Munich Personal RePEc Archive

The rising wage inequality in Mexico, 1984-2000: A distributional analysis

Gurleen Popli

2006

Online at *http://mpra.ub.uni-muenchen.de/399/*
MPRA Paper No. 399, posted 11. October 2006 / 16:26

Rising Wage Inequality in Mexico, 1984-2000: A Distributional Analysis

Gurleen K. Popli*
Department of Economics
University of Sheffield
9 Mappin Street
Sheffield, S1 4DT, UK
g.popli@shef.ac.uk
Tel: ++44-114-2223485
Fax: ++44-114-2223558

ABSTRACT

In this paper we look at the distribution of wages to examine the extent and cause of the increasing wage inequality in Mexico over the last two decades (1984 to 2000). To understand the causes of the increase in inequality over time we do a counterfactual analysis. We find that over the last two decades not only did the inequality increase, there also was an erosion of real wages, and it's the middle class which was affected the most. Main reason for the decrease in real wages was the declining unionization in the country. While the main reason for the rise in inequality was the changing distribution of skills.

Keywords: kernel density estimation, counterfactual distribution , unions, trade liberalization, changing distribution of skills

JEL Classification: C14, J31

March 2006

*I am highly indebted to Aman Ullah, David Fairris and Robert R. Russell for comments and suggestions on the successive drafts of the paper. I would also like to thank Indrnail Dutta, Zuzana Janko, Sarah Brown, the editor, three anonymous referees, and seminar participants at the University of California, Riverside, and the University of East Anglia, for their helpful comments. All errors are mine.

1 Introduction

Mexico signed GATT in 1986 and NAFTA in 1994, with two free trade agreements in as many decades it was expected that as the Mexican product markets integrate with the rest of the world the factor prices would also converge, particularly with its rich neighbor, the US (Lederman *et al.*, 'Lessons from NAFTA', World Bank, 2003). Standard neoclassical trade theories suggested that as the country, which is abundant in unskilled labour, opens itself to trade the relative returns to the unskilled workers would increase, lowering the wage inequality in the economy. Instead two decades on, we observe a rise in the relative returns to the skilled workers leading to increased wage inequality (Hanson, 2003).

In this paper we look at the distribution of wages to examine the extent and cause of the increasing wage inequality in Mexico over the last two decades. We bring the institutional and trade arguments in one framework, and compare their relative contribution to the rise in inequality. To quantify the contribution of different factors to increase in inequality we do counterfactual analysis in the spirit of work done by Oaxaca (1973).

Oaxaca's work focused on 'means' and was based on questions such as, 'how much would a worker, with the mean characteristics of 1984 workforce, have been paid in 2000?' Generalization of Oaxaca's method to decompose distributions gives us the 'counterfactual distributions'. This involves asking questions like, 'what would the distribution of wages be if the distribution of the individual attributes had remained as in 1984 and workers had been paid according to the wage schedule observed in 2000?'. DiNardo, Fortin and Lemieux (1996), DFL henceforth, proposed a semiparametric method to estimate these counterfactual distributions, which we will be using here.

The literature on Mexico has typically attributed the increased inequality

to two broad factors: (1) trade liberalization and foreign direct investment, and (2) changing labor market institutions. The first, argues that trade liberalization and increased foreign direct investment have lead to an increase in the relative demand for skilled labor, leading to increased relative returns to education, which in turn has lead to an increase in wage inequality in the country (Cragg and Epelbaum, 1996; Feenstra and Hanson, 1997; Hanson and Harrison, 1999). The second view relies on the declining power of unions (Panagides *et al.*, 1994; Fairris, 2003) and falling real value of the minimum wage (Cortez, 2001; Fairris *et al.*, 2005) as the causes of rising inequality.

There are three main limitations of the current papers. First, most of the studies rely on summary measures for their analysis - measures of inequality, mean wage differentials by industries and skill levels. Summary measures of inequality, though good starting points, are limited in the information they convey; they often give conflicting results;[1] further, decomposition of inequality indices into different factors, contributing to the change in inequality, can be sensitive to the measure used.[2]

Second, while the changes in the returns to education and the increasing relative demand for the skilled labor has been analyzed in the literature, nothing has been said about the changing distribution of the skills, and its impact on the wage distribution.

Third, most studies in this area are limited both in the horizon they cover and the data set they use. Majority of the studies have looked at only the 1980s and the early half of the 1990s (exceptions being, Hanson, 2003; and Esquivel *et al.*, 2003). Feenstra and Hanson (1997) and Hanson

[1] "...they [inequality measures] will rank the same set of distributions in different ways, simply because of their differing sensitivity to incomes in different parts of the distributions." Litchfield (1999).

[2] Shorrocks (1982) points out the sensitivity of decomposition analysis to the choice of the inequality index. There are regression based methods also available to decompose inequality, they too suffer from the same limitations as the decomposition of inequality statistics (see Fields, 2002).

and Harrison (1999) in their study used macro-survey data of manufacturing plants; Cragg and Epelbaum (1996) in their analysis use micro level data for only 16 urban areas; Esquivel *et al.* (2003) look only at the manufacturing sector; Hanson (2003) in the only exception, he uses a 1% random sample from 1990 and 2000 census data. The implications that one can draw from these studies are thus limited in scope.

This paper is different from the other studies not only in looking at a longer horizon but also in the methodology used. Looking at the longer horizon helps us study the impact of both the liberalization of the 1980s and free trade agreements of 1990s, on the wage inequality. We use micro (household) level data, which is nationally representative, covers larger share of population, and has more detailed information on the workers characteristics. More importantly, instead of looking only at the summary measures of inequality we analyze the entire distribution of wages.

Looking at how the distributions have changed over time, and an indication of which segments of the distribution have been effected the most over time, can be important in understanding the factors responsible for the observed changes. By focusing on the distributions we are able to comment on, and observe, multiple features of the wage distribution and not just the mean, variance, or some other summary statistic.[3] While there are a number of methods available to decompose the summary measures that characterize a distribution, DFL paper is seminal in providing a way to decompose the whole distribution.[4]

In section 2 we give details of the data used in this paper and an explanation of how the actual and the counterfactual distributions are estimated. The results of the empirical exercise are put forth in section 3 of the paper, we focus on three explanatory factors: (i) changing levels of unionization;

[3]Jenkins (1995) and Cameron (2000) provide empirical evidence in support of using distributions as opposed to summary measures of inequality.

[4]Lemieux (2002) and Jenkins and van Kermer (2005) offer alternative approaches, derived from the original DFL approach.

(ii) changing industrial affiliations of the workers; and (iii) the changing distribution of skills. Section 4 concludes the paper.

