Labour income inequality in Mexico

Puzzles solved and unsolved

Raymundo M. Campos-Vazquez[1] and Nora Lustig[2]

November 2017

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: We analyse the evolution and proximate determinants of labour income inequality in Mexico between 1989 and 2017. Labour income inequality increased between 1989 and 1994 and declined between 1994 and 2006. What happened after 2006 is subject to uncertainty. The national labour force survey shows a steady decline and the income expenditure survey suggests that inequality increased. We correct for high and rising item (labour income) non-response and under-representation of high-wage earners through a 'hot deck' imputation method and post-survey weight adjustments. We obtain the new weights for formal workers from tabulations recently released by the Mexican social security administration. For informal workers, we correct using the hot deck imputation method. With corrected data, inequality no longer declines between 2006 and 2017.

[1] El Colegio de México, corresponding author: rmcampos@colmex.mx; [2] Tulane University, New Orleans, USA.

1   Introduction

Mexico is an upper-middle-income country with high levels of labour income inequality: in 2014, the Gini coefficient was in the range of .45 to .52 (depending on the source). In the last thirty years, the country opened up to international trade (through both unilateral liberalization and free trade agreements) and its labour force became considerably more educated: the proportion of individuals with primary education or less was 67 per cent in 1990 and equalled 33 per cent in 2015; and the proportion of individuals with a college education more than doubled between 1990 and 2015, when it reached around 15 per cent. With the advent of digitalization, the last thirty years also witnessed sweeping changes in production technologies. In the light of these developments, how did labour income inequality change? We find that between the late 1980s and 1994, inequality increased. This increase was driven, primarily, by an increase in the skill premium (returns to education, in particular) associated with higher demand for skills and falling minimum wages and unionization rates. After NAFTA came into effect (1994) and up to the mid-2000s, labour income inequality declined. In stark contrast with the earlier period, the skill premium fell. Since institutional factors such as the minimum wage and the unionization rate remained constant, the fall in the skill premium was associated with the fact that the supply of skilled workers outpaced demand.

After 2006, the evolution of labour income inequality is far from clear. According to the Mexican labour force survey (herein ENOE, the Spanish acronym), labour income inequality continued its steady decline: the Gini coefficient went from .424 in 2006 to .382 in 2017. However, the Household Income Expenditure National Survey (herein ENIGH, the Spanish acronym) shows a slight increase: the Gini coefficient went from .511 in 2006 to .523 in 2014. Which of the two is correct? A serious problem with the labour force survey is that labour income (item) non-response (did not answer question on labour income) is high; and it steadily increased from 2006, reaching about a third of all workers (both formal and informal) by 2017. Using administrative data, we carry out a post-survey weight adjustment to equate the distribution of workers by category in the survey to that observed in the administrative source (assumed to be the 'true' distribution).[1] Since we do not have an external administrative source to generate 'true' weights, we estimate the missing labour incomes for *informal* workers by applying the so-called 'hot deck' imputation method, as proposed by Little and Rubin (2002). The corrected labour force survey data no longer show the sharp fall mentioned above, but they still show a slight decline. With the original survey, ENIGH shows an increase of about 1 Gini point and, with the corrected survey, the labour force survey data show a decline of about the same size. If one were to apply the same correction method to ENIGH, the increase in labour income inequality would be of 4 Gini points. In sum, the evolution of inequality in the period after 2006 remains a puzzle unsolved. The most one can say with certainty is that labour income inequality did not continue to decline.

Labour income inequality is affected by two main factors: the distribution of (observable and unobservable) characteristics of workers (education, experience, gender, etc.) and the returns to those characteristics. Workers' characteristics, in turn, are affected by 'fate' (gender, race, talent, and so on), households' decisions (e.g. to enrol or not in post-secondary education) and policy (e.g. expanding access to education). Returns to households' characteristics depend on market

---

[1] The theory behind this post-survey weight adjustment has a long tradition in statistics. For details see, for example, Biemer and Christ (2008).

forces (i.e. demand and supply of workers of different skills and experience) and institutional/policy factors (e.g. minimum wage policy and the unionization rate). In order to separate the contribution of characteristics and returns, we employ the 're-centered influence function' (RIF) procedure proposed by Firpo et al. (2009) to decompose effects into characteristics or composition and returns effects. Relative returns (also known as wage structure) are affected by demand and supply of workers of different skills and by institutional factors such as the minimum wage and unionization rate. In order to examine the effect of supply and demand on relative wages, we follow Bound and Johnson's (1992) approach.[2]

Research shows that in Mexico changes in labour income inequality can be largely linked to changes in the relative wage between skilled and unskilled workers, that is, in the returns to skill. In particular, the rise in inequality during this period was associated with an increase in returns to schooling.[3] Applying the RIF method proposed by Firpo et al. (2009) and the Oaxaca-Blinder decomposition method, Campos-Vazquez et al. (2014) show that the increase in earnings inequality between 1989 and 1994 is primarily driven by a rise in the returns to characteristics (schooling and experience). The effect of changes in the distribution of characteristics was almost flat. Using the framework proposed by Bound and Johnson (1992), Campos-Vazquez et al. (2014) conclude that both institutional factors and, surprisingly, the increase in relative demand for skilled workers (workers with high-school education and more) explained the increase in hourly wage inequality between 1989 and 1994. Hence, a key question is why demand for higher-educated individuals increased at a time when the Stolper-Samuelson theorem would have predicted the opposite.[4]

The decline in earnings inequality between 1994 and 2006 is primarily driven by a fall in the returns to characteristics (schooling and experience). The effect of changes in the distribution of characteristics (education, experience, female and urban) was unequalizing. In other words, if the returns to characteristics had remained unchanged in this period, the change in characteristics in the population would have resulted in higher levels of inequality. Why were changes in characteristics unequalizing during a period in which there was substantial educational upgrading and the distribution of years of schooling became more equal? This seemingly contradictory result is a mathematical consequence of increasing returns to skill and was first noted by Bourguignon et al. (2005), who called it the 'paradox of progress'.

Campos-Vazquez et al. (2014) show that the change in the skill premium during this period is the result of a combination of a rising supply of workers with college education and a slow-down in demand for skilled workers. After 1996 the real minimum wage and the unionization rate were fairly stable. Hence, these key institutional factors could not explain the decline in returns. Behind

---

[2] We attempt to estimate a model similar to Bound and Johnson (1992) and Manacorda et al. (2010). However, as pointed out by the latter, the relevant elasticities of substitution for the case of Mexico cannot be precisely estimated. In order to estimate the structural parameter σ, these authors use a sample of workers from Argentina, Brazil, Chile, Colombia, and Mexico; they mention that 'Mexico does not really contribute to the identification of the regression parameters' (Manacorda et al.: footnote 1, page 314).

[3] This result was found in many other studies. For example, Bouillon (2000a, 2000b), Bouillon et al. (1999), Lopez-Acevedo (2004, 2006), Meza Gozález (1999), Popli (2011), and Campos et al. (2014).

[4] In this paper, returns to skill and returns to schooling are used interchangeably.

the latter is the effect of the post-NAFTA intensification of integration with the United States, which has favoured the relatively low-skill sectors.

As for the period after 2006, we find that—with the labour force survey—the effect of characteristics continued to be unequalizing and the returns effect, equalizing. With ENIGH, however, this is no longer the case.

The paper is organized as follows. Section 2 describes the data and the methods used to correct for item non-response and under-representation. Section 3 summarizes the methodology to identify proximate determinants of the changes in inequality. Section 4 analyses the evolution of labour income inequality for 1989–2006 and reviews results from previous research on its determinants. Section 5 analyses the evolution of inequality with corrected data and decomposes the change into its proximate determinants. Section 6 concludes.

2 Data and data corrections

Research on labour income dynamics usually relies on labour force surveys such as the *Current Population Survey* in the USA. Labour force surveys contain detailed information on employment, unemployment, hourly wages, self-employment income, and sociodemographic characteristics in a sample that is large enough for the study of patterns by sociodemographic characteristics, regions, and so on. However, in Mexico, the labour force survey reached national coverage only in 2000; before that, it covered urban areas only.[5] Thus, for the analysis between 1989 and 2000, we rely entirely on ENIGH.

2.1 ENIGH, ENOE, and the puzzle of rising item non-response

ENIGH is available for 1984 and 1989, at intervals of two years between 1992 and 2014, and also for 2005. The survey includes detailed information on income measures (wages and salaries, self-employed income, interests, rents, dividends, transfers, remittances, and income from other sources), sociodemographic characteristics of household members (age, education, gender, education), and expenditures by type, wealth, and assets at the household level. However, the methodology used to generate labour income was modified in 2006, which creates challenges for comparisons over time, and this is one of the reasons we carry out the analysis for the period after that year separately.[6] The labour income measure from ENIGH that we use in this paper includes wage and self-employment earnings, in both the formal and informal sectors.[7] ENIGH is publicly

---

[5] The labour force survey was first implemented in 1987 for a sample of cities and was called the National Survey of Urban Employment (ENEU in its Spanish acronym). The number of cities gradually increased. Between 2000 and 2004, the labour force survey was called the National Employment Survey (ENE) and thereafter it became the National Survey of Occupation and Employment (ENOE). For simplicity, we use the last acronym throughout to refer to the labour force survey.

[6] For example, starting in 2008, the question regarding the *aguinaldo* (annual bonus) generated a figure for bonus income for the whole of the previous year, while before 2008 individuals were asked to give the amount of any bonus received the previous month. To avoid comparability issues, we removed the variable 'Bonus' from labour income in ENIGH. ENOE does not report bonus income, which also helps cross-survey comparability.

[7] Formality here is defined by whether the worker contributes to social security.

available and can be downloaded from the website of the Mexican Statistical Office (INEGI, www.inegi.org.mx/).

From 2000 onwards, our analysis uses both ENIGH and ENOE. ENOE is a quarterly survey of over 100,000 households (see Table A2 in Appendix) that follows the same household for five consecutive quarters. ENOE is used to calculate total employment, unemployment rates, and other labour market outcome characteristics at the national and state levels. In general, the methodology used to generate these surveys has remained largely the same, but the sampling has changed, as have some parts of the questionnaire. In particular, in 2005, ENOE was modified to homogenize the sociodemographic questionnaire with other surveys and to measure more precisely job search and employment duration. This year roughly coincides with the change in ENIGH mentioned above, so in both surveys it makes sense to do the analysis from 2006 onwards separately from before. As we shall see below, the year 2006 also (roughly) coincides with the point at which the decline in labour income inequality that had started in the mid-1990s stopped. (ENOE is also publicly available and can be downloaded from the INEGI website).

Figure 1: Workers with item non-response for labour income: 1988–2017 (ENOE and ENIGH; in % of total workers)
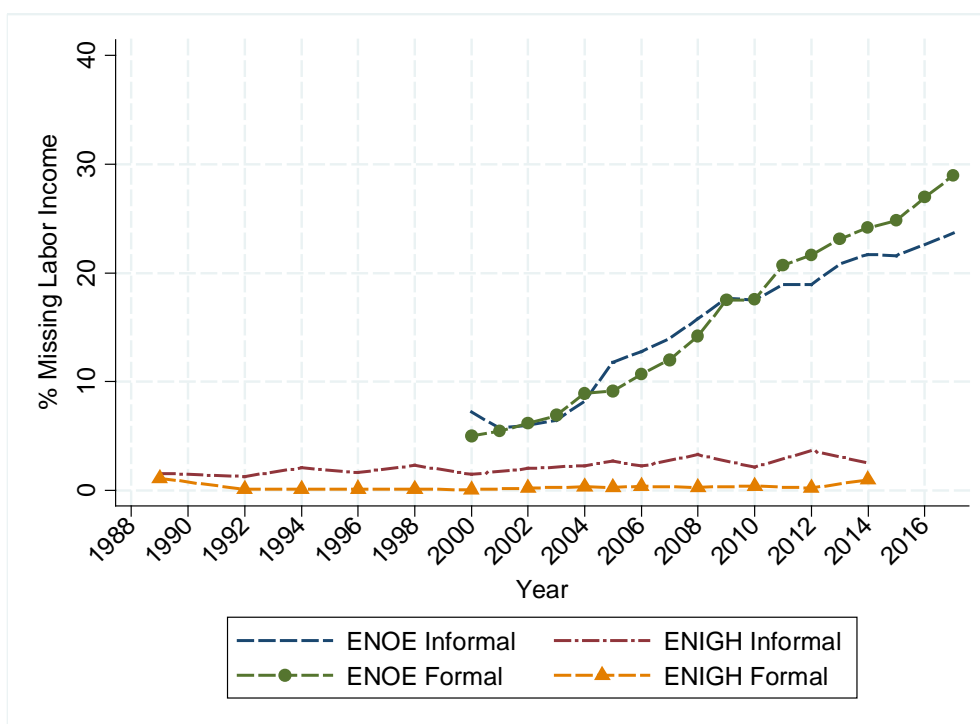
Notes: Sample restricted to workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year.

Source: Authors' construction.

Although ENOE should be preferred to ENIGH given its sample size, it suffers from an egregious drawback: item non-response for labour income is higher and, worse, it has been increasing over time (Figure 1).[8] In 2017, item non-response reached a whopping 30 per cent of all workers in

---

[8] The sample size for ENIGH is 19,419 households in 2014 (the last year we use here), much smaller than the labour force survey, which includes 107,274 households in 2014.

ENOE. Why there has been such an increase in non-response for labour income in ENOE (but not in ENIGH) is a puzzle that remains unsolved.

Some authors have attempted to correct for item non-response and under-reporting (which is also present in ENIGH) by adjusting (mainly, scaling up) the survey-based information to match totals in national accounts (Campos et al. 2015; Del Castillo 2015). Taking advantage of the fact that in 2016 the Mexican Institute of Social Security (IMSS) released (monthly) disaggregated data for formal sector workers by age, sex, and multiples of the minimum wage from 2000, here we follow a statistically more robust approach to correct for the rise in item non-response.[9] In essence, we make a post-survey weight adjustment (i.e. change the expansion factors) in ENIGH and ENOE so that the frequency of individuals for categories of formal workers defined by age, sex, and multiples of the minimum wage in the surveys equals the frequency distribution observed in tabulations from the IMSS.[10] In other words, we correct the distribution of labour income for formal sector workers on the assumption that the distribution of labour income in the IMSS data is the true distribution. We describe the method in more detail below. The theory behind this post-survey weight adjustment has a long tradition in statistics. For details see, for example, Biemer and Christ (2008).

2.2 Correcting for rising item non-response: post-survey reweighting and 'hot deck' imputation

As mentioned above, IMSS data include information on the number of workers in the formal sector by multiples of the minimum wage since 2000, and it also reports the brackets by gender and age group. Thus, it is possible to compare the distribution of workers by multiples of the minimum wage between IMSS and both ENIGH and ENOE (in the latter two datasets, we include the units with response on labour income only). Figure 2 and Table 1 show such a comparison for the years 2000, 2005, 2010, and 2017 (2014 in the case of ENIGH).

---

[9] IMSS data can be downloaded from http://datos.imss.gob.mx/.

[10] Formal workers also include workers that are employed in the public sector. Although information on them is not included in the IMSS data, here we include these workers in the formal sector categories.

