COLUMBIA | SIPA
Center for Development Economics and Policy

CDEP-CGEG WORKING PAPER SERIES

CDEP-CGEG WP No. 64

Exports and Wage Premia: Evidence from Mexican Employer-Employee Data

Judith A. Frias, David S. Kaplan, Eric Verhoogen, and David Alfaro-Serrano

September 2018

COLUMBIA | SIPA
Center on Global Economic Governance

Exports and Wage Premia: Evidence from Mexican Employer-Employee Data*

Judith A. Frías†

David S. Kaplan‡

Eric Verhoogen§

David Alfaro-Serrano**

Sept. 2018

Abstract

This paper draws on employer-employee and longitudinal plant data from Mexico to investigate the impact of exports on wage premia, defined as wages above what workers would receive elsewhere in the labor market. We decompose plant-level average wages into a component reflecting skill composition and a component reflecting wage premia. Using the late-1994 peso devaluation interacted with initial plant size as a source of exogenous variation in exports, we find that exports have a significant positive effect on wage premia, and that the effect on wage premia accounts for essentially all of the medium-term effect of exporting on plant-average wages.

*This paper began many years ago, and has accumulated many debts along the way. We would like to thank, without implicating, Rita Almeida, Paula Bustos, David Card, Bo Honoré, Leonardo Iacovone, Wojciech Kopczuk, Thomas Lemieux, Arik Levinson, Steven Lich-Tyler, Leigh Linden, Osmel Manzano, Marc Melitz, Marcelo Moreira, Marc-Andreas Muendler, Daniel Parent, Cristian Pop-Eleches, Jesse Rothstein, Matthias Schuendeln, Matthew Slaughter, John Strauss, John Sutton, Catherine Thomas, Miguel Urquiola, Till von Wachter, and many seminar participants for helpful comments. We would also like to thank Gerardo Durand, José Antonio Mejía, Jorge Alberto Reyes, Gerardo Leyva, Abigail Durán, Adriana Ramírez, and Gabriel Romero of INEGI for help with the first versions of the paper, and Natalia Volkow and other staff of the INEGI Microdata Laboratory for assistance after a long hiatus due in part to a change in data-access regime. All analysis of confidential data was carried out on-site and in accordance with Mexican confidentiality laws. Guru Sethupathy, Hitoshi Shigeoka, Kensuke Teshima and Matt Wai-Poi provided excellent research assistance. Verhoogen is indebted to the National Science Foundation (SES-0721068) for funding. An earlier version of this paper was presented at an NBER ITI meeting in 2005, a UC Berkeley labor lunch in 2009, and a number of other seminars. The views expressed are the opinions of the authors and do not reflect the official positions of the Instituto Mexicano del Seguro Social, the Secretaría de Hacienda y Crédito Público, or the Inter-American Development Bank.

†Unidad de Productividad Económico, Secretaría de Hacienda y Crédito Publico, judith_frias@hacienda.gob.mx.

‡Inter-American Development Bank, dkaplan@iadb.org

§Columbia University, eric.verhoogen@columbia.edu. Corresponding author.

**Columbia University, da2628@columbia.edu.

1 Introduction

In textbook models of perfectly competitive labor markets, the wage a worker earns does not depend on the particular firm in which she is employed; she would receive the same wage in any firm that would employ her in equilibrium. A long-established and recently resurgent literature has challenged this view. Following the seminal work of Abowd, Kramarz and Margolis (1999, hereafter AKM), a number of recent studies have fit simple models with individual and firm fixed effects and have found (under a conditional random mobility assumption discussed below) that the firm effects account for a substantial fraction of overall wage variation (Card et al., 2013, 2015; Barth et al., 2018; Alvarez et al., 2018; Song et al., 2016). In other words, firms appear to pay different wage premia, defined as wages above what individuals would earn elsewhere on the labor market, and understanding how these premia evolve appears to be important for understanding the broader evolution of wage structures.

Despite the recent advances, there has been relatively little empirical work on *why* firms pay wage premia. What leads some firms to pay more than others to similar workers? Much of the recent literature has aimed to characterize the extent of rent-sharing with different groups of workers or the contribution of firms to overall wage inequality, but not to identify the causal determinants of wage premia.[1] The recent review by Card et al (2018) concludes with a call for research to fill this gap.[2]

In this paper, we draw on a combination of employer-employee and longitudinal plant data from Mexico to investigate one possible determinant of wage premia: firms' engagement in export markets. It is well documented that exporting firms tend to pay higher wages on average, within narrow industries.[3] There is growing evidence that this relationship is causal — that exogenous increases in exporting lead firms to pay higher wages on average (Verhoogen, 2008; Álvarez and López, 2009; Kandilov, 2009; Bustos, 2011; Brambilla et al., 2012). But it remains unclear to what extent these plant-level results reflect changes in wage premia as opposed to changes in workforce

[1]See for example Card et al. (2015) on differences in rent-sharing for different groups, and Card et al. (2013), Barth et al. (2016, 2018), and Song et al. (2016) on characterizing the role of firms in overall wage inequality. Lazear and Shaw (2018) provide a useful overview.

[2]"The field continues to rely almost exclusively on observational studies predicated on plausible, but ultimately debatable, identifying assumptions. More research is needed applying (quasi-)experimental research designs that convincingly tease out the mechanisms through which firm shocks are transmitted to workers." (Card et al., 2018, p. 24)

[3]The seminal papers are Bernard and Jensen (1995, 1999). Brambilla et al. (2017) review the literature and present evidence from 61 developing countries of a robust positive relationship between exporting and firm-level wages.

composition. Several studies (discussed briefly below) have examined the relationship between exporting and wages in employer-employee data, but to our knowledge none has employed quasi-experimental variation to estimate the effect of exporting on wage premia as we have defined them.

This paper proceeds in two steps. First, following Card, Heining and Kline (2013, hereafter CHK), we fit simple AKM-type models in different periods and decompose plant-level wages into a plant component, which we interpret as a plant-specific average wage premium, and a person component reflecting individual characteristics, including individual effects capturing time-invariant ability.[4] Because of data constraints, discussed below, we focus on the periods 1992-1994, 1996-1998, and 2000-2002 in our baseline specification. We conduct diagnostic exercises similar to CHK and find that the data appear to be well described by a model with additive individual and plant fixed effects within several-year periods.

Second, following Verhoogen (2008), we use the late-1994 peso devaluation as a source of exogenous variation in the incentive to export, to identify the effect of exporting on the estimated wage premia. Verhoogen (2008) motivated this approach with a model of quality upgrading in which firms in a developing country are heterogeneous in entrepreneurial ability (or "capability"), sell higher-quality varieties when exporting in order to appeal to richer consumers in the export market, and pay higher wages when producing higher-quality goods. A devaluation leads more-capable firms to increase export share and hence to raise average product quality, average wages, and (possibly) wage premia more than less-capable firms in the same industry.[5] In this paper, we rely on the same reduced-form prediction for the relationship between exporting and wage outcomes without taking a stand on the underlying theoretical mechanism. (We discuss several possible mechanisms below.) As in the earlier paper, we use plant size as the primary proxy for the underlying plant heterogeneity and interpret the interaction of initial size and an indicator for the devaluation period as a proxy for the differential inducement to export generated by the peso devaluation. We estimate an instrumental-variables (IV) regression of within-plant changes in wage premia on changes in export share instrumented by this interaction. We also present the corresponding reduced-form regressions of the changes in wage premia on the interaction term

[4] In Mexican manufacturing, the vast majority of firms are single-establishment. For instance, the 2009 Economic Census in Mexico reports that more than 75% of manufacturing firms with 51 or more employees, and more than 90% of manufacturing firms with 11 or more employees, are single-establishment. The data we use are at the plant level and we will typically refer to plants as decision-makers.

[5] Verhoogen (2008, Section III.B) suggested that the plant-level wage effects could reflect changes either in wage premia or in workforce composition.

and other covariates.

Our main finding is that exporting has a robust, positive, statistically significant effect on wage premia. The magnitude is economically significant: a 1% increase in a plant's export share is associated with approximately a 3% increase in wage premia at the plant. Our estimates suggest that essentially all of the effect of exporting on average wages at the plant level is explained by the increase in wage premia, rather than by changes in workforce composition, at least for the medium-term time frame we are able to study. These results are robust to using two alternative proxies for plant capability, total factor productivity (TFP) and predicted export share, and to including an earlier pre-crisis period (1988-1990) in the analysis. We conclude that exporting is an important determinant of wage premia among Mexican manfacturing firms.

Two broader implications of these findings seem particularly salient. First, for the recently resurgent literature on the role of firms in wage-setting, they underline the causal role of product markets in shaping the evolution of wage distributions through their effect on firms' wage policies. Product-market shocks are potentially related to, but are conceptually distinct from, shocks to productivity (i.e. technical efficiency), which have been the focus of much of the rent-sharing literature.[6] Second, the paper contributes to a growing body of evidence about how firms' responses to trade shocks contribute to wage inequality. Previous work has argued that trade liberalization can increase wage dispersion across firms within industries in developing countries, as larger, more-productive firms, which already tend to be higher-wage, take greater advantage of export opportunities, which leads them to raise wages further (Verhoogen, 2008; Helpman et al., 2017). This paper finds that the increase in across-plant dispersion in our setting is largely driven by plants' wage policies, rather than sorting on worker ability. The limited impact on workforce composition we find suggests that the export shock had limited effects on the general-equilibrium return to skill, at least in the short to medium run. The trade-induced increase in wage dispersion may be confined to manufacturing (and other tradable) sectors.

As mentioned above, a number of theoretical mechanisms are consistent with a causal relationship between exports and wage premia. Differential quality upgrading is one possibility, if producing higher-quality products for richer consumers in the export market requires a particularly motivated workforce, which in turn requires paying high efficiency wages (Verhoogen, 2008). A second possibility is fair-wage concerns: if the exogeous increase in exporting is associated with

[6] See e.g. Van Reenen (1996), Kline et al. (2017), and the review by Card et al. (2018). These papers are discussed further below.

an increase in profitability (as in many Melitz (2003)-type models), firms may share profits in order to induce workers to reciprocate with effort (Akerlof and Yellen, 1990; Egger and Kreickemeier, 2009, 2012; Amiti and Davis, 2012). A third possibility is that the labor market is characterized by search frictions and firm-worker bargaining; several papers have posited mechanisms through which the bargained wage may increase with exports in the context of heterogenous-firm models.[7] A fourth possibility is that managers share rents with workers in order to ensure themselves a quiet life, rather than mazimizing profits (Bertrand and Mullainathan, 2003).[8] These mechanisms (and others) are often classified under the general heading of "rent-sharing." In this paper, we do not take a stand on precisely which mechanism is correct. Nevertheless, we believe that our results carry an important implication for trade theory, by calling into question the still-common approach of assuming perfectly competitive, frictionless labor markets, in which any relationship between exporting and wages at the firm level is explained only by the sorting of workers by skill.

In addition to the research cited above, this paper is related to several streams of literature. Early papers investigating the effects of exogenous shocks on rent-sharing at the firm level include Abowd and Lemieux (1993) and Van Reenen (1996). In firm-level data from collective bargaining contracts in Canada, Abowd and Lemieux (1993) employed industry-level import and export prices as instruments for quasi-rents per worker (a measure of profitability when labor is valued at its alternative wage), and found that greater quasi-rents led to higher wages. In panel data on British firms, Van Reenen (1996) related major innovations, which are arguably exogenous, to wage changes at the firm level, and found that quasi-rents due to innovation were passed through to higher wages. In both cases, the researchers lacked data on individual workers and hence were not able to definitively answer questions about whether the wage changes reflect changes in wage premia or changes in skill composition.[9]

A recent paper by Kline et al. (2017) also examines innovation and rent-sharing, with employer-employee data and rich patent information from the U.S. Under the identifying assumption that

[7]In Helpman et al. (2010), firms employ a fixed-cost screening technology; as scale increases with increased exports, firms screen more intensively, which increases the revenue per worker to be bargained over, which in turn increases wages. In Coşar et al. (2016), firms face labor adjustment costs and the benefit of filling a position is greater for expanding firms than shrinking firms, leading the bargained wage to be higher in such firms; to the extent that increased exporting leads firms to expand, wages will also increase. See also Davidson et al. (2008), Felbermayr et al. (2011), Fajgelbaum (2016), and Bellon (2017).

[8]This list of possible mechanisms is not exhaustive. Another is simply that workers are unionized and are able to bargain for a share of firm profits.

[9]The current paper is also related to the earlier literature on *industry* wage differentials (Krueger and Summers, 1988; Katz and Summers, 1989) and in particular to Dickens and Katz (1987), which examined correlations between various industry chacteristics and industry wage differentials.

the U.S. Patent Office's initial decision on a patent application is as good as random, conditional on observable characteristics of the application and firm, the authors find that each patent-induced additional dollar of operating surplus yields a 29-cent increase in a firm's wage bill. Using a matching estimator in Finnish data, Aghion et al. (2018) find that patents are associated with higher wages for co-workers of inventors. They also make the important point that the higher wages for co-workers may reflect compensation for efforts to operationalize and commercialize an invention, rather than a pure shock to profitability or productivity.[10] The current paper is complementary to these papers and focused on a very different type of shock — a shock to product-market conditions rather than to the technical capabilities of a firm.

As noted above, this paper is also related to a number of papers that have related exporting to wages in employer-employee data, including Schank et al. (2007), Munch and Skaksen (2008), Davidson et al. (2014), Baumgarten (2013), Klein et al. (2013), Irarrazabal et al. (2013), Hummels et al. (2014), Krishna et al. (2014), Araújo and Paz (2014), Macis and Schivardi (2016), Helpman et al. (2017), and Barth et al. (2018).[11] To our knowledge, this paper is the only one to relate quasi-experimental variation in exports to differences in wage premia defined as wages above what a given individual would earn elsewhere. Schank et al. (2007), Munch and Skaksen (2008), Hummels et al. (2014), Araújo and Paz (2014), and Krishna et al. (2014) estimate models with job-spell effects, which absorb wage differences across firms and hence do not address the question of why a given individual earns a higher wage at one firm than another.[12] Klein et al. (2013), Macis and Schivardi (2016), and Barth et al. (2018) relate firm (or plant) effect estimates to exports (and other firm-level variables), but do not employ a design-based strategy to isolate the

[10]In this sense, it is not clear that a patent can be considered a pure shock to firm productivity or profitability; commercializing an invention may require changes in how a firm allocates workers' time or organizes production.

[11]An earlier version of the current paper (Frías et al., 2009) estimated a model that allowed more flexibly for changes over time in the return to individual ability, using an approach pioneered by Holtz-Eakin et al. (1988) and found results broadly similar to those reported here. In this version, delayed in part by a change in data-access regime, we focus on simple AKM-type specifications estimated in separate periods, because the method is arguably more transparent and is more comparable to the literature as it is developing after the important contribution of CHK.

[12]The paper by Hummels et al. (2014), perhaps the highest-profile in this set, focuses on the wage effects of offshoring in Denmark but also examines the wage effects of exporting. As an instrument for exports, the authors construct a firm-specific weighted average of imports of particular goods by a firm's trading partners, using the firm's initial export shares as weights. As the authors acknowledge, this strategy is subject to the concern that demand shocks are correlated in Denmark and the trading partners. In addition, the paper estimates the effect of exporting only within job spells, as noted above.

effect of exogenous variation in exports.[13,14]

The next section describes our econometric strategy in more detail. Section 3 describes the data and briefly provides background on the peso crisis. Section 4 presents results from the first step of our econometric procedure, estimating wage premia at the plant level. Section 5 presents results from the second step, estimating the effect of exports on the estimated wage premia. Section 6 concludes.

2 Econometric Strategy

Our estimation strategy has two parts. We first use the employer-employee data to decompose plant-level wages into a plant component due to wage premia and a person component due to skill composition. We then relate changes in those components to the export shock brought about by the peso devaluation, linking the employer-employee results to longitudinal data on manufacturing plants.

