Sheffield Economic Research Paper Series

SERP Number: 2008006

ISSN 1749-8368

Gurleen Popli

Gender wage discrimination in Mexico: A distributional approach

April 2008

Department of Economics
University of Sheffield
9 Mappin Street
Sheffield
S1 4DT
United Kingdom
www.shef.ac.uk/economics

Abstract:

This paper examines the observed wage differentials by gender in Mexico over the last two decades (1984 to 2002). To estimate and understand the wage gap the paper uses a nonparametric-distributional approach, and compares the results with the other parametric approaches. The paper finds evidence of labour market discrimination against women. The average discrimination in the first decade fell considerably, but then started to increase again. The distribution of discrimination also changed over time. The average fall in discrimination over time has resulted largely from a fall in the estimated discrimination at the lower tail of the wage distribution, with little to no change at the upper tail.

Key Words:. gender wage gap, nonparametric-distributions, Mexico

JEL Classification: J16, C14, O15

Acknowledgements: The research was supported by a grant from Social Sciences Devolved Fund, University of Sheffield. I would like to thank Sarah Brown and Andy Dickerson for helpful comments, and Syed Hamid for research assistance. All errors are mine.

Gender wage discrimination in Mexico: A distributional approach

Gurleen K. Popli*
Department of Economics
University of Sheffield

ABSTRACT

This paper examines the observed wage differentials by gender in Mexico over the last two decades (1984 to 2002). To estimate and understand the wage gap the paper uses a nonparametric-distributional approach, and compares the results with the other parametric approaches. The paper finds evidence of labor market discrimination against women. The average discrimination in the first decade fell considerably, but then started to increase again. The distribution of discrimination also changed over time. The average fall in discrimination over time has resulted largely from a fall in the estimated discrimination at the lower tail of the wage distribution, with little to no change at the upper tail.

JEL classification code: J16, C14, O15
Keywords: gender wage gap, nonparametric-distributions, Mexico
April 2008

Acknowledgements: The research was supported by a grant from Social Sciences Devolved Fund, University of Sheffield. I would like to thank Sarah Brown and Andy Dickerson for helpful comments, and Syed Hamid for research assistance. All errors are mine.

*9 Mappin Street, Sheffield, UK, S1 4DT. Email: g.popli@sheffield.ac.uk; Tele. +44(0)114-2223485; Fax 44(0)114-2223458.

1 Introduction

This paper examines the observed wage differentials, by gender, in Mexico over the last two decades (1984 to 2002). To estimate and understand the gender wage gap the paper uses a nonparametric-distributional approach. Analyses of labor market discrimination largely rely on Blinder-Oaxaca (BO henceforth) kind of decompositions (Blinder, 1973 and Oaxaca, 1973). Though a good starting point and a very good tool, BO decomposition has certain limitations. First, it is a summary measure, it gives us the behavior around the mean. To illustrate the problem associated with focussing only on the mean, consider the following example:

> *"Suppose two sets of sample survey data about working women's wages relative to men's are available. Each reveals an average wage differential attributed to discrimination of 10%, but in the first survey all working women are underpaid by 10%, while in the second half are underpaid by 20% and half are equally paid."* (pp.82, Jenkins, 1994.)

Looking at the mean alone also misses out other useful information. For e.g. there is evidence of a relationship between the wage gap and the wage levels (Wood *et al.*, 1993), which will not be captured by just looking at the behavior around the mean. Knowing which segment of the distribution is affected the most can be important for policy.

The second limitation of the BO decomposition is that it requires parametric specification of the conditional expectation function linking earnings with the individual characteristics. The mis-specification of this conditional function however can give misleading results.[1]

[1] The two limitations mentioned here are not the only criticisms of the BO approach. For a discussion of the other, related, criticisms of the BO approach see Neumark (2004, pages 9-10).

There have been attempts in the literature to correct for these two limitations. Jenkins (1994) and DiNardo *et al.* (1996) give distributional alternatives to the BO decomposition. Barsky *et al.* (2002) and Racine *et al.* (2004) propose nonparametric alternatives to the BO method, which do not require parametric specification of the conditional expectation function.[2] The nonparametric-distributional approach used in this paper is in the spirit of the work done by these authors.[3]

Previous research on gender wage gap in Mexico largely relies on the BO kind of decompositions. Brown *et al.* (1999) use the National Urban Employment Survey (ENEU) for 1987-1993, their findings suggest that most of the male-female difference can be explained by the differences in the human capital endowments. Sanchez *et al.* (2001) focus on the earnings gap between the female and male-owned micro-enterprises, and reach the same conclusion as the Brown *et al.* (1999) paper. Pagan and Ullibari (2000) in their paper use the ENEU data for 1995 and the measure proposed by Jenkins (1994) to analyze the gender wage gap in Mexico.[4] They conclude that the unexplained wage gap is largest for those with very low or very high level of education.