Figure 2: Frequency of formal sector workers by multiple of the minimum wage in: 2000, 2005, 2010, and 2017/2014 (ENOE, ENIGH, and IMSS, in %; total formal workers = 100%)
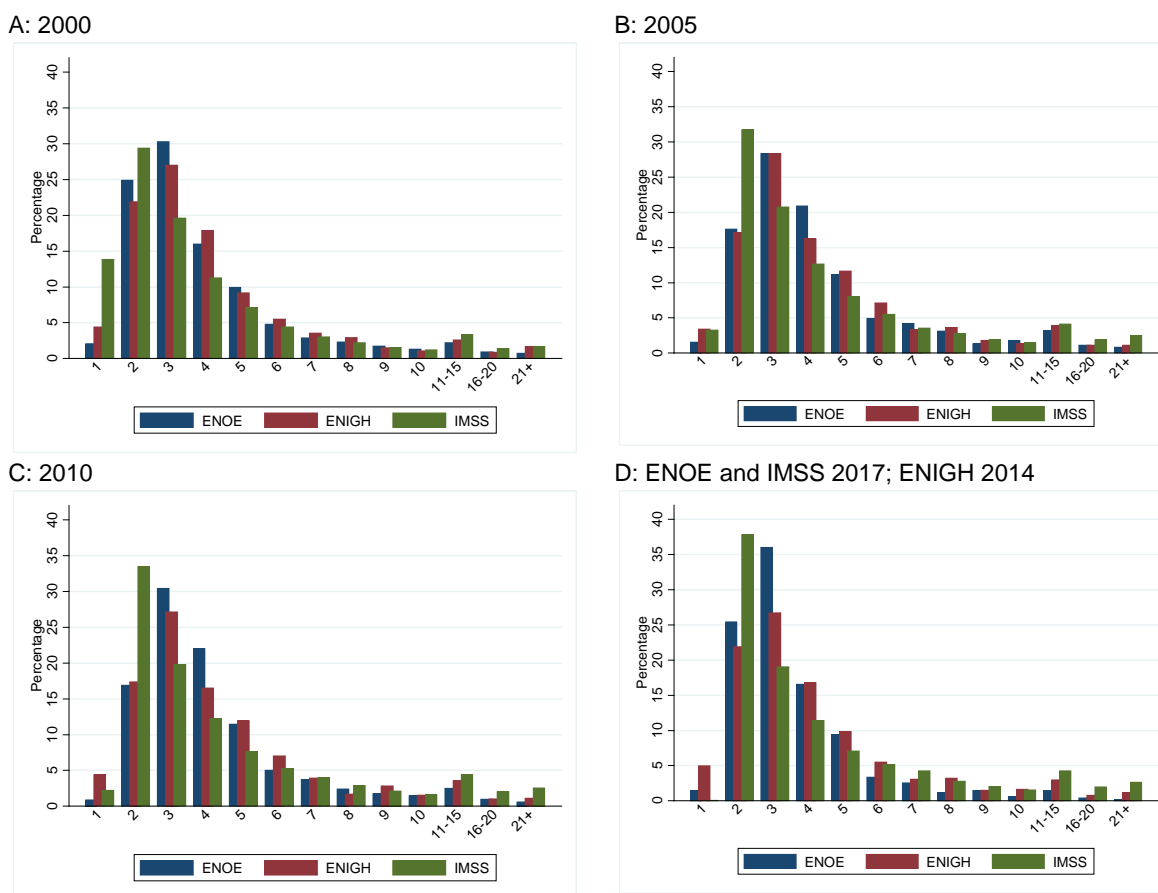
Notes: Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000, when we report the value for the month of December. Panel D for ENIGH refers to year 2014 because it was the most recent available at the time the analysis was completed.

Source: Authors' construction.

As shown in Table 1, IMSS data include a larger proportion of workers with two minimum wages and at the top of the distribution—those who earn more than ten and twenty times the minimum wage—than ENOE or ENIGH. Of relevance to our purposes, the under-representation of workers at the top increased over time, especially for ENOE. In 2000, the frequency of workers earning more than 10 MW equalled 3.8 per cent in ENOE and 6.5 per cent in IMSS. By 2017, the share was 2.1 per cent in ENOE and 8.9 per cent in IMSS. Workers earning more than 20 MW were 0.7 per cent (ENOE) and 1.7 per cent (IMSS) in 2000, while they were 0.2 per cent (ENOE) and 2.7 per cent (IMSS) in 2017.[11] The extent of under-representation is lower in ENIGH but still substantial for workers earning at least ten times the minimum wage.

---

[11] As the IMSS data are censored at twenty-five or more multiples of the minimum wage, the difference is in reality even greater.

Table 1: Frequency of formal sector workers by multiple of the minimum wage in: 2000, 2010, and 2017/2014 (ENOE, ENIGH and IMSS, in %; total formal workers = 100%)

	Multiples MW	ENOE	ENIGH	IMSS
2000	1–5	83.2	80.3	81.2
	6–10	13.0	14.5	12.4
	11–20	3.1	3.4	4.8
	21+	0.7	1.7	1.7
2010	1–5	81.6	77.3	75.3
	6–10	14.4	17.0	15.8
	11–20	3.5	4.6	6.4
	21+	0.6	1.1	2.5
2017/2014*	1–5	88.8	80.2	75.4
	6–10	9.1	14.9	15.7
	11–20	1.8	3.7	6.2
	21+	0.14	1.2	2.7

* ENIGH data are for 2014.

Notes: Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December.

Source: Authors' construction.

To correct for rising item non-response, we assume that the distribution of labour income of formal workers in the IMSS data is the true distribution and we reweigh individuals in ENOE and ENIGH so that the distribution of workers in the formal sector replicates the distribution of workers observed in the IMSS data. The new weights are calculated as follows. We classify workers in IMSS into categories by sex, age group (20–29, 30–39, 40–49, 50+), and multiples of the minimum wage (in groups: 1–10, 11–15, 16–20, 21+)[12], or 104 (2 x 4 x 13) categories. We obtain the percentage of workers in the formal sector within each category and then replace these weights in ENOE and ENIGH so that the surveys end up with the same distribution of workers by category as in the administrative data. In other words, for each of the categories, we multiply the original expansion factors by the ratio

$$\frac{\%\,Formal\,Workers\,in\,category\,i\,in\,IMSS}{\%\,Formal\,Workers\,in\,category\,i\,in\,survey}.$$

Figure 3 shows this reweighting process graphically for ENOE. Although due to data availability this correction method can be applied only from 2000 onwards, it is reassuring to note that the rise in item non-response in ENOE starts in the mid-2000s (Figure 1).
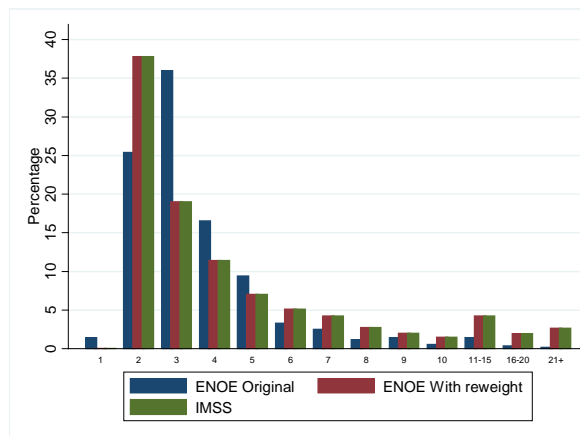
---

[12] The IMSS data are censored at twenty-five or more multiples of the minimum wage.

Figure 3: Frequency of formal sector workers by multiple of the minimum wage: 2005 and 2017 (ENOE and IMSS: original, reweighting only, imputation only, and reweighting and imputation combined; in %; total formal workers = 100%)
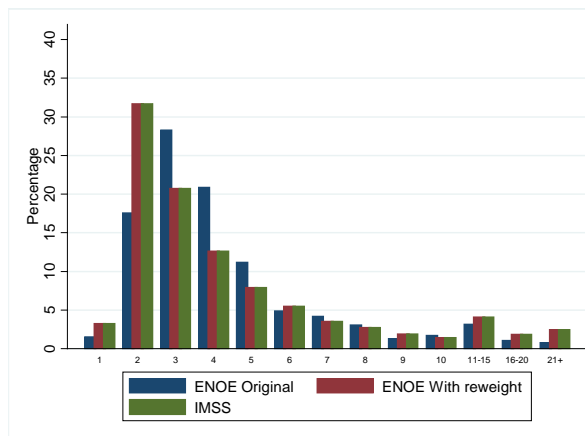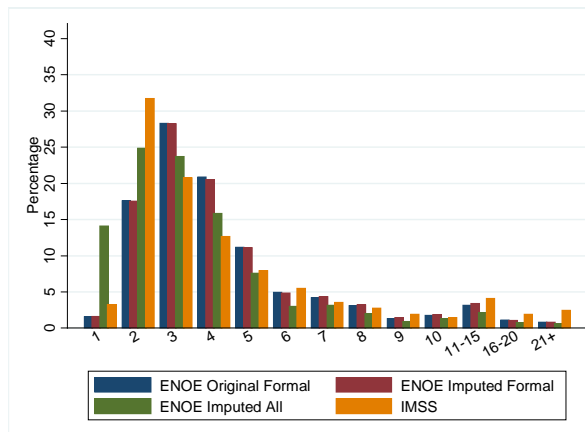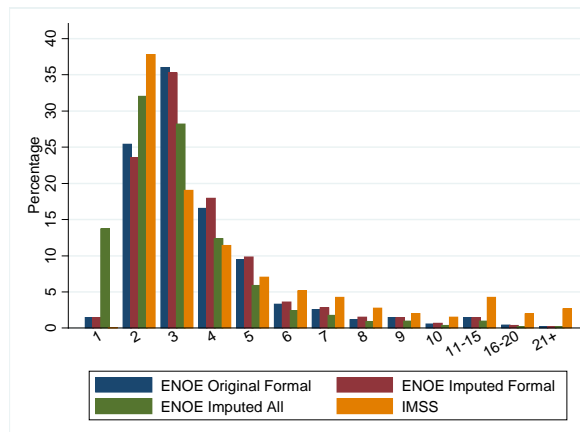

Notes: Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December. For reweighting we use IMSS data.

Source: Authors' construction.

The post-survey reweighting method is applied to formal sector workers. As shown in Figure 1, however, rising item non-response occurs among informal workers as well. Since we do not have an external administrative source to generate 'true' weights, we estimate the missing labour incomes for informal workers by applying the so-called 'hot deck' imputation method, as proposed by Little and Rubin (2002).

We also applied the hot deck imputation method to formal workers and reweighted after the imputation. We generated a new distribution of workers (which no longer has no response because of the hot deck imputation) by the 104 categories in ENOE which equals that in IMSS. The new comparisons are shown in Figure 3 (Panels C and D). As can be observed, hot deck imputation yields a different distribution of workers by multiples of the minimum wage than the one observed in IMSS. To note is the fact that the imputation method yields a much lower proportion of workers

earning more than ten minimum wages (still lower for those earning 21 MW or more). This under-representation of high-wage workers is higher in 2017 than in 2005.

In this paper, we refer to the post-survey reweighted data as 'reweighted' and the data with post-survey hot deck imputations as 'imputed'. The data that include both corrections are called 'corrected'. Given that all methods have their limitations, we carried out a series of robustness checks, which are discussed below. The results are shown in the Appendix.

2.3   Comparing original with reweighted and imputed data

Figure 4 shows average monthly labour income from the original and reweighted surveys (for formal workers) and imputed wages (for informal workers) for ENIGH and ENOE. For formal workers, both the original and reweighted averages show a similar trend. The main differences are: (i) the decline between 2007 and 2010 (the so-called Great Recession) is smaller when using the reweighted measure and (ii) there is an increase in average labour income since 2015 with the reweighted measure but not with the original. Unsurprisingly given the reweighting method, the pattern of the reweighted surveys is similar to that observed with IMSS data, as shown in Figure 5.

Figure 4: Average monthly labour income for formal and informal workers: 2000–2017 (ENOE, ENIGH, IMSS: original and reweighted surveys; in constant Mexican pesos of August 2015)
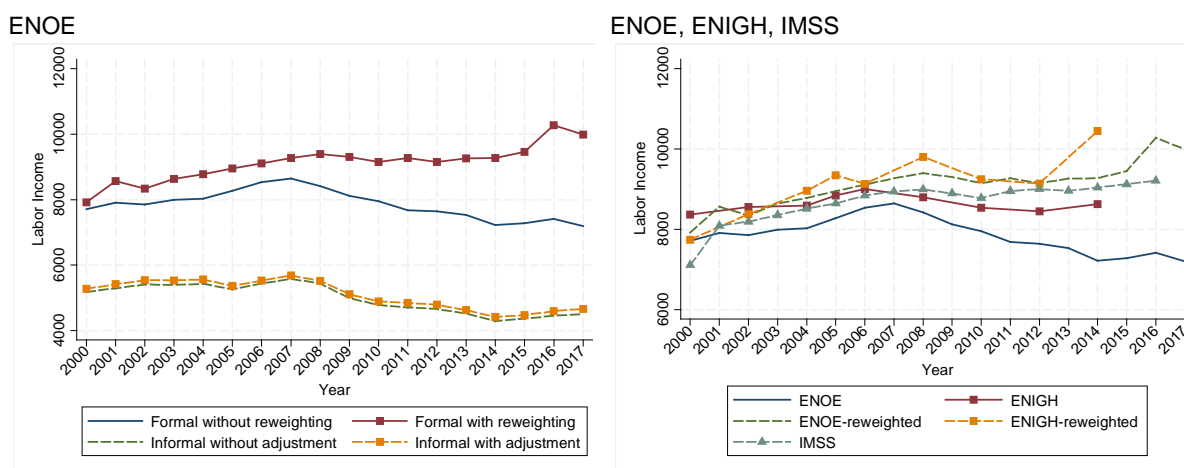
Notes: Sample restricted to workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For reweighting we use IMSS data. Wage adjustment in the informal sector follows a hot deck imputation procedure using gender, region, informal status, age, and education group.