Within a given (several-year) period, we assume that the log wage of person i in time t is given by:

$$w_{it} = \alpha_i + \psi_{J(i,t)} + X'_{it}\beta + \varepsilon_{it} \tag{1}$$

where α_i is a time-invariant individual effect; $J(i,t)$ indicates the plant j in which person i is employed in year t; $\psi_{J(i,t)}$ is the corresponding plant effect; X_{it} is a vector of time-varying observables; and ε_{it} is a mean-zero error term. This specification follows AKM and CHK.[15]

We interpret the individual effect α_i as portable ability, invariant during a given several-year period, compensated equally in all firms. The plant effect $\psi_{J(i,t)}$ reflects a premium (or discount) paid by plant j to all employees. In X_{it}, we include tenure and its square, year effects, and a polynomial in age. Including a linear term for age is not possible because it would be collinear with the person and year effects. Following CHK and Card et al. (2015), we drop the linear term

[13]Macis and Schivardi (2016) attempt to implement a design similar to Verhoogen (2008) and this paper using the 1992 devaluation of the Italian lira, but they find no evidence of a differential effect of the devaluation by firm size. In the end, they regress estimates of wage premia on current export share, without instrumenting, letting the coefficient on export share differ pre- and post-devaluation.

[14]In a related paper using a similar design as this paper, Frías et al. (2012) show (without controlling for individual effects) that exports increase within-plant wage inequality.

[15]Card et al. (2018) provide a theoretical justification for this specification, in a model in which workers have idiosyncratic attachments to particular plants and either different worker skill groups are substitutes in the production function or the plants face the same supply elasticity from these different groups.

and include the square and cube of recentered age (age - 40). We define the "person" component to be the individual effect plus the contribution of the other observables: $s_{it} \equiv \alpha_i + X'_{it}\beta$. We will use the term "average person component" to refer to the plant-level mean of the (individual-level) person component. (We will use the terms plant effect and plant component interchangeably.)

The identifying assumption is that the error term ε_{it} is uncorrelated with the other covariates in all years (within a given several-year period), a sufficient condition for which is:

$$E(\varepsilon_{it}|X_{i1}...X_{iT}, \psi_1...\psi_J, \alpha_i) = 0 \tag{2}$$

In the employer-employee literature, (2) is referred to as a *conditional random mobility* assumption, since it requires that, conditional on observables, an individual's current-period idiosyncratic shock is uncorrelated with which plant she is employed in.[16] While (2) rules out sorting across plants on the basis of contemporaneous shocks to individual ability, it allows for many salient forms of sorting. For instance, as CHK point out, it allows for sorting on time-invariant ability (e.g. higher turnover among lower-ability workers) as well as for the possibility that workers are more likely on average to move from low- to high-wage plants than vice-versa (since we condition on $\psi_1...\psi_J$). In Section 4 below we present evidence, following CHK, that the model (1)-(2) appears to summarize well the wage and mobility patterns we observe.

We face an important choice about how to define the several-year periods in which to estimate this model. As discussed in Section 3 below, the employer-employee data are available for 1985-2005, but the plant panel with export information is available on a consistent basis only for 1993-2003. The peso devaluation occured in December 1994, and the crisis played out over the next several months. We need the first period to be clearly pre-devaluation, but we also need information on export status, the earliest of which is from 1993. Our preferred solution is to use 1992-1994 as the pre-crisis period. We do not use the crisis year 1995. To maintain equally spaced periods, we use 1996-1998 as the first post-crisis period and 2000-2002 as the second post-crisis period. We will refer to 1992-1994, 1996-1998, and 2000-2002 as periods 1, 2, and 3 respectively.[17]

As discussed in Abowd et al. (2002), unique solutions for the estimates of the person and plant

[16] In the terminology of the panel-data literature, (2) requires the covariates to be strictly exogeneous, i.e. uncorrelated with all past and future error terms. Intuitively, after a within transformation, the transformed covariates and errors contain within-individual averages across years, which must be uncorrelated.

[17] In a robustness check using the employer-employee data alone, we will also present estimates for 1988-1990, which we refer to as period 0.

effects can only be obtained within a set of plants linked by worker switchers during the period (a "connected set"). As has become standard in the literature, we focus on the largest connected set of plants in each period. We will see below that the largest connected sets capture a smaller share of plants and workers than has typically been the case in developed-country settings, but also that they capture almost all of the plants in the plant panel (for which we have export information).

Under assumption (2), for plant $j = j'$ in year $t = t'$ the expected plant-level wage can be expressed as the sum of two components:

$$\begin{aligned} E(w_{it}|j=j',t=t') &= \psi_{j'} + E(\alpha_i + X'_{it}\beta|j=j',t=t') \\ &= \psi_{j'} + E(s_{it}|j=j',t=t') \end{aligned} \tag{3}$$

where the conditioning on $X_{i1}...X_{iT}, \psi_1...\psi_J, \alpha_i$ is omitted by should be understood. The first term is the plant component, the wage premium paid by the plant. The second term is the expected person component, which captures average workforce skill. To generate sample analogues of these components, we fit model (1) and recover the estimated plant effects and a plant-year-level average of the estimated person components, which we refer to as the average person component. Computationally, we estimate (1) using the pre-conditioned conjugate gradient (PCG) algorithm in Matlab, following CHK. We estimate the model separately by (3-year) period.

Once we have recovered these estimates, the next step is to determine the effect of exporting on the plant and average person components. As mentioned above, our approach is motivated by the Melitz (2003)-type theoretical framework of Verhoogen (2008). In this framework, firms differ in an underlying "capability" parameter, λ, there is a fixed cost of exporting, and firm capability and worker effort are complements in determining product quality. In equilibrium, more-capable firms are larger, pay higher wages, and produce higher-quality goods in cross-section. In response to an exogenous inducement to export such as a devaluation, more-capable firms increase exports, increase average product quality, and raise wages relative to less-capable firms in the same industry. If producing high-quality products requires especially motivated workers, which in turn requires paying high efficiency wages, then the increase in exports will be accompanied by an increase in wage premia. As noted in the introduction, subsequent papers have developed heterogenous-firm models that emphasize different mechanisms but that generate similar predictions for the relationship between firm size, exporting and wages in response to trade shocks, and a subset of these predict similar effects on wage premia.

Motivated by this framework, we adopt an econometric specification of the following form:

$$\triangle y_{jp} = \theta \triangle e_{jp} + \gamma \widehat{\lambda}_{jp-1} + \xi_{kp} + \psi_{rp} + u_{jp} \tag{4}$$

where j indexes establishments, p indexes periods; $\triangle y_{jp}$ is the change in an outcome variable (e.g. wages, the plant component or average person component from our AKM-type model, or the capital-labor ratio) between period $p-1$ and period p; $\triangle e_{jp}$ is the change in export share between period $p-1$ and p; $\widehat{\lambda}_{jp-1}$ is a proxy for the capability parameter, λ, discussed below; and ξ_{kp}, ψ_{rp} and u_{jp} are an industry-period effect, a region-period effect and a mean-zero disturbance, respectively. For variables that vary by year, we average over years within a period. Note that (4) is effectively in first-differences, with time-invariant firm characteristics differenced out.

There are a number of reasons why the change in export share, $\triangle e_{jp}$, may be correlated with the error term, u_{jp}, and OLS estimation of (4) may plausibly lead to either positive or negative bias in the estimate of the coefficient of interest, θ. On one hand, positive productivity shocks at the plant level may lead both to greater exports and to higher wages, generating a positive bias in the OLS coefficient. On the other hand, a positive labor supply shock to a plant would be expected to lead to lower wages and greater exports, generating a negative bias in OLS. Other biases are also possible.[18] To address the endogeneity of export changes, we instrument $\triangle e_{jp}$ with the interaction between an indicator for the period immediately following the peso crisis ($p = 2$) and the value of the capability proxy in the previous period: $\widehat{\lambda}_{jp-1} * T_2$. This term can be interepreted as a proxy for the differential inducement to export created by the peso devaluation, which we take to be exogenous. Note that any stable relationship between the capability proxy and the change in wage outcomes will be captured by the uninteracted $\widehat{\lambda}_{jp-1}$ term. The IV estimate of θ will reflect only the effect of the change in exporting induced by the differential impact of the devaluation on more- vs. less-capable firms between period 1 and period 2.

A potential concern with our IV approach is that exporting and wages can be thought of as simultaneous outcomes of the same firm/plant optimization problem. In this view it is not clear it makes sense to estimate the effect of one choice on the other. This difficulty can be avoided by

[18] For instance, a positive demand shock specific to the domestic market would likely lead to greater investment, a higher capital-labor ratio, higher wages, and a *lower* export share (since domestic sales appear in the denominator of export share), generating a negative bias in OLS.

focusing on the reduced form corresponding to the IV model:

$$\triangle y_{jp} = \varphi(\widehat{\lambda}_{jp-1} * T_2) + \widetilde{\gamma}\ \widehat{\lambda}_{jp-1} + \widetilde{\xi}_{kp} + \widetilde{\psi}_{rp} + \widetilde{u}_{jp} \tag{5}$$

where $\widetilde{\xi}_{kp}$, $\widetilde{\psi}_{rp}$, and $\widetilde{u}_{jp}$ are again an industry-period effect, a region-period effect, and a mean-zero error, respectively. Here the $\widehat{\lambda}_{jp-1} * T_2$ term, the proxy for the differential inducement to export, is clearly external to the firm, and the coefficient on that term can be interpreted as capturing the effect of the export inducement on the outcome variable.[19]

A key step in the implementation is to choose the proxy for the capability parameter, $\widehat{\lambda}$. Following Verhoogen (2008), we use plant size as our primary proxy. Intuitively, since capability leads plants to become large, we can infer plants' underlying capability from their size. Plant size has the advantage that it is directly observed, is relatively well measured, and does not require the strong assumptions imposed by standard residual-based methods of estimating total factor productivity (TFP). Two natural measures of plant size are domestic sales and employment. In Melitz (2003)-type theoretical frameworks (such as the one used by Verhoogen (2008), which we rely on here), domestic sales have the attractive theoretical property that they bear a smooth, continuously differentiable relationship to the latent capability term; by contrast, employment (like total sales) jumps discontinuously at the extensive margin for exports. On the other hand, domestic sales have the empirical disadvantage that measurement error can generate spurious correlation with the export share, since domestic sales appear in the denominator of export share. Measurement error in domestic sales in period $p-1$ will generate a spurious negative relationship with the level of the export share in $p-1$ and a spurious positive relationship with the change in export share from $p-1$ to p.[20] Because our primary specification is an IV specification with export share on the left-hand side of the first stage, we prefer (log) employment as our primary proxy for the capability parameter, $\widehat{\lambda}$. But we also present results using (log) domestic sales in

[19]Note that the reduced form is equivalent to the approach of Verhoogen (2008). That paper estimated OLS models of the form:

$$\triangle y_{jp} = \mu + \widehat{\lambda}_{jp-1}\beta_p + \widetilde{\xi}_k + \widetilde{\psi}_r + v_j \tag{6}$$

separately by period, with $p = 1993, 1997, 2001$, and then compared β_{1997} (for the peso crisis period 1993-1997) to β_{2001}. Except for the fact that here we look at three-year periods 1992-1994, 1996-1998, and 2000-2002 in place of individual years 1993, 1997, 2001, the difference in coefficients $\beta_{2001} - \beta_{1997}$ is equivalent to the OLS estimate of φ in our reduced-form model, (5).

[20]A similar issue arises when log employment is used as the proxy and the capital-labor ratio or average hourly wage (from the EIA data), which are calculated by dividing by employment, are dependent variables.

the reduced form, (5), where export share is not on the left-hand side and spurious correlation due to measurement error is not a concern.[21]

To check robustness, we also present results using two additional proxies: a measure of total factor productivity using the Levinsohn and Petrin (2003) method, and an index of export propensity, calculated as the predicted values from a tobit regression of export share on log sales, log employment and log capital-labor ratio. Since both rely in part on variation in sales, they are subject to the spurious correlation issue discussed above; for this reason, we only estimate the reduced-form model when we use them. Qualitatively, the results do not depend on the choice of proxy.

We conduct two further robustness checks. First, using only the IMSS data we include an earlier period, 1988-1990 (which we refer to as period 0), in addition to the three periods mentioned above, 1992-1994 (period 1), 1996-1998 (period 2), and 2000-2002 (period 3). This allows us to compare changes between two pre-shock periods (periods 0 and 1) to changes over periods that span the shock (periods 1 and 2). Because we do not observe export share when using only the IMSS data, we again use the reduced-form model for this check. Second, we estimate the effect of exporting on the wages of stayers, workers who are continuously employed in a given firm over two periods. These results do not use the AKM methodology and in particular do not require the conditional random mobility assumption, (2). They are analogous to estimating a model with job-spell fixed effects, as in several papers cited above. We will see that both of these robustness checks yield results qualitatively similar to our baseline estimates.

In both the IV and reduced-form approaches, our interpretation relies on the assumption that the devaluation affected wage outcomes differentially within industries only through its impact on the incentives of plants to export. Verhoogen (2008) considered a number of reasons why this assumption might be violated, if for instance the devaluation affected larger and smaller firms within industries differently for reasons unrelated to exporting. Readers are referred to that paper for more extensive discussion, but two points are worth re-emphasizing. First, the assembly-for-export (*maquiladora*) sector in Mexico provides a sort of placebo test. The *maquiladora* sector exported essentially all of its output both before and after the devaluation, and hence we would not expect the devaluation to have generated a differential within-industry shock to exporting. At the

[21] Verhoogen (2008) used log domestic sales as the preferred proxy. But that paper used a reduced-form framework (see footnote 19) where estimates did not rely on a first stage with export share as the dependent variable. (When export share was on the left-hand side, that paper instrumented log domestic sales with its lag.)

same time, if the macro shock had a differential within-industry impact through a channel other than exporting, one would expect it to show up in the *maquiladora* sector as well. Consistent with our interpretation, there was no differential change in wages between larger and smaller *maquiladora* plants during the peso-crisis period. Second, it does not appear that differential access to credit markets can explain the empirical patterns. While there is evidence that exporting plants faced a lower cost of capital than non-exporters, likely due to greater access to foreign capital, they also had a greater share of dollar-denominated loans before the crisis and hence their balance sheets were more adversely affected by the devaluation; these effects appear to offset. There was no differential within-industry change in the cost of capital that would explain the differential wage changes.

3 Data and Background

The employer-employee data we use are from the administrative records of the *Instituto Mexicano del Seguro Social (IMSS)*, the Mexican social security agency.[22] In principle, all private Mexican employers are required to report wages for their employees to IMSS and to pay social-security taxes on the basis of their reports. In practice, only about half of private-sector remunerated employees are reported to IMSS (and hence considered formal), and half are considered informal; see Appendix Table A.1.[23] While the size of the informal sector seems large by developed-country standards, it is not out of line for countries at Mexico's income level.[24] At the level of individuals, the IMSS data contain information on age, sex, daily wage (including benefits), and state and year of the individual's first registration with IMSS. Unfortunately, they do not contain information on individuals' education levels. At the establishment level, the data contain only industry and location. We have access to the IMSS data from 1985 to 2005.

An important practical issue is that the bottom- and top-codes in the IMSS data have changed over time. Prior to 1991, IMSS allowed establishments to report wages below the corresponding regional minimum wage;[25] beginning in 1991, this practice was disallowed (even if actual wages

[22]Previous papers using these data include (Castellanos et al., 2004; Kaplan et al., 2004, 2005).

[23]Most public-sector workers and employees of the state-run petroleum company are covered by separate systems.

[24]See e.g. Schneider and Enste (2000). The employment figures for manufacturing in the IMSS data differ by less than 10% from independently reported figures in the 1993 Industrial Census, suggesting that underreporting of employment is not especially severe in the social security data (Kaplan et al., 2004, 2005).