The period since 1984 has been one of trade liberalization and increased foreign direct investment for Mexico (Lustig 1998 and 2001). The effects of free trade on women are not clear. One argument is, as competition increases, returns to employer discrimination fall, reducing the discrimination faced by women – both in terms of increasing their previously depressed wages and by feminization of high paid jobs, particularly in the industrial

[2]Another approach is the use of quantile regressions; while this does allow us to look at more points in the earnings distribution it still does not give a complete distributional experience and requires a parametric specification of the conditional expectation function (see Gardeazabal *et al.*, 2005, for an application).

[3]A similar approach is used by Breunig and Rospabe (2005) to study the male-female wage gap in France.

[4]While using Jenkins' measure takes care of the first criticism mentioned above the second limitation still applies.

sector. The potential negative impacts of free trade include: 'masculinization' of typical female jobs, as seen in maquiladoras in Mexico and a decrease in the prices of commodities produced by women, particularly in the agricultural sector (Artecona *et al.*, 2002). The actual effect, however, is an open empirical question.

There have been attempts in the literature to look directly at the gender specific impacts of structural adjustments in Mexico. Results are mixed: Alarcon-Gonzalez *et al.* (1999) and Artecona *et al.* (2002) find trade is detrimental to women. Aguayo-Tellez *et al.* (2006) find no impact of trade on the gender wage gap. Anderson and Dimon (1995) use data from 1988 and focus on two female dominated industries (export processing and tourism) in two Mexican cities (Tijuana and Torrean), their findings suggests that the overall wage gap decreased as the export processing activity (as a result of trade liberalization) increased.

This paper is different from the previous literature both in terms of the methodology used and the longer horizon covered. All the studies mentioned above cover the period till the mid-1990s. Most of them, like this study, note that the unexplained gender wage gap in Mexico fell from the early 1980s to the mid 1990s, what they fail to notice is the subsequent upward trend in the unexplained wage gap. Aguayo-Tellez *et al.* (2006) is the most recent contribution and covers the same time horizon as this paper, but the focus of their study is mainly on the employment changes faced by women.

The paper is structured as follows: the second section explains the different measures of discrimination used – the BO decomposition, the Jenkins (1994) method and the nonparametric-distributional approach. The third section describes the data used in the study and the empirical findings from all three approaches. The last section concludes the paper.

2 Measures of Discrimination

The earnings equation is estimated separately for men (M) and women (W):

$$\begin{aligned} \log y_i &= X_i\beta^m + \varepsilon_i^m, \quad \text{for all } i \in M, \\ \log y_i &= X_i\beta^w + \varepsilon_i^w, \quad \text{for all } i \in W, \end{aligned} \tag{1}$$

where, y_i is the wage of individual i; M and W are the sets of men and women, respectively; X_i is the vector of individual characteristics, e.g. human capital; β^m and β^w are the vectors of unknown rates of returns to individual characteristics, to be estimated for men and women, respectively; ε_i^m and ε_i^w are independent random residual terms for men and women, respectively.

2.1 Blinder-Oaxaca Decomposition

The average wage gap between men and women can be decomposed as:

$$\overline{\log y}^m - \overline{\log y}^w = (\overline{X}^m - \overline{X}^w)\widehat{\beta}^m + \overline{X}^w(\widehat{\beta}^m - \widehat{\beta}^w), \tag{2}$$

where the bar over the variable indicates the mean of the variable. The first term on the right hand side of equation (2) is referred to as the explained wage gap, the wage differential arising due to the differences in the endowments. The second term is the part of the wage gap due to differences in returns to similar endowments, and is an indirect measure of discrimination.[5] The summary measure of discrimination can be written as:

[5]There is a debate in the literature, about which wage structure – male or female – should be used as the non-discriminatory one. All the measures presented here assume that the male wage structure is the non-discriminatory wage structure. Another assumption inherent in this analysis is that the distribution of X's is not impacted by the β's. For discussion of the issues raised see Neumark (1988).

$$D_{BO} = 100[\exp(\overline{S}) - 1], \qquad (3)$$

where $\overline{S} = \overline{X}^w(\widehat{\beta}^m - \widehat{\beta}^w)$. D_{BO} can be interpreted as the wage increase women would have if there was no direct discrimination in the labor market.