Source: Authors' construction.
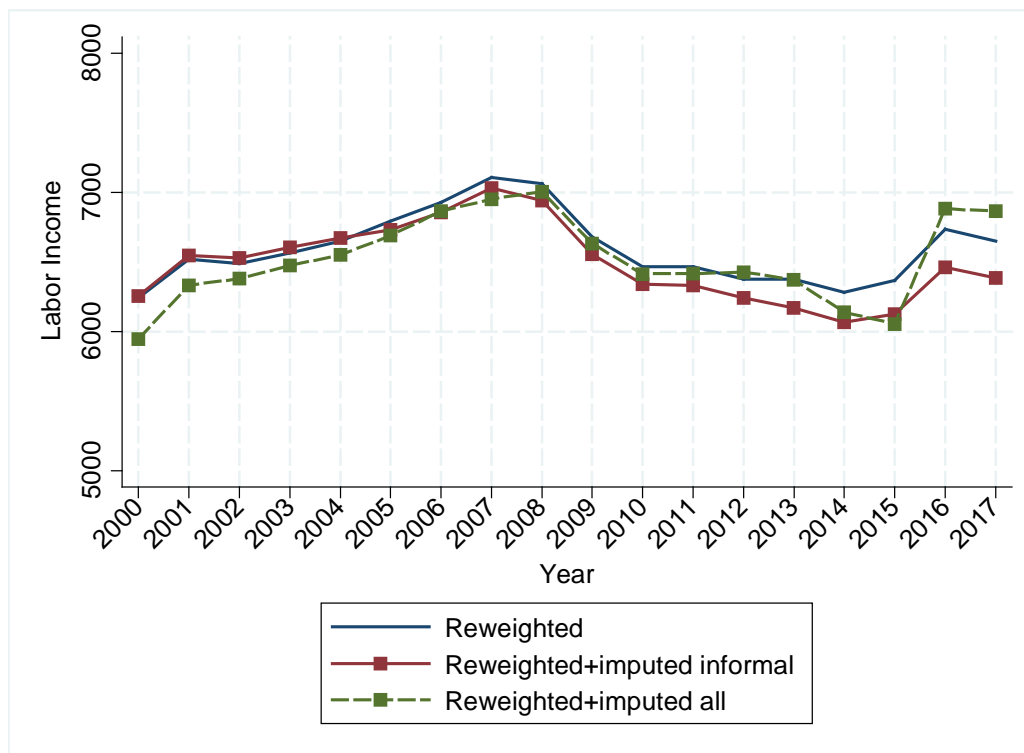
Figure 5: Average monthly labour income for all workers: 2000–2017 (ENOE: original and corrected surveys; in constant Mexican pesos of August 2015)
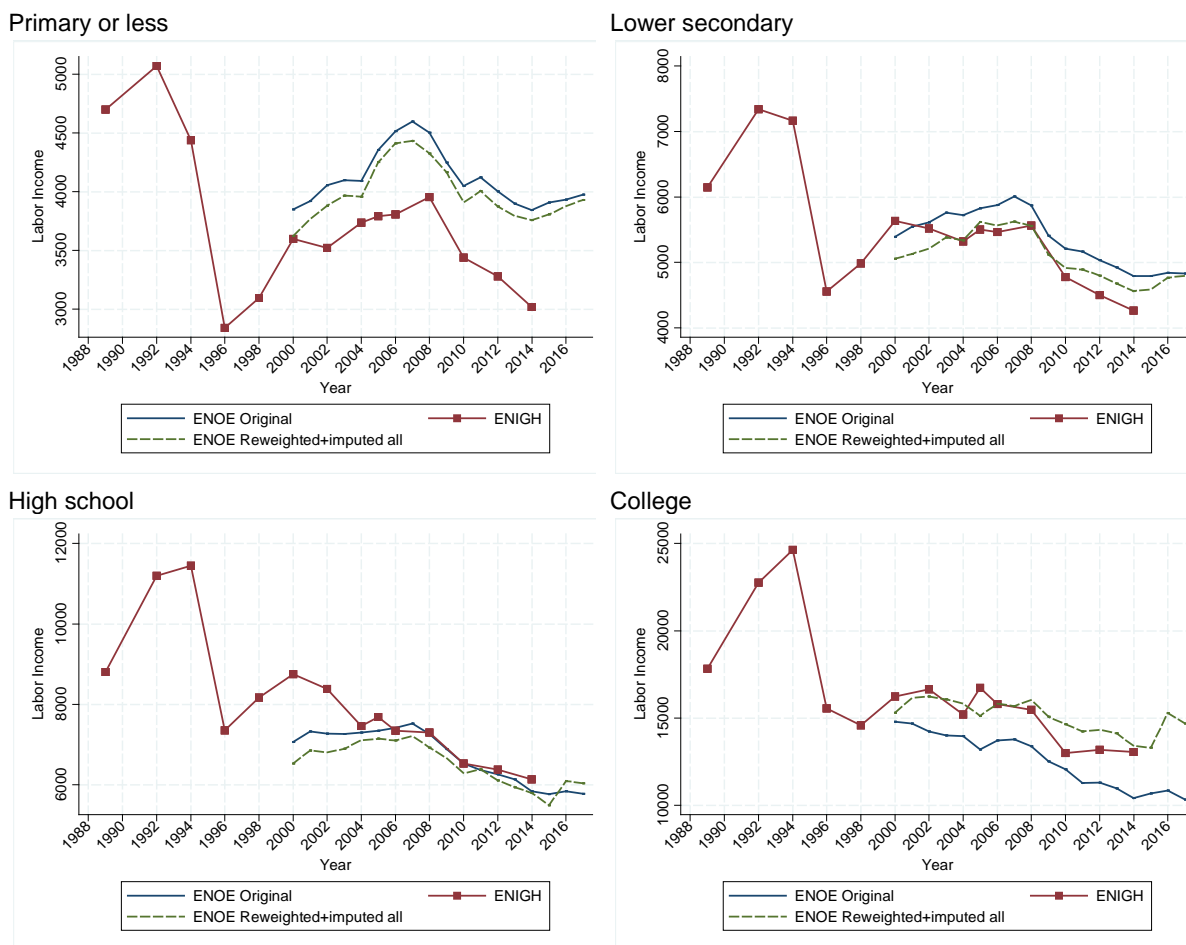

Notes: Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. To correct for item non-response in ENOE, we follow a hot deck imputation procedure using gender, region, informal status, age, and education group. For post-survey reweighting for formal sector workers in ENOE we use IMSS data.

Source: Authors' construction.

In Figure 6, we show average monthly labour income for the original, reweighted, and imputed surveys by education. The first thing to notice is that the reweighting procedure mainly affects the trend for formal workers with a college education: while the original data show a steady decline, the reweighted surveys show that average labour incomes have remained largely unchanged since 2008 with an upswing after 2015. Since the reweighting method increases primarily the percentage of formal workers earning between one and two minimum wages and of those earning more than ten minimum wages, the reweighted averages tend to be slightly lower for those with less than a college education and higher for those with a college education. Moreover, in recent years the gap between the original average labour income and the reweighted measures has increased, especially in ENOE since 2008. When survey data have been reweighted (for formal workers) *and* imputed (for informal workers), the patterns are almost identical to those observed with reweighted-only surveys.

10

Figure 6: Average monthly labour income for all workers by education: 1989–2017 (ENOE: original, reweighted, and imputed surveys; ENIGH: original surveys; in constant Mexican pesos of August 2015)

Notes: Sample restricted to all workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For reweighting we use IMSS data.

Source: Authors' construction.

2.4 Robustness checks

In order to check the robustness of the correction methods followed here, we reweighted the surveys on the assumption that the true distribution in IMSS starts with the group earning five times the minimum wage or more and that below this threshold the data in the surveys are accurate. As shown in the Appendix, this correction yields almost exactly the same outcomes as those observed above. We also tried the rescaling method which has been applied by others (Piketty et al. 2016). That is, we did not change the weights but we replaced the average labour income for formal workers by centile in ENOE with the corresponding one in IMSS. Since the latter requires the entire population in IMSS and not just tabulations by multiples of the minimum wage, we were able to do this only for 2010 and 2012, the two years that these data were made available to our team.[13] As we shall discuss further below, the most important difference with rescaled surveys versus reweighted surveys is that, using ENOE, from 2008 onwards the Gini coefficient rises

---

[13] We thank Facundo Alvaredo from the Paris School of Economics for sharing the rescaling factors for the total population for 2010 and 2012.

instead of remaining roughly constant. In other words, with rescaled labour incomes, the trend in inequality is in even more stark contrast to the one observed with the original survey, which shows a steady decline. Finally, we also checked what might happen if we corrected the IMSS data for the fact that they are censored at the top using a Pareto approximation. While this correction increases the level of inequality, trends remain largely the same.

3   Methodology[14]

Labour income inequality is affected by two main factors: the distribution of (observable and unobservable) characteristics of workers (education, experience, gender, etc.) and the returns to those characteristics. Workers' characteristics, in turn, are affected by 'fate' (gender, race, talent, and so on), households' decisions (e.g. to enrol or not in post-secondary education), and policy (e.g. expanding access to education). Returns to households' characteristics depend on market forces (i.e. demand and supply of workers of different skills and experience) and institutional/policy factors (e.g. minimum wage policy and the unionization rate).

In order to separate the contribution of characteristics and returns, research on the proximate determinants of labour income inequality relies on decomposition techniques. Many decomposition procedures are employed in the literature.[15] Most of them are variations on the Oaxaca-Blinder decomposition.[16] In this paper, we follow the same approach. We employ the 're-centered influence function' (RIF) procedure proposed by Firpo et al. (2009) to decompose effects into characteristics or composition and returns effects.

The RIF procedure is very similar to the typical OB decomposition.[17] The main difference is that the dependent variable, Y, is replaced by the RIF.[18] Firpo et al. (2009) demonstrate that the RIF procedure is equivalent to a simple unconditional quantile regression. They show that $E[RIF(v,y)|X] = X\beta^v$, where the coefficient β^v represents the marginal effect of X on the dependent variable statistic v.[19] The main difference from the basic Oaxaca-Blinder decomposition

---

[14] This section draws almost entirely on Campos-Vazquez et al. (2014).

[15] See the excellent review by Fortin et al. (2011).

[16] We can divide the decomposition into four groups: (i) reweighting procedures (DiNardo et al. 1996), (ii) residual-imputation procedures (Almeida dos Reis and Paes de Barros 1991; Juhn et al. 1993), (iii) quantile decomposition procedures (Machado and Mata 2005), and (iv) RIF procedures (Firpo et al. 2009).

[17] See Firpo et al. (2009) and Fortin et al. (2011) for more details of the RIF procedure.

[18] Define $RIF(v,y)$ as the re-centred influence function with distributional statistic of interest $v(F_y)$ and observed wage y. Then it can be shown that $RIF(v,y) = v(F_y) + IF(v,y)$, where IF denotes the influence function such that $\int RIF = v(F_y)$. For the case of quantiles, it can be shown that the influence function is equal to $(Q_\tau, Y) = \frac{\tau - 1\{Y \le Q_\tau\}}{f_Y(Q_\tau)}$. Each statistic $v(F_y)$ refers to a specific quantile in the distribution of Y or to the Gini coefficient or the variance.

[19] For example, if v represents quantile 0.50, then $\beta^{v=0.5}$ represents the effect of X on the wage quantile 0.50. It can also be applied to scalar indicators of inequality such as the Gini or the variance. In order to estimate the RIF regression, we first estimate the sample $\widehat{RIF}(v,y)$. In practice, we follow the ado file *rifreg* in Stata published by Fortin et al. (2011) and provided by N. Fortin (www.faculty.arts.ubc.ca/nfortin/datahead.html). The *RIF* dependent variable is estimated using kernel methods. We use the following explanatory variables: dummy variables of female, urban, and education categories and a cubic polynomial in age. We also estimate a more flexible model that included interactions among all variables, but the difference in explained and unexplained components was minimal.

is that, because of its statistical properties, the RIF approach allows you to decompose the contributions for the entire distribution rather than just having to use the mean. Moreover, the RIF approach has an advantage over other methods that permit decomposition for the entire distribution, in that it does not 'suffer' from path dependency.[20]

We start our analysis by calculating the difference in average labour income for each quantile between the initial and end years for every quantile in segments of 1 per cent (that is, from the 1st to the 99th percentile). Then we estimate the RIF regression for each quantile and the initial and end years. Once the parameters β^v are estimated, we proceed to apply the basic Oaxaca-Blinder decomposition for each quantile (1st–99th percentile). That is, we calculate $\hat{v}(Y_t) - \hat{v}(Y_s) = \hat{\beta}_s^v(\bar{X}_t - \bar{X}_s) + \bar{X}_t(\hat{\beta}_t^v - \hat{\beta}_s^v)$, where t is the final year and s is the initial year.[21] Note that the $\bar{X}$ are for the entire sample, as in the traditional Oaxaca-Blinder. In our application, we set up the initial years as 1989, 1994, and 2006 and the final years as 1994, 2006, and 2017, respectively. The term $\hat{\beta}_s^v(\bar{X}_t - \bar{X}_s)$ refers to the characteristics effects and the term $\bar{X}_t(\hat{\beta}_t^v - \hat{\beta}_s^v)$ refers to the return or price effects to observable characteristics included in X as well as unobservable ones (which is why this term is often referred to as the 'unexplained component'). We use as reference the wage distribution in the initial year (for each decomposition). With this information, we can decompose all the labour income growth incidence curves into two curves: the characteristics component and the returns component or relative returns.

We use Bound and Johnson's (1992) decomposition method to describe changes in inequality.[22] Changes in relative wages (for example college-educated workers against the rest) come from changes in competitive and non-competitive forces, and unexplained changes. Competitive forces relate to changes in supply and demand. If there are the same number of jobs over time but the number of college-educated workers is increasing in proportion to all workers (i.e. there is an increase in their relative supply), we expect lower wages for college-educated and higher wages for non-college-educated workers. On the other hand, demand relates to how firms are using certain types of worker. If labour demand is increasing for higher-educated workers and nothing else is changing, we expect a higher relative wage for that group. Changes in demand could come from changes in trade patterns or developments in technology, for example. However, they are unobserved. Data show only labour supply but not real labour demand. Researchers then traditionally use the residual (after accounting for all possible changes) as an approximation to changes in demand. The second part of the decomposition comes from non-competitive forces. Among those are changes in the unionization rate and the value of the real minimum wage. If the real minimum wage or the unionization rate falls, we expect lower bargaining power among workers, which may imply that wages do not grow as fast at the bottom of the distribution as at

---

[20] For a discussion and application of such methods and their limitations see, for example, Bourguignon et al. (2005).

[21] See equation 35 in Fortin et al. (2011).

[22] For more details see Campos-Vazquez (2013) and Campos-Vazquez et al. (2014). Assuming a simple constant elasticity substitution production function with only two inputs: workers with high-school (or less than college) and

college education $Y = [S^\rho + C^\rho]^{1/\rho}$, using the first order conditions we get $\ln\left(\dfrac{w^S}{w^C}\right) = -\dfrac{1}{\sigma}\ln\left(\dfrac{S}{C}\right)$, where $\sigma = \dfrac{1}{1-\rho}$

is the elasticity of substitution. Hence, if we know the elasticity of substitution, we can calculate the effect on relative wages due to relative supply. Bound and Johnson (1992) extend this framework to add changes in demand, and relative changes in non-competitive forces like the unionization rate and the minimum wage. The residual is generally assumed to be equal to skill-biased technical change.

the top. Finally, the residual refers to everything that is unaccounted for. As mentioned, in most cases changes in demand are not observed independently from residual changes. We keep this theoretical framework to explain previous literature.

Relative returns (also known as wage structure) are affected by demand and supply of workers of different skills and by institutional factors such as the minimum wage and unionization rate. In order to examine the effect of supply and demand on relative wages, we follow the Bound and Johnson (1992) method.[23] Assuming a simple CES (constant elasticity of substitution) production function with elasticity of substitution, σ, constant across skills, it is possible to determine the effect of supply and demand on relative wages:[24]

$$\Delta\%\left(\frac{\overline{w}^c}{\overline{w}^s}\right) = \frac{1}{\sigma}\Delta\%(Demand) - \frac{1}{\sigma}\Delta\%(Supply) + \xi$$

The residual term ξ contains the effect of skill-biased technical change and institutional factors such as the minimum wage and unionization rate (sometimes called non-competitive factors).