[25]There are three minimum-wage regions in Mexico, with the minimum wage in Mexico City and other urban areas generally 10-20% higher than in poorer rural areas.

presumably continued to be below the legal minimum in some cases). The changes in the top-code are illustrated in Figure 1, which also displays several wage quantiles from the IMSS data.[26] (Note also that average real wages dropped significantly following the peso devaluation in late 1994; we will see below that they simply dropped less in plants that increased exports during the crisis period.) To reduce biases due to changes in top- and bottom-codes, we "winsorize" the wage data, by replacing wages below the 10th percentile by the wage at the 10th percentile, and wages above the 90th percentile by the wage at the 90th percentile. This process also reduces the influence of outliers due to misreporting and other forms of measurement error.[27] Another potentially important measurement issue is that employers' payroll tax burdens depend on the wages they report to IMSS and hence they have an incentive to under-report wages. Kumler et al. (2015) document such under-reporting. But they also show that that the under-reporting appears to be minimal for the set of larger manufacturing plants that can be linked to the EIA plant panel; see Figure 10 of that paper and the corresponding discussion in the main text. Moreover, as long as any under-reporting is constant over time within plant, it will be differenced out in our estimation procedure.

Our cleaning procedure for the IMSS data is described in detail in Appendix A.1. Briefly, we focus on workers classified by IMSS as permanent, who earn a positive wage, in an establishment for which we know municipality and industry, and who are age 14-64. If wages for more than one job are reported for a particular worker, we select only the highest-wage job. Although in principle we observe a wage for every individual for every day of every year, we select the wage on a single day of each year, Sept. 30. After cleaning, we observe between 5.26 and 10.15 million individuals per year over the period. Summary statistics for the cleaned IMSS data are reported in Appendix Table A.2.

The plant-level data are from the *Encuesta Industrial Anual (EIA)* [Annual Industrial Survey], conducted by the *Instituto Nacional de Estadísticas y Geografía (INEGI)*, the Mexican statistical agency.[28] The variables are standard for plant surveys: employment, total wage bill, total hours

[26] Prior to 1993, the top-code was 10 times the minimum wage in Mexico City; in 1994, it was 18 times; and since 1995 it has been 25 times the minimum wage in Mexico City.

[27] Our procedure for dealing with the top- and bottom-coding differs from that of CHK, who impute wages above the top-codes in their data using a series of Tobit models. Given the large changes in the top-codes over time, we worry that such an imputation would run the risk of introducing significant errors in our setting. Given also that our aim is not to characterize firms' contributions to overall wage inequality, the advantages of such imputation seem small relative to the risk of errors they might introduce.

[28] The name of INEGI's main plant panel has changed over time; for recent years, it is referred to as the *Encuesta Anual de la Industria Manufacturera (EAIM)* [Annual Survey of Manufacturing Industry].

worked, investment, capital stock, domestic and export sales, among others. The sampling design is less standard. INEGI has periodically drawn a deterministic sample of plants from a subset of manufacturing industries and followed those plants over time, with no refreshing of the sample. Plants with more than 100 workers are included with certainty in the initial sample. INEGI has created separate panels for 1984-1994, 1993-2003, 2003-2009, and 2009-present.[29] Although it is possible to link some plants across panels, the number that can be linked both across panels and to the IMSS data is prohibitively small for our purposes. We therefore focus on the 1993-2003 panel. During this period, the EIA did not include information on *maquiladora* plants, assembly-for-export plants located mainly along the U.S. border, which were covered by a different dataset. In our cleaning procedure, we do imputations for missing information following the procedure described in Appendix II of Verhoogen (2008). We then require plants to have complete information at the plant level on employment (hours and number employed), hourly wage, total sales, export share, and capital-labor ratio. After cleaning (also following Appendix II of Verhoogen (2008)), there are 3,529 plants in a balanced panel with complete EIA information in every year over the 1993-2003 period. We refer to this balanced panel as the "EIA panel." Summary statistics by export status for the EIA panel for 1993 are in Appendix Table A.3. The differences between exporters and non-exporters are similar to those documented for the U.S. by Bernard and Jensen (1999), and subsequently for many other countries: exporters are larger, more capital-intensive, and higher-wage than non-exporters, and they make up a minority of plants in each industry.

The EIA data have been linked to the IMSS employer-employee data using establishment name, location (municipality and state) and street address. Although in principle all plants appearing in the EIA should also appear in the IMSS data, it has only been possible to link approximately 2,800 of the 3,529 plants from the EIA panel to the IMSS data in each year.

As noted above, we estimate wage premia using the largest connected set in each period. In our setting, the largest connected set does not include as large a share of establishments as has typically been the case in developed countries. Table 1 reports summary statistics from the IMSS data for the largest connected sets in each period — period 1 (1992-1994), period 2 (1996-1998), and period 3 (2000-2002), as well as period 0 (1988-1990) which will be used in a robustness check below. Comparing to Table A.2, we see that the largest connected sets typically include roughly

[29] INEGI has also conducted a monthly version of the survey, formerly called the *Encuesta Industrial Mensual (EIM)* [Monthly Industrial Survey] and now called the *Encuesta Mensual de la Industria Manufacturera (EMIM)* [Monthly Survey of Manufacturing Industry], with slightly different periodization. Wage bill and hours for two occupational categories, white-collar (*empleados*) and blue-collar (*obreros*), are available from this companion survey.

half the number of establishments in the full IMSS dataset in each year. This is in part because of the high rate of informality in Mexico and in part because we focus on three-year periods, in contrast to AKM (and Abowd et al. (2002)), who focus on a single 12-year period, and CHK, who focus on 7-year periods; both factors tend to reduce the number of observable connections and hence the size of the largest connected sets. In addition, the establishment size distribution in Mexico is skewed to the left relative to, for instance, the U.S.;[30] this in itself tends to reduce the share of establishments in the largest connected set. But it does not appear that the limited coverage of the largest connected sets is a severe problem for our purposes, for two reasons. First, although they cover a minority of establishments, the largest connected sets nonetheless cover a large majority of workers — approximately 90%. Second, and perhaps more importantly, the largest connected sets include almost all of the larger plants in manufacturing that appear in the EIA panel, for which we observe exporting behavior. Appendix Table A.4 reports the number of EIA panel plants that can be linked to the IMSS data, by connected set status, for each of our three main periods. Of the EIA plants that can be linked to the IMSS data, 98% or more are in the largest connected set in each year.

Once the EIA panel has been linked to the IMSS data, we impose the requirement that plants be in the largest connected set in periods 1-3. The resulting panel contains 2,625 plants. We refer to this balanced panel as the "EIA-IMSS panel." Table 2 reports summary statistics for the EIA-IMSS panel for 1993. Comparing to Appendix Table A.3, this panel is on average quite similar to the EIA panel. As mentioned above, once the plant and person components have been calculated in the IMSS data, we collapse the EIA-IMSS panel to the plant-period level, averaging within plant-period.[31]

It is important to recognize that our estimates of the effect of exporting on wage premia are valid conditional on firms being in the EIA-IMSS panel, in which large firms are over-represented. Moreover, the wage premia estimates themselves are identified by workers who switch between formal plants (i.e. plants that are registered and report regularly to the social security agency); these are likely to be workers with more stable attachment to the formal sector. At the same time, larger manufacturing plants and workers with strong attachment to the formal sector are those for whom exporting is a realistic possibility. In this sense, the EIA-IMSS panel is a relevant

[30]See e.g. Figure XIV of Hsieh and Klenow (2014), which plots the establishment size distribution for Mexico, the U.S., and India.

[31]For the plant and person components, we average over 1992-1994, 1996-1998, and 2000-2002; for the EIA variables, we average over 1993-1994, 1996-1998, and 2000-2002, since the EIA data we use only begin in 1993.

sample for investigating the effect of exporting on wages. We believe that the results are likely to generalize to larger manfacturing plants in other countries, especially those approximately at Mexico's level of industrial development.

As discussed in Verhoogen (2008), the devaluation of the peso in December 1994 represented an enormous shock to the Mexican economy. The Mexican peso lost approximately 50% of its nominal value in a few days. Figure 2 plots the real exchange rate over the 1989-2004 period. GDP fell by 6.7% from 1994 to 1995. Exports rose sharply, with approximately 85% destined for the U.S. market. Using the EIA panel of 3,529 plants, Figure 3 illustrates the shift toward the export market: the export share for the panel as a whole jumped sharply, and the number of plants with positive exports rose from approximately 30% to 45% of the sample.[32] Nominal wages, perhaps surprisingly, remained nearly constant on average through the crisis; the sudden increase in the price of imports, and the inflation that it generated, meant that real wages fell dramatically. As noted above, the wage gains among exporting plants relative to other plants in the same industry should be understood in this context of large real wage declines overall; real wages in exporting plants fell by less than in non-exporting plants.

4 Estimating Wage Premia

As described in Section 2 above, the first step of our approach is to estimate the AKM-type model (1) in the IMSS individual-level data, separately for periods 1992-1994, 1996-1998, and 2000-2002. Before presenting the estimates, we provide evidence for the validity of the conditional random mobility assumption, (2). Following CHK, we show simple event-study plots of average wage changes between firms in different quartiles of the distribution of wages paid to co-workers. For these plots, to ensure that we can observe two years before and two years after a job transition (again following CHK), we focus on four-year periods, 1992-1995 and 2000-2003. (The requirement that we observe two years before and two years after means that we must focus on transitions between the middle two years of each period, i.e. 1993-1994 or 2001-2002.) Figures 4 and 5 show the mean real wages of movers for these periods; to reduce visual clutter, we focus on workers

[32] The peso crisis was a much larger shock than the North American Free Trade Agreement (NAFTA), which took effect in January 1994. Mexico's main trade liberalization came with its entrance into the General Agreement on Tariffs and Trade in the mid-1980s, and by 1994 the vast majority of Mexican imports were covered by tariffs of 20% or less. Average U.S. tariffs on goods from Mexico were on the order of 3-5%. In the majority of cases, NAFTA phased out existing tariffs slowly over time. Relative to the exchange-rate devaluation, the year-by-year tariff changes were quite small.

leaving firms in quartiles 1 and 4. As in CHK, two features stand out. First, the wage trends before and after job switches are parallel across the different types of transitions; there are no Ashenfelter (1978)-type dips or rises before transitions. If the individual time-varying productivity shocks, ε_{it}, were determining job transitions, we would expect positive wage changes prior to movement to a higher-quartile plant, and negative changes prior to movement to a lower-quartile plant. Second, the wage change for workers moving from one quartile to another has approximately the same magnitude and opposite sign of the change for workers moving in the opposite direction.[33] If the ε_{it} contained an employer-employee match-specific effect that also affected workers' mobility, we would expect the gain for workers moving to a higher quartile to be larger than the loss for those making the opposite move.[34] Both features suggest that the data are well described by a model with a fixed individual component, a fixed plant component, time-varying observables, and a random shock uncorrelated with mobility.

We now turn to estimation of the AKM-type model, (1). We estimate plant and person components for all individuals in the largest connected sets described by Table 1, separately by period, for periods 1-3. Table 3 reports key statistics. There are several points to notice. First, despite being parsimonious, the model has a high in-sample predictive power, with adjusted R^2 above 90%, as in other studies in the literature. Second, we see already that the standard deviation of the individuals effects remains relatively constant but the standard deviation of the plant component increases over time, foreshadowing our results below that the plant-level response to the export shock is explained primarily by changes in wage premia. Third, there is little evidence of negative correlation between estimated person and plant effects.

One might be concerned that the use of three-year periods and the high rate of informality in our context might limit the number of observed plant-to-plant switches and exacerbate what Abowd et al. (2004) call "limited mobility bias." Small numbers of switchers can generate a negative bias in the correlation between individual and plant effects (Andrews et al., 2008) which can be substantial (Maré and Hyslop, 2006; Andrews et al., 2012). Intuitively, if plant effects are calculated first, and individual effects are calculated by substracting the plant effects and contributions of observables from log wages (as is common), then over-estimation of plant effects

[33]Appendix Figures A.1 and A.2 make the symmetry clear visually. Appendix Figures A.3 and A.4 are similar to figures 4 and 5, but use nominal wages instead of real wages. The patterns are similar.

[34]As Card et al. (2018) point out, models with a worker-firm match component tend to predict that workers who move will see wage increases, irrespective of whether they move to higher or lower quartiles.

will tend to lead to under-estimation of individual effects and vice-versa.[35] The smaller is the number of switchers, the less precisely the plant effects are estimated and the more severe is the bias. Andrews et al. (2008) suggest a correction, but (as CHK and Card et al. (2018) note) the correction requires strong assumptions on the covariance structure of the time-varying error terms. Following CHK, we instead rely on the assumption that the limited mobility bias is constant across periods and will be differenced out when we look at within-firm changes in the plant and average person components. The fact that the correlation between plant and individual effects is close to zero in our data (Table 3), suggests that the bias, while it may exist, does not appear to be severe.

5 Estimating the Effect of Exporting on Wage Premia

This section examines the relationship between exporting and the estimated plant and average person components. Before turning to regression tables, we provide a visual illustration of the key reduced-form relationships. Using the EIA-IMSS panel, Figures 6a-c present cross-sectional non-parametric regressions for outcome variables from the EIA — export share, log capital-labor ratio, and log plant-average hourly wage — against log plant size for period 1 (1992-1994). Because of the potential spurious correlation due to measurement error discussed in Section 2 above, we use log domestic sales as the capability proxy when export share is the y-axis variable (Figure 6a) and log domestic sales as the proxy when log capital-labor ratio and log hourly wage are the y-axis variables (Figures 6b-c). All variables have been deviated from industry means. All three graphs display a clear positive relationship: within industries, larger plants are more likely to export, are more capital intensive, and pay higher wages. The graphs in the second row, Figures 6d-f, plot non-parametric regressions of *changes* in the same variables between two consecutive periods against log employment in the first of those periods (i.e. the x-axis variable is log plant size in period 1 (1992-1994) for the period 2 vs. period 1 curve, and log plant size in period 2 (1996-1998) for the period 3 vs. period 2 curve). Given that all variables have been deviated from industry means, the key features are the relative slopes of the curves. We see that larger plants saw a greater increase in export share, capital intensity, and average hourly wages than smaller plants from period 1 (1992-1994) to period 2 (1996-1998), and these differential changes were greater

[35] In our case, the Matlab PCG algorithm solves simultaneously for the plant effects and the individual effects for switchers, and the individual effects for non-switchers are then calculated using the estimated plant effects.

than the corresponding changes from period 2 (1996-1998) to period 3 (2000-2002). The graphs show that the peso crisis affected plants differently within industries, with more positive (or less negative) effects on larger plants. These figures are similar to Figures A4-A5 of Verhoogen (2008), which used a broader EIA panel without linking to the IMSS data.

Turning to outcomes from the IMSS employer-employee data, Figures 7a-c presents non-parametric regressions similar to those in Figures 6a-f, with average log daily wage from IMSS, the plant component, and the average person component as y-axis variables and log employment (hours) from the EIA as the x-axis variable.[36] In the first row (Figures 7a-c), we see that all three variables bear a positive relationship to log domestic sales. Recall that in expectation the plant component and average person component sum to the plant-average log wage (refer to equation (3)); the equality very nearly holds true for the sample analogues. The y-axes have the same scale in all graphs and the slope for the plant component is approximately twice the slope for the average person component. Intuitively, this suggests that approximately two thirds of the within-plant *cross-sectional* correlation between plant-average wages and plant size can be attributed to wage premia. Figures 7d-f plot regressions of changes against log hours, similarly to Figures 6d-f. The similarity of Figure 7d, which uses average log daily wage from the IMSS data, to Figure 6f, which uses log average hourly wage from the EIA data, suggests that there was not much differential change in hours worked in response to the devaluation. Figures 7e and 7f illustrate the main point of the paper: the lion's share of the difference in differential changes in wages between periods 1-2 and 2-3 can be explained by the difference in differential changes in wage premia. There is little difference in differential changes in the average person component, our measure of skill composition.