2.2 Jenkins Measure

To estimate a distributional experience of discrimination, I start by constructing the Generalized Lorenz Curve (GLC) for $\widehat{y}_i^w = \exp(\log \widehat{y}_i^w)$, where $\log \widehat{y}_i^w$ is the estimated wage for women obtained from the estimated wage equation in (1). Next, using the same ranking as for the GLC a Generalized Concentration Centre (GCC) is obtained for $\widehat{r}_i^w = \exp(\log \widehat{r}_i^w)$, where $\log \widehat{r}_i^w = X_i^w \widehat{\beta}^m$. r_i is the reference wage, the wage that women would receive if their attributes are rewarded at rates β^m; this is the wage that women would receive if there was no direct discrimination. As long as $\widehat{y}_i^w < \widehat{r}_i^w$ for all $i \in W$, the GCC must lie above the GLC. The further apart the two curves are, the higher is the degree of discrimination. The area between the two curves gives us the average discrimination index:

$$D_J = [1 + 1/(2n_w)](\overline{r}^w - \overline{y}^w)/\overline{y}^w - (1/n_w)^2 \sum_{i \in w} i(\widehat{r}_i - \widehat{y}_i)/\overline{y}^w, \qquad (4)$$

where the first term is the difference between the means (hence comparable to D_{BO} above), and the second term is the weighted sum of deflated wage gaps where the weights are a woman's rank in the predicted earnings distribution.

While D_J takes care of the limitation of looking only at the mean, i.e. with this measure we are able to see the discrimination across the whole distribution, it still has the problem of potential mis-specification of the conditional expectation function (given by equation (1)). To correct for both issues raised in the introduction I look at the nonparametric measure.[6]

2.3 Nonparametric Measure

The distribution of log earnings, $h(\log y|m)$ for men and $h(\log y|w)$ women, respectively, can be written as:

$$\begin{aligned} h(\log y|m) &= \int f(\log y|m,x)g(x|m)dx, \qquad (5) \\ h(\log y|w) &= \int f(\log y|w,x)g(x|w)dx, \end{aligned}$$

where $f(\log y|m,x)$ is the conditional log wage distribution given individual characteristics (X) and gender; and $g(x|m)$ is the distribution of individual characteristics, given gender. The empirical counterpart of (5) is:

$$\begin{aligned} \widehat{h}(\log y|m) &= \sum_{i=1}^{n} \frac{\theta_i}{nh} K\left(\frac{\log y - \log y_i}{h}\right), \quad \text{for all } i \in M, \qquad (6) \\ \widehat{h}(\log y|w) &= \sum_{i=1}^{n} \frac{\theta_i}{nh} K\left(\frac{\log y - \log y_i}{h}\right), \quad \text{for all } i \in W, \end{aligned}$$

where θ are the sample weights; $K(.)$ is the kernel function; and h is the window width.

BO decomposition and the Jenkins measure focus on the counterfactual *mean* log wage, $\log r_i^w = X_i^w \beta^m$, the corresponding counterfactual *distribution* would be:

[6] One advantage of the Jenkins measure, not central to the analysis done here but worth noting, is that it satisfies a set of axioms (normative properties) often used in the poverty and deprivation literature (del Rio, Gradin and Canto, 2006).

$$
\begin{aligned}
h_w(\log y|m) &\equiv \int f(\log y|m,x)g(x|w)dx \qquad (7)\\
&= \int \omega(x)f(\log y|m,x)g(x|m)dx,
\end{aligned}
$$

where $\omega(x)$ is the re-weighting function defined as $\omega(x) \equiv g(x|w)/g(x|m)$. $h_w(\log y|m)$ is the re-weighted distribution of men, such that the distribution of individual characteristics is as that of women, but they are paid as men would be. This is in the spirit of the work done by DiNardo *et al.* (1996). To estimate the counterfactual distribution we need an estimate of the re-weighting function, which using Baye's rule can be written as:

$$
\omega(x) = \frac{\Pr(w|x)/\Pr(w)}{\Pr(m|x)/\Pr(m)}. \qquad (8)
$$

The conditional probabilities in the re-weighting function can be estimated using a probit model. Once the re-weighting function has been estimated the empirical counterpart of equation (7) is:

$$
\widehat{h}_w(\log y|m) = \sum_{i=1}^{n} \frac{\theta_i}{nh}\widehat{\omega}(x)K\left(\frac{\log y - \log y_i}{h}\right), \quad \text{for all } i \in M. \qquad (9)
$$

The difference between the two distributions, $h_w(\log y|m)$ and $h(\log y|w)$, gives us the extent of discrimination faced by women. To make the estimates obtained here comparable with the summary measures above, I define:

$$
D_{NP} = 100[\exp(\overline{S_d}) - 1], \qquad (10)
$$

where $\overline{S}_d = E_w[\log y|m] - E[\log y|w]$; $E[\log y|w] = \int \log y h(\log y|w) d\log y$ and $E_w[\log y|m] = \int \log y h_w(\log y|m) d\log y$.