4   Labour income inequality 1989–2006: an overview of research findings

In Figure 7, we show the evolution of the Gini coefficient using the original ENIGH for 1989–2014, ENOE-urban (the labour survey for urban areas only) for 1988–2015, and the original (uncorrected) ENOE for 2000–2017.[25] Recall that the labour force survey did not have national coverage before 2000, so our analysis for the period 1989–2000 must rely on data from ENIGH only. For the period before 2000 we do not have external administrative data to correct for non-response and under-representation, but the proportion of workers with item non-response in ENIGH is relatively low and constant (see Figure 1). Thus, in principle, it should be fine to use the original ENIGH survey for this period.[26] We can observe two distinct periods. Between 1989 and 1994, labour income inequality as measured by the Gini coefficient rose. Between 1994 and 2006, it declined. We now proceed to analyse the determinants of this evolution, reviewing previous research that applied the methodological framework that was briefly discussed above.

---

[23] We attempt to estimate a model similar to those of Bound and Johnson (1992) and Manacorda et al. (2010). However, as pointed out by Manacorda et al. (2010), the relevant elasticities of substitution for the case of Mexico cannot be precisely estimated. In order to estimate the structural parameter σ, Manacorda et al. use a sample of workers from Argentina, Brazil, Chile, Colombia, and Mexico; they mention that 'Mexico does not really contribute to the identification of the regression parameters' (footnote 1, page 314).

[24] See formula 3 on page 377 and formula A8 on page 390 of Bound and Johnson (1992).

[25] Although throughout this paper we present our inequality analysis using the Gini coefficient, other inequality measures present the same trend. See Appendix, for example.

[26] The problem of the under-reporting of labour incomes and under-representation of high-wage earners may persist, but we do not address it here.

Figure 7: Gini coefficient for labour income: 1989–2017 (ENOE, ENOE-urban, and ENIGH)
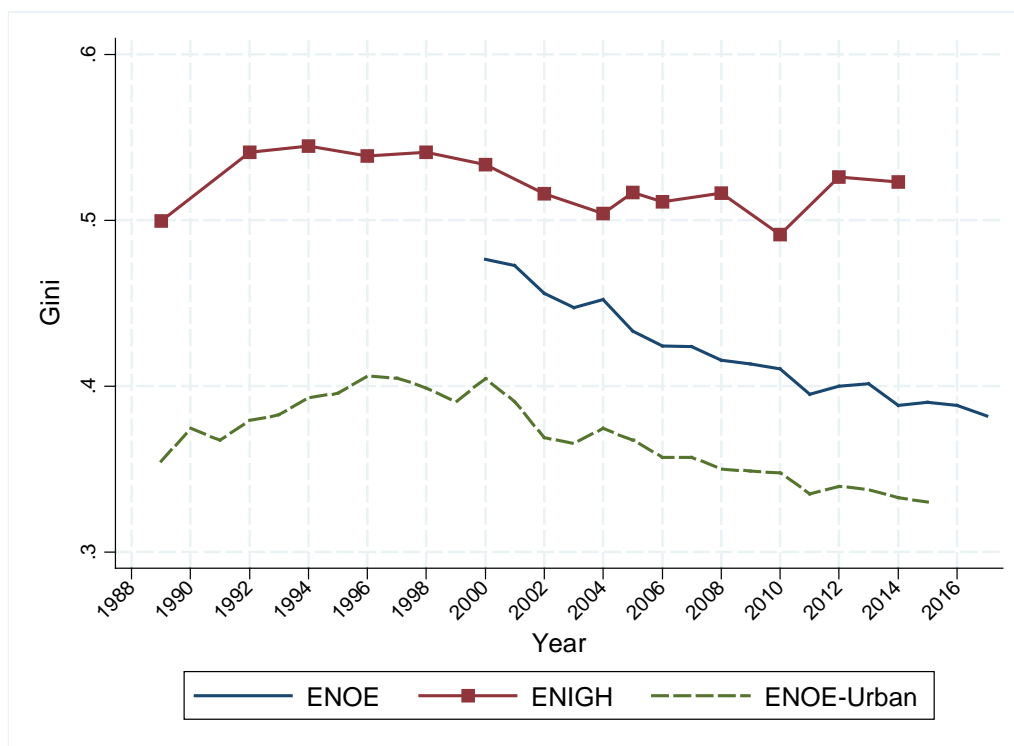
Notes: Workers aged 20–64 years and with positive labour income and working hours. For ENOE we use the second quarter of each year. We exclude households whose head reports zero monetary income.

Source: Authors' construction.

4.1 Rising labour income inequality: 1989–1994

Research shows that in Mexico changes in labour income inequality can be largely linked to changes in the relative wage between skilled and unskilled workers, that is, in the returns to skill. In particular, the rise in inequality during this period is associated with an increase in returns to schooling.[27] Applying the RIF method proposed by Firpo et al. (2009) and the Oaxaca-Blinder decomposition method, Campos-Vazquez et al. (2014) show that the increase in earnings inequality between 1989 and 1994 is primarily driven by a rise in the returns to characteristics (schooling and experience), as shown by the upward sloping curve in Figure 8. The distribution of characteristics remains almost flat.

---

[27] This result was found in many other studies, including Bouillon et al. (1999), Meza González (1999), Bouillon (2000a, 2000b), Lopez-Acevedo (2004, 2006), Popli (2011), and Campos-Vazquez et al. (2014).

Figure 8. Decomposition of differences in the distribution of earnings: 1989–1994
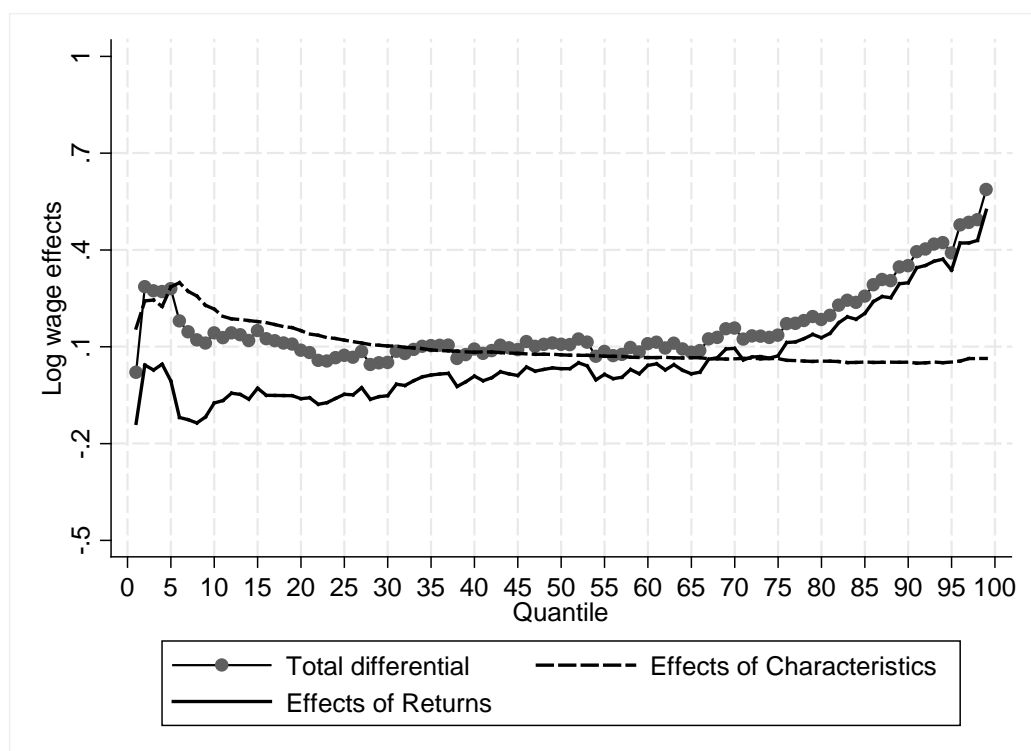

Notes: Calculations using ENIGH. Total differential is the total change in hourly wages (in logs); Effects of characteristics and Effects of returns are the portions that can be ascribed to changes in characteristics and returns, respectively.

Source: Campos-Vazquez et al. (2014: figure 7.4).

Changes in returns to schooling, in turn, can be due to changes in the relative demand and supply of workers of different characteristics such as education and experience, and/or changes in institutional factors such as the minimum wage and the unionization rate. The supply of college-educated workers increased during this period but at a lower pace than in subsequent years (Figure 9).[28] Applying the methodology proposed by Bound and Johnson (1992), Campos-Vazquez et al. (2014) conclude that both institutional factors and, surprisingly, the increase in relative demand for skilled workers (workers with high-school education and more) explained the increase in hourly wage inequality between 1989 and 1994. Hence, a key question is why demand for higher-educated individuals increased at a time when theory would have predicted the opposite.

---

[28] The slow pace of supply growth of college-educated workers has been attributed to the consequences of the debt crisis and what has become known in Latin America as the 'lost decade', characterized by a reduction in education budgets (Campos-Vazquez 2013) and in the proportion of students who continued to the next stage after graduating from primary and secondary schooling (Friedman et al. 1995). In the USA, a slow-down in the growth of workers with post-secondary education has also been identified as an important driver of the rise in the returns to skill (Autor and Katz 1999; Goldin and Katz 2009).

Figure 9: Relative returns and relative supply, 1989–2010 (high school and more vs. lower secondary or less)
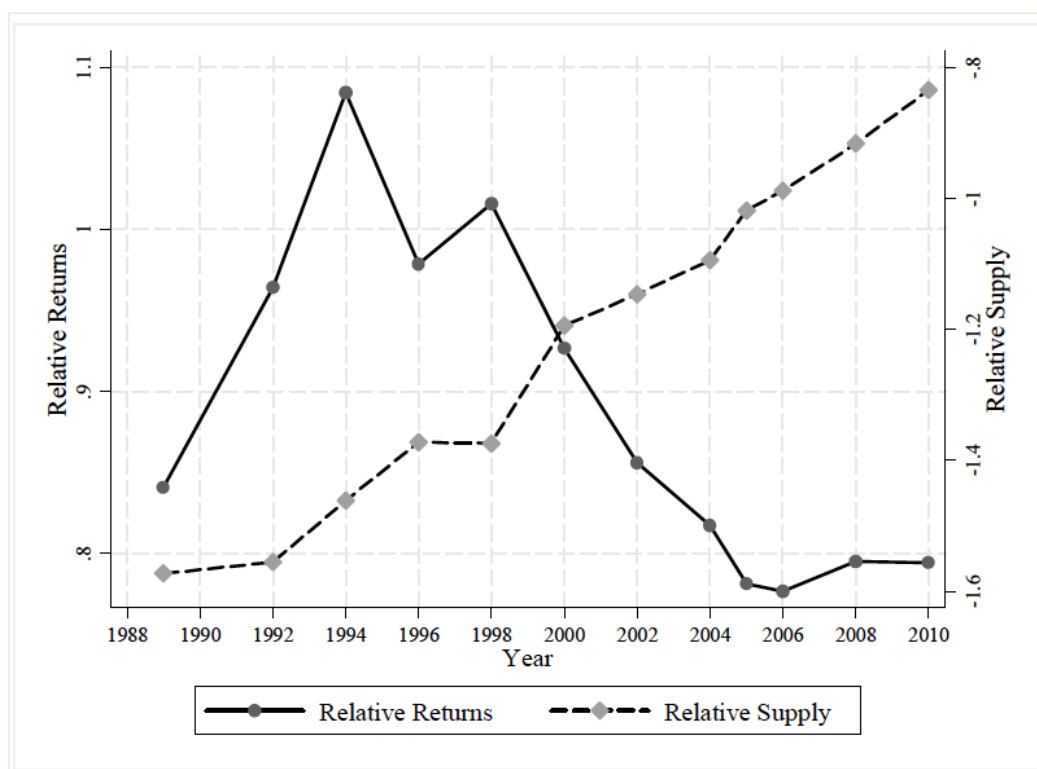
Notes: Sample restricted to workers 18–65 years old in ENIGH. Relative returns are obtained from a regression of log hourly wages. Relative supply is equal to the log of the ratio of proportion of workers with high-school or college education over the proportion of workers with lower secondary or less. For more details, see Campos-Vazquez et al. (2014).

Source: Campos-Vazquez et al. (2014: figure 7.5).

Mexico experienced a large opening of its economy in 1986 when it joined the precursor of the World Trade Organization, the General Agreement on Tariffs and Trade or GATT. Due to the relative abundance of less-skilled labour, the Stolper-Samuelson theorem would have predicted that liberalization would lead to a decrease in the relative wage of high-skilled workers and, therefore, a fall in inequality. However, as noted by Cragg and Epelbaum (1996) and Esquivel and Rodríguez-Lopez (2003), the opposite occurred. What drove this seemingly contradictory outcome? There are several persuasive explanations that offer an answer to the puzzle.

First, there is evidence that the most protected industries during the previous period were low-skill intensive sectors (e.g. textiles) and, thus, trade liberalization reduced the relative price of these industries and, as a consequence, the relative wage of the low-skilled (Robertson 2004, 2007).[29] Second, there is evidence that during this period there was skill-biased technical change and a change in the composition of output that gave skill-intensive industries a higher share.[30] Third,

---

[29] Hanson and Harrison (1999) and Feliciano (2001) both find a decline in protection for less-skilled industries, but are unable to find a link to wages. Using more detailed and varied price data, however, Robertson (2004, 2007) finds evidence of changes in output prices and wages. Also, Revenga (1997) finds that trade liberalization after GATT caused a decline in wages of 10–14 per cent in the most protected industries.

[30] Cragg and Epelbaum (1996) argue that the export-oriented sector became more skill-intensive, causing an increase in demand for skilled workers relative to unskilled workers. A similar pattern is observed by Esquivel and Rodríguez-

changes in the pattern of foreign direct investment (FDI) favoured skill-intensive firms. With trade liberalization, FDI also increased primarily through the expansion of maquiladoras. These establishments import most of their inputs and assemble the product to export (mainly to the USA). Maquiladoras are important both as a source of employment and in the share of exports to the USA: the employment share in maquiladoras (within manufacturing) grew from 5 to 25 per cent in the period 1980–1997, and in 1995 the share of exports coming from maquiladoras was 40 per cent (Hanson 2003). Hence, if this sudden change is correlated with skill intensiveness, it might explain the increase in wage inequality prior to NAFTA being in place. Using industry and state-level data from 1975 to 1988, Feenstra and Hanson (1997) track the impact of FDI on employment and wages. They find that the outsourcing of US multinationals caused an increase in the number of establishments in Mexico that favoured skill-intensive industries.[31]

In other words, there is no real contradiction with the standard Stolper-Samuelson theorem: trade opening benefited skill-intensive industries relatively more because, contrary to expectations, low-skill industries had been relatively more protected before. This change, combined with skill-biased technical change and the change in the composition of output towards more skill-intensive sectors, favoured wages of skilled workers and increased labour income inequality.[32]

In addition to the positive impact on skilled workers' wages stemming from trade liberalization and skill-biased technical change, the evolution of the minimum wage and unionization rate might have played a role. Figure 10 shows the evolution of the real value of the minimum wage (2015 MXP) and of the unionization rate. From 1988 to 1996 the real minimum wage lost close to 50 per cent of its value and the unionization rate declined by roughly 40 per cent.