We now turn to the estimation of our parametric IV model, equation (4). Table 4 presents our baseline results, with the first stage in Panel A Column 1, the reduced-form results in Panel A Columns 2-6, and the IV results in Panel B. As discussed in Section 2 above, we use log employment (hours) as the capability proxy in our baseline IV estimates.[37] In the first stage (Panel A, Column 1), the coefficient on the excluded instrument, initial log employment × devaluation dummy (written as $\widehat{\lambda}_{p-1} * T_2$ above), is significant at the 1% level. The coefficient indicates that comparing two plants that differ in initial employment by a factor of 2.7 (100 log points), the larger

[36]The spurious correlation issue is less of a concern here, because the IMSS-derived wage variables are not calculated by dividing by the EIA employment variable.

[37]The spurious-correlation concern discussed in Section 2 above still applies for the capital-labor ratio and log average hourly wage from the EIA. We present similar results using alternative proxies below.

plant experienced approximately a 1% greater increase in export share due to the devaluation, above and beyond the trend of .3% greater increase between periods. This is a modest but highly statistically significant difference. Consistent with Figures 6-7, the reduced form results show that the relationships between the changes in the capital-labor ratio, log average hourly wage, average log daily wage, and the plant component are steeper for periods 1-2 than for periods 2-3, and are statistically highly significant. There is no significant difference in these slopes for the person component. The IV estimates in Panel B indicate a positive effect of exporting on the capital-labor ratio, log average hourly wage, average log daily wage, and the plant component, but not on the person component. The coefficient in Column 4, for the change in average log daily wage, indicates that an .01 (1%) increase in export share generates a 3 log point (3.07%) increase in the average log daily wage. This response is large but not implausible.

Comparing the estimates in Columns 4 and 5, we see that the effect on the plant component is not statistically different from that of the effect on plant-level average wages. The point estimate for the plant component (Column 5) is larger than the coefficient for the plant-level average wage (Column 4) in this specification, but this difference will not persist across specifications below. The robust fact is that the plant component effect is as large as the effect on plant-level average wages. Since the plant and person components sum to the plant-level average in expectation, and very nearly in our sample, the coefficients in Columns 5 and 6 very nearly sum to the Column 4 coefficient, and the point estimate for the person component (Column 6) is negative in this specification, and will be robustly close to zero across specifications. Our interpretation of the IV results is that essentially all of the within-plant effect of exporting on plant-level average wages driven by the peso devaluation can be explained by the effect on wage premia, at least over the time frame we are able to study.

Weakness of our instrument is a potential concern. The robust first-stage F-statistic, reported in Column 1 of Panel A, is 7.320, which is below the Staiger and Stock (1997) rule-of-thumb level of 10. But two arguments suggest that the weak-instrument concern is not severe in this context. First, in the presence of weak instruments, we would expect the IV estimates to be biased toward the OLS estimates. (See e.g. Angrist and Pischke (2009, Section 4.6.4).) Appendix Table A.5 reports the OLS estimates, and reveals that the OLS estimates are smaller that the IV estimates, suggesting that weakness of the instrument would bias the IV estimates down in our case, and make it more difficult to reject the null of no effect of exporting on our outcomes.[38] Second, in this just-identified context with a single endogenous regressor and a single instrument, the Anderson and Rubin (1949) Wald test, which is robust to weak instruments, is equivalent to a test that the coefficient on the instrument is zero when we include it in place of the endogenous covariate in the second stage. We effectively do that in the reduced-form specifications of Columns 2-6, and t-tests comfortably reject the hypothesis that the coefficients on the instrument are zero in Columns 2-5.

To examine robustness of these results, we present results using three alternative proxies: log domestic sales, log TFP, and predicted export share. To estimate log TFP, we use a version of the Levinsohn and Petrin (2003) methodology, with value-added as the outcome, employment (white collar and blue collar separately) and log capital as covariates, and log materials and log electricity as proxies, separately by 2-digit sector. We predict export share using a tobit regression of export share on log employment (hours), log total sales, and log capital-labor ratio. Coefficient estimates from the TFP and predicted export share estimation are reported in Appendix Table A.6. In Melitz (2003)-type theoretical frameworks, all three proxies are expected to be positively correlated to underlying plant capability. Unsuprisingly, the proxies are highly (but not perfectly) correlated with one another (Appendix Table A.7). Table 5 reports cross-sectional regressions of the outcome variables from Figures 6-7 and Table 4 against these three proxies, as well as log employment (hours) for comparison. The results are almost entirely consistent with the positive relationships we observed in Figures 6a-c and 7a-c; the one exception is the relationship between export share and domestic sales, which is subject to the spurious negative bias (in levels) due to measurement error discussed in Section 2 above.

Table 6 presents first-stage, reduced-form, and IV results using the three alternative proxies, similar to Columns 1, 4, 5, and 6 of Table 4. The results are very consistent with the baseline results. The first-stage coefficient remains highly significant, and the IV estimates are statistically indistinguishable from those of Table 4. The reduced-form estimates naturally vary are of different magnitudes, because of the different scaling of the proxies, but the qualitative patterns are similar.

[38] Panel A of Appendix Table A.5 reports estimates omitting the initial level of the capability proxy (log employment). The coefficients for the changes in capital-labor ratio, plant-average wages, and the plant component are positive but not significant. Once the initial level of the capability proxy is included, in Panel B, the estimates are even smaller. Comparing the Panel B results to the IV results in Table 4, we see that the OLS coefficients are significantly smaller than the IV estimates. This is consistent with the hypothesis that plants are subject to idiosyncratic labor- (or capital-) supply shocks that generate a negative correlation between exporting and wages (or capital-intensity), generating a negative bias in the OLS coefficients.

These results reinforce the finding that essentially all of the differential increase in wages due to the peso devaluation can be attributed to changes in the plant component (i.e. wage premia), rather than sorting on skill. It is reassuring that the key results do not depend on the choice of capability proxy.

The share of the plant-average wage effect that can be explained by the effect on wage premia is large, indeed larger than we originally expected, and merits further discussion. Before the devaluation, we observe a positive, significant relationship between plant size (a strong predictor of export status) and the average person component in cross-section, which presumably reflects a long-run equilibrium. The contrast with the zero differential response of the average person component in response to the devaluation suggests that there may be a difference between the short- to medium-term and long-term responses to an exogenous increase in exporting. As discussed in the introduction, in models with search frictions in the labor market, an expansion of exporting may raise the value to firms of avoiding vacancies, and may lead firms to raise wage premia to retain existing workers and/or to attract new ones. These wage premia are then expected to fall over time, as firms replace existing workers with higher-skill workers from outside the firm or simply as the value of avoiding vacancies falls (Coşar et al., 2016; Fajgelbaum, 2016; Bellon, 2017). There is suggestive evidence of this dynamic in Table 4: the coefficients on initial log employment for the average person component in Column 6, Panels A and B (which are identified by differential trends between smaller and larger plants in the post-devaluation periods (between period 2 (1996-1998) and period 3 (2000-2002)) are positive and significant, while the corresponding coefficients for the plant component are close to zero, suggesting that the person component catches up in the post-devaluation period. However, this finding is not robust to using log domestic sales or predicted export share as alternative capability proxies in Table 6. It seems difficult to draw strong conclusions about the overshooting phenomenon in this setting.

As a further robustness check, we consider a longer panel and include an additional period before the devaluation. Due to the data constraints discussed in Section 3, we cannot use the EIA plant-level data for this exercise, and hence we cannot estimate the IV model with export share as an endogenous covariate. However, using only the IMSS employer-employee data, we are able to estimate the corresponding reduced form over the longer period. We define period 0 to be 1988-1990 and estimate the AKM-type model (1) for the largest connected set for this period. We then select establishments in the EIA-IMSS panel which which AKM estimates are also available in period 0. There are 2,316 plants in this balanced panel, observed over four periods. We use log employment (headcount) as the capability proxy, since we do not observe hours in the IMSS data. Panel A of Table 7 reports cross-sectional correlations similar to Table 5 for this modified sample and capability proxy; the results are similar to those in Table 5. To show the time pattern across periods, Panel B of Table 7 reports regressions period-by-period, of changes in the dependent variables on the level of the capability proxy separately by period. As discussed in footnote 19 above, the difference in coefficients between Panels B.2 and B.3 is precisely the estimate we would obtain of the coeffcient φ in the reduced-form model, (5).[39] The advantage of viewing the results period-by-period is that we can also compare the coefficient in Panel B.2 (for the change from period 1 (1992-1994) to period 2 (1996-1998)) to the coefficient in Panel B.1 (for the change from period 0 (1988-1990) to period 1 (1992-1994)). Panel C reports the differences in coefficients and their standard errors. Overall, the table is consistent with our story. For plant-average log wages and the plant component, the coefficients for the peso-crisis period, β_2, are significantly larger than the ones for the earlier period without a devaluation, β_1, as well as the ones for the later period, β_3. For the person component, the coefficient for the earlier periods, β_1, is marginally significant, but of a similar magnitude to the other periods (and not statistically distinguishable from the baseline estimate in Table 4).

As a final robustness check, we consider the effect of the export shock on the wages for stayers, employees continuously employed in a given plant for consecutive periods. Table 8 reports first-stage, reduced-form, and IV results similar to Columns 1 and 4 of Table 4, but where the change in the average log daily wage of stayers is the outcome of interest. As mentioned in Section 2 above, this is analogous to estimating a model with job-spell fixed effects, as in several existing papers.[40] The estimates are identified by variation in wages within firm-worker matches in response to changes in export share at the plant level. Despite the very different approach, the reduced-form and IV coefficients in Column 2 are very similar to the baseline estimates for changes in wage premia in Table 4, Column 5. The coefficient from this table falls within the range of the point estimates for the four capability proxies from Table 4, Column 5 and Table 6 Column 3. It is consistent with our expectations, and reassuring, that the estimates for stayers are similar to the estimates for switchers between plants.

[39] The coefficients in Panels B.1-B.3 correspond to estimates of β_p in equation (6) in footnote 19.

[40] This specification is not exactly equivalent to an individual-year-level specification with job-spell effects, because we have collapsed the data at the period level, and define as a stayer any worker who appears in a given plant in at least one year of two consecutive periods.

6 Conclusion

This paper has investigated the role of exporting in shaping Mexican plants' wage policies – in particular the payment of wage premia, defined as wages above what individuals would earn elsewhere on the labor market. Following CHK, we have estimated simple AKM-type models separately by three-year periods, to decompose plant-average wages into a plant component (i.e. wage premium) and an average person component. Following Verhoogen (2008), we have argued that the peso devaluation of late 1994 generated a greater effective export inducement to larger than to smaller plants within each industry. Using this differential shock as a source of exogenous variation in exports, we find that exporting led plants to increase wage premia.[41] The increase in the plant component can explain essentially all of the differential increase in plant-average wages arising from the export shock over the period we study. The findings highlight the causal importance of trade shocks — and product-market shocks more generally — in determining wage premia at the firm level.

Two questions seem particularly worthy of further research. First, how persistent are the effects of exporting on wage premia? We have found that the effects persist over 6-8 years, but it remains an open question whether firms will gradually reduce premia over a longer period, as several recent models with search frictions suggest. This question awaits an empirical setting with a longer data series following the trade shock.

Second, to what extent did the differential increase in wage premia in response to the peso devaluation contribute to aggregate wage inequality in Mexico? Our paper suggests a mixed response to this question. On one hand, the export shock led initially larger plants within each industry, which already paid higher wages and wage premia, to increase wages and wage premia further. In this sense, the effect of exports on wage premia tended to increase inequality. On the other hand, the fact that we find limited effects on skill composition suggests that the effect of the export shock on the general-equilibrium return to skill is likely to be limited. The within-industry disequalizing effect of the trade shock may be confined to manufacturing (and possibly other tradable) sectors. A fuller accounting of the role of firms' wage policies in overall inequality in Mexico is a topic for future work.

[41] For readers who prefer the reduced-form specification over the IV, for the reasons discussed in Section 2 above, the results can be interpreted as showing that the differential inducement to export created by the devaluation increased wage premia.

References

Abowd, John and Thomas Lemieux, "The Effects of Product Market Competition on Collective Bargaining Agreements: The Case of Foreign Competition in Canada," *Quarterly Journal of Economics*, Nov. 1993, *108* (4), 983–1014.

_ , Francis Kramarz, and David Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, 1999, *67* (2), 251–333.

_ , _ , Paul Lengermann, and Sébastien Pérez-Duarte, "Are Good Workers Employed By Good Firms? A Test of a Simple Assortative Matching Model for France and the United States," Jan. 2004. Unpub. paper, Cornell.

_ , Robert Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," 2002. LEHD Technical Paper 2002-06.

Aghion, Philippe, Ufuk Akcigit, Ari Hyytinen, and Otto Toivanen, "On the Returns to Invention within Firms: Evidence from Finland," *American Economic Review Papers and Proceedings*, 2018, *108* (2), 208–212.

Akerlof, George and Janet Yellen, "The Fair Wage-Effort Hypothesis and Unemployment," *Quarterly Journal of Economics*, 1990, *105* (2), 255–83.

Alvarez, Jorge, Felipe Benguria, Niklas Engbom, and Christian Moser, "Firms and the Decline in Earnings Inequality in Brazil," *American Economic Journal: Macroeconomics*, 2018, *10* (1), 149–89.

Álvarez, Roberto and Ricardo A. López, "Skill Upgrading and the Real Exchange Rate," *The World Economy*, 2009, *32* (8), 1165–1179.

Amiti, Mary and Donald R. Davis, "Trade, Firms, and Wages: Theory and Evidence," *Review of Economic Studies*, 2012, *79* (1), 1–36.

Anderson, T. W. and Herman Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 1949, *20* (1), 46–63.

Andrews, Martyn J., Leonard Gill, Thorsten Schank, and Richard Upward, "High Wage Workers and Low Wage Firms: Negative Assortative Matching or Limited Mobility Bias?," *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 2008, *171* (3), 673 – 697.

Andrews, Martyn J, Leonard Gill, Thorsten Schank, and Richard Upward, "High Wage Workers Match with High Wage Firms: Clear Evidence of the Effects of Limited Mobility Bias," *Economics Letters*, 2012, *117* (3), 824–827.

Angrist, Joshua D. and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Araújo, Bruno César and Lourenço S. Paz, "The Effects of Exporting on Wages: An Evaluation Using the 1999 Brazilian Exchange Rate Devaluation," *Journal of Development Economics*, 2014, *111*, 1–16.

Ashenfelter, Orley, "Estimating the Effect of Training Programs on Earnings," *Review of Economics and Statistics*, 1978, *60* (1), 47–57.

Barth, Erling, Alex Bryson, James C. Davis, and Richard B. Freeman, "It's Where You Work: Increases in the Dispersion of Earnings Across Establishments and Individuals in the United States," *Journal of Labor Economics*, 2016, *34* (S2), S67–S97.

_ , James C. Davis, and Richard B. Freeman, "Augmenting the Human Capital Earnings Equation with Measures of Where People Work," *Journal of Labor Economics*, 2018, *36* (S1), S71–S97.

Baumgarten, Daniel, "Exporters and the Rise in Wage Inequality Evidence from German Linked Employer-Employee Data," *Journal of International Economics*, 2013, *90* (1), 201 – 217.

Bellon, Matthieu, "Trade Liberalization and Inequality: A Dynamic Model with Firm and Worker Heterogeneity," 2017. Unpub. paper, Columbia University, Dec.

Bernard, Andrew B. and J. Bradford Jensen, "Exporters, Jobs, and Wages in U.S. Manufacturing: 1976-1987," *Brookings Papers on Economic Activity: Microeconomics*, 1995, pp. 67–112.

_ and _ , "Exceptional Exporter Performance: Cause, Effect, or Both?," *Journal of International Eco-*

nomics, Feb. 1999, *47*, 1–25.

Bertrand, Marianne and Sendhil Mullainathan, "Enjoying the Quiet Life? Corporate Governance and Managerial Preferences," *Journal of Political Economy*, 2003, *111* (5), 1043–1075.

Brambilla, Irene, Daniel Lederman, and Guido Porto, "Exports, Export Destinations and Skills," *American Economic Review*, 2012, *102* (7), 3406–3488.

_ , Nicolas Depetris Chauvin, and Guido Porto, "Examining the Export Wage Premium in Developing Countries," *Review of International Economics*, 2017, *25* (3), 447–475.