2 Data and methodology

The data used for the analysis is from Encuesta National de Ingresos y Gastos de los Hogares (ENIGH). These are the national household surveys that began in 1984 and continued in 1989, 1992, and every two years thereafter. The number of households interviewed by ENIGH is different for each year of the survey; the sample analyzed here, for each year, ranges from 15% to 20% of the total ENIGH sample. This is the only Mexican data set that gives information on the union status of the workers. The survey employs a 'stratified sampling' technique, so we use sample weights made available by ENIGH in the analysis below.

The sample utilized in this study is only of the working individuals from the surveyed households. We look only at the wage earners and drop the self-employed from our sample. The main reasons for dropping the self-employed are: the focus of this study is on the labor market institutions (unions) and returns to skills. Concept of unions is only relevant for the wage earners (the survey does not ask the self-employed about their union status). Further, one cannot distinguish between the returns to skills and the returns to capital for the self-employed.[5]

The earnings variable is the hourly wage, and is computed from the reported earnings (net of taxes) during the month before the survey and reported hours of work in the week before the survey. To insure an accurate measure of the wage all those who are working without pay or those who hold more than one job are also excluded.[6] In the estimate of the wage, no

[5] According to McKenzie and Woodruff (2003) returns to investment are substantial in Mexico, particularly at low levels (15% a month for investment levels below US$ 200).

[6] In the earlier years of the survey information for the union status was collected only for the primary job of the respondent. This is another reason to exclude those with more

fringe benefits, tips, bonuses or commissions are included. Throughout the study we use real wages. To obtain real wages nominal wages are deflated by the National Consumer Price Index, reported in the Annual Report (1996 and 2001), published by Banco de Mexico.

In Table 1 we report the descriptive statistics for the years 1984 to 2000.[7] In 1984 the median wage was 3.4% higher than the mean wage; by year 2000, the median wage was 6% below the mean wage, indicating a change in the skewness of the distribution from negatively skewed to positively skewed. The average real wages over this period fell by 7% and the median real wages fell by 15%; most of the fall in real wages is a result of the currency crisis of 1994.

Table 1

Descriptive statistics

	1984	**1989**	**1992**	**1994**	**1996**	**1998**	**2000**
Mean log real wages	3.743	3.722	3.768	3.854	3.464	3.495	3.674
Median log real wages	3.777	3.687	3.737	3.657	3.399	3.411	3.615
SD of log real wages	0.765	0.780	0.806	0.838	0.827	0.845	0.815
Gini coefficient	0.113	0.114	0.121	0.124	0.137	0.136	0.127
Coefficient of variation	0.204	0.210	0.214	0.217	0.239	0.242	0.222
Number of observations	3644	10358	8925	10891	11935	9299	8824

Three summary measures of inequality are reported here, standard deviation (SD) of log real wages, Gini coefficient and coefficient of variation. Whatever measure of inequality we use, wage inequality in Mexico increased between 1984 and 2000. Between 1984 and 2000 the SD of log real wages increased by 6.5%, coefficient of variation increased by 8.8% and the Gini

than one job.

[7]The broad trends in inequality are robust to the inclusion of the self-employed in the sample. This is supported by the findings of Airola and Juhn (2005).

coefficient increased by 12%. Increase in inequality seems to relent in the mid 1990s, though the cut off period of when the increase stopped is different depending on the measure of inequality used.[8]

In the subsequent analysis we focus only on the two end periods – 1984 and 2000; the horizon is long enough to consider the effect of almost all major changes faced by the economy – the 1982 debt crisis, signing of GATT in 1986 and NAFTA in 1994, and three currency crisis. Choosing 2000 as the end date means we are well beyond the last big currency crisis of 1994 and the severe recession of 1995[9]; it also means we do not have to arbitrarily pick the date when the inequality stabilized, as all measures indicate some stabilization by 2000.

In Figure 1 the kernel density estimates of log real wages of 1984 and 2000 are shown. A clear change in the shape of the distribution's is evident. Both the distributions are unimodal but the peak density of the distribution falls dramatically from 1984 to 2000. The reduced peak indicates fewer individuals working in the middle group, giving an indication of 'shrinking middle-class' in Mexico.[10] The 2000 distribution is to the left of the 1984 distribution and shows higher dispersion. Decrease in mean and increase in inequality by itself should lead to an increase in lower tail of the distribution, what we see here instead is an increase in mass in the upper tail – indicating that the distribution has also become positively skewed over time. Any explanation for increase in inequality should also be able to explain the above mentioned changes in the distribution of wages over time.

Before we embark on the analysis of what explains this increase in inequality an explanation of notation and the estimation of the nonparametric distributions is provided.

[8]Gonzalez and McKinley (1997) also point out the conflicting results given by the different summary measures of inequality, for Mexico.

[9]In 1995 GDP fell by 6.2%. Source: Banco de Mexico.

[10]Decline of the middle class is a hypothesis studied for both the UK and the US, using parametric and non-parametric approaches, see Jenkins (1995), Bradbury (1986), and references therein.

2.1 Estimating the counterfactual distributions

Consider a vector (w, z), where w is wages and z is a vector of individual attributes, such as age, gender, union status, education etc.. Let the joint distribution of w and z at any time t be represented as $F(w, z; t_w = t_z = t)$, where t_w and t_z indicate the date of w and z (here t). The density of wages at any point in time,$f_t(w)$, can be written as:

$$\begin{aligned} f_t(w) &= \int_z dF(w, z; t_w = t_z = t) \\ &= \int_z f(w|z; t_w = t)dF(z; t_z = t) \\ &\equiv f(w; t_w = t_z = t). \end{aligned} \tag{1}$$

The last identity in (1) is notational, it shows that the distribution of w is defined in period t, conditional on the distribution of z in the period t. Using the notation in (1), the actual density of wages in year 2000 will be written as: $f(w; t_w = t_z = 2000)$. Estimate of the distribution as expressed in (1) is given by,

$$\widehat{f_h}(w_j) = \frac{1}{n}\sum_{i=1}^{n}\frac{\theta_i}{h}K\left(\frac{w_j - W_i}{h}\right), \tag{2}$$

where (2) represents the kernel density estimate of a univariate distribution based on a random sample $(W_1, ..., W_n)$ of size n. θ_i are the sample weights, $K(.)$ is the kernel function that depends on the distance of W_i from w_j, and the sample size through h, which is the window width. In the empirical work, we use the Gaussian kernel and $h = 1.06(\sigma_w)n^{-1/5}$, where σ_w is the standard deviation of the random sample.[11]

[11]Estimation of the distribution, given in equation (2), depends on the choice of h and K. Of the two its the choice of h that is more critical. For further details on nonparametric estimation of density functions and the choice of different h and K refer to Silverman (1986) and Chapter 2, Pagan and Ullah (1999).