If the sharp decline in minimum wages and the unionization rate is correlated with workers' bargaining power, they could affect the distribution of labour income because of their downward pressure on the wages of the low-skilled. In the case of unionization, there is evidence that its decline before NAFTA affected the wage structure. Using ENIGH, Fairris (2003) and Fairris and Levine (2004) conclude that the fall in unionization rate from 1984 to 1996 explains 11 per cent of the increase in wage inequality. In terms of minimum wages, Bosch and Manacorda (2010) analyse the effect of the minimum wage on the wage structure and wage inequality during the


López (2003) using an industry-level dataset as opposed to a household survey. Esquivel and Rodríguez-López are able to separate the channels of pure trade vs technology, and find that while the trade channel acted to decrease inequality, this was more than offset by the skill bias in technological change, resulting in an increase in inequality. Although inequality increased within observable characteristics, there is evidence that between-industry shifts among tradable industries explain part of the increase in wage inequality (Airola and Juhn 2005).

[31] As shown by Kurokawa (2011), industries in the USA that shift production to Mexico are low-skilled within the USA but high-skilled within Mexico. Hanson (2003) reviews the evidence of FDI and trade, concluding that 'Mexico's economic opening thus appears to have raised the relative demand for skilled labour, and tariff and quota reductions have altered inter-industry wage differentials' (p. 3). There have also been changes in regional inequality. For an analysis of these patterns and their determinants see, for example, Hanson (2003), Cortez (2005), Garduño-Rivera (2010), Borraz and Lopez-Cordova (2007), and De Hoyos (2013). Another possible channel is the upgrading quality hypothesis (Verhoogen 2008), which states that in the case of the devaluation in the Mexican peso, capital-intensive industries have incentives to increase exports (upgrade the quality of their products) and, as capital is complementary to skilled labour, wage inequality occurs.

[32] In addition to the impact of trade liberalization and its implications on the demand for skills, there may be an adverse effect on their supply: e.g. more job opportunities available in the maquiladoras could cause a higher high-school dropout rate. For instance, Atkin (2016) finds that for every twenty-five jobs created one student dropped out of school at grade 9 (final year of middle school).

1989–1994 period and in later years. They find that all of the increase in inequality in the bottom part of the distribution is caused by the fall in the real minimum wage. This is mainly due to the fact that the minimum wage affects other wages close to the minimum wage (lighthouse effect). In particular, Kaplan and Novaro (2006) argue that although the minimum wage binding process has declined over time (at least until 1996), it affects other wages in the distribution (a similar result is provided by Fairris et al. 2008). Cortez (2001) analyses both aspects (unionization and minimum wages) and concludes that the increase in wage inequality can be fully explained by the decline in institutional forces.

Figure 10: Unionization rate and the real minimum wage: 1988–2017
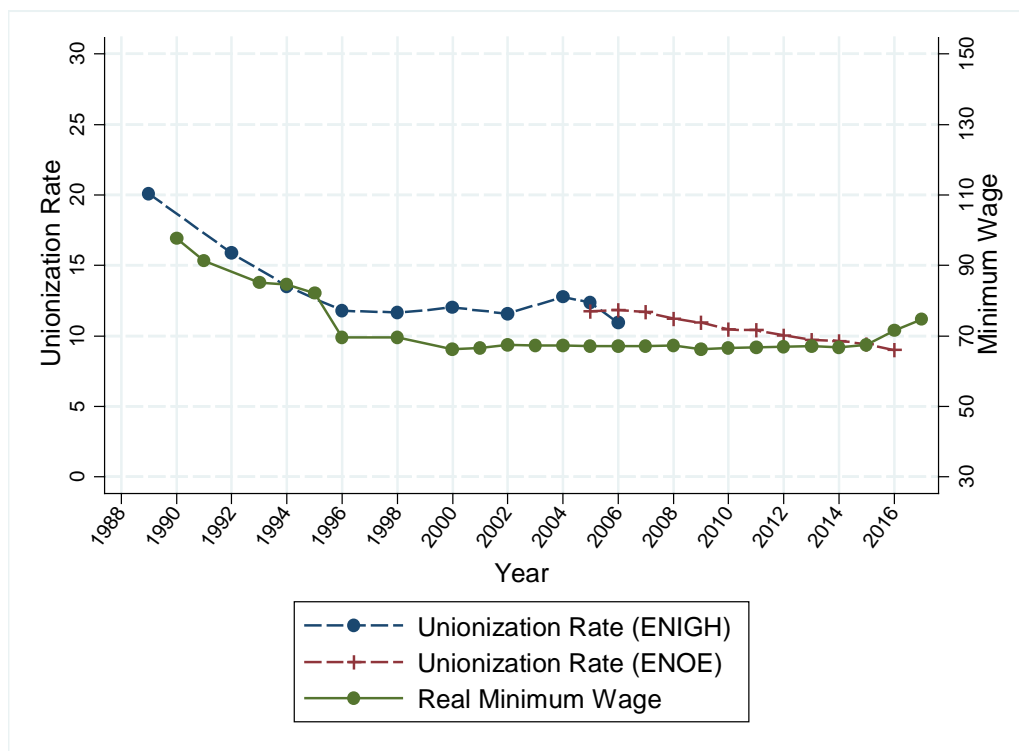
Notes: Unionization rate calculated for workers aged 20–65 years. Figures from 1989–2004 obtained from ENIGH and from 2005–2016 from ENOE. Minimum wage figures obtained from the Central Bank (http://www.banxico.org.mx/estadisticas/index.html).

Source: Authors' construction.

4.2   Falling labour income inequality: 1996–2006

From the mid-1990s up to the mid-2000s, labour income inequality steadily declined (Figure 7).[33] Applying the RIF method, Campos-Vazquez et al. (2014) show that the decline in earnings inequality between 1994 and 2006 is primarily driven by a fall in the returns to characteristics (schooling and experience), as shown by the downward curve in Figure 11. The effect of changes in the distribution of characteristics (education, experience, female, and urban) was, in contrast, unequalizing, as shown by the upward curve for the effect of characteristics in Figure 11. If returns had remained unchanged in this period, the change in characteristics in the population would have

---

[33] Not only the Gini coefficient and other summary indicators for hourly wages and labour income decline; firm data, for instance, show a decline in the relative wage of white- over blue-collar workers (Esquivel 2011).

resulted in higher levels of inequality. As shown in Figure 11, the effect of returns to those characteristics contributed to equalize the labour income distribution by such an amount that they compensated for the inequality-increasing effects induced by characteristics. The puzzle is why changes in characteristics were unequalizing during a period in which, for example, there was substantial educational upgrading and the distribution of years of schooling became more equal. This seemingly contradictory result was first noted by Bourguignon et al. (2005), who called it the 'paradox of progress'. These authors show that this puzzling result is the mathematical consequence of the convexity in (i.e. increasing) returns to skill.[34]

Figure 11: Decomposition of differences in the distribution of earnings: 1994–2006
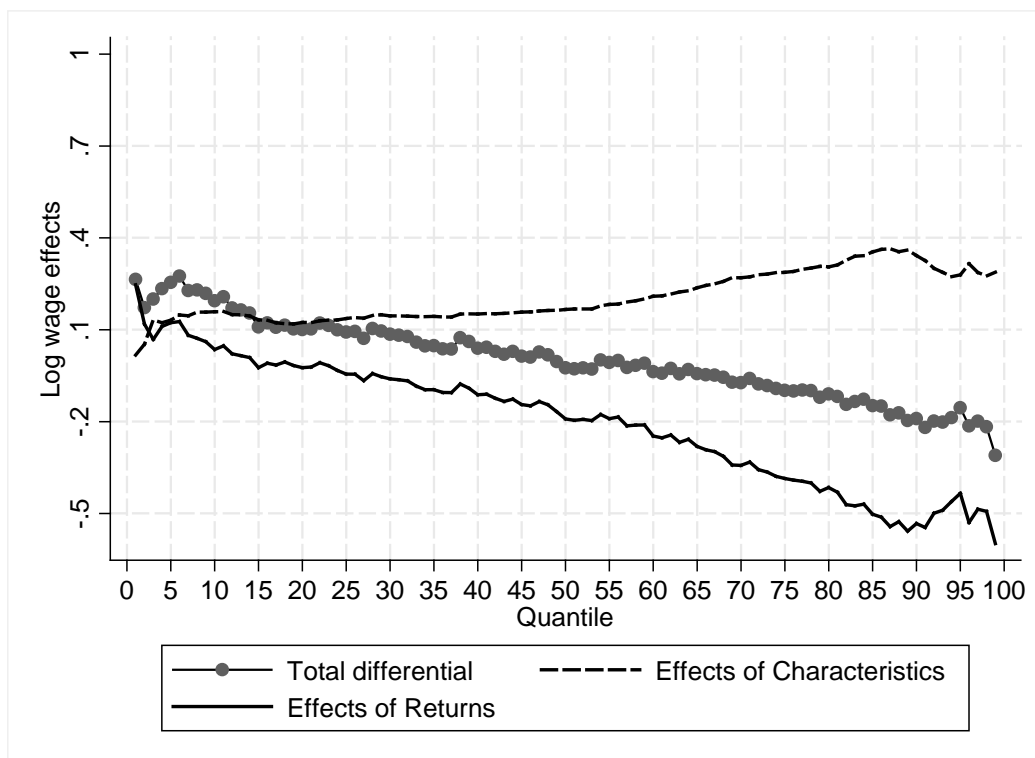
Notes: Calculations using ENIGH. Total differential is the total change in hourly wages (in logs); Effects of characteristics and Effects of returns are the portions that can be ascribed to changes in characteristics and returns, respectively.

Source: Campos-Vazquez et al. (2014: figure 7.4).

---

[34] As explained by Campos-Vazquez et al. (2014): 'Although there was significant educational upgrading and the distribution of the stock of education became more equal over the entire period under study, whether this change was equalizing or unequalizing depends on the extent of convexity in the returns to education and at what point of the education equalization process the country found itself. Bourguignon, Ferreira and Lustig (2005) were among the first to notice that a reduction in the inequality of education—in the presence of increasing returns to education—could lead to a rise in earnings inequality. They call this result the "paradox of progress" alluding to the fact that a more equal stock of education can be inequality-increasing (at least during part of the educational upgrading process) if the returns to education increase at an increasing rate with the level of attainment (convexity in the returns). As Gasparini et al. (2015) write, the "paradox of progress" has been quite a pervasive phenomenon in Latin American labour markets in the last couple of decades.' The 'paradox of progress' had also been found in the analysis for Mexico by Bouillon et al. (2005).

Although the RIF method does not disaggregate the returns into their various components, the result shown in Figure 11 is consistent with the fall in the relative returns to education shown in Figure 9, where it can be seen that the relative supply of college-educated (skilled) workers rose substantially during this period while the relative returns declined. This means that: (i) supply of skilled labour during this period outpaced demand; (ii) institutional factors moved in favour of the unskilled; or (iii) both.[35] Figure 10 shows that the real minimum wage and the unionization rate remained largely constant during this period. Thus, changes in institutional determinants cannot drive the decrease in wage inequality.[36] As Campos-Vazquez et al. (2014) suggest, the change in the skill premium during this period is the result of a combination of a rising supply of workers with college education and a slow-down in demand for skilled workers. So, what drove the fall in demand for skilled workers?

Robertson (2004, 2007) argues that although trade benefited more skilled-intensive industries in the 1980s and early 1990s, with NAFTA this process was reversed. After NAFTA, the relative price of tradable goods continued to decline over time. This potentially explains the decline in the skill premium given that NAFTA favoured skill-intensive industries. Thus, this process drives in part the decline in wage inequality.

Other explanations that have been proposed for the decline in inequality include international migration, labour market distortions, and technical change that may have hurt older skilled workers. Migration increased during this period, probably due to the large negative effects of the 1995 crisis. Mishra (2007) shows that the increase of migration to the USA by low-skilled workers caused a decrease in their relative supply (holding everything else constant), which in a traditional supply and demand model would increase their wages. Other research has shown that misallocation across firms induced by labour market distortions may have contributed to the decline in labour income inequality.

Levy and López-Calva (2016) argue that these distortions limit the growth of the high-productivity sectors, which are also more skill-intensive. As a result, there is a 'surplus' of workers with post-secondary education, who end up having to work in low-productivity firms, where their wages are lower. The misallocation of workers with high levels of education into low-productivity firms may be one of the drivers of the fall in absolute wages for college-educated workers and the stagnation of wages at the bottom.[37] Campos-Vazquez et al. (2016) explore the reasons behind the decline in absolute wages for college-educated workers. They observe that older cohorts are worse affected than younger cohorts and argue that the displacement of older educated workers may have been a result not only of technological change making skilled workers redundant but also of younger workers, who can be paid lower wages, being more adept in the use of the new technologies.

Hence, it seems that the changes in the composition of output induced by NAFTA, the misallocation of skilled labour because of labour market distortions, and the characteristics of technological change were behind the slow-down in demand growth for skilled workers.

---

[35] The gross enrolment rate almost doubled in the period 1994–2004 (Campos-Vazquez 2013).

[36] There is evidence that the minimum wage is currently not binding, and has not been binding since the mid-1990s. For a detailed and recent explanation of the role of minimum wages in Mexico see Escobar Toledo (2014).

[37] Halliday et al. (2016) obtain similar results using firm heterogeneity.

5   Labour income inequality and proximate determinants: 2006–2017

As shown in Figure 7, from 2006 onwards the Gini coefficient for labour income steadily declined with the original (uncorrected) ENOE survey but rose slightly with ENIGH. Thus, the two sources lead to almost opposite narratives as to the evolution of labour income inequality in this period. As shown in Figure 1, however, there has been a sharp rise in item non-response in ENOE since 2006 for both formal and informal workers (roughly a third of workers in 2017). In section 2 we explained how the ENOE data were corrected through a combination of the hot deck imputation method to 'assign' an income to non-respondents and a post-survey weight adjustment using IMSS tabulations. Although ENIGH's item non-response is lower and constant, we also produced results correcting with the same methods, which are available upon request.

The effect of the correction on labour income inequality trends is notable. Figure 12 shows the growth incidence curves for ENOE and ENIGH. With the corrected ENOE we no longer observe that the decline in average labour income rises with income from decile 6 onwards. The decline remains broadly the same for deciles 6–8, and becomes smaller afterwards. Labour income in decile 10 actually grows instead of declining.