Bustos, Paula, "The Impact of Trade Liberalization on Skill Upgrading: Evidence from Argentina," July 2011. Unpub. Paper, Universitat Pompeu Fabra.

Card, David, Ana Rute Cardoso, and Patrick Kline, "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women," *Quarterly Journal of Economics*, 2015, *131* (2), 633–686.

_ , _ , Jörg Heining, and Patrick Kline, "Firms and Labor Market Inequality: Evidence and Some Theory," *Journal of Labor Economics*, 2018, *36* (S1), S13–S70.

_ , Jörg Heining, and Patrick Kline, "Workplace Heterogeneity and the Rise of West German Wage Inequality," *Quarterly Journal of Economics*, 2013, *128* (3), 967–1015.

Castellanos, Sara, Rodrigo Garcia-Verdu, and David Kaplan, "Wage Rigidities in Mexico: Evidence from Social Security Records," *Journal of Development Economics*, 2004, *75* (2), 507–33.

Coşar, A Kerem, Nezih Guner, and James Tybout, "Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy," *American Economic Review*, 2016, *106* (3), 625–663.

Davidson, Carl, Fredrik Heyman, Steven Matusz, Fredrik Sjöholm, and Susan Chun Zhu, "Globalization and Imperfect Labor Market Sorting," *Journal of International Economics*, 2014, *94* (2), 177 – 194.

_ , Steven J. Matusz, and Andrei Shevchenko, "Globalization and Firm Level Adjustment with Imperfect Labor Markets," *Journal of International Economics*, 2008, *75* (2), p295 – 309.

Dickens, William and Lawrence Katz, "Inter-Industry Wage Differences and Industry Characteristics," in Kevin Lang and Jonathan Leonard, eds., *Unemployment and the Structure of Labor Markets*, New York, NY: Blackwell, 1987.

Egger, Hartmut and Udo Kreickemeier, "Firm Heterogeneity and the Labour Market Effects of Trade Liberalisation," *International Economic Review*, 2009, *50* (1), 187–216.

_ and _ , "Fairness, Trade, and Inequality," *Journal of International Economics*, 2012, *86* (2), 184–196.

Fajgelbaum, Pablo D, "Labor Market Frictions, Firm Growth, and International Trade," 2016. Unpub. paper, UCLA.

Felbermayr, Gabriel, Julien Prat, and Hans-Jörg Schmerer, "Globalization and Labor Market Outcomes: Wage Bargaining, Search Frictions, and Firm Heterogeneity," *Journal of Economic Theory*, 2011, *146* (1), 39 – 73.

Frías, Judith A., David S. Kaplan, and Eric A. Verhoogen, "Exports and Wage Premia: Evidence from Mexican Employer-Employee Data," 2009. Unpub. paper, Columbia University.

_ , _ , and Eric Verhoogen, "Exports and Within-Plant Wage Distributions: Evidence from Mexico," *American Economic Review Papers and Proceedings*, 2012, *102* (3), 435–440.

Helpman, Elhanan, Oleg Itskhoki, and Stephen Redding, "Inequality and Unemployment in a Global Economy," *Econometrica*, July 2010, *78* (4), 1239–1283.

_ , _ , Marc-Andreas Muendler, and Stephen J Redding, "Trade and Inequality: From Theory to Estimation," *The Review of Economic Studies*, 2017, *84* (1), 357–405.

Holtz-Eakin, Douglas, Whitney Newey, and Harvey S. Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, Nov. 1988, *56* (6), 1371–1395.

Hsieh, Chang-Tai and Peter J. Klenow, "The Life Cycle of Plants in India and Mexico," *Quarterly Journal of Economics*, 2014, *129* (3), 1035–1084.

Hummels, David, Rasmus Jørgensen, Jakob R. Munch, and Chong Xiang, "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data," *American Economic Review*, 2014, *104* (6), 1597–1629.

Irarrazabal, Alfonso, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Heterogeneous Firms or Heterogeneous Workers? Implications for Exporter Premiums and the Gains from Trade," *Review of Economics and Statistics*, 2013, *95* (3), 839–849.

Kandilov, Ivan, "Do Exporters Pay Higher Wages? Plant-Level Evidence from an Export Refund Policy in Chile," *World Bank Economic Review*, 2009, *23* (2), 269–294.

Kaplan, David S., Gabriel Martinez Gonzalez, and Raymond Robertson, "Worker- and Job-Flows in Mexico," 2004. Unpub. paper, ITAM.

_ , _ , and _ , "What Happens to Wages After Displacement?," *Economía*, 2005, *5* (2), 197–242.

Katz, Lawrence F. and Lawrence H. Summers, "Industry Rents: Evidence and Implications," *Brookings Papers on Economic Activity*, 1989, pp. 209–275.

Klein, Michael W, Christoph Moser, and Dieter M. Urban, "Exporting, Skills and Wage Inequality," *Labour Economics*, 2013, *25*, 76–85.

Kline, Patrick, Neviana Petkova, Heidi Williams, and Owen Zidar, "Who Profits from Patents? Rent-Sharing at Innovative Firms," 2017. Unpub. paper, UC Berkeley.

Krishna, Pravin, Jennifer P. Poole, and Mine Zeynep Senses, "Wage effects of trade reform with endogenous worker mobility," *Journal of International Economics*, 2014, *93* (2), 239–252.

Krueger, Alan and Lawrence Summers, "Efficiency Wages and the Inter-Industry Wage Structure," *Econometrica*, March 1988, *56* (2), 259–93.

Kumler, Todd, Eric Verhoogen, and Judith A. Frías, "Enlisting Workers in Improving Payroll-Tax Compliance: Evidence from Mexico," 2015. NBER working paper no. 19385, updated April 2015.

Lazear, Edward P. and Kathryn L. Shaw, "Introduction: Firms and the Distribution of Income: The Roles of Productivity and Luck," *Journal of Labor Economics*, 2018, *36* (S1), S1–S12.

Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, April 2003, *70*, 317–341.

Macis, Mario and Fabiano Schivardi, "Exports and Wages: Rent Sharing, Workforce Composition, or Returns to Skills?," *Journal of Labor Economics*, 2016, *34* (4), 945–978.

Maré, David C. and Dean R. Hyslop, "Worker-Firm Heterogeneity and Matching: An Analysis Using Worker and Firm Fixed Effects Estimated from LEED," 2006. Unpub. paper.

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

Munch, Jakob Roland and Jan Rose Skaksen, "Human Capital and Wages in Exporting Firms," *Journal of International Economics*, 2008, *75* (2), 363 – 372.

Schank, Thorsten, Claus Schnabel, and Joachim Wagner, "Do Exporters Really Pay Higher Wages? First Evidence from German Linked Employer-Employee Data," *Journal of International Economics*, 2007, *72*, 52–74.

Schneider, Friedrich and Dominik H. Enste, "Shadow Economies: Size, Causes, and Consequences," *Journal of Economic Literature*, 03// 2000, *38* (1), 77–114.

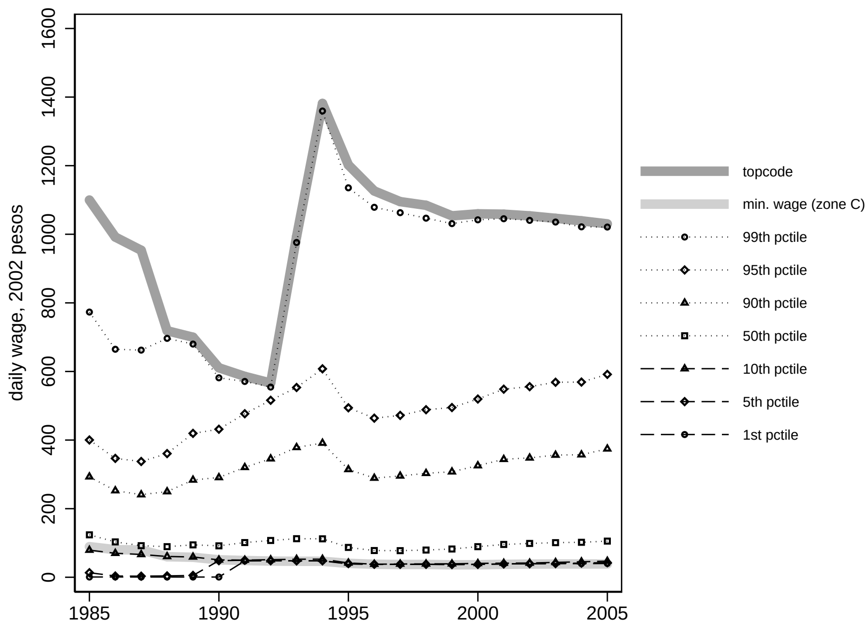
Song, Jae, David J. Price, Fatih Guvenen, Nicholas Bloom, and Till von Wachter, "Firming Up Inequality," 2016. Unpub. paper, Stanford University, Oct.

Staiger, Douglas and James H. Stock, "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 1997, *65* (3), 557–586.

Van Reenen, John, "The Creation and Capture of Rents: Wages and Innovation in a Panel of U.K. Companies," *Quarterly Journal of Economics*, 1996, *111* (1), 195–226.

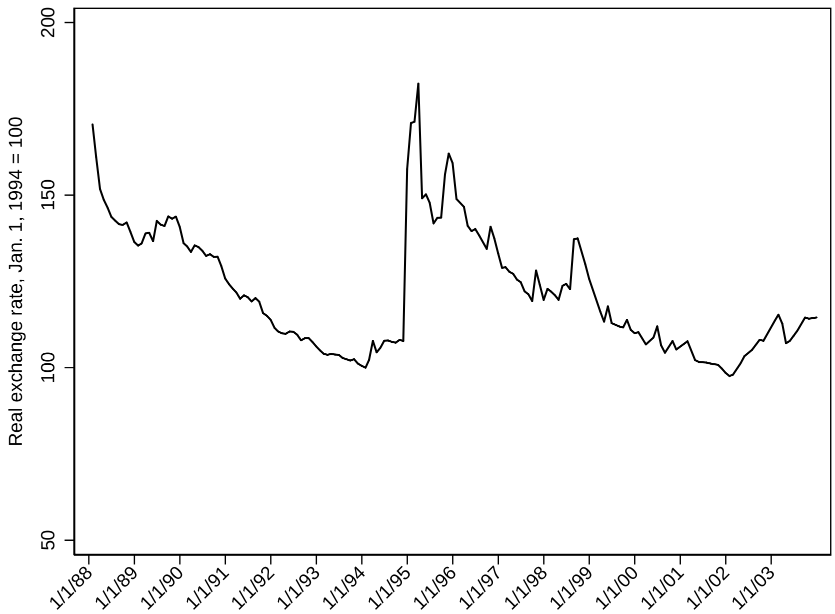
Verhoogen, Eric, "Trade, Quality Upgrading, and Wage Inequality in the Mexican Manufacturing Sector," *Quarterly Journal of Economics*, 2008, *123* (2), 489–530.

Figure 1. Wage percentiles, top- and bottom-codes, IMSS data, 1985-2005



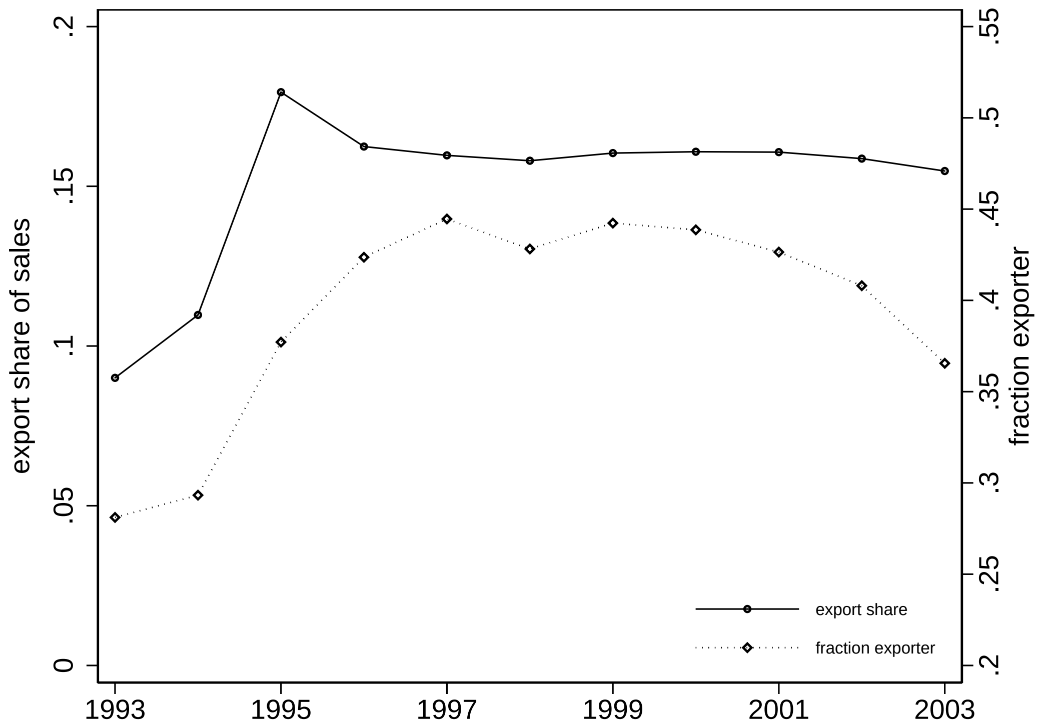
Notes: Wage percentiles calculated from raw IMSS data, after very basic cleaning (steps 1-3 of cleaning procedure described in Appendix A.1). There are three minimum wages in Mexico, corresponding to different geographic regions (zones A, B, C). Displayed is the minimum wage for Zone C, the lowest of the three. The top-code was 10 times the minimum wage in Mexico City (Zone A) from 1985-1993, 18 times in 1994, and 25 times from 1995-2005. Prior to 1991, establishments were allowed to report wages below the corresponding minimum wage to IMSS. Beginning in 1991, this practice was disallowed. Average 2002 exchange rate: 9.60 pesos/US$1.

Figure 2. Real exchange rate



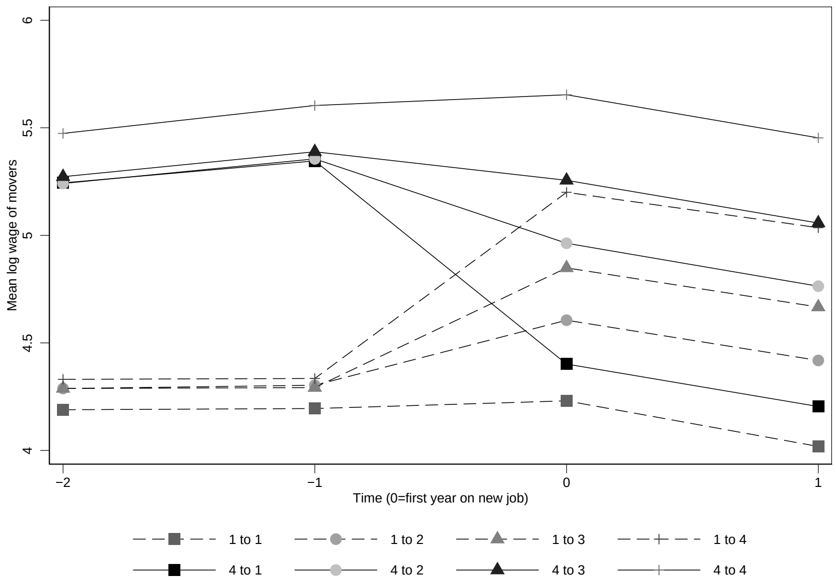
Notes: Real exchange rate calculated as $e \times$CPI(US)/CPI(Mexico), where e is peso/US$ nominal exchange rate. Data from IMF International Financial Statistics.