Estimation of the counterfactual distribution involves holding z or different components of it (such as union status) at the earlier year (1984) levels - i.e., to estimate what the distribution of wages in 2000 would be if the distribution of z had remained the same as in 1984. In terms of notation in (1), we are interested in:

$$f^c(w) \equiv f(w; t_w = 2000, t_z = 1984). \tag{3}$$

The distribution as expressed in (3) is the hypothetical counterfactual distribution, where superscript c is used to denote the counterfactual distribution. The hypothetical density $f^c(w)$ is given as,

$$f^c(w) = \int_z f(w|z; t_w = t) dF(z; t_z = s), \tag{4}$$

where s is any time period different from t ($s \neq t$).

To estimate the counterfactual distribution re-write (4) as:

$$f^c(w) = \int_z f(w|z; t_w = t) \Psi(z) dF(z; t_z = t), \tag{5}$$

where $\Psi(z)$ is a reweighting function defined as:

$$\Psi(z) = \frac{dF(z; t_z = s)}{dF(z; t_z = t)}. \tag{6}$$

Equation (5) is now identical to (1), with the exception of the reweighting function. Once the estimate of this reweighting function is available, the counterfactual distribution can be estimated as:

$$\widehat{f}^c_h(w_j) = \frac{1}{n} \sum_{i=1}^{n} \frac{\theta_i \widehat{\Psi}_i}{h} K\left(\frac{w_j - W_i}{h}\right). \tag{7}$$

What exactly does this reweighting function do? What is the economic meaning of this function? Before answering these questions we introduce

some more notation. In our analysis below we will focus on the following individual characteristics: u a union dummy, where 1 represents union members and 0 otherwise; e indicates industrial affiliation, where $e = 1$ if the individual is employed in the industrial sector and 0 otherwise; x is a vector of individuals age (proxy for experience) and education, which taken together constitute labor market skills.[12] We split the vector of individual characteristics: $z \equiv (u, e, x, g, r)$, where g, the gender dummy, equals 1 for males and 0 otherwise, and r indicates the region of residence.[13] z_{-k} denotes the vector of individual characteristics z without the variable k. For example, z_{-u} indicates vector z without u, so $z_{-u} \equiv (e, x, g, r)$.[14]

To illustrate what the reweighting function does consider the impact of changing levels of unionization on the wage distribution. For the impact of de-unionization we generate a counterfactual distribution that would have prevailed in 2000, had the union levels remained as they were in 1984, keeping the distribution of all other individual characteristics as in 2000, and the workers getting paid according to the wage schedule observed in 2000. In terms of notation in (4) the distribution we want is:

$$f(w; t_w = t_{z_{-u}} = 2000, t_{u|z_{-u}} = 1984) \tag{8}$$
$$= \int\int f(w|u, z_{-u}; t_w = 2000) dF(u|z_{-u}; t_{u|z_{-u}} = 1984) dF(z_{-u}|t_{z_{-u}} = 2000).$$

Which can be written as,

[12] While age is a good proxy for labor market experience for men, it is likely to overestimate the actual experience for women. Alternate definition of skill was used, where skill is captured only by the education, the results are not qualitatively different from those presented here.

[13] Details of all the individual characteristics with their definition are reported in Table (A1), Appendix A of the paper.

[14] Elements of vector z_{-k} are control variables. Analysis was done separately by gender but the results are not qualitatively different from those presented here for the whole sample.

$$\int\int f(w|u, z_{-u}; t_w = 2000)\Psi_u(z)dF(u|z_{-u}; t_{u|z_{-u}} = 2000)dF(z_{-u}|t_{z_{-u}} = 2000), \tag{9}$$

where the reweighting function $\Psi_u(z)$ is defined as,

$$\begin{aligned} \Psi_u(z) &\equiv \frac{dF(u|z_{-u}; t_{u|z_{-u}} = 1984)}{dF(u|z_{-u}; t_{u|z_{-u}} = 2000)} \\ &= u\left(\frac{Pr(u = 1|z_{-u}; t_{u|z_{-u}} = 1984)}{Pr(u = 1|z_{-u}; t_{u|z_{-u}} = 2000)}\right) \\ &\quad +(1-u)\left(\frac{Pr(u = 0|z_{-u}; t_{u|z_{-u}} = 1984)}{Pr(u = 0|z_{-u}; t_{u|z_{-u}} = 2000)}\right). \end{aligned} \tag{10}$$

The first line identity in (10) is obtained by substituting the expression on the right hand side into (9) and canceling out the denominator. The second line equality is derived by noting that u only takes values 0 or 1.

This weight represents the change in the probability between 1984 and 2000 that an individual, defined by characteristics z_{-u}, is a union member or not. The reason we control for the individual characteristics z_{-u} is to keep the relationship between unionization and the individual characteristics at the 1984 level. All this reweighting function does is, it up-weights the individuals in the union sector by a factor that is proportional to the decrease in unionization in the economy and similarly down-weights the non-union sector employees.

The conditional probabilities in the reweighting function are estimated using a probit model:

$$Pr(u = 1|z_{-u}; t_{u|z_{-u}} = t) = Pr(\epsilon > -\beta' H(z_{-u})) = 1 - \Phi(-\beta' H(z_{-u})). \tag{11}$$

The above equation is estimated for both 1984 and 2000. The estimated

coefficients $\widehat{\beta}$ for each year are retained. To obtain the fitted probabilities in (10): for the numerator the 1984 coefficients and for the denominator the 2000 coefficients are used with the 2000 sample, respectively.

3 Explaining the inequality

3.1 Institutions: De-unionization

We start with the institutional explanation proposed in the literature. The institutional factors cover the role of unions and the impact of minimum wage on the wages. In our analysis we will focus only on changing unionization levels. This is not to say that analysis of minimum wage is not important. Analysis of minimum wage in distributional setting is not straightforward for Mexico. Minimum wages in Mexico are set by regions and occupations, there is no unique national minimum wage that one can work with. Thus the analysis of minimum wage deserves a separate analysis by itself (see Fairris *et al.* 2005; Maloney *et al.*, 2001).

Impact of decline of labor unions in Mexico has been explored by Fairris (2003 and 2005), where the author finds declining unionization and the decreasing bargaining power of the unions to be an important factor in explaining the increase in inequality in Mexico. The decreasing union membership and declining union power can not only help explain the increasing inequality but also shed light on the decline in middle class jobs. The argument given by Davis and Huston (1992) is that "unions convert jobs that probably weren't middle-income jobs based on the skills required, into middle-income jobs." Decline in the unions pushes people from the middle of the distribution to the lower tail of the distribution, as these people were in the middle only because of the union action.

Another effect that the unions have is to reduce the overall dispersion in the wage distribution. Unions usually take wages out of competition and

depress the wage structure with respect to the productive characteristics (such as education and occupation) thus lowering the overall dispersion. Once unions lose their power and/or the unionization rates go down, the wage structure is less leveled, thus increasing dispersion. In either case while some workers move from the middle to the lower tail of the distribution the others might move from the middle to the upper tail.