3 Empirical Findings

3.1 Data and Descriptive Statistics

The data used for the study is from *Encuesta National de Ingresos y Gastos de los Hogares* (ENIGH). ENIGH is the national household survey, which started in 1984, continued in 1989, 1992 and every two years thereafter. The analysis in this paper focuses on four years – 1984, 1994, 1996 and 2002. To look at the period before the trade reforms were introduced I analyze the data for 1984. The majority of reforms were implemented by 1994, hence comparison of the results for 1984 with 1994 give us some idea about the 'immediate' impact of the changing macroeconomic environment in the country. December 1994, for Mexico, was marked by a major currency crisis followed by an economic recession.[7] To see what impact this crisis had on gender wage differentials, I analyze the data for 1996. Since 1996, the Mexican economy has seen relatively stable growth, analysis of the 2002 data allows us to capture the impact of stabilization and growth on the gender wage gaps.

Sample Selection: The sample is only of wage earners, over the age of 16 years. All those individuals who have more than one job are excluded; this is done to insure that workers can be correctly classified as wage earners only, it will rule out individuals who are wage earners, but also help out in the family business and receive income from that. Students, retired people, and those who give their reason for not working as 'inability to work due to disability' are also dropped from the sample. ENIGH interviews a different number of households each year, the sample selected here ranges from 30% to 35% of the total ENIGH sample.

ENIGH employs a 'stratified sampling' technique making the data nationally representative. Sample weights made available by ENIGH are used

[7]In 1995 the GDP growth rate for Mexico was -6.2%. *Source*: Table 7. The Mexican Economy 1999, Banco de Mexico.

in the analysis below. The earnings measure is hourly real wage (net of taxes), in 2002 Mexican Pesos, calculated from last months income and usual hours worked last week.

Descriptive Statistics: The labor force participation of women has increased over time in Mexico. In 1984 women were 28% of the sample, this number increased steadily to 36% by 2002.[8] Descriptive statistics for the sample used here are reported in Tables 1a and 1b, for women and men respectively.

Real wages for men and women increased between 1984 and 1994; the currency crisis of 1994 and the recession following that are associated with a decrease in real wages, which recovered only by year 2002.[9] The average age of the wage earners has increased over time; women however continue to be on average two years younger then men. Hours worked per week are also lower for women, by an average of seven hours per week. More men tend to be heads of their households, relative to women. On average the unionization rates among women tend to be higher than the rates for men, though the rates of unionization have declined for both over time. Wage earning women, on average, tend to have higher educational qualifications than men, and this gap has widened over time.[10]

Table 1c reports the occupational distribution by gender across the years. The occupations are ranked from the lowest to the highest paying, the ranking is based on the 1984 average wages for the whole sample in each oc-

[8]Though the labor force participation (LFP) of women in Mexico has increased over the last two decades, it is still much lower than what is prevalent in the developed countries. According to OECD statistics (accessed on 31/10/2007) in 2002 LFP of women in Mexico was 40% and for US and UK it was 66%.

Since the majority of the work force in Mexico is male, it is not unreasonable to assume that the male wage structure is closest to the productivity related wages, and hence the non-discriminatory wage structure. (Refer back to footnote 5)

[9]The fall in the real wages, during the period of the currency crisis, is true not only for the sample considered here but for the whole economy. Following the 1994 crisis, the real wages in the manufacturing sector fell by 12.5%, and the real minimum wage (nationwide average) fell by 12.3% (*Source*: Banco de Mexico).

[10]Working women in Mexico are more educated than working men - this finding is consistent with the findings of Pagan and Sanchez (2000) and Anderson and Dimon (1995).

cupational category; the ranking has not changed much across the years. There is some evidence of occupational segregation, which is persistent over the years. Women are largely employed as personal service workers (domestic and establishment), education workers or as administrators in the industrial sector. Men are largely employed as agricultural workers, and as operators and less-skilled workers in the industrial sector. In most of the occupations, that are female or male dominated, men on an average earn more than women. Three occupations where women, across the years, on average earned more than men are - vendors, peddlers with no business representation, workers in arts, entertainment and sports, and police and armed forces. However, a small proportion of the workforce, both men and women, work in these three occupations.