Figure 12: Growth incidence curves: 2006–2017 (ENOE original and corrected; ENIGH original; in constant Mexican pesos of August 2015)
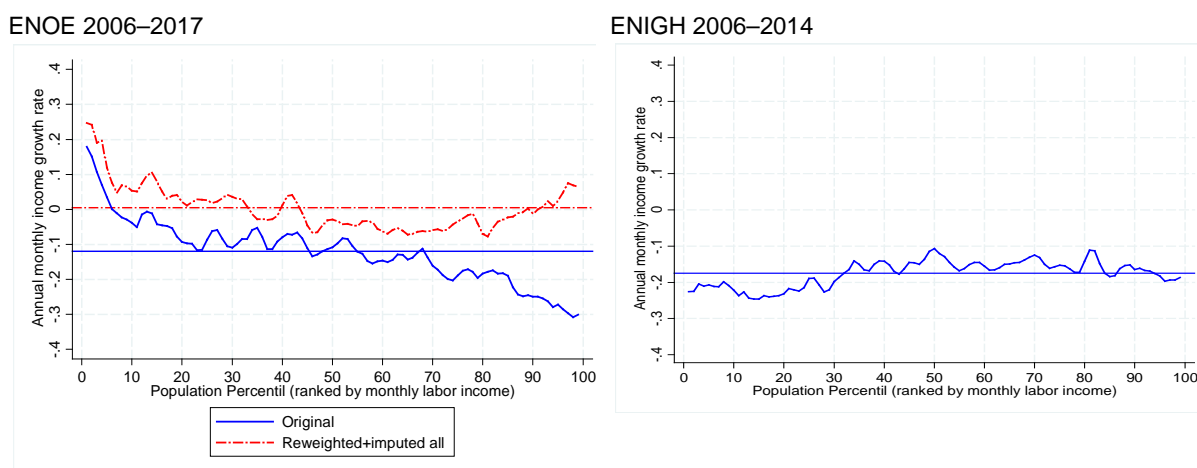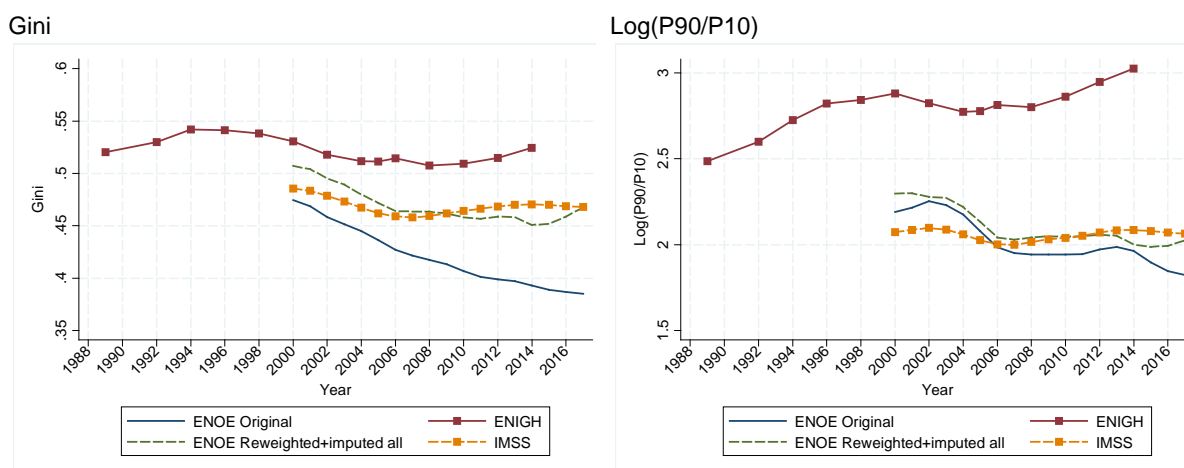

Notes: Sample restricted to all workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. Smoothed lines with a simple moving average with weights 0.4 for the current observation and 0.3 for the lead and lag. Horizontal lines are the respective averages. For post-survey reweighting for formal sector workers in ENOE we use IMSS data. To correct for item non-response for informal workers in ENOE, we follow a hot deck imputation procedure using gender, regions, informal status, age, and education groups.

Source: Authors' construction.

As expected, summary inequality indicators change as well. In Figure 13 and Table 2, it can be observed that labour income inequality is higher for the corrected ENOE survey (with both the Gini coefficient and Log 90/10 ratio). In 2017, the Gini coefficient with the original ENOE survey equals .382, while it equals .464 with corrected data. An unsolved puzzle is that even for the corrected survey, ENOE shows much lower inequality than ENIGH. In 2014, ENIGH's Gini coefficient for labour income for all workers equals .523, while it is .451 with ENOE corrected (Table 2).

Figure 13: Labour income inequality: 1989–2017 (Original ENIGH and ENOE and corrected ENOE surveys)

Notes: Sample restricted to all workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. Smoothed lines with a simple moving average with weights 0.4 for the current observation and 0.3 for the lead and lag. For post-survey reweighting for formal sector workers in ENOE we use IMSS data. To correct for item non-response for informal workers in ENOE, we follow a hot deck imputation procedure using gender, regions, informal status, age, and education groups.

Source: Authors' construction.

For our purposes, an important result to note is the effect of the correction on labour income inequality. The corrected ENOE data no longer show the sharp fall observed in Figure 7 between 2006 and 2017 (Table 2). Between 2006 and 2014, ENIGH shows an increase of about 1 Gini point and, with the corrected survey, ENOE shows a decline of about the same size between 2006 and 2017. (If one were to apply the same correction method to ENIGH, however, the increase in labour income inequality during this period would be of 4 Gini points and the difference between the two sources would reach roughly 10 Gini points!) In sum, the evolution of inequality in the period after 2006 remains a puzzle unsolved. The most one can say with certainty is that labour income inequality definitely did not continue to decline.[38]

---

[38] We calculated inequality using all workers and only salaried workers and found that the difference stems primarily from the self-employed because the inequality levels for salaried workers are much more similar.

Table 2: Gini coefficient for labour income for all workers: 1989–2017 (ENIGH, ENOE (original and corrected) and IMSS)

	ENIGH	ENOE				IMSS (from tabulations for formal workers only)
		Original	Reweighted	Reweighted + imputed informal	Reweighted + imputed all	
1989	0.500					
1992	0.541					
1994	0.545					
1996	0.539					
1998	0.541					
2000	0.534	0.476	0.523	0.519	0.513	0.466
2001		0.473	0.513	0.507	0.501	0.482
2002	0.516	0.456	0.507	0.503	0.498	0.479
2003		0.447	0.493	0.489	0.485	0.474
2004	0.504	0.452	0.496	0.490	0.485	0.466
2005	0.517	0.433	0.480	0.473	0.467	0.463
2006	0.511	0.424	0.472	0.468	0.464	0.460
2007		0.424	0.475	0.468	0.461	0.459
2008	0.516	0.416	0.475	0.468	0.466	0.461
2009		0.413	0.475	0.465	0.461	0.463
2010	0.491	0.410	0.471	0.462	0.458	0.465
2011		0.395	0.469	0.458	0.454	0.470
2012	0.526	0.400	0.469	0.459	0.458	0.471
2013		0.402	0.477	0.464	0.464	0.474
2014	0.523	0.388	0.473	0.459	0.451	0.473
2015		0.390	0.468	0.455	0.438	0.473
2016		0.388	0.482	0.468	0.471	0.472
2017		0.382	0.472	0.459	0.464	0.470

Notes: Sample restricted to workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December, and 2017, when we report the value for June. IMSS data are censored at twenty-five minimum wages or more. For post-survey reweighting for formal sector workers in ENOE we use IMSS data. To correct for item non-response for informal workers in ENOE, we follow a hot deck imputation procedure using gender, regions, informal status, age, and education groups.

Reweighted: frequency distribution of formal workers who responded the question on labour income is adjusted to match that found in IMSS.

Reweighted + imputed informal: in addition to reweighting, we corrected for item non-response using the hot deck imputation method for informal workers only.

Reweighted + imputed all: we first used the hot deck imputation method on both formal and informal workers and then applied the reweighting to formal sector workers. For more details see Section 2.

Source: Authors' construction.

In order to assess the contribution of proximate determinants such as the changes in characteristics and the changes in returns to the changes in labour income inequality, we estimated the RIF regression described in Section 3 for the ENOE corrected survey and ENIGH original survey. Results are shown in Figure 14. With ENOE corrected, the effect of characteristics continues to be unequalizing, as in the previous two periods. The returns effect continues to be equalizing but is no longer monotonical; it now mimics what happened along the income distribution. With ENIGH, the characteristics effect is no longer unequalizing but flat. The returns effect is slightly unequalizing because it is negative—and more strongly so for the lower centiles. The difference in results depending on the source is another puzzle that remains unsolved. Until the evolution of labour incomes along the distribution can be better assessed, this puzzle will remain.

Figure 14: RIF decomposition of inequality: 2006–2017 (corrected ENOE and original ENIGH)
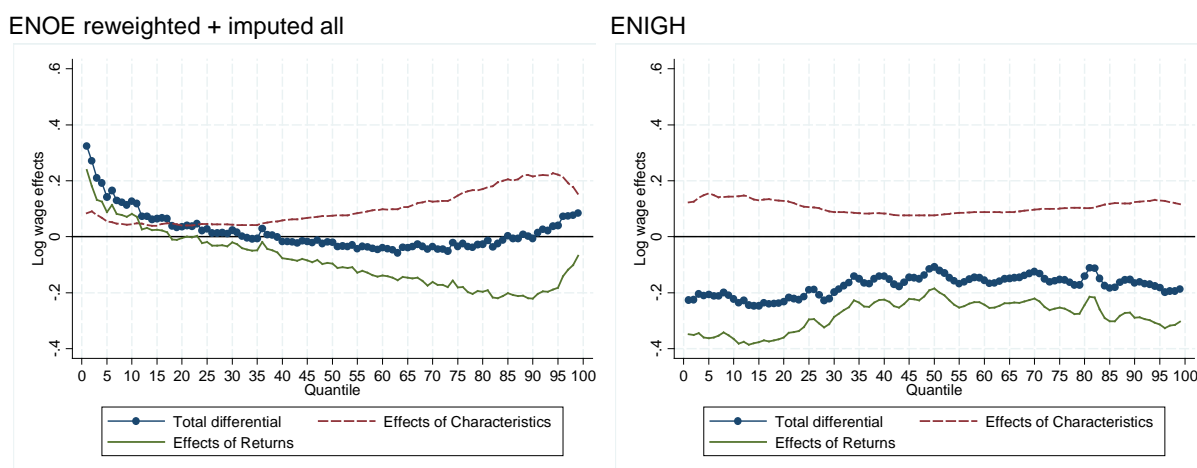
Notes: RIF decomposition method proposed by Firpo et al. (2009), which decomposes the change in the monthly wage in characteristics and returns. The y-axis is the change in log wage between 2006 and 2017 and the x-axis shows the effect for each quantile. Smoothed lines with a simple moving average with weights 0.4 for the current observation and 0.3 for the lead and lag. For corrections of surveys, see Section 2.

Source: Authors' construction.

6   Conclusions

In this paper, we analyse the evolution of labour income inequality and its determinants in Mexico between 1989 and 2017. For the period prior to 2006, we identify two periods: 1989–1994, when inequality increased, and 1994–2006, when inequality declined. We review existing research on the role played by the skill premium (returns to education, in particular) in explaining these trends, and the extent to which trade patterns and technology affected relative returns.

For the period 1989–1994, there is a consensus that trade benefited high-skilled workers. The key channel through which trade affected wages was the establishment of maquiladoras and the decline or elimination of tariffs in low-skill-intensive industries. This combined with the erosion of institutional forces, including a declining rate of unionization and a falling minimum wage. During this period, although the share of workers with college education increased, it did so at a slower

pace than in subsequent periods. After NAFTA came into effect, inequality declined up to 2006.[39] In contrast with the earlier period, the skill premium fell. Since institutional factors such as the minimum wage and the unionization rate remained constant during the entire period, the fall in the skill premium was associated with the fact that the supply of skilled workers outpaced demand.

After 2006, however, the measurement of labour inequality by itself becomes a challenge because the household income expenditure survey (ENIGH) shows a slight increase (Gini increases from .511 in 2006 to .523 in 2014) while the labour force survey (ENOE, without any corrections) shows a steady and sharp decline (from .424 to .388, in the same period). There is reason to believe that ENOE's results are not accurate because the survey has not only a higher but also an increasing rate of item non-response (i.e. workers who do not report earnings). The proportion of workers who do not report their labour income reaches about a third of surveyed workers in 2017. We attempt to correct this problem by using a combination of the hot deck imputation method for all workers to get rid of non-response and a post-survey weight adjustment for formal sector workers. The post-survey reweighting uses recently released tabulations by the Mexican social security administration, IMSS. The new weights are such that the distribution of workers by categories of age, gender, and multiples of the minimum wage are the same in IMSS and ENOE.

The corrected ENOE no longer shows a steady decline in inequality: the Gini coefficient in both 2006 and 2017 is equal to .464. Thus, the most one can say is that after 2006 inequality stopped its downward trend. An unsolved puzzle is that even for the corrected survey, ENOE shows much lower inequality than ENIGH. In 2014, ENIGH's Gini coefficient for labour income for all workers equals .523, while it is .451 with corrected ENOE. If we were to compare corrected ENOE with an equally (same methods) corrected ENIGH, the difference would be around a whopping 10 Gini points owing to the inclusion of the self-employed.

Using the RIF decomposition we find that during this period, the effect for characteristics is still unequalizing and the wage structure effect (returns) is equalizing. Given the remaining uncertainty over the inequality trends during this period, we prefer to refrain from further speculating on the underlying causes of a change whose size remains unknown.

---

[39] NAFTA was signed between Canada, Mexico, and the United States at the end of 1993 and it came into effect on 1 January 1994.

References

Airola, J., and C. Juhn (2005). 'Wage Inequality in Post-reform Mexico'. IZA Discussion Papers 1525. Bonn: Institute for the Study of Labour (IZA).

Almeida dos Reis, J., and R. Paes de Barros (1991). 'Wage Inequality and the Distribution of Education: A Study of the Evolution of Regional Differences in Inequality in Metropolitan Brazil'. *Journal of Development Economics*, 36(1): 117–43.

Atkin, David (2016). 'Endogenous Skill Acquisition and Export Manufacturing in Mexico'. *American Economic Review*, 106(8): 2046–85.

Autor, David H., and Lawrence Katz (1999). 'Changes in the Wage Structure and Earnings Inequality'. In Orley Ashefelter and David Card (eds), *Handbook of Labour Economics*, Volume 3A. Amsterdam: Elsevier, 1463–555.

Biemer, Paul P., and Sharon L. Christ (2008). 'Weighting Survey Data'. In Edith D. de Leeuw, Joop J. Hox, and Don A. Dillman (eds), *International Handbook of Survey Methodology*. Hove, UK: Psychology Press, Taylor and Francis Group.

Borraz, Fernando, and Jose Ernesto Lopez-Cordova (2007). 'Has Globalization Deepened Income Inequality in Mexico?' *Global Economy Journal*, 7(1): 1–57.

Bosch, Mariano, and Marco Manacorda (2010). 'Minimum Wages and Earnings Inequality in Urban Mexico'. *American Economic Journal: Applied Economics*, 2(4): 128–49.

Bouillon, Cesar Patricio (2000a). 'Inequality and Mexico's Labour Market after Trade Reform'. Inter-American Development Bank and Georgetown University Technical Report.

Bouillon, Cesar Patricio (2000b). 'Returns to Education, Sector Premiums, and Male Wage Inequality in Mexico'. Technical Report, Inter-American Development Bank.