In our sample 25.4% of the individuals were unionized in 1984. This number decreased to 16% by 2000. We compare the average wages and the standard deviation of wages across the two groups in Table 2.

Average wage in the union sector is higher than the average wage in the non-union sector in both the years, with gap remaining almost the same over time. Dispersion in wages, as measured by standard deviation of wages, is higher for the non-union sector compared with the union sector, this gap however has decreased (almost halved) over time. Increase in inequality was also higher for the union sector. Standard deviation of wages in the union sector increased by almost 29% from 1984 to 2000, that for the nonunion sector increased by only 3.87%.

Table 2

Union vs. Non-union workers

Mean and Standard deviation (SD) of log real wages

	Union Members		Non-union Members	
	Mean	SD	Mean	SD
1984	4.199	0.506	3.588	0.776
2000	4.192	0.652	3.575	0.806

Based on the evidence of the standard deviation of wages, unions are not as strong a force of equalizing wages in 2000 as they were in 1984, it would seem we have evidence not only for the declining union membership,

but also of declining union power.[15] To see whether, and to what extent, de-unionization has lead to the observed changes in the distribution of wages we generate a counterfactual distribution.[16]

Figure 2, gives the counterfactual distribution adjusted for the union effects. The counterfactual distribution of 2000 is to the right of the actual 2000 distribution. If the union levels had remained as they were in 1984 and the workers were paid according to the wage schedule of 2000, then the real wages in year 2000 would have been higher than those actually observed in 2000.

3.2 Trade liberalization

Starting from 1985 Mexico has seen radical trade reforms and liberalization (Lustig, 1998 and 2001). Import licensing has almost been eliminated, tariff and quota restrictions have been substantially reduced. Adjustments in the labor market, as a result of these reforms, are reflected through changes in the employment and wages. For Mexico most of the adjustment took place in form of declining real wages. Downward flexibility of real wages meant less aggregate employment effects.[17] Here we will look at the employment and wage effects separately.

3.2.1 Employment shifts

Decreasing employment share of the industrial (particularly manufacturing) sector jobs is often cited as the leading cause of decline in the middle-class jobs and an increase in inequality (Bluestone and Harrison, 1988; Davis and

[15] Over the 1980s and the 1990s government intervention in the labor market weakened the labor movement in Mexico, which contributed to the rise in wage inequality, see Gonzalez and McKinley (1997) and Fairris (2005).

[16] Looking at the de-unionization rates will not be able to fully capture the selection bias or the general equilibrium (spillover) effect of unions, see Lewis (1986) and Card (1996) for details.

[17] Refer to Revenga (1997) and Feliciano (2001) for the details of the employment and wage effects of trade liberalization in Mexico.

Huston, 1992; Valletta, 1997). The argument is that as the high technology and service jobs increase the overall inequality also increases; these service sector jobs, which tend to have high variance, are replacing the basic manufacturing jobs, which are the mainstay of the blue-collar middle class and which tend to have low variance.[18]

De-industrialization (or the decreasing employment in the industrial sector) is often associated with high-income countries, so this argument might make sense in countries like US, but not so in developing countries like Mexico. This argument is particularly difficult to make for Mexico, given the trade liberalization that the economy faced over this period. If anything the employment in the industrial sector, which is also largely the tradeable sector of the economy, should have increased.

We start by looking at the employment shares, average wages and standard deviation of wages across the different sectors of the economy in our sample, reported in Table 3. The share of employment in the industrial sector over this period increased by 6.2%. All sectors saw a decrease in the real wages over this period, however the largest decline in the real wages was observed for the industrial sector with real wages decreasing by 13%. Inequality in wages also increased for the industrial and the services sectors by 10% and 8.5%, respectively. The standard deviation of wages in the services sector is higher compared to that in the industries and this gap increased over time.

[18]The academic debate on the issue is however far from over. For example, for US, Murphy and Welch (1993) and Juhn, Murphy and Pierce (1993) find no effect of industry shifts on average wages and the variance of the wage distribution respectively.

Table 3

Descriptive statistics by sectors

Mean and Standard deviation (SD) of log real wages;

Employment shares (%)

	1984			2000		
	Mean	SD	Share	Mean	SD	Share
Agricultural	3.047	0.673	12.60	2.904	0.623	9.22
Industrial	3.798	0.672	33.34	3.665	0.739	35.42
Services	3.871	0.754	54.07	3.808	0.818	55.36

We do have some evidence of industrial shift in Mexico. Trade liberalization has lead to growth in employment in the industrial sector. This by itself should mean more people are employed in the middle class jobs, resulting in lower inequality. However when we look at the wage distribution we see a decline in the middle class jobs. To see where exactly in the distribution the changing nature of industrial employment has its impact we generate a counterfactual distribution.

The question we ask is, 'what would the density of wages be if the industrial employment shares had remained at their 1984 level and workers had been paid according to the wage schedule observed in 2000?'. The distribution we are interested in is:

$$f(w; t_w = t_{z_{-e}} = 2000, t_{e|z_{-e}} = 1984). \tag{12}$$

The detailed expression for this counterfactual distribution will be similar to that in equations (8) and (9), with the reweighting function defined similarly to that in (10), except instead of u, now we will have e.

Figure 3 gives the counterfactual distribution adjusted for the industrial employment shares. The two distributions are not very different, though the counterfactual distribution shows slight decreased mass in the middle.

The decreased mass in the middle for the 2000 counterfactual distribution is consistent with the lower levels of employment in the industrial sector in 1984. The employment effect, however, is not consistent with the overall shift observed in the distribution.

3.2.2 Changing skills

The major impact of liberalization, as mentioned above, was however not in employment adjustments but the wage adjustments. The increased dispersion of wages is largely linked to the increased relative demand for the skilled labor over the last two decades. The increased demand for skilled labor could be due to liberalization and/or skill biased technological changes.[19] While the increasing returns to skill and the demand for skills has been analyzed, not much has been said about the change in the distribution of skills itself.

In our sample over the last two decades the average years of education in Mexico increased from 6.9 years in 1984 to 8.7 years in 2000, over the same period the median years of education increased from 6 to 9 years. The standard deviation of years of schooling increased by only 2% over this period. The average age of the work force over these two decades has gone up from 29 to 31 years. This increase in supply of skilled labor has been noted by others as well, see Hanson (2003) and Airola and Juhn (2005). The increased supply of skilled labor could be due to a number of factors: supply response to demand shifts, government polices and the household decisions to invest in human capital.[20]

[19] Esquivel et al (2003) look at inequality in Mexico over the period 1988-2000. Over the entire period they suggests that it's the skill-biased technological change that has caused an increase in inequality; trade liberalization in fact was an equalizing force before 1994 and had no effect on dispersion of wages after 1994. Given their findings there should be little to no change in inequality prior to 1994, and a large increase after 1994 – this however is not what we observe in the data.

