five age-group and firm-size indicators, in indicated columns; details of category definitions are in Appendix C (online). *** 1%, ** 5%, * 10% level.

Table 9. Differential effects of pension reform within metro area/sector/firm size cell, men

	wage gap (medians)		wage gap (means)		exc. mass (below 25^{th} perc.)	
	(1)	(2)	(3)	(4)	(5)	(6)
(init. avg. age)*1989	-0.017**	-0.012	-0.013*	-0.008	-0.006	-0.003
	(0.008)	(0.008)	(0.007)	(0.008)	(0.004)	(0.004)
(init. avg. age)*1990	-0.007	-0.006	-0.016**	-0.014*	-0.005	-0.004
	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)	(0.004)
(init. avg. age)*1991	-0.004	-0.002	-0.013	-0.012	-0.002	-0.001
	(0.008)	(0.008)	(0.008)	(0.009)	(0.004)	(0.004)
(init. avg. age)*1992	0.004	0.004	-0.002	-0.002	0.002	0.002
	(0.009)	(0.010)	(0.009)	(0.010)	(0.004)	(0.004)
(init. avg. age)*1993	-0.005	-0.005	-0.009	-0.009	-0.006	-0.007*
	(0.008)	(0.008)	(0.008)	(0.008)	(0.004)	(0.004)
(init. avg. age)*1994	0.007	0.007	-0.000	0.002	0.003	0.002
	(0.007)	(0.008)	(0.008)	(0.008)	(0.003)	(0.003)
(init. avg. age)*1996	0.018**	0.019**	0.009	0.008	0.009**	0.009**
	(0.008)	(0.008)	(0.009)	(0.009)	(0.004)	(0.005)
(init. avg. age)*1997	0.021**	0.019**	0.011	0.008	0.007*	0.005
	(0.009)	(0.009)	(0.009)	(0.009)	(0.004)	(0.004)
(init. avg. age)*1998	0.028***	0.029***	0.013	0.012	0.014***	0.014***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.003)	(0.004)
(init. avg. age)*1999	0.043***	0.042***	0.035***	0.034***	0.014***	0.013***
	(0.007)	(0.008)	(0.008)	(0.008)	(0.003)	(0.003)
(init. avg. age)*2000	0.043***	0.039***	0.028***	0.024***	0.011***	0.009**
	(0.008)	(0.008)	(0.008)	(0.009)	(0.003)	(0.003)
(init. avg. age)*2001	0.043***	0.039***	0.028***	0.025***	0.012***	0.011***
	(0.008)	(0.008)	(0.008)	(0.008)	(0.003)	(0.004)
(init. avg. age)*2002	0.047***	0.044***	0.035***	0.032***	0.013***	0.013***
	(0.010)	(0.011)	(0.010)	(0.010)	(0.004)	(0.004)
(init. avg. age)*2003	0.036***	0.037***	0.029***	0.029***	0.010***	0.010***
	(0.010)	(0.011)	(0.009)	(0.010)	(0.003)	(0.003)
firmsize-metro-sector effects	Y	Y	Y	Y	Y	Y
metro-year effects	Y	Y	Y	Y	Y	Y
sector-year effects	N	Y	N	Y	N	Y
R-squared	0.81	0.82	0.75	0.76	0.84	0.85
N	3255	3255	3255	3255	3255	3255

Notes: Data are from IMSS and ENEU baseline samples, collapsed to metro area/firm size/sector/year level. Wage gap (medians), wage gap (means), excess mass calculated as in Table 4. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details.

Table 10. Differential effects of pension reform within metro area/sector/firm size cell, men