Bouillon, César, Arianna Legovini, and Nora Lustig (1999). 'Can Education Explain Income Inequality Changes in Mexico?' Inter-American Development Bank Draft.

Bound, John, and George Johnson (1992). 'Changes in the Structure of Wages in the 1980s: An Evaluation of Alternative Explanations'. *The American Economic Review*, 82(3): 371–92.

Bourguignon, François, Francisco H.G. Ferreira, and Nora Lustig (2005). *The Microeconomics of Income Distribution Dynamics in East Asia and Latin America*. Washington, DC: Oxford University Press.

Campos-Vazquez, Raymundo M. (2013). 'Why Did Wage Inequality Decrease in Mexico after NAFTA?' *Economía Mexicana Nueva Época*, 22(2): 245–78.

Campos-Vazquez, Raymundo, Gerardo Esquivel, and Nora Lustig (2014). 'The Rise and Fall of Income Inequality in Mexico, 1989–2010'. In Giovanni Andrea Cornia (ed.), *Falling Inequality in Latin America: Policy Changes and Lessons*. New York: Oxford University Press.

Campos-Vazquez, Raymundo, Emmanuel Chávez, and Gerardo Esquivel (2015). 'A New Approach for Top Income Shares: Mexico Since the 1990s'. Paper presented to the Latin American and Caribbean Economic Association, Santa Cruz de la Sierra, Bolivia.

Campos-Vazquez, Raymundo, Luis F. López-Calva, and Nora Lustig (2016). 'Declining Wages for College-educated Workers in Mexico: Are Younger or Older Cohorts Hurt the Most?' *Revista de Economía Mundial*, 43: 93–112.

Cortez, Willy Walter (2001). 'What is Behind Increasing Wage Inequality in Mexico?' *World Development*, 29(11): 1905–22.

Cortez, Willy Walter (2005). 'Dispersión y Estabilidad de las Diferencias Salariales Interestatales en México, 1984-2000'. *Investigación Económica*, 64(253): 123–58.

Cragg, Michael Ian, and Mario Epelbaum (1996). 'Why Has Wage Dispersion Grown in Mexico? Is It the Incidence of Reforms or the Growing Demand for Skills?' *Journal of Development Economics*, 51(1): 99–116.

De Hoyos, Rafael E. (2013). 'The Effects of Trade Expansion: Poverty and Inequality in Post-NAFTA Mexico'. *Journal of CENTRUM Cathedra: The Business and Economics Research Journal*, 6(1): 103–27.

Del Castillo, Miguel (2015). 'La magnitud de la desigualdad en el ingreso y la riqueza en México: una propuesta de cálculo'. *Serie Estudios y Perspectivas*, 167.

DiNardo, J., N.M. Fortin, and T. Lemieux (1996). 'Labour Market Institutions and the Distribution of Wages, 1973–1992: A Semiparametric Approach'. *Econometrica*, 64(5): 1001–44.

Escobar Toledo, Saúl (2014). 'Salarios Mínimos: Desiguladad y Desarrollo'. *Economía UNAM*, 11(3): 94–109.

Esquivel, Gerardo (2011). 'The Dynamics of Income Inequality in Mexico since NAFTA'. *Economía*, 12(1): 155–79.

Esquivel, Gerardo, and José Antonio Rodríguez-López (2003). 'Technology, Trade, and Wage Inequality in Mexico Before and After NAFTA'. *Journal of Development Economics*, 72(2): 543–65.

Fairris, David (2003). 'Unions and Wage Inequality in Mexico'. *Industrial and Labour Relations Review*, 56(3): 481–97.

Fairris, David, and Edward Levine (2004). 'La disminución del poder sindical en Mexico'. *El Trimestre Económico*, 71(284): 847–76.

Fairris, David, Gurleen Popli, and Eduardo Zepeda (2008). 'Minimum Wages and the Wage Structure in Mexico'. *Review of Social Economy*, 66(2): 181–208.

Feenstra, Robert, and Gordon Hanson (1997). 'Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras'. *Journal of International Economics*, 42(3): 371–93.

Feliciano, Zadia M. (2001). 'Workers and Trade Liberalization: the Impact of Trade Reforms in Mexico on Wages and Employment'. *Industrial and Labour Relations Review*, 55(1): 95–115.

Firpo, S., N. Fortin, and T. Lemieux (2009). 'Unconditional Quantile Regressions'. *Econometrica*, 77(3): 953–73.

Fortin, N., T. Lemieux, and S. Firpo (2011). 'Decomposition Methods in Economics'. In D. Card and O. Ashenfelter (eds), *Handbook of Labour Economics*. Amsterdam: North Holland, 2–104.

Friedman, Santiago, Arianna Legovini, and Nora Lustig (1995). 'Mexico: Social Spending and Food Subsidies During Adjustment in the 1980s'. In Nora Lustig (ed.), *Coping with Austerity*. Washington, DC: Brookings Institution Press.

Garduño-Rivera, Rafael (2010). 'Effect of NAFTA on Mexico's Income Distribution in the Presence of Migration'. Paper presented at the Joint Annual Meeting, Agricultural & Applied Economics Association, Denver, Colorado.

Gasparini, Leonardo C., and Mariana Marchionni (eds) (2015). *Bridging Gender Gaps? The Rise and Deceleration of Female Labour Force Participation in Latin America.* Centro de Estudios Distributivos, Labourales y Sociales, Faculdad de Ciencias Económicas, Universidad de la Plata.

Goldin, Claudia, and Lawrence F. Katz (2009). 'The Race between Education and Technology'. Cambridge, MA, and London: Harvard University Press.

Halliday, Timothy, Daniel Lederman, and Raymond Robertson (2016). 'Tracking Wage Inequality Trends with Prices and Different Trade Models: Evidence from Mexico'. IZA Discussion Paper 10156.

Hanson, Gordon H. (2003). 'What Has Happened to Wages in Mexico Since NAFTA? Implications for Hemispheric Free Trade'. NBER Working Paper 9563.

Hanson, Gordon, and Ann Harrison (1999). 'Trade Liberalization and Wage Inequality in Mexico'. *ILR Review*, 52(2): 271–88.

Juhn, C., K. Murphy, and B. Pierce (1993). 'Wage Inequality and the Rise in Returns to Skill'. *Journal of Political Economy*, 3(3): 410–48.

Kaplan, David S., and Francisco Pérez Arce Novaro (2006). 'El Efecto de los Salarios Mínimos en los Ingresos Labourales en México'. *El Trimestre Económico*, 73(289): 139–73.

Kurokawa, Yoshinori (2011). 'Is a Skill Intensity Reversal a Mere Theoretical Curiosum? Evidence from the US and Mexico'. *Economics Letters*, 112(2): 151–54.

Levy, Santiago, and Luis Felipe López-Calva (2016). 'Labour Earnings, Misallocation, and the Returns to Education in Mexico'. IDB Working Paper Series 671.

Little, Roderick J.A., and Donald B. Rubin (2002). *Statistical Analysis with Missing Data*, Second Edition, Wiley Series in Probability and Statistics. New Jersey: John Wiley & Sons.

López-Acevedo, Gladys (2004). 'Mexico: Evolution of Earnings Inequality and Rates of Returns to Education (1988–2002)'. *Estudios Económicos*, 38(19): 211–84.

López-Acevedo, Gladys (2006). 'Mexico: Two Decades of the Evolution of Education and Inequality'. World Bank Policy Research Working Paper 3919.

Machado, J.A.F., and J. Mata (2005). 'Counterfactual Decomposition of Changes in Wage Distributions using Quantile Regression'. *Journal of Applied Econometrics*, 20(4): 445–65.

Manacorda, M., C. Sánchez-Páramo, and N. Schady (2010). 'Changes in Returns to Education in Latin America: The Role of Demand and Supply of Skills'. *ILR Review*, 63(2): 307–26.

Meza Gozález, Liliana (1999). 'Cambios en la Estructura Salarial de México en el Período 1988-1993 y el Aumento en el Rendimiento de la Educación Superior'. *El Trimestre Económico*, 66(262): 189–226.

Mishra, Prachi (2007). 'Emigration and Wages in Source Countries: Evidence from Mexico'. *Journal of Development Economics*, 82(1): 180–99.

Piketty, Thomas, Li Yang, and Gabriel Zucman (2016). 'Capital Accumulation, Private Property and Rising Inequality in China 1978–2015,' WID.world Working Paper.

Popli, Gurleen K. (2011). 'Changes in Human Capital and Wage Inequality in Mexico'. *Oxford Development Studies*, 39(3): 369–87.

Revenga, Ana (1997). 'Employment and Wage Effects of Trade Liberalization: The Case of Mexican Manufacturing'. *Journal of labour Economics*, 15(S3), S20–S43.

Robertson, Raymond (2004). 'Relative Prices and Wage Inequality: Evidence from Mexico'. *Journal of International Economics*, 64(2): 387–409.

Robertson, Raymond (2007). 'Trade and Wages: Two Puzzles from Mexico'. *The World Economy*, 30(9): 1378–98.

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

Appendix

Table A1: Question on labour income: ENEU, ENE, ENOE, and ENIGH

Survey	Years	Labour income or total income	Hours	Question
ENEU	1989–1999	Labour income	Yes	In your last week's principal source of employment: How often do you get paid? How much did you get paid or how much do you estimate were your earnings? If you didn't work last week, how much do you earn regularly? (En el trabajo principal de la semana pasada; ¿cada cuánto obtiene sus ingresos o le pagan? ¿Cuánto ganó o en cuánto calcula sus ingresos? Si no trabajo la semana pasada, ¿cuánto gana normalmente?)
ENE	2000–2004	Labour income	Yes	In your last week's principal source of employment: How often do you get paid? How much did you get paid or how much do you estimate were your earnings? If you didn't work last week, how much do you earn regularly? (En el trabajo principal de la semana pasada; ¿cada cuánto obtiene sus ingresos o le pagan? ¿Cuánto ganó o en cuánto calcula sus ingresos? Si no trabajo la semana pasada, ¿cuánto gana normalmente?)
ENOE	2005–2016	Labour income	Yes	How often do you get paid? How much did you get paid or how much do you estimate were your earnings? (¿Cada cuándo obtiene... sus ingresos o le pagan? ¿Cuánto ganó o en cuánto calcula sus ingresos?)
ENIGH	1989, 1992, 1994, 1996, 1998, 2000	Both	Yes	How much income did you receive last month for … ? (El mes pasado ¿cuánto percibió por … ?)
ENIGH	2002	Both	Yes	How much money did you receive for (SOURCE OF INCOME), in the month of … ? (¿Cuánto dinero recibió por (CONCEPTO) en el mes de … ?)
ENIGH	2004, 2005, 2006, 2008, 2010, 2012, 2014	Both	Yes	How much income did you receive for (SOURCE OF INCOME), in the month of … ? (¿Cuánto dinero recibió por (CONCEPTO) en el mes de … ?)

Notes: ENEU: Spanish acronym for National Survey of Urban Employment; ENE: Spanish acronym for National Employment Survey; ENOE: Spanish acronym for National Occupation and Employment Survey; ENIGH: Spanish acronym for Household Income Expenditure National Survey.

Source: Authors' construction based on information in the surveys' questionnaires.
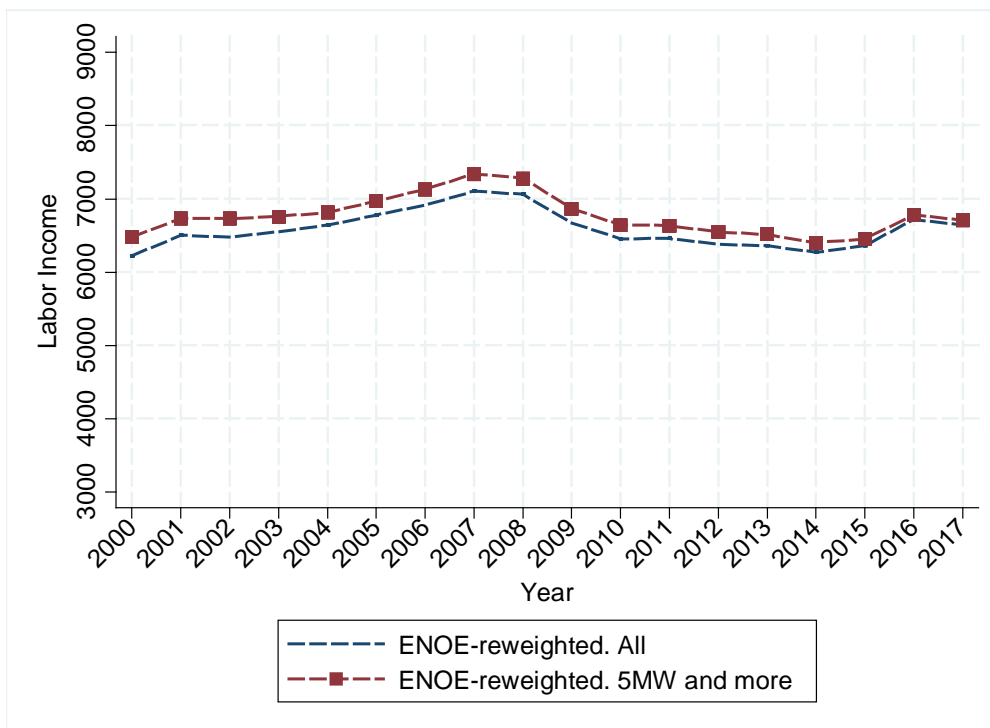
Table A2: Sample size ENOE and ENIGH: 1989–2017 (number of households surveyed)

Year	ENOE	ENIGH
1989		11,531
1992		10,530
1994		12,815
1996		14,042
1998		10,952
2000	143,444	10,108
2001	149,427	
2002	147,246	17,167
2003	138,014	
2004	106,954	22,595
2005	104,083	23,174
2006	106,854	20,875
2007	107,468	
2008	106,145	29,468
2009	104,355	
2010	105,342	27,655
2011	104,842	
2012	105,236	9,002
2013	103,905	
2014	107,274	19,419
2015	107,960	
2016	107,318	
2017	108,757	

Notes: ENOE: Spanish acronym for National Occupation and Employment Survey; ENIGH: Spanish acronym for Household Income Expenditure National Survey. For ENOE we use the second quarter of each year.

Source: Authors' construction.