[20] If the demand changes had taken place without the supply response the gap between the returns to high skilled and low skilled workers would have been even greater. Thus any simulation done will actually underestimate the effect of the change in the skill dis-

As the workforce ages and gets more educated the changing skill distribution will change the wage distribution. Higher education and experience means higher returns in the labor market, which means there will be an upward mobility of the people - moving them from the lower tails to the middle and from the middle to the upper tail - the entire distribution shifts to the right. Such a movement of the workforce may or may not cause an increase in dispersion of wages. Robinson (1976) presents a theoretical model where, as the proportion of the educated people in the society increases, the inequality first increases and then starts decreasing (inverted U-shaped relationship). This however holds only if we consider two groups, skilled and unskilled. In case of more than two groups, as in our case with different levels of education and labor market experience, the relationship cannot be predicted analytically.

How has the changing distribution of skills impacted the wage distribution? To capture the impact of the changing skill distribution we generate another counterfactual distribution. We want a distribution of wages such that the individuals have the skill distribution of 1984 but are paid according to the wage schedule of 2000. Such a counterfactual is given by:

$$f(w; t_w = t_{z_{-x}} = 2000, t_{x|z_{-x}} = 1984) \tag{13}$$
$$= \iint f(w|x, z_{-x}; t_w = 2000) dF(x|z_{-x}; t_{x|z_{-x}} = 1984) dF(z_{-x}|t_{z_{-x}} = 2000)$$
$$= \iint f(w|x, z_{-x}; t_w = 2000) \Psi_x(z) dF(x|z_{-x}; t_{x|z_{-x}} = 1984) dF(z_{-x}|t_{z_{-x}} = 2000),$$

where x as defined earlier, is the vector of age, age squared and education dummies. In this case the reweighting function is obtained by applying the Baye's Law:

tribution. (We thank one of the referees for bringing this point to our attention.)

$$\begin{aligned}\Psi_x(z) &\equiv \frac{dF(x|z_{-x};t_{x|z_{-x}}=1984)}{dF(x|z_{-x};t_{x|z_{-x}}=2000)} \\ &= \frac{\Pr(t_x=2000)}{\Pr(t_x=1984)}\frac{\Pr(t_x=1984|x)}{\Pr(t_x=2000|x)}.\end{aligned} \tag{14}$$

This function represents the relative probability of observing an individual with characteristics (skills) x in the 1984 versus 2000, normalized by the conditional probability of being in either sample. To estimate the conditional probabilities, observations from 1984 and 2000 are pooled, a dummy variable for time is generated and a probit model is estimated. The unconditional probabilities are simply the number of observations in year t, divided by the total number of observations.

Figure 4 gives the counterfactual distribution generated by adjusting for the skill distribution. The changes in the distribution of skills over 1984 to 2000 have been such as to move people from the lower tail and the middle of the distribution to the upper tail. This is not surprising, given the increase in skill levels from 1984 to 2000. If the relative returns to the different skills do not change, then as the population gets more educated (skilled) - wages will increase. This increase in wages will result in an upward mobility of workers along the wage distribution.

3.3 Quantitative measures

In this section we discuss the quantitative contributions of the different factors to the changing wage distribution. All the reported statistics are calculated using the estimated actual and counterfactual distributions The results are reported in Table 4 below. The first two rows of Table 4 give the descriptive statistics from the actual distributions for 1984 and 2000; rows 3 to 5 give the descriptive statistics from the counterfactual distributions for 2000; row 6 is the total observed change from 1984 to 2000; and the

subsequent rows give the effect of the different factors. We look at changes in mean, standard deviation and quantile differences. Three quantile differences are analyzed. While 90-10 difference captures the overall dispersion in wages the 50-10 and 90-50 differences capture the effect on the lower and the upper tail of the distribution, respectively. Over this period while the 90-50 wage differential increased the 50-10 wage differential actually decreases.

Table 4

Descriptive statistics from actual and counterfactual log wage distribution

	Mean	SD	90-10[a]	90-50	50-10
Distribution:					
1. Actual - 1984	3.743	0.780	1.904	0.924	0.980
2. Actual - 2000	3.674	0.822	2.023	1.139	0.884
Counterfactual distribution - 2000, controlling for:					
3. Union levels	3.722	0.810	1.999	1.099	0.900
4. Employment shares	3.682	0.838	2.047	1.163	0.884
5. Skill distribution	3.496	0.772	1.856	1.012	0.844
6. Total change (row 2-row 1)	-0.069	0.042	0.119	0.215	-0.096
Effect of:					
7. Unions (row 2-row 3)	-0.049	0.012	0.024	0.040	-0.016
8. Employment (row 2-row 4)	-0.008	-0.016	-0.024	-0.024	0
9. Skills (row 2-row 5)	0.178	0.050	0.167	0.127	0.040

[a] Difference between the 90th and the 10th percentiles of the log wage distribution.

The 50-10 and 90-50 statistics are defined similarly.

The analysis done here is not sequential, when we generate a counterfactual distribution controlling for a particular factor, i.e. keeping it at its 1984 level, all the other factors impacting the wage distribution are assumed

to be at their 2000 level. For example, when we generate a counterfactual, where the distribution of skills is at the 1984 level, everything else - the union levels, employment shares, gender distribution, regional distribution, returns to all these covariates including the returns to skills, are assumed to be at the 2000 level.[21]

The counterfactual for 2000 - controlling for union effects - gives a higher mean and a lower dispersion for 2000, compared to the actual values for 2000. Over the period 1984 to 2000 the mean and the standard deviation of log real wages decreased by 0.069 and 0.042 points respectively. 70% of the decrease in mean and the 28% of the increase in standard deviation can be attributed to de-unionization, *ceteris paribus*. If the union levels had remained as they were in 1984 and the workers were paid according to the wage schedule of 2000, then the real wages in year 2000 would have been higher than those actually observed in 2000; the overall dispersion of wages would have been lower; the 90-10 wage differential also would have been lower; but the 50-10 wage differential would have been higher. While the unions appear to compress the upper tail of the distribution, they seem to increase the distance of the lower tail from the middle.[22]

To disentangle the effect of the changing returns to skills and the increasing levels of skills we look at the counterfactual for 2000, controlling for the skill distribution. Comparing the 2000 actual distribution with the 2000 counterfactual distribution gives us the impact of change in the distribution of skills. As the workers get older and more educated we see an increase in inequality (0.772 to 0.822). This increase in inequality, independent of the changing returns to skills, reflects the higher residual wage dispersion among the older and more educated workers; also reflected in a bigger increase in

[21]Since the analyses are not not sequential in nature, nor are all the factors considered orthogonal to each other, we cannot sum up the contribution of the different factors.