Figure 3. Shift to export market, EIA panel



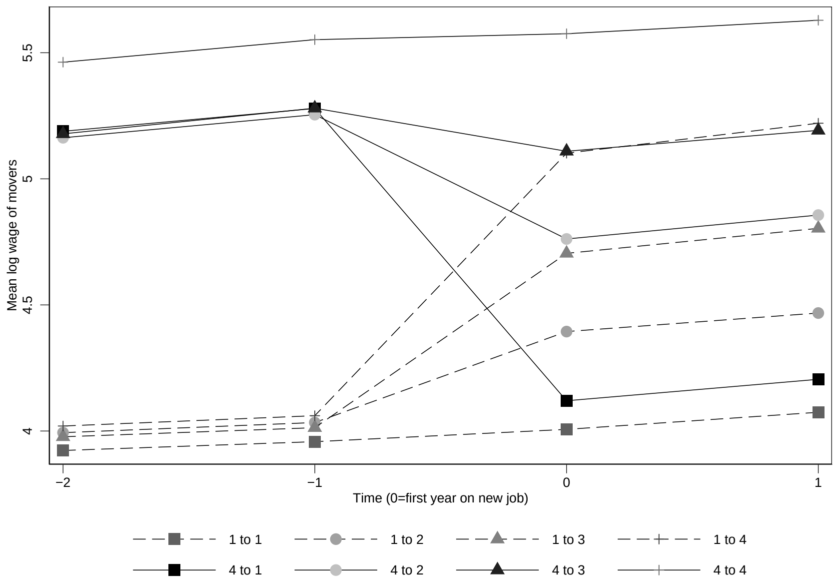
Notes: Data from 1993-2003 EIA (balanced) panel of 3,529 plants. Export percentage of sales calculated as (total exports for all plants)/(total sales for all plants). Plants with exports greater than zero classified as exporters. See data appendix for details.

Figure 4. Movers' mean real wages, IMSS data, 1992-1995



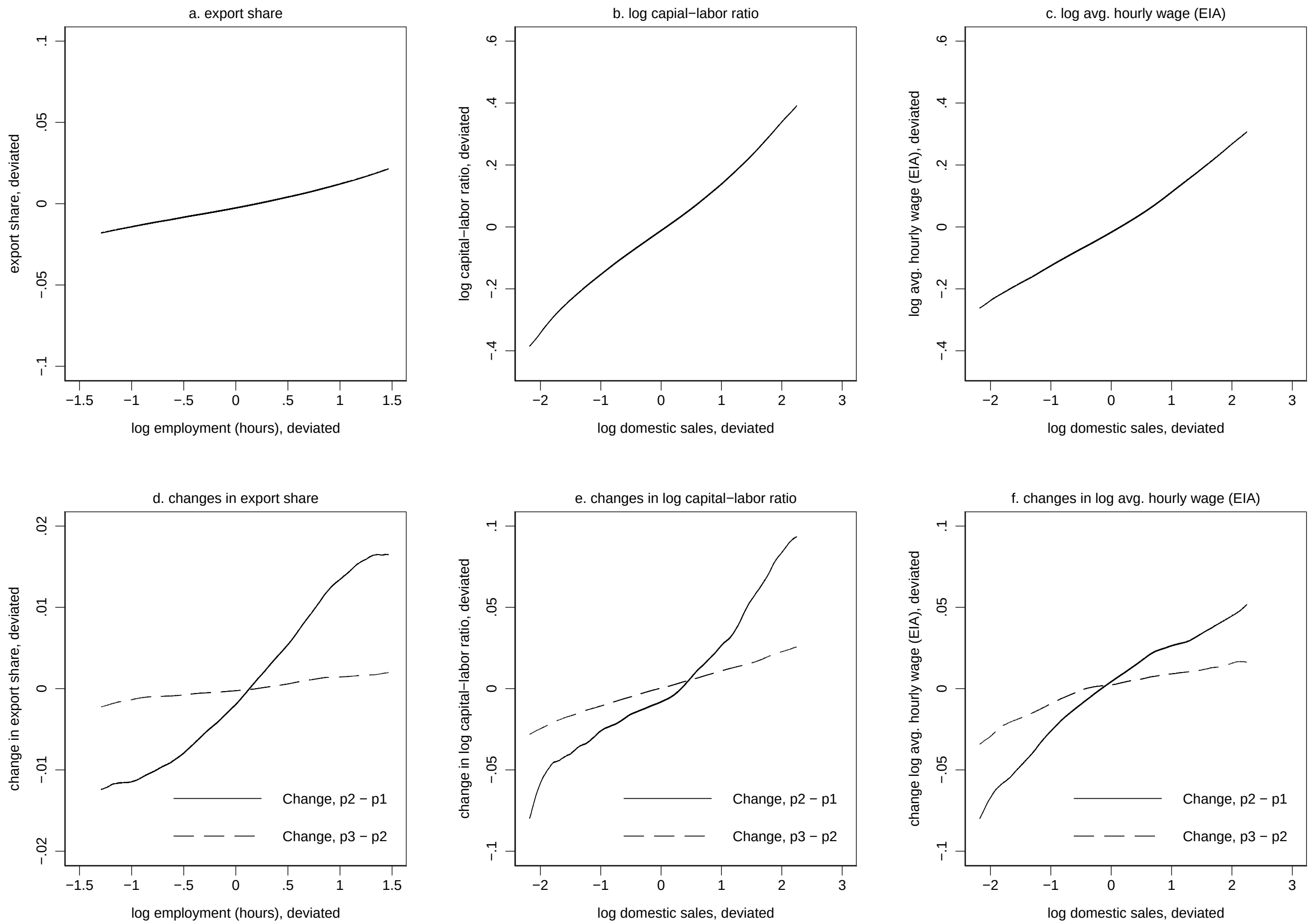
Notes: Sample is all workers observed in 1992-1995 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 1993 and 1994 and held both the preceding and new job for at least two years. Each line corresponds to a transition between types of firms classified by quartiles of the average coworkers' wage.

Figure 5. Movers' mean real wages, IMSS data, 2000-2003



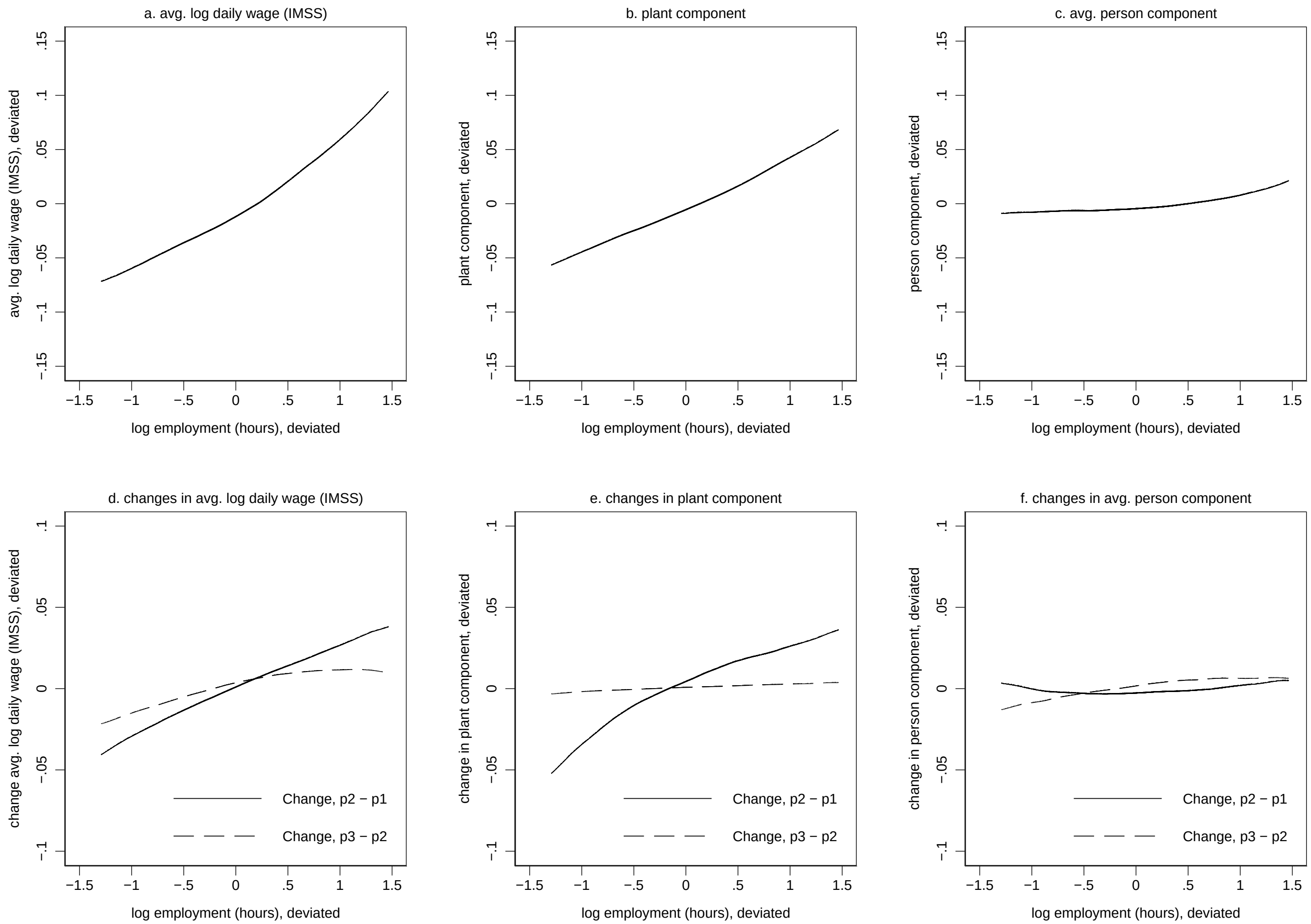
Notes: Sample is all workers observed in 2000-2003 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 2000 and 2001 and held both the preceding and new job for at least two years. Each line corresponds to a transition between types of firms classified by quartiles of the average coworkers' wage.

Figure 6. Non-parametric regressions, variables from EIA plant data



Notes: Non-parametric regressions are estimated using Stata lpoly command, Epanechnikov kernel, and the default rule-of-thumb bandwidth. Figures 6a-c plot non-parametric regressions in levels for period 1 (1992-1994). Figures 6d-f plot non-parametric regressions of changes between two consecutive periods against log employment in the earlier period (i.e. the x-axis variable is log plant size in period 1 (1992-1994) for the "Change, p2 - p1" curve, and log plant size in 1996-1998 for the "Change, p3 - p2" curve. Sample is EIA-IMSS panel (with variables averaged by plant-period). Variables are from EIA, and are deviated from their industry-period means. Log avg. hourly wage calculated as (total wage bill/total hours worked). Further data details are in Section 3 and Appendix A.

Figure 7. Non-parametric regressions, wage components from IMSS employer-employee data



Notes: Non-parametric regressions are estimated using Stata lpoly command, Epanechnikov kernel, and the default rule-of-thumb bandwidth. Figures 7a-c plot non-parametric regressions in levels for period 1 (1992-1994). Figures 7d-f plot non-parametric regressions of changes between two consecutive periods against log employment in the earlier period (i.e. the x-axis variable is log plant size in period 1 (1992-1994) for the "Change, p2 - p1" curve, and log plant size in 1996-1998 for the "Change, p3 - p2" curve. Sample is EIA-IMSS panel (with variables averaged by plant-period). Plant component is $\hat{\psi}_{J(i,t)}$ and average person component is the plant-level average of $\hat{s}_{it} = \hat{\alpha}_i + X_{it}'\hat{\beta}$ from estimation of equation (1) in IMSS data for corresponding period. Further data details are in Section 3 and Appendix A.

Table 1. Summary statistics, IMSS individual-level data, largest connected sets

year	# individuals	# establishments	avg. age	fraction male	avg. daily wage (raw, 2002 pesos)		avg. daily wage (winsorized, 2002 pesos)	
					mean	std. dev.	mean	std .dev.
1988	4,813,280	219,015	31.44	0.72	152.90	727.68	119.24	63.05
1989	5,474,761	235,765	30.97	0.70	158.50	640.48	130.00	74.59
1990	6,202,204	253,573	30.76	0.69	152.04	427.17	128.59	79.09
1992	7,007,431	298,743	30.81	0.68	169.47	271.40	149.05	95.53
1993	6,929,394	301,639	31.11	0.68	190.29	254.97	159.40	106.09
1994	7,109,067	302,213	31.29	0.68	201.12	266.48	162.78	110.33
1996	7,260,643	296,641	31.58	0.67	147.44	193.03	116.93	81.88
1997	7,816,910	309,762	31.54	0.67	148.37	201.05	118.07	84.10
1998	8,205,521	316,307	31.66	0.67	150.85	182.16	120.80	86.50
2000	9,396,001	360,718	31.93	0.65	160.49	185.94	130.33	92.57
2001	9,261,411	374,332	32.47	0.65	169.14	191.87	138.23	97.93
2002	9,303,348	380,401	32.84	0.65	171.96	193.72	140.95	99.00

Notes: Sample is from IMSS employer-employee records after cleaning steps 1-7 in Appendix A.1. (Sample includes individuals who earn a positive wage in an establishment for which municipality and industry are observed, are age 14-64, and are employed in an establishment in the largest connected graph of establishments in their respective period (among 1988-1990, 1992-1994, 1996-1998 and 2000-2002). Winsorization is at 10th and 90th percentiles. Wages are reported both in "raw" (i.e. pre-winsorized) and winsorized form. See Section 3 and Appendix A.1 for further details. Average 2002 exchange rate: 9.60 pesos/US$1.

Table 2. Summary statistics, EIA-IMSS plant panel, 1993

	non-exporters (1)	exporters (2)	all plants (3)
Total revenue	153.10	417.40	232.02
	(9.55)	(58.12)	(18.74)
Employment	188.32	333.43	231.66
	(5.55)	(13.40)	(5.73)
K/L	140.14	177.64	151.34
	(5.35)	(9.19)	(4.66)
Export share of sales		0.14	0.04
		(0.01)	(0.00)
Avg. hourly wage (EIA)	42.75	58.14	47.35
	(0.55)	(1.15)	(0.53)
N	1841	784	2625

Notes: Table reports statistics using 1993 data for the EIA-IMSS (balanced) panel of plants. Standard errors of means in parentheses. Exporter defined as export sales > 0. Export share is fraction of total sales derived from exports. Sales are measured in millions of 2002 Mexican pesos, capital-labor ratio in thousands of 2002 pesos, and average daily wage in 2002 pesos. Average 2002 exchange rate: 9.60 pesos/US$1. For further details, refer to Section 3 and Appendix A.2.

Table 3. Estimation results for AKM-type model, per period

	(1) Period 1 1992-1994	(2) Period 2 1996-1998	(3) Period 3 2000-2002
Number of individuals	10,121,284	11,155,022	13,137,161
Number of plants	394,672	402,213	479,088
Summary of parameter estimates			
Std. dev. individual effect	0.451	0.441	0.445
Std. dev. plant effect	0.362	0.390	0.423
Std. dev. Xb	0.144	0.196	0.164
Corr. individual/plant effects	0.007	0.060	0.054
Corr. individual effect/Xb	0.145	0.019	0.068
Corr. plant effect/Xb	0.116	0.112	0.092
RMSE of AKM residual	0.184	0.189	0.189
Adjusted R-squared	0.916	0.919	0.923
Additional statistics			
Std. dev. of log wages	0.637	0.665	0.682
N	21,045,892	23,283,074	27,960,760

Notes: The table shows statistics from the estimation of model (1) for periods 1992-1994, 1996-1998 and 2000-2002 separately. Sample is described in notes to Table 1. Individual effects are estimates of α_i, plant effects are estimates of $\psi_{J(i,t)}$, and Xb are estimates of $X_{it}'\beta$ in (1). Covariates included in X_{it} are age squared, age cubed (both recentered at 40), tenure, tenure squared, and year effects.