	wage gap (medians)		wage gap (means)		exc. mass (below 25^{th} perc.)	
	(1)	(2)	(3)	(4)	(5)	(6)
(init. share > 55)*1989	-1.265	-0.907	-0.698	-0.350	0.212	0.418
	(0.871)	(0.847)	(0.802)	(0.817)	(0.447)	(0.431)
(init. share > 55)*1990	-1.076	-1.004	-1.798**	-1.651**	-0.107	-0.058
	(0.789)	(0.798)	(0.791)	(0.820)	(0.418)	(0.429)
(init. share > 55)*1991	-0.735	-0.613	-1.442*	-1.338	0.166	0.272
	(0.858)	(0.856)	(0.822)	(0.844)	(0.473)	(0.454)
(init. share > 55)*1992	0.976	0.974	0.373	0.404	0.865*	0.860*
	(1.057)	(1.092)	(1.106)	(1.130)	(0.506)	(0.510)
(init. share > 55)*1993	-0.387	-0.364	-0.878	-0.828	-0.188	-0.212
	(0.832)	(0.865)	(0.766)	(0.787)	(0.431)	(0.438)
(init. share > 55)*1994	1.767**	1.758*	0.803	0.983	0.740*	0.697
	(0.872)	(0.921)	(0.826)	(0.878)	(0.421)	(0.446)
(init. share > 55)*1996	2.652***	2.739***	2.086***	2.062**	1.357***	1.389***
	(0.788)	(0.804)	(0.802)	(0.824)	(0.415)	(0.431)
(init. share > 55)*1997	2.887***	2.718***	1.862**	1.651*	1.391***	1.262***
	(0.871)	(0.847)	(0.924)	(0.906)	(0.434)	(0.417)
(init. share > 55)*1998	4.193***	4.255***	2.435***	2.359***	2.284***	2.265***
	(0.842)	(0.874)	(0.852)	(0.898)	(0.414)	(0.435)
(init. share > 55)*1999	5.483***	5.364***	3.938***	3.804***	2.062***	2.025***
	(0.825)	(0.831)	(0.773)	(0.810)	(0.426)	(0.430)
(init. share > 55)*2000	5.902***	5.553***	3.849***	3.562***	1.792***	1.622***
	(0.840)	(0.858)	(0.836)	(0.872)	(0.407)	(0.406)
(init. share > 55)*2001	5.301***	4.935***	3.365***	3.104***	1.683***	1.569***
	(0.911)	(0.925)	(0.813)	(0.839)	(0.412)	(0.415)
(init. share > 55)*2002	6.105***	5.816***	4.472***	4.250***	1.957***	1.935***
	(1.015)	(1.002)	(0.867)	(0.887)	(0.420)	(0.429)
(init. share > 55)*2003	5.176***	5.202***	3.771***	3.716***	1.603***	1.561***
	(1.007)	(1.048)	(0.912)	(0.967)	(0.405)	(0.410)
firmsize-metro-sector effects	Y	Y	Y	Y	Y	Y
metro-year effects	Y	Y	Y	Y	Y	Y
sector-year effects	N	Y	N	Y	N	Y
R-squared	0.81	0.82	0.75	0.76	0.84	0.85
N	3255	3255	3255	3255	3255	3255

Notes: Data are from IMSS and ENEU baseline samples, collapsed to metro area/firm size/sector/year level. Wage gap (medians), wage gap (means), excess mass calculated as in Table 4. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details.

Table 11. Differential effects on age composition within metro area/sector/firm size cell, men

	avg. age		share age 56-65	
	(1)	(2)	(3)	(4)
(init. age measure)*1989	0.221***	0.266***	0.195***	0.216***
	(0.043)	(0.044)	(0.056)	(0.055)
(init. age measure)*1990	0.162***	0.196***	0.055	0.067
	(0.051)	(0.050)	(0.054)	(0.054)
(init. age measure)*1991	0.033	0.067*	-0.022	-0.009
	(0.037)	(0.036)	(0.046)	(0.047)
(init. age measure)*1992	-0.004	0.028	-0.014	-0.002
	(0.042)	(0.041)	(0.054)	(0.055)
(init. age measure)*1993	-0.002	0.019	-0.002	0.006
	(0.037)	(0.038)	(0.049)	(0.049)
(init. age measure)*1994	-0.043	-0.014	-0.092*	-0.070
	(0.037)	(0.035)	(0.053)	(0.052)
(init. age measure)*1996	-0.016	-0.016	-0.034	-0.034
	(0.037)	(0.038)	(0.048)	(0.048)
(init. age measure)*1997	-0.037	-0.035	-0.039	-0.040
	(0.048)	(0.050)	(0.049)	(0.049)
(init. age measure)*1998	-0.002	-0.012	0.018	0.010
	(0.032)	(0.031)	(0.045)	(0.045)
(init. age measure)*1999	-0.036	-0.052	-0.032	-0.047
	(0.035)	(0.034)	(0.046)	(0.046)
(init. age measure)*2000	-0.029	-0.038	-0.040	-0.058
	(0.042)	(0.042)	(0.050)	(0.052)
(init. age measure)*2001	-0.048	-0.066*	-0.005	-0.026
	(0.036)	(0.036)	(0.050)	(0.051)
(init. age measure)*2002	-0.055	-0.072*	-0.021	-0.041
	(0.038)	(0.037)	(0.053)	(0.052)
(init. age measure)*2003	-0.091**	-0.117***	-0.042	-0.070
	(0.040)	(0.042)	(0.057)	(0.057)
firmsize-metro-sector effects	Y	Y	Y	Y
metro-year effects	Y	Y	Y	Y
sector-year effects	N	Y	N	Y
R-squared	0.89	0.90	0.81	0.82
N	3255	3255	3255	3255

Notes: The initial age measure in Columns 1-2 is the mean age in 1988; in Columns 3-4 it is the share in the oldest age group in 1988. Data are from IMSS and ENEU baseline samples, collapsed to metro area/firm size/sector/year level. Wage gap (medians), wage gap (means), excess mass calculated as in Table 4. *** 1%, ** 5%, * 10% level. See Section 4 and Appendix C (online) for further details.