[22]This is similar to the findings for the US, see DFL (Table III) and Freeman and Medoff (1984).

the 90-50 wage differential compared to the 50-10 wage differential.[23]

If the skill levels had remained at their low 1984 levels, and workers paid according to the wage schedule of 2000 the real wages would have fallen even more. The fact that skill levels increased and there was an upward mobility of workers, prevented some of the losses in average real wages.

4 Conclusion

In this paper we presented a distributional analysis of the rising wage inequality in Mexico over the two decades of 1980s and the 1990s. While the distributional approach is an improvement over the summary measures, it too has some limitations. In particular the approach used here ignores the general equilibrium effects of the explanatory variables. This limitation is shared by other approaches as well (for example regression based decompositions), however looking at distributions does further our understanding of the rising wage inequality in Mexico.

To do our analysis we used nonparametric econometrics and focused on entire distribution of wages under different counterfactual scenarios. In this sense this paper is different from other studies on Mexico, which rely only on the summary measures. We considered two main hypotheses proposed in the literature – institutions and trade liberalization.[24]

We find declining unionization to be significant in explaining the decrease in real wages. If the union membership had remained as it was in 1984 and workers were paid according to the wage setting mechanisms of 2000, the

[23] Higher residual variance among the more educated and more experienced workers can be explained by a Mincerian human capital model, see Lemieux (2003) for details. Juhn et al (1993) interpret this residual variance as the 'price' for unmeasured human capitals, which are assumed to be higher for the more educated and experienced workers. We observe higher within-group inequality for other groups as well – inequality within union and nonunion members (Table 2) and within different sectors (Table 3) has also increased; this is consistent with the findings of Gonzalez and McKinley (1997).

[24] As there are no self-employed workers in our sample, all findings are relevant only for the 'wage earners', who form about 65% of the workforce in Mexico (World Bank).

decline in real wages would have been much less. While number of studies in the literature have noted the importance of unions in explaining the rise in dispersion there role in explaining the fall in mean wages has not been noted.

Though the shifts in employment have little effect, they imply a decrease in dispersion. Both the findings are consistent with the literature – the small magnitude of the effect is supported by Feliciano (2001) and Revenga (1997); the direction of the effect is supported by Esquivel et al. (2003).

Numbers of studies have convincingly associated the rise in inequality in Mexico with the changing returns to skills. What we suggest in this paper is that over this period the distribution of skills in Mexico has also changed, and its effect on the inequality should be considered. We find that changes in skill distribution contribute in a big way to the rise in inequality, and deserve further exploration.

References

Airola, Jim and Juhn, Chinhui, 2005. "Wage Inequality in Post-Reform Mexico." IZA Discussion Paper No. 1525.

Banco de Mexico, Annual Report, 1996 and 2001.

Bluestone, Barry, and Bennett Harrison, 1988. "The Growth of Low-Wage Employment: 1963-86." *American Economic Review*, vol. 78, pp. 124-128.

Bradbury, Katharine, 1986. "The Shrinking Middle Class." *New England Economic Review* (September/October).

Cameron, L. A., 2000. "Poverty and Inequality in Java: Examining the Impact of the Changing Age, Educational and Industrial Structure." *Journal of Development Economics*, vol. 62, pp. 149-180.

Card, D., 1996. "The Effects of Unions on the Structure of Wages: A Longitudinal Analysis." *Econometrica*, vol. 64, pp. 957-979.

Cortez, Willy W., 2001. "What is Behind Increasing Wage Inequality in Mexico?" *World Development*, vol. 29, pp. 1905-1922.

Cragg, Michael, and Mario Epelbaum, 1996. "Why has Wage Dispersion grown in Mexico? Is it the Incidence of Reforms or the Growing Demand for Skill?" *Journal of Development Economics*, vol. 51, pp. 99-116.

Davis, Joe C., and John H. Huston, 1992. "The Shrinking Middle-Income Class: A Multivariate Analysis." *Eastern Economic Journal*, vol. 18, pp. 277-285.

DiNardo, John, Nicole M. Fortin, and Thomas Lemieux, 1996. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach." *Econometrica*, vol. 64, pp. 1001-1044.

ENIGH, Encuesta National de Ingresos y Gastos de los Hogares, Mexico City: INEGI. 1984, 1989, 1992, 1994, 1996, 1998, 2000.

Esquivel, Gerardo, and Jose Antonio Rodriguez-Lopez, 2003. "Technology, Trade and Wage Inequality in Mexico before and after NAFTA." *Journal of Development Economics*, vol. 72, pp. 543-565.

Fairris, David, 2003. "Unions and Wage Inequality in Mexico." *Industrial and Labor Relations Review*, vol. 56, pp. 481-497.

Fairris, David, 2005. "What do Unions do in Mexico." Working Paper, University of California, Riverside.

Fairris, David, Gurleen Popli, and Eduardo Zepeda, 2005. "Minimum Wages and the Wage Structure in Mexico." Working Paper, University of California, Riverside.

Feenstra, Robert C., and Gordon H. Hanson, 1997. "Foreign Direct Investment and Relative Wages: Evidence from Mexico's Maquiladoras." *Journal of International Economics*, vol. 42, pp. 371-393.

Feliciano, Zadia M., 2001. "Workers and Trade Liberalization: The Impact of Trade Reforms in Mexico on Wages and Employment." *Industrial and Labor Relations Review*, vol. 55, pp. 95-115.

Fields, Gary, 2002. "Accounting for Income Inequality and its Change: A New Method, with Application to the Distribution of Earnings in the United States," Working Paper, Cornell University.

Freeman, R. B. and J. L. Medoff, 1984. *What do Unions Do?* New York: Basic Books.

Gonzalez, D.A. and T.McKinley, 1997. "The Paradox of Narrowing Wage Differentials and Widening Wage Inequality in Mexico." *Development and Change*, vol. 28, pp. 505-530.

Hanson, Gordon H., 2003. "What has happened to Wages in Mexico since NAFTA? Implication for Hemispheric Free Trade." Working Paper # 9563, NBER.

Hanson, Gordon H., and Ann Harrison, 1999. "Trade Liberalization and Wage Inequality in Mexico." *Industrial and Labor Relations Review*, vol. 52, pp. 271-288.

Jenkins, Stephen, 1995. "Did the Middle Class Shrink during the 1980s? UK Kernel Density Estimates." *Economic Letters*, vol. 49, pp. 407-413.

Jenkins, Stephen, and Philippe van Kerm, 2005. "Accounting for Income Distribution Trends: A Density Function Decomposition Approach." *Journal of Economic Inequality*, vol. 3, pp. 43-61.

Juhn, Chinhui, Kevin M. Murphy, and Brooks Pierce, 1993. "Wage Inequality and the Rise in Returns to Skill." *Journal of Political Economy*, vol. 101, pp. 410-442.