Table 4. Baseline estimates, exports and plant-level outcomes

	(1) △ export share	(2) △ log K/L	(3) △ log avg. hourly wage (EIA)	(4) △ avg. log daily wage (IMSS)	(5) △ plant component	(6) △ avg. person component
A. First stage and reduced form						
init. log employ. × devaluation	0.009***	0.040**	0.045***	0.029***	0.040***	-0.011
	(0.003)	(0.018)	(0.009)	(0.006)	(0.008)	(0.007)
init. log employ.	0.004*	0.053***	0.024***	0.012***	0.002	0.010**
	(0.004)	(0.053)	(0.024)	(0.012)	(0.002)	(0.010)
F-stat	7.320					
B. IV						
△ export share		4.449*	5.041**	3.172**	4.425**	-1.237
		(2.495)	(2.008)	(1.280)	(1.755)	(0.852)
init. log employ.		0.037*	0.006	0.001	-0.014	0.015**
		(0.021)	(0.016)	(0.010)	(0.014)	(0.007)
6-digit industry × period effects	Y	Y	Y	Y	Y	Y
region (state) × period effects	Y	Y	Y	Y	Y	Y
N (plants)	2625	2625	2625	2625	2625	2625
N (obs)	5250	5250	5250	5250	5250	5250

Notes: The dependent variables (at top) are changes between periods (after averaging within periods). The excluded instrument is the interaction between initial log employment (i.e. in the earlier period) and an indicator for the devaluation that equals 1 for change from period 1 to period 2, and 0 otherwise. Panel A reports the first stage in column (1) and reduced-form results in columns (2)-(6). Panel B reports the corresponding IV regressions. Sample is EIA-IMSS panel. Export share is fraction of total sales derived from exports. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table 5. Cross-sectional correlations, alternative proxies, period 1 (1992-1994)

Capability proxy	(1) export share	(2) log K/L	(3) log avg. hourly wage (EIA)	(4) avg. log daily wage (IMSS)	(5) plant component	(6) person component
log employment (hours)	0.029***	0.182***	0.126***	0.095***	0.074***	0.020***
	(0.003)	(0.026)	(0.011)	(0.007)	(0.006)	(0.004)
log domestic sales	0.004	0.364***	0.190***	0.119***	0.082***	0.037***
	(0.003)	(0.020)	(0.007)	(0.005)	(0.004)	(0.003)
log TFP (Levinsohn-Petrin)	0.020***	0.097***	0.184***	0.111***	0.076***	0.035***
	(0.003)	(0.029)	(0.011)	(0.007)	(0.005)	(0.004)
predicted export share	0.274***	6.150***	2.005***	1.257***	0.893***	0.360***
	(0.029)	(0.211)	(0.084)	(0.056)	(0.044)	(0.037)
6-digit industry effects	Y	Y	Y	Y	Y	Y
region (state) effects	Y	Y	Y	Y	Y	Y

Notes: Each panel reports six regressions, all corresponding to period 1 (1992-1994). Log TFP estimated using Levinsohn and Petrin (2003) method, with value-added as outcome, employment (white collar and blue collar separately) and log capital as covariates, and log materials and log electricity as proxies, separately by 2-digit sector. Predicted export share extimated from tobit regression of export share on log sales, log employment and log capital-labor ratio. Sample is EIA-IMSS panel. All regressions have N=2,625 except those with log TFP proxy, which have N=2,588 (because observations with non-positive value-added have been dropped and only plants with TFP estimates in all periods included). See Appendix Table A.6 for coefficient estimates from estimation of log TFP and predicted export share. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table 6. First stage, reduced-form, and IV estimates, alternative proxies

	(1) △ export share	(2) △ avg. log daily wage	(3) △ plant comp.	(4) △ avg. person comp.
A. Log domestic sales as proxy				
1. First stage and reduced form				
init. log domestic sales × devaluation	0.008***	0.032***	0.030***	0.002
	(0.003)	(0.004)	(0.006)	(0.005)
init. log domestic sales	0.006***	0.006***	0.004	0.002
	(0.006)	(0.006)	(0.004)	(0.002)
2. IV				
△ export share		3.773***	3.554***	0.247
		(1.266)	(1.269)	(0.585)
init. log domestic sales		-0.016	-0.017	0.001
		(0.012)	(0.012)	(0.006)
B. Log TFP (Levinsohn-Petrin) as proxy				
1. First stage and reduced form				
init. log TFP (Levinsohn-Petrin) × devaluation	0.009***	0.028***	0.035***	-0.007
	(0.003)	(0.005)	(0.007)	(0.006)
init. log TFP (Levinsohn-Petrin)	0.001	0.009***	-0.001	0.010**
	(0.001)	(0.009)	(-0.001)	(0.010)
2. IV				
△ export share		2.989***	3.740***	-0.731
		(1.067)	(1.377)	(0.678)
init. log TFP (Levinsohn-Petrin)		0.007	-0.002	0.010**
		(0.007)	(0.009)	(0.005)
C. Predicted export share as proxy				
1. First stage and reduced form				
init. Predicted export share × devaluation	0.111***	0.333***	0.346***	-0.012
	(0.030)	(0.046)	(0.064)	(0.060)
init. Predicted export share	0.027	0.067**	0.020	0.050
	(0.027)	(0.067)	(0.020)	(0.050)
2. IV				
△ export share		3.013***	3.132***	-0.113
		(0.874)	(0.967)	(0.515)
init. Predicted export share		-0.015	-0.066	0.053
		(0.072)	(0.085)	(0.050)
6-digit industry × period effects	Y	Y	Y	Y
region (state) × period effects	Y	Y	Y	Y

Notes: Each panel is similar to Columns 1, 4, 5 and 6 of Table 4 for indicated alternative capability proxy. See notes to Table 5 for construction of log TFP and predicted export share. Sample is EIA-IMSS panel. Regressions in Panels A and C have 2,625 plants and 5,250 observations; Regressions in Panel B have 2,588 plants and 5,176 observations (because observations with negative value-added have been dropped and only plants with TFP estimates in all periods included). Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table 7. Differential effect of devaluation, IMSS 1988-2002 panel

	(1) avg. log daily wage (IMSS)	(2) plant component	(3) person component
A. Cross-sectional correlations, period 1			
log employment, period 1	0.091***	0.077***	0.013***
	(0.007)	(0.005)	(0.004)
	△ avg. log daily wage (IMSS)	△ plant component	△ person component
B.1 Outcomes: change from period 0 to 1			
log employment, period 0 (β_0)	0.026***	0.017***	0.009*
	(0.004)	(0.005)	(0.005)
B.2 Outcomes: change from period 1 to 2			
log employment, period 1 (β_1)	0.042***	0.037***	0.005
	(0.004)	(0.005)	(0.005)
B.3. Outcomes: change from period 2 to 3			
log employment, period 2 (β_2)	0.009***	0.003	0.006
	(0.003)	(0.005)	(0.005)
C. Differences in coefficients			
$\beta_1 - \beta_0$	0.017***	0.021***	-0.004
	(0.006)	(0.007)	(0.007)
$\beta_2 - \beta_1$	-0.033***	-0.034***	0.002
	(0.005)	(0.007)	(0.007)
6-digit industry effects	Y	Y	Y
region (state) effects	Y	Y	Y
N	2316	2316	2316

Notes: Sample is plants in EIA-IMSS panel that are also in largest connected set in period 0 (1988-1990). Variables are from IMSS data. Period 1 is 1992-1994, period 2 is 1996-1998, period 3 is 2000-2002. Panel A reports three separate cross-sectional regressions of the dependent variable at top on average log employment in period 1, analogous to Table 5. Panels B1-B3 report three separate regressions each, of the dependent variable at the top (i.e. change for indicated period) on average log employment in the previous period. See discussion in text (Section 5) for details. All regressions include industry and state effects and have sample size indicated at bottom. Panel C reports differences in coefficients from Panel B and the standard errors on the differences. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table 8. Differential effect of devaluation, IV estimates, stayers only

	(1) △ export share	(2) △ avg. log daily wage, stayers
A. First stage and reduced form		
initial log employ. × devaluation	0.009***	0.024***
	(0.003)	(0.005)
initial log employ.	0.004*	0.019***
	(0.004)	(0.019)
B. IV		
change in export share		2.703**
		(2.703)
initial log employ.		0.010
		(0.010)
6-digit industry × period effects	Y	Y
region (state) × Period effects	Y	Y
N	2623	2623

Notes: Table is similar to Columns 1 and 4 of Table 4, but here (in Column 2) the plant-level average wage has been computed using only workers who did not change jobs between periods (stayers). Sample is plants in EIA-IMSS panel for which a change in average log wage for stayers could be constructed both between periods 1 and 2 and between periods 2 and 3. (Of the 2,625 EIA-IMSS panel plants, two did not have stayers between one pair of these periods.) Export share is fraction of total sales derived from exports. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Exports and Wage Premia: Evidence from Mexican Employer-Employee Data*

Judith A. Frías

David S. Kaplan

Eric A. Verhoogen

David Alfaro-Serrano

Sept. 2018

Appendix

For Online Publication

A Data Appendix

A.1 IMSS individual-level data

All private Mexican employers are legally required to report wages for their employees to the Mexican social security agency, *Instituto Mexicano del Seguro Social (IMSS)*. Not all employers comply; those that do not are commonly considered to be in the informal sector. The raw IMSS data can thus be considered a census of private, formal-sector establishments and their workforces for 1985-2005. (Most public-sector workers and employees of the state-run oil company are covered by other insurance programs.)

The IMSS data contain information on the daily wage of individuals. The wages are a measure of total compensation, called the *salario base de cotización*, which includes both earnings and benefits, including payments made in cash, bonuses, commissions, room and board, overtime payments, and in-kind benefits. The data are reported as a sequence of spells for each worker, with beginning and end dates. In principle it is possible to recover a wage for every individual for every day of every year. We extracted data for September 30 for each year. At the level of individuals, the data also contain information on age, sex, and state and year of the individual's first registration with IMSS. At the establishment level, the data contain information only on location and industry (using the IMSS's own 4-digit industrial categories, of which there are 276.)

We impose the following criteria in cleaning the data. (1) In its internal records, IMSS classifies wage records by types referred to as *modalidades*. We use only *modalidades* corresponding to permanent workers and for which consistent, reliable wage figures are available.[1] (2) We require that an individual have a positive wage. (3) We require that municipality and industry are reported for establishment. (4) We winsorize wages within year, assigning wages above the 90th percentile to the 90th percentile and wages below the 10th percentile to the 10th percentile, for the reasons discussed in Section 3. (5) If wages for more than one establishment are observed simultaneously for a given individual, we keep only the highest-wage observation. (6) We require that individuals be 14 years or older and 64 years or younger. (7) We require that workers be employed in an establishment in the largest connected set of establishments, as described in Section 2 above.

The total number of workers with wage data in the "raw" IMSS files (i.e. the sample size after step 3 of the cleaning procedure described in the previous paragraph) ranges from approximately 4 million in 1985 to approximately 10 million in 2005. The numbers of individuals in the cleaned data, after step (6) above but before limiting to the largest connected sets, are in Appendix Table A.2. The numbers after limiting to the largest connected sets are in Table 1. Additional details on the IMSS data are available in Castellanos et al. (2004) and Kaplan et al. (2005, 2007).

A.2 EIA plant-level data

The cleaning procedure for the plant-level data from the *Encuesta Industrial Anual (EIA)* [Annual Industrial Survey] is the same as described in Appendix II of Verhoogen (2008), and rather than repeat the entire description we focus here on key points.

For the reasons discussed in Section 3, we focus in this paper on the EIA data from 1993-2003. The sample was drawn in 1993, to include the largest plants in 205 of the 309 6-digit industries (*clases*) in the Mexican industrial classification system, covering 85% of the value of production in each industry. These plants were followed over time, with minimal refreshing of the sample.

Capital stock was constructed using the perpetual-inventory method. Capital was classified into three types: machinery and equipment, land and buildings, and transportation equipment and other fixed assets. Following Olley and Pakes (1996), each type of capital was assumed to evolve according to $K_{jt} = (1 - \delta_j)K_{jt-1} + i_{jt-1}$, where j indexes the three types of capital. Following Levinsohn and Petrin (2003), the depreciation rates, δ_j for machinery and equipment, land and buildings, and transportation equipment

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

were assumed to be 10%, 5% and 20% respectively. Total capital stock is the sum of the three types of capital. The book value of capital stock in 1993 was taken as the initial value.

The following cleaning procedures were implemented. (1) Plants in multi-plant firms for which complete information was not reported separately by plant were dropped. (2) Plants owned in whole or in part by government entities were dropped. (3) Establishments that appeared to be *maquiladoras*, because they derived more than 95% of their income from exports or subcontracting, were dropped. (4) Variables that changed within a plant by more than a factor of 10 from one year to the next were set to missing. (5) Missing values of variables were imputed following the procedure described in Appendix II of Verhoogen (2008). (6) After imputation, plants with incomplete information on any key variable (employment, hours, wage bill, total costs, domestic sales, total sales, capital stock) were dropped. (7) The key variables listed in the previous point were "winsorized" at the 1st and 99th percentiles, following a suggestion Angrist and Krueger (1999).

We then selected a balanced panel of plants with complete data in all years 1993-2003, which we refer to as the EIA panel. 3,529 plants are included in this balanced panel. We then linked the EIA panel to the IMSS data and collapsed to the period level (period 1 is 1992-1994, period 2 is 1996-1998, period 3 is 2000-2002; see Section 3 for justification), averaging variables within period.[2] We then selected plants with estimated plant and average person components for all three periods. 2,625 plants satisfied this requirement. We refer to this balanced plant-period-level panel as the EIA-IMSS panel.

References

Angrist, Joshua D. and Alan B. Krueger, "Empirical Strategies in Labor Economics," in Orley C. Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3A, Elsevier Science, 1999.

Castellanos, Sara, Rodrigo Garcia-Verdu, and David Kaplan, "Wage Rigidities in Mexico: Evidence from Social Security Records," *Journal of Development Economics*, 2004, *75* (2), 507–33.

Kaplan, David S., Gabriel Martinez Gonzalez, and Raymond Robertson, "What Happens to Wages After Displacement?," *Economía*, 2005, *5* (2), 197–242.

_ , _ , and _ , "Mexican Employment Dynamics: Evidence from Matched Firm-Worker Data," 2007. World Bank Policy Research working paper No. 4433.

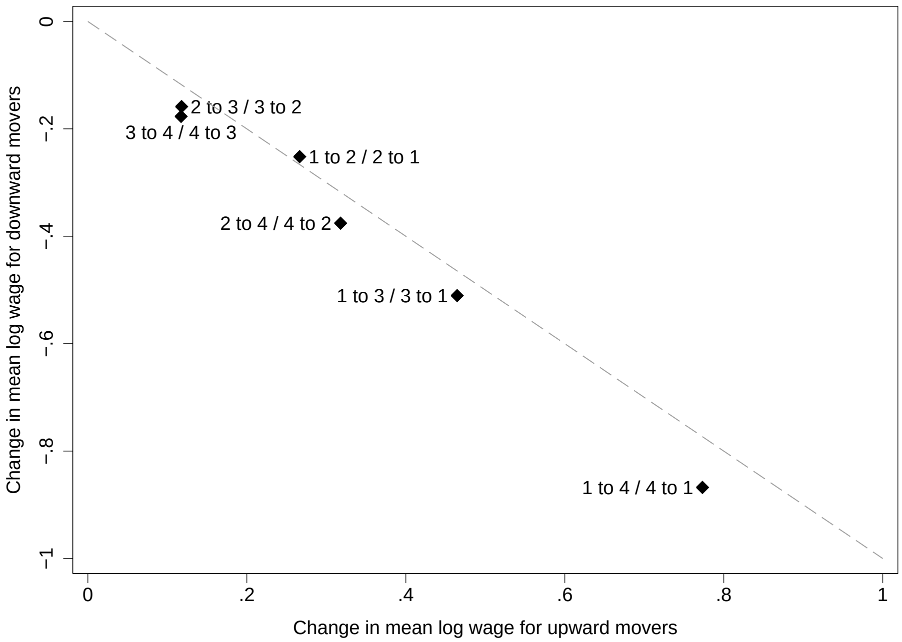
Levinsohn, James and Amil Petrin, "Estimating Production Functions Using Inputs to Control for Unobservables," *Review of Economic Studies*, April 2003, *70*, 317–341.