Figure A1: Robustness check: average monthly labour income for formal sector workers reweighting all vs reweighting for workers earning five minimum wages or more: 2000–2017

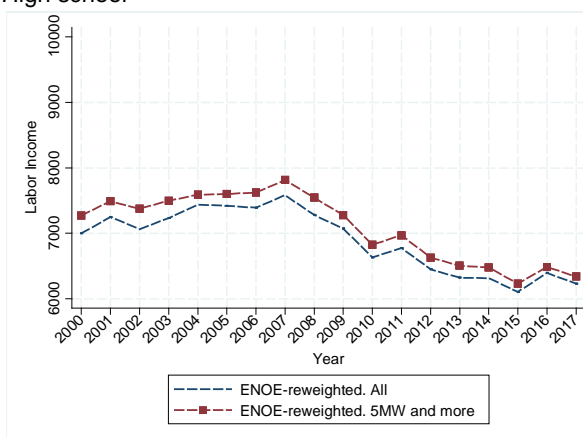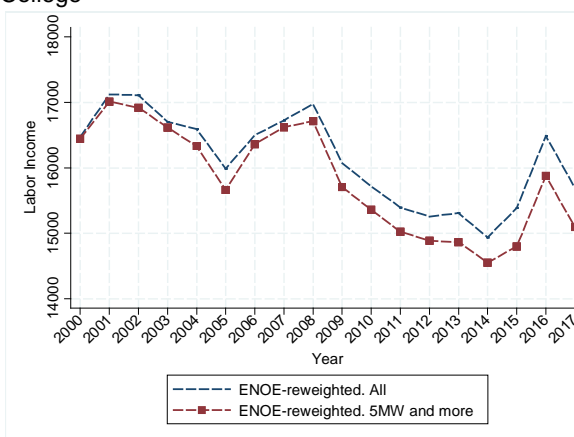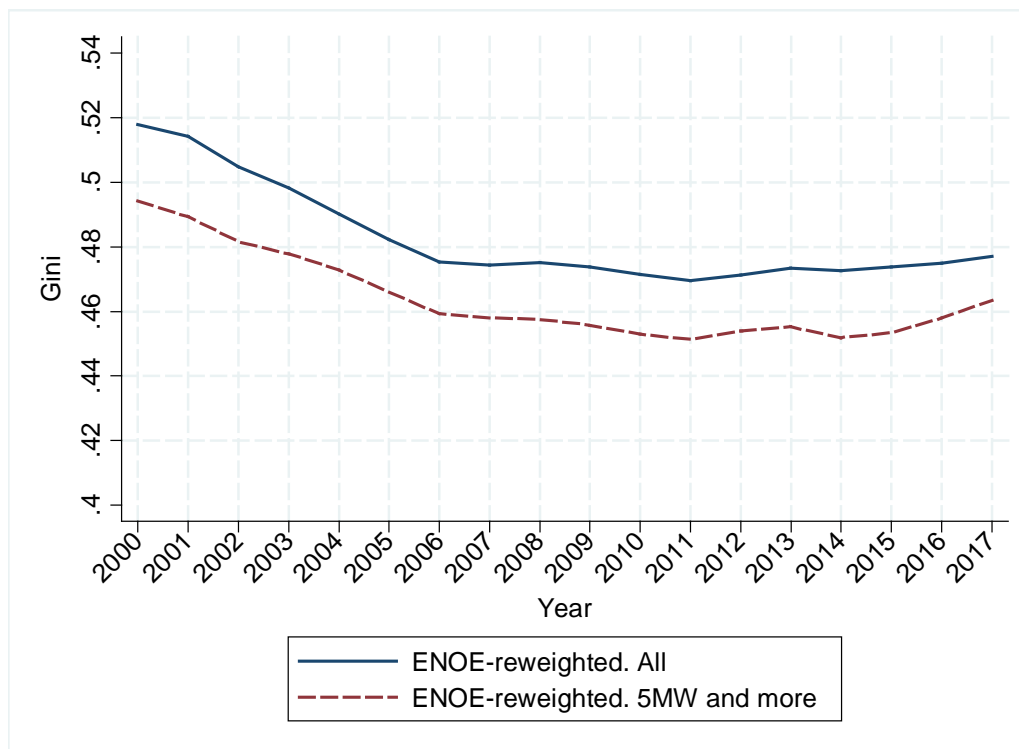Notes: The line with the legend 'All' uses the reweighting procedure for all formal sector workers as described in Section 2 of main text. The line with the legend '5MW and more' reweights according to IMSS data only for workers with labour income at least five times the minimum wage. In other words, we do not modify the sampling weights for workers with labour income less than five times the minimum wage in ENOE, assuming that for poorer workers ENOE is the 'true' distribution. Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December. For reweighting we use IMSS data.

Source: Authors' construction.

Figure A2: Robustness check: average monthly labour income for formal sector workers reweighting all vs reweighting for workers earning five minimum wages or more, by education level: 2000–2017


Notes: The line with the legend 'All' uses the reweighting procedure for all formal sector workers as described in Section 2 of main text. The line with the legend '5MW and more' reweights according to IMSS data only for workers with labour income at least five times the minimum wage. In other words, we do not modify the sampling weights for workers with labour income less than five times the minimum wage in ENOE, assuming that for poorer workers ENOE is the 'true' distribution. Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December. For reweighting we use IMSS data.

Source: Authors' construction.

Figure A3: Robustness check: Gini coefficient for average monthly labour income for formal sector workers reweighting all vs reweighting for workers earning five minimum wages or more: 2000–2017

Notes: The line with the legend 'All' uses the reweighting procedure for all formal sector workers as described in Section 2 of main text. The line with the legend '5MW and more' reweights according to IMSS data only for workers with labour income at least five times the minimum wage. In other words, we do not modify the sampling weights for workers with labour income less than five times the minimum wage in ENOE, assuming that for poorer workers ENOE is the 'true' distribution. Sample restricted to formal sector workers aged 20–64 years with positive labour income and working hours. For ENOE we use the second quarter of each year. For IMSS data we report the value for the month of April of each indicated year except for 2000 and 2001, when we report the value for the month of December. For reweighting we use IMSS data.

Source: Authors' construction.

Figure A4: Robustness check: average monthly labour income and inequality measures for formal sector workers. comparing reweighting vs. rescaling: 2000–2017
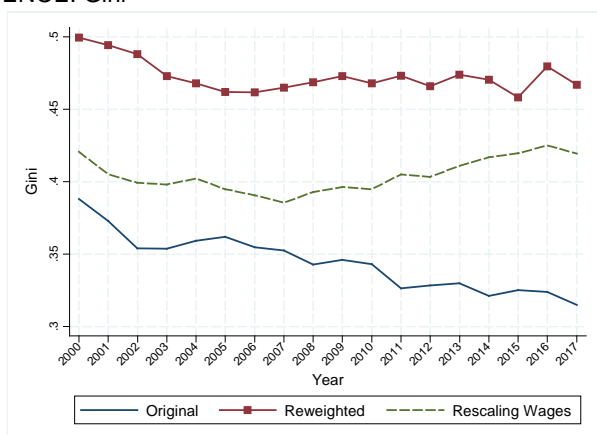
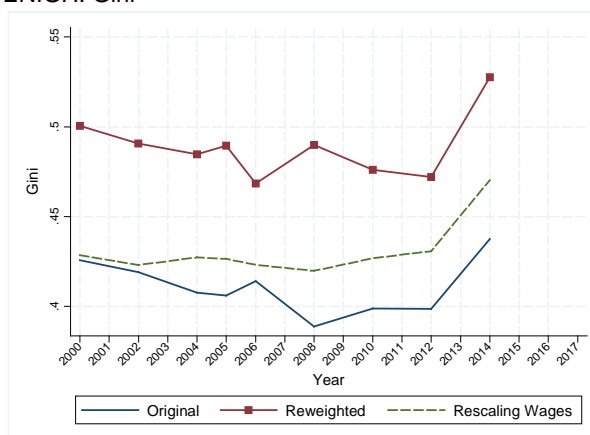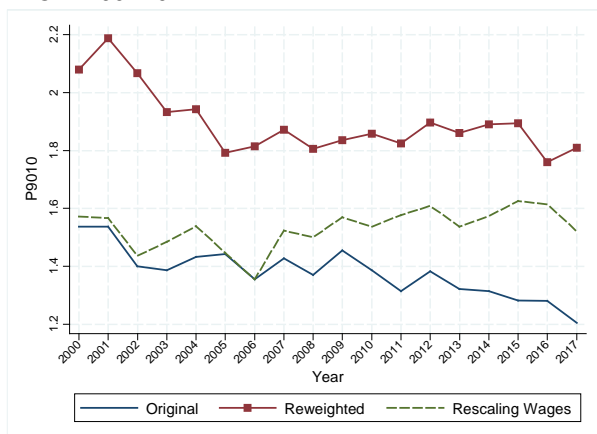ENOE: Average labour income
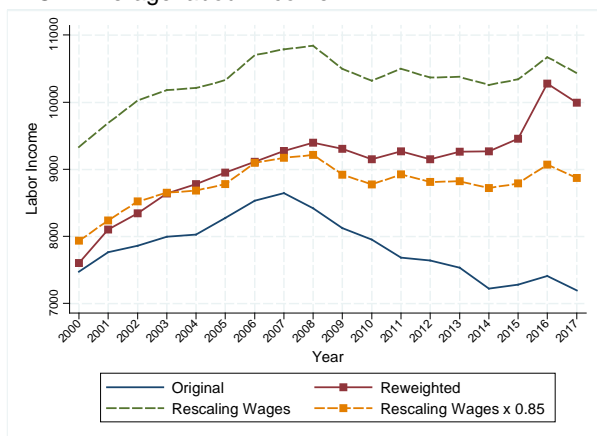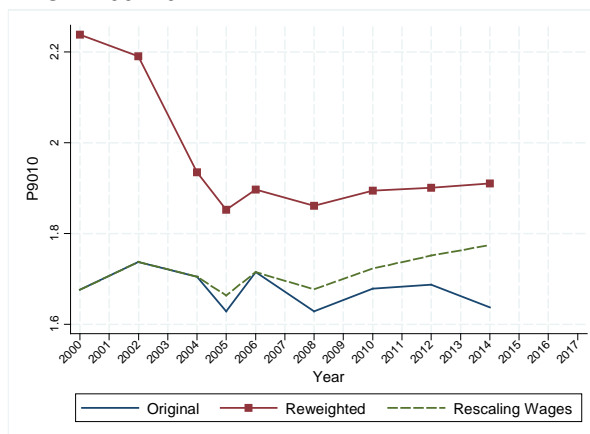

ENIGH: Average labour income
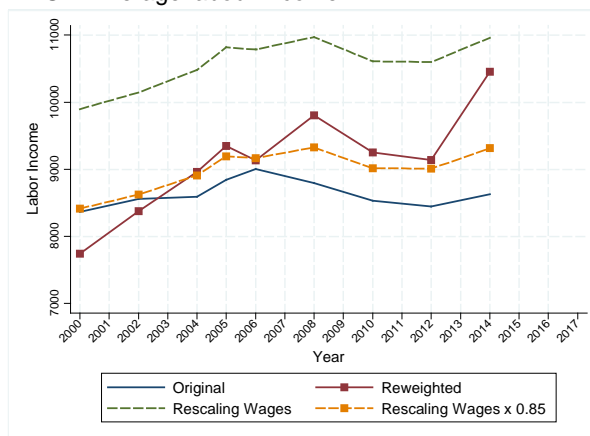
ENOE: Gini

ENIGH: Gini



ENOE: P90P10

ENIGH: P90P10

Notes: Samples restricted to formal sector workers. Reweighting procedure explained in text. 'Rescaling wages' procedure consists in the following: first, we obtain the average wage by percentile from administrative data and household surveys (ENOE and ENIGH); second, we obtain the ratio of those earnings by percentile (wages in administrative data over wages in each household survey); third, we multiply by 0.85 to account for social security contributions so we are left with wages net of contributions in both sources; fourth, we multiply this ratio by the observed wage in each percentile and household survey; fifth, if the ratio is less than 1 we leave it as equal to 1 (in other words, for these cases, it is assumed that the information in the survey is accurate).
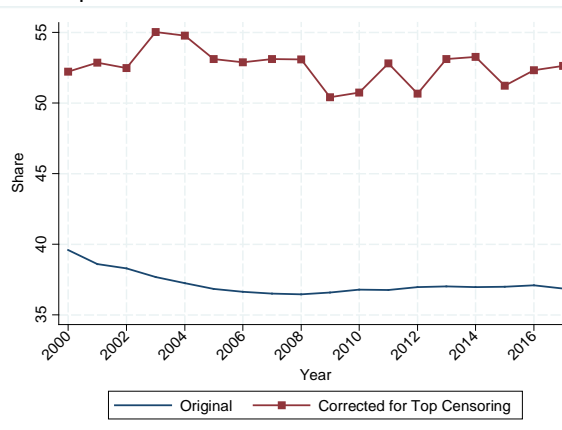
Source: Authors' construction.

36

Figure A5: Robustness check: IMSS monthly labour income, top share, Gini coefficient and Theil index: original data and data corrected for censoring at the top
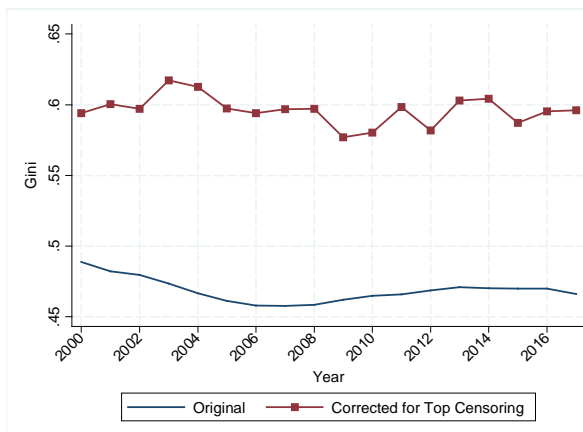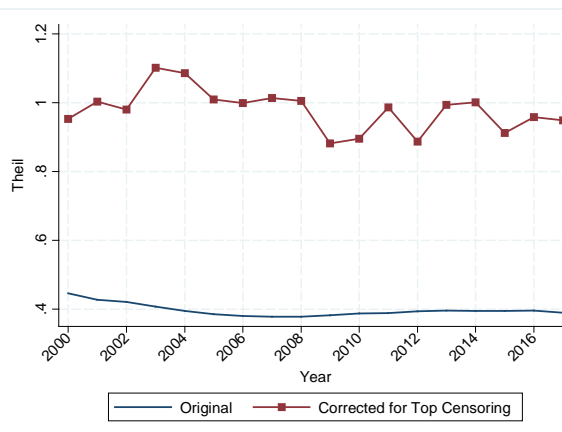
Labour income
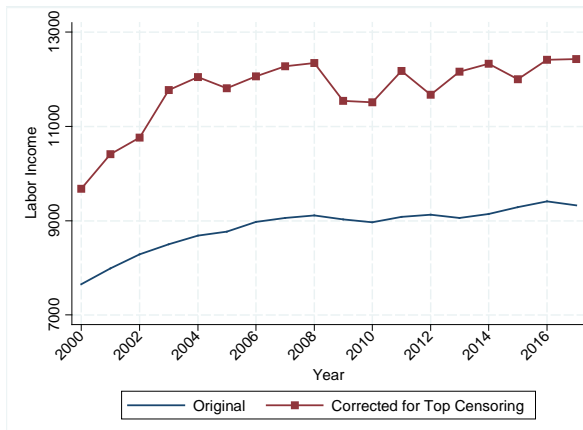
Share top 10%

Gini coefficient

Theil

Notes: Sample restricted to workers aged 20–64 years. Data from IMSS. 'Original' refers to estimates using the data with no correction. 'Corrected for Top Censoring' includes a correction using a Pareto distribution at the top. Wages in the administrative data are censored at twenty-five times the minimum wage. We divided the censored wages into ten groups and imputed an average wage according to the Pareto distribution. We obtained the Pareto coefficient using the workers earning more than twenty times the minimum wage and those in the censoring cut-off.

Source: Authors' construction.