Lederman, Daniel, William F. Maloney and Luis Serven, 2003. "Lessons from NAFTA for Latin American and the Caribbean Countries: A Summary of Research Findings." World Bank.

Lemieux, Thomas, 2002. "Decomposing Changes in Wage Distributions: A Unified Approach." *Canadian Journal of Economics*, vol. 35, pp. 646-688.

Lewis, H. G., 1986. *Union Relative Wage Effects: A Survey.* Chicago, University of Chicago Press.

Litchfield, Julie A., 1999. "Inequality: Methods and Tools." Text for the World Bank PovertyNet website: http://www.worldbank.org/poverty.

Lustig, Nora, 1998. Mexico: The Remaking of an Economy, Second Edition (Brookings Institution Press, Washington DC).

Lustig, Nora, 2001. "Life is not Easy: Mexico's Quest for Stability and Growth." *Journal of Economic Perspectives*, vol. 15, pp. 85-106.

Maloney, W., and Jairo Nunez, 2001. "Measuring the Impact of Minimum Wage: Evidence from Latin America." Working Paper, World Bank.

McKenzie, David and Christopher Woodruff, 2003. "Do Entry Costs Provide an Empirical Basis for Poverty Traps? Evidence from Mexican Microenterprises." BREAD Working Paper #020, Februrary.

Murphy, Kevin M., and Finis Welch, 1993. "Industrial Change and the Rising Importance of Skills." In Sheldon Danzinger and Peter Gorrschalk (eds), Uneven Tides: Rising Inequality in America, (Russell Sage Foundation, New York).

Oaxaca, Ronald, 1973. "Male-Female Wage Differentials in Urban Labor Markets." *International Economic Review*, vol. 14, pp. 693-709.

Pagan, Adrian, and Aman Ullah, 1999. Nonparametric Econometrics (Cambridge University Press).

Panagides, Alexix, and Harry Anthony Patrines, 1994. "Union-Non-union Wage Differentials in the Developing World: A Case Study of Mexico." Policy Research Working Paper # 1269, World Bank.

Revenga, Ana, 1997. "Employment and Wage Effects of Trade Liberalization: The Case of Mexican Manufacturing." *Journal of Labor Economics*, vol. 15, pp. S20-S43.

Robinson, Sherman, 1976. "A Note on the U-Hypothesis Relating Income Inequality and Economic Development." *American Economic Review*, vol. 66, pp. 437-440.

Shorrocks, Anthony, 1982. "Inequality Decomposition by Factor Components." *Econometrica*, vol. 50, pp. 193-211.

Silverman, B.W., 1986. Density estimation for Statistics and Data Analysis (New York, Chapman and Hall).

Valletta, Robert G., 1997. "The Effects of Industry Employment Shifts on the U.S. Wage Structure, 1979-1995." *Economic Review*, Federal Reserve Bank of San Francisco (1), pp. 16-32.

Appendix A

Table A1: Variable Definitions

Gender (g): 1 for male,0 otherwise
Union (u): 1 for union member, 0 otherwise
Employment (e): 1 if employed in the industrial sector, 0 otherwise
Industrial Sector includes: mineral and fuel extraction; electricity and water; manufacturing industries including maquiladores; and construction.
Components of vector (x)
Age: age of the respondent; **Age**^{2}: square of age
Education Dummies
No formal education - base category
Primary incomplete
Primary complete
Junior high incomplete
Junior high complete
High school incomplete
High school complete
Some college
College complete
More than college
Regional Dummies (r)
Southern states (Base): Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, Yucatan
Central states: Aquascalientes, Colima, Jalisco, Guanajuato, Hidalgo, Mexico, Michoacan, Morelos, Puebla, Queretaro, Tlaxcala
Northern states: Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nayarit, Nuevo Leon, San Luis Potosi, Sinaloa, Sonora, Tamaulipas, Zacatecas
Federal District