Olley, G. Steven and Ariel Pakes, "The Dynamics of Productivity in the Telecommunications Industry," *Econometrica*, 1996, *64* (6), 1263–1297.

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

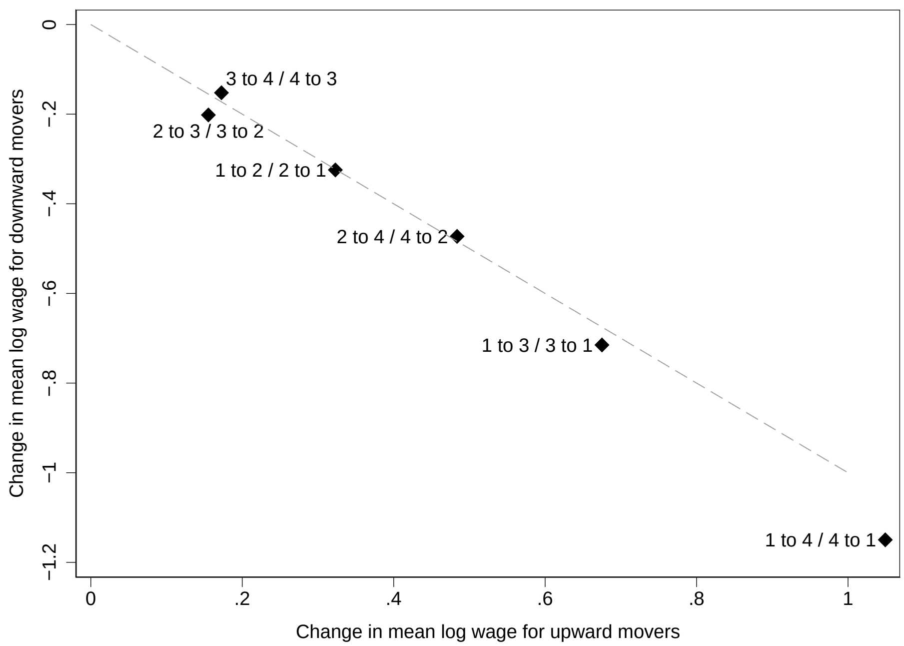
[2]For period 1, we averaged EIA variables for 1993-1994, since the EIA variables are not available in 1992.

Figure A.1. Comparing upward and downward moves, IMSS data, 1992-1995



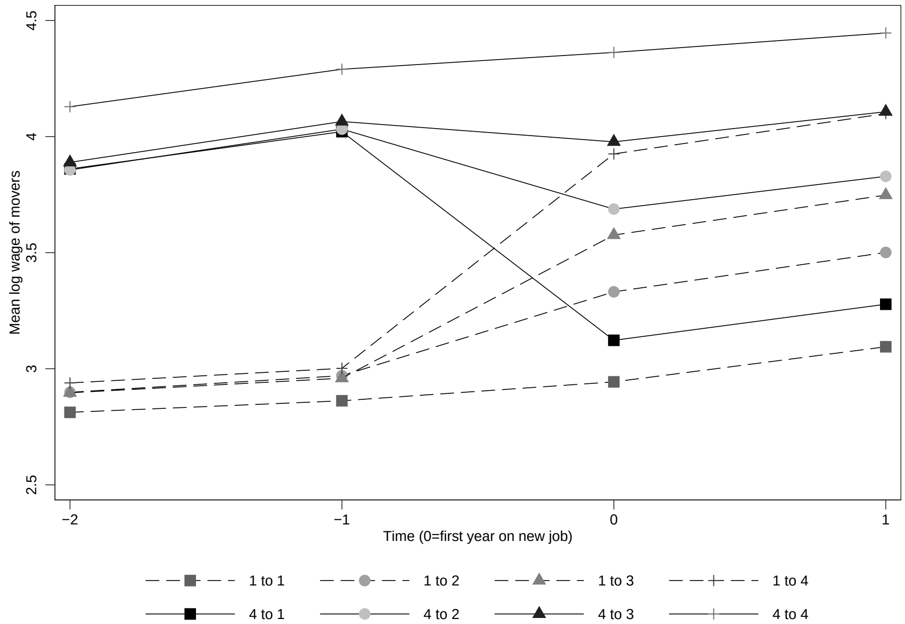
Notes: Sample is all workers observed in 1992-1995 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 1993 and 1994 and held both the preceding and new job for at least two years. The dashed line is at −45 degrees. Each dot plots upward and downward transitions between two types of firms, classified according to quartiles of average coworkers' wage. Wage changes are changes in log real wage, averaged over workers making same transition, between 1993 and 1994.

Figure A.2. Comparing upward and downward moves, IMSS data, 2000-2003



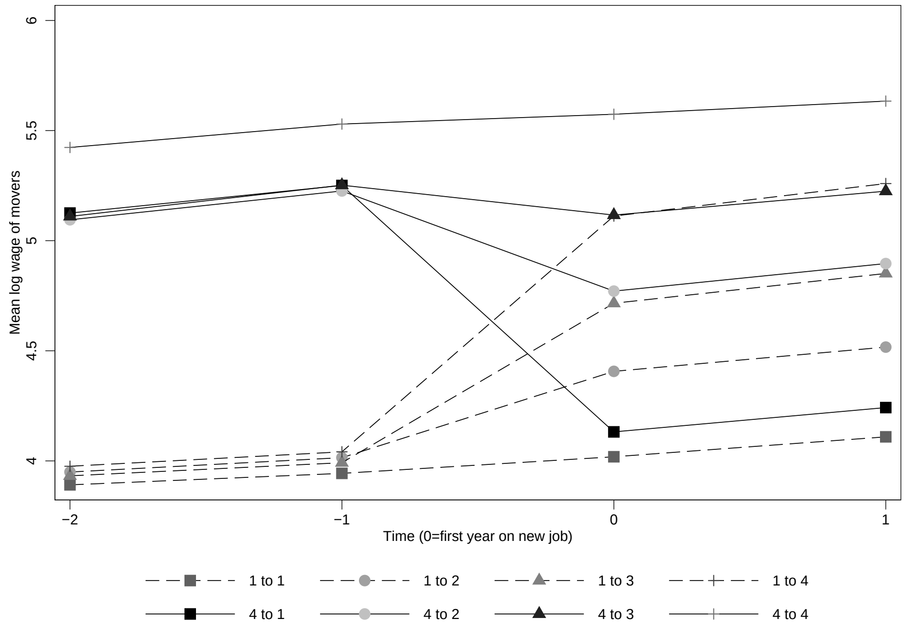
Notes: Sample is all workers observed in 2000-2003 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 2000 and 2001 and held both the preceding and new job for at least two years. The dashed line is at −45 degrees. Each dot plots upward and downward transitions between two types of firms, classified according to quartiles of average coworkers' wage. Wage changes are changes in log real wage, averaged over workers making same transition, between 2000 and 2001.

Figure A.3. Movers' mean nominal wages, IMSS data, 1992-1995



Notes: Figure is similar to Figure 4 but shows *nominal* wage changes. Sample is all workers observed in 2000-2003 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 2001 and 2002 and held both the preceding and new job for at least two years. Each line corresponds to a transition between types of firms classified by quartiles of the average coworkers' wage.

Figure A.4. Movers' mean nominal wages, IMSS data, 2000-2003



Notes: Figure is similar to Figure 5 but shows *nominal* wage changes. Sample is all workers observed in 2000-2003 in the IMSS database (after cleaning steps 1-6 described in Appendix A.1) who changed job between 2001 and 2002 and held both the preceding and new job for at least two years. Each line corresponds to a transition between types of firms classified by quartiles of the average coworkers' wage.

Table A.1. Aggregate labor force statistics

	1990	2000
Total population	81.25	97.48
Eonomically active pop. age > 14	31.23	40.16
Remunerated workers	25.96	32.01
Remunerated workers, private sector	21.27	27.20
Workers registered in IMSS	10.76	15.24
Workers registered in IMSS, permanent	9.53	13.53

Notes: Numbers in millions. Figures drawn from *Anuario Estadístico de los Estados Unidos Mexicanos* [Statistical Yearbook of Mexico], 2005, which draws in turn from the following: decennial population censuses (total population), 1991 *Encuesta Nacional de Empleo* [National Employment Survey] (economically active population age > 14), and INEGI Banco de Información Económica (remunerated employees), IMSS *Memoria Estadística*.

Table A.2. Summary statistics, IMSS individual-level data, before limiting to largest connected sets

					avg. daily wage (raw, 2002 pesos)		avg. daily wage (winsorized, 2002 pesos)	
year	# individuals	# establishments	avg. age	fraction male	mean	std. dev.	mean	std .dev.
1988	5,257,200	426,570	31.76	0.72	146.81	702.83	115.61	62.21
1989	5,993,961	469,018	31.29	0.70	151.86	618.46	125.50	73.70
1990	6,869,806	538,274	31.02	0.69	144.82	417.67	123.23	78.16
1991	7,546,628	596,124	31.01	0.68	153.03	235.67	134.68	86.85
1992	7,756,268	621,246	31.10	0.68	161.18	264.05	142.49	94.67
1993	7,659,363	615,684	31.41	0.68	180.30	249.97	152.09	105.07
1994	7,843,005	619,991	31.58	0.67	190.37	259.41	155.34	109.18
1995	7,413,728	600,015	32.01	0.67	152.28	202.47	122.94	87.77
1996	7,998,174	617,721	31.95	0.67	139.70	187.39	111.67	80.87
1997	8,592,365	640,381	31.92	0.67	140.61	194.91	112.75	83.02
1998	9,001,372	653,151	32.03	0.67	142.99	177.21	115.34	85.46
1999	9,578,857	674,710	32.17	0.66	145.03	176.33	117.56	86.67
2000	10,203,195	711,176	32.32	0.65	153.09	181.65	125.09	91.72
2001	10,103,668	736,849	32.85	0.65	160.87	187.25	132.29	97.03
2002	10,151,601	748,620	33.20	0.65	163.62	189.03	134.94	98.13

Notes: Sample is from IMSS employer-employee records after cleaning steps 1-6 in Appendix A.1 (before restricting to largest connected sets). Winsorization is at 10th and 90th percentiles. Wages are reported both in "raw" (i.e. pre-winsorized) and winsorized form. See Section 3 and Appendix A.1 for further details. Average 2002 exchange rate: 9.60 pesos/US$1.

Table A.3. Summary statistics, EIA panel, 1993

	non-exporters (1)	exporters (2)	all plants (3)
Total revenue	151.88	417.78	226.55
	(7.96)	(47.51)	(14.65)
Employment	184.56	370.31	236.78
	(5.66)	(21.78)	(7.48)
K/L	146.63	194.80	160.17
	(4.82)	(8.96)	(4.30)
Export share of sales		0.15	0.04
		(0.01)	(0.00)
Avg. hourly wage (EIA)	43.64	60.98	48.51
	(0.80)	(1.16)	(0.67)
N	2537	992	3529

Notes: Table reports statistics using 1993 data from EIA panel (before linking to IMSS data). Standard errors of means in parentheses. Exporter defined as export sales > 0. Export share is fraction of total sales derived from exports. Sales are measured in millions of 2002 Mexican pesos, capital-labor ratio in thousands of 2002 pesos, and average daily wage in 2002 pesos. Average 2002 exchange rate: 9.60 pesos/US$1. For further details, refer to Section 3 and Appendix A.2.

Table A.4. Number of EIA plants linked to IMSS data, by connected set status

	EIA panel plants	EIA panel plants linked to IMSS			EIA panel plants not linked to IMSS	EIA-IMSS panel plants
		Total	Connected	Not connected		
Period 1 (1992-1994)	3,529	2,769	2,746	23	760	2,625
Period 2 (1996-1998)	3,529	2,903	2,868	35	626	2,625
Period 3 (2000-2002)	3,529	2,872	2,812	60	657	2,625

Notes: Data from IMSS employer-employee records and EIA plant panel as described in Section 3. "Connected" means contained in the largest connected set, as described in Section 2.

Table A.5. OLS estimates, exports and plant-level outcomes

		(1) △ log K/L	(2) △ log avg. hourly wage (EIA)	(3) △ avg. log daily wage (IMSS)	(4) △ plant component	(5) △ avg. person component
A. Not including initial value of capability proxy						
△ export share		0.043	0.042	0.034	0.058**	-0.024
		(0.082)	(0.039)	(0.023)	(0.026)	(0.024)
B. Including initial value of capability proxy						
△ export share		0.001	0.016	0.019	0.046*	-0.028
		(0.082)	(0.038)	(0.023)	(0.026)	(0.024)
initial log employ.		0.072***	0.045***	0.026***	0.021***	0.005
		(0.009)	(0.004)	(0.003)	(0.004)	(0.003)
6-digit industry × period effects	Y	Y	Y	Y	Y	Y
region (state) × period effects	Y	Y	Y	Y	Y	Y
N (plants)		2625	2625	2625	2625	2625
N (obs)		5250	5250	5250	5250	5250

Notes: Table reports OLS regressions corresponding to equation (4) in the main text. Panel A omits the initial value of the capability proxy, $\widehat{\lambda}_{jp-1}$; Panel B includes it. Export share is fraction of total sales derived from exports. Robust standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A.6. Construction of alternative proxies, TFP and predicted export share

	(1) log VA ind. 31	(2) log VA ind. 32	(3) log VA ind. 33	(4) log VA ind. 34	(5) log VA ind. 35	(6) log VA ind. 36	(7) log VA ind. 37	(8) log VA ind. 38	(9) log VA ind. 39	(10) export share
log w.c. empl.	0.176***	0.122***	0.126***	0.365***	0.311***	0.073**	0.223***	0.156***	0.291**	
	(0.024)	(0.036)	(0.038)	(0.034)	(0.031)	(0.028)	(0.061)	(0.018)	(0.135)	
log b.c. empl.	0.259***	0.295***	0.510***	-0.041	0.038	0.261***	0.141	0.157***	0.014	
	(0.030)	(0.047)	(0.102)	(0.028)	(0.036)	(0.053)	(0.090)	(0.032)	(0.247)	
log capital	0.341***	0.054	0.128	0.145**	0.299***	0.514**	0.165	0.128*	0.247	
	(0.055)	(0.050)	(0.109)	(0.059)	(0.052)	(0.227)	(0.100)	(0.068)	(0.185)	
log empl.										0.053***
										(0.003)
log sales										0.040***
										(0.003)
log K/L										0.027***
										(0.002)
N (plants)	611	575	123	312	743	226	81	814	40	3529
N (obs)	6519	6117	1317	3364	8003	2407	870	8769	429	38819

Notes: Columns 1-9 report coefficients from Levinsohn and Petrin (2003) TFP estimation, with log value-added as outcome, log employment (white-collar and blue-collar separately) and log capital as covariates, and log materials and log electricity as proxies, separately by 2-digit industry. The industries (indicated at top) are: food, beverages, tobacco (31); textiles, apparel, leather goods (32); wood products, including wood furniture (33); paper, papers, products, publishing (34); chemical products (35); non-metallic mineral products (36); basic metal products (37); metal products, machinery, equipment (38); other manufacturing (39). Column 10 reports the coefficients of a tobit model of the export share using log employment, log sales, and log capital-labor ratio as covariates, and including 4-digit sector fixed effects. Observations with negative or zero value-added omitted in Columns 1-9. For all columns, if capital or employment variable has value zero, log is set to zero. Standard errors in parentheses. *10% level, **5% level, ***1% level.

Table A.7. Correlations between proxies, EIA-IMSS panel, pooling periods

	log empl. (hours)	log dom. sales	log TFP (L-P)	pred. exp. share
log employment (hours)	1.0000			
log domestic sales	0.8142	1.0000		
log TFP (Levinsohn-Petrin)	0.5578	0.6979	1.0000	
predicted export share	0.8893	0.9306	0.6247	1.0000

Notes: Table reports bilateral correlation coefficients using the EIA-IMSS panel, pooling periods 1 (1992-1994), 2 (1996-1998), and 3 (2000-2002).