Figure1: Wage Distribution 1984 and 2000

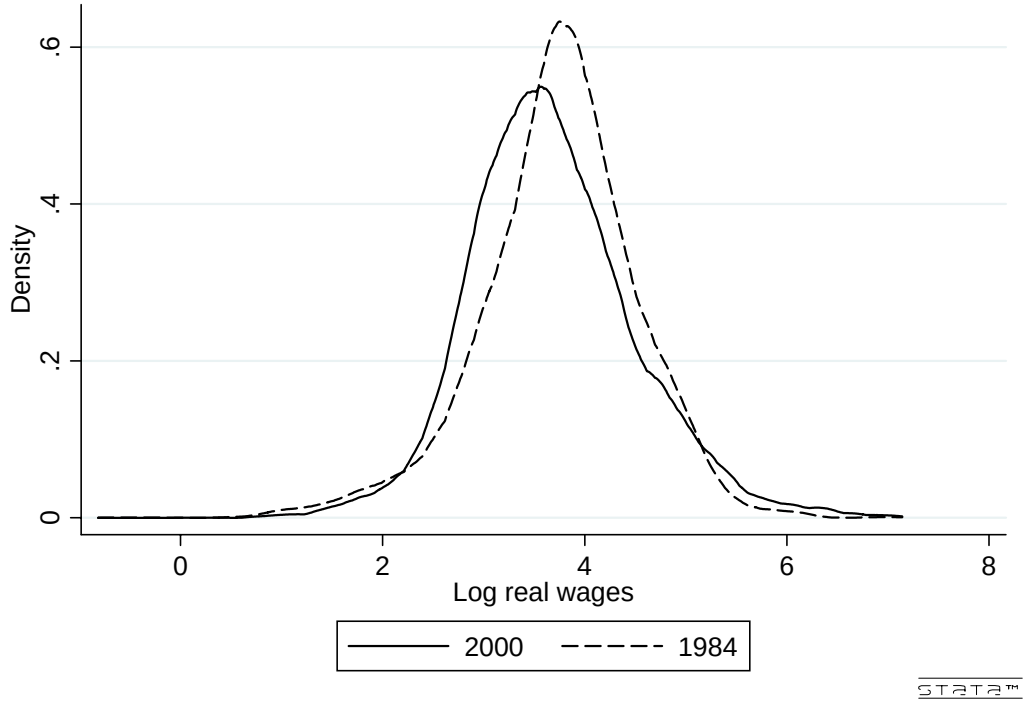

Figure2: Wage Distribution of 2000 adjusted for union effects

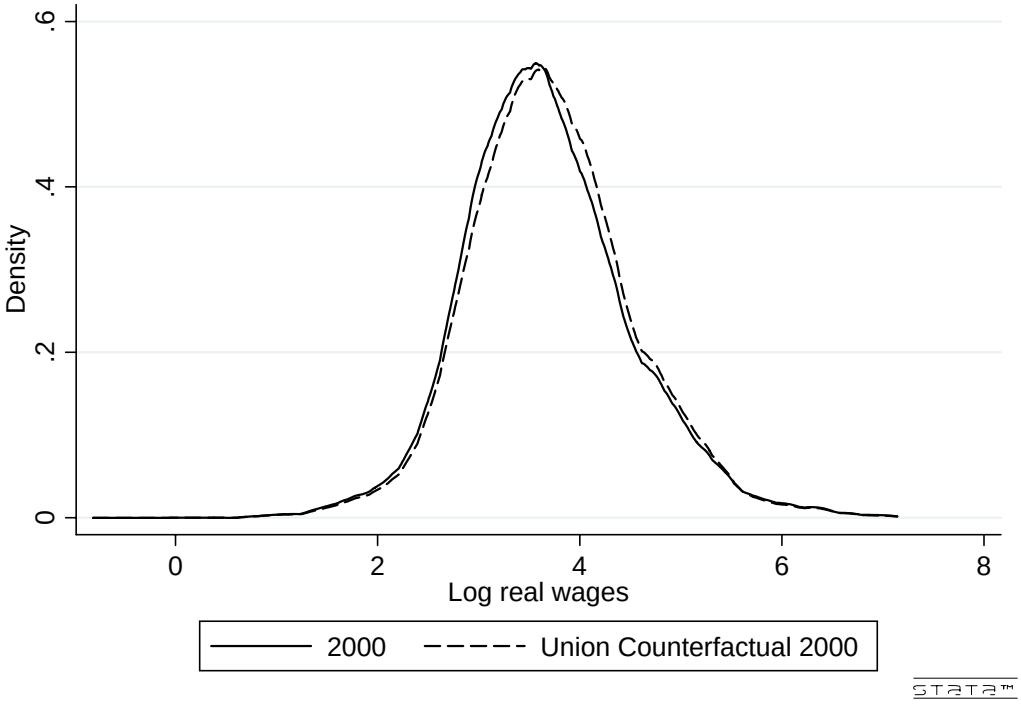

Figure3: Wage Distribution of 2000 adjusted for employment effects

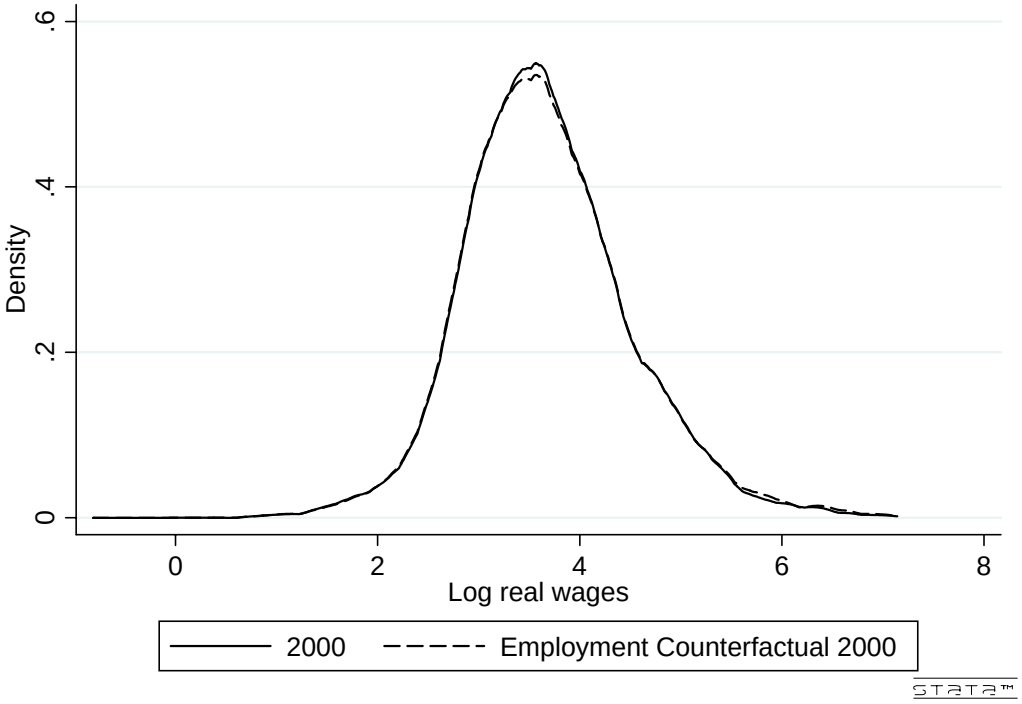
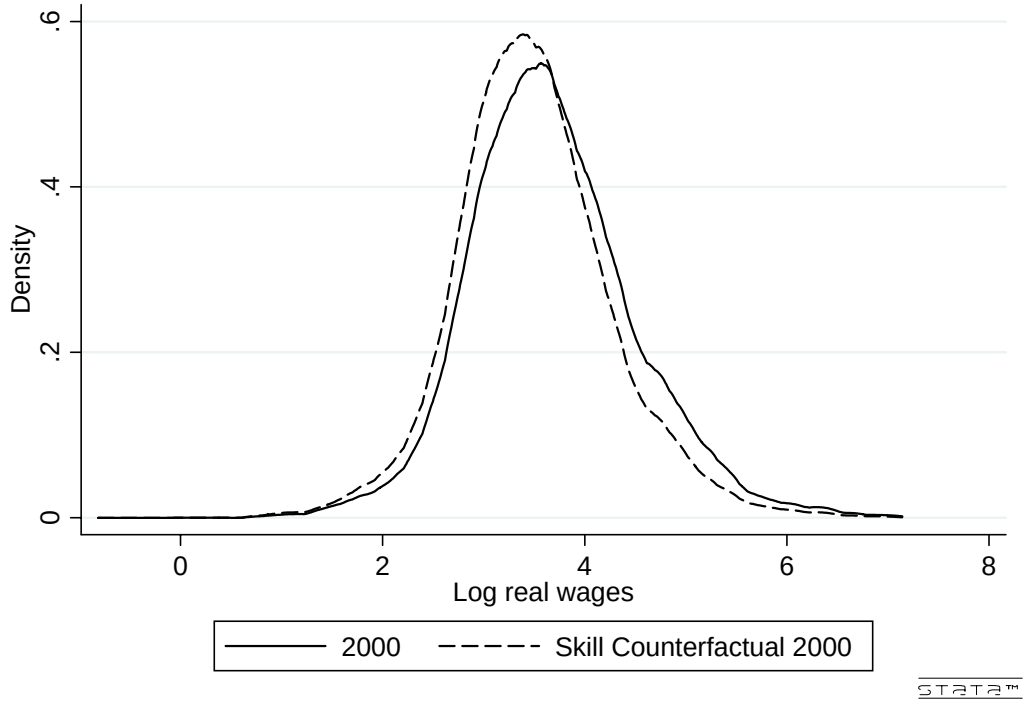
Figure4: Wage Distribution of 2000 adjusted for skill distribution