3.2 Discrimination

Overall trend: The average raw log wage differential[11] between men and women is not very high. On average, over the entire period under consideration female wages were about 99% of the male wages. This difference is not statistically significant. The average conditional log wage difference[12] however is higher and statistically significant (see Table 3). In 1984, female wages were 21% lower than the male wages, this proportion decreased to 17% by 2002.

An earnings equation is estimated for both men and women separately[13], results are reported in Table 2. There is debate on what should/should not be included as the control variable. Some argue that the control variables

[11]The average raw log wage differential is simply the difference between the average log wages for men and women.

[12]The average conditional log wage differential is obtained by running an OLS regression on the pooled sample of men and women. The dependent variable in the regression is log wages and the independent variables are: age and age squared, dummy variables for gender, education, union status, region of residence and occupation. The conditional log wage difference is the coefficient on the 'gender' dummy.

[13]A Chow test rejects the null hypothesis of a single regression for the pooled male and female sample.

(vector X) should not include factors that themselves can be affected by discrimination - like industry and occupation (Blau and Ferber, 1987). Results presented here are for the specification including the following control variables: age, age squared, dummy variables for union status, education (no education is the base category), region of residence (south is the base category), and occupation (agricultural workers are the base category). Alternate specifications were also estimated to check the robustness of the results, which are commented on later.

All coefficients are jointly significant. Age has a significant and a positive effect, at a decreasing rate, on log wages. Education at all levels has a positive impact on the wage, with premiums increasing as the level of education increases. Returns to education are higher for women relative to men, particularly at mid levels of education (more than primary, but less than college); though the gap has narrowed over time. The union dummy is significant, with the effect being larger for women than for men.

Regional dummies are significant and positive, with significant premiums for those (both men and women) in the capital region or in the north (closer to US) of the country. Returns in almost all the occupations are higher for men relative to women, with a few exceptions, notable exceptions being the three occupations where women on average earn more than men.

Table 3 reports the summary discrimination measures from all the three approaches. Whichever summary measure we look at (D_{BO}, D_J, D_{NP}), discrimination was the highest in 1984, declined in the mid 1990s, since then it has been rising again.

In 1984, if there were no discrimination, wages of women, depending on the measure used, would have been 21% to 25% higher than their observed wages. Given that to begin with average female wages were almost the same as those of men, this would mean that the female wages, on average, would be 21% to 25% higher than the male wages. Over the next decade, the levels of discrimination faced by women fell, with almost all of the decline happening after the 1994 currency crisis. Subsequent growth and stabilization of the economy are associated with increased levels of discrimination faced by women.

Next we look at each measure individually and see what additional information they provide.

BO decomposition: The entire raw wage differential can be explained by discrimination. Not surprisingly the component attributable to 'endowments' is negative – women have better human capital endowments.

Jenkins Measure: For all years, the GCC lies above the GLC everywhere, indicating that discrimination is present across the whole distribution (see Figures 1a to 1d). But how has the experience of discrimination across the years changed? Discrimination dominance checks across the years are reported in Table 4. The mean predicted reference (i.e. counterfactual) wage for each year is higher than the mean predicted wage. Both the 1984 and 1994 distributions lie above the 1996 and 2002 distributions, indicating that the 1984 and the 1994 distribution of wages had more discrimination. Of interest is the comparison of the 1984 and the 1994 distributions. While the summary statistics indicate the two distributions have a similar degree of discrimination, the nature of this discrimination is different for the two years. Below the median, the 1994 distribution has less discrimination than the 1984 distribution, above it the discrimination was higher in 1994 relative to 1984.

Nonparametric Measure: Figures 2a to 2d, give us the distributions of men and women for all the four years.[14] The dispersion of wages is higher for women relative to men in all years. For the years 1994, 1996 and 2002, the actual distributions for men and women look similar at the lower tail;

[14]To estimate the nonparametric distributions, I used the Gaussian kernel and window width, $h = 1.06\sigma n^{-1/5}$, where σ is the standard deviation of the log wages and n is the sample size.

differences are in the upper tail where there seems to be another, smaller, peak for women. Imposed on these graphs are also the counterfactual wage distributions for women (i.e. the distribution for men reweighted). If women had the endowments that they do and were paid as men are, then their distribution would be as men's but slightly to the right (with the exception of 1996). Though the peak on the upper tail still remains for the later years.

Using the estimated actual and the counterfactual distributions, the distributional experience of discrimination can be summarized using the following equation:

$$\log y_p^m - \log y_p^w = \underset{TermA}{(\log y_p^m - \log r_p^w)} + \underset{TermB}{(\log r_p^w - \log y_p^w)}, \tag{11}$$

where $\log y_p^m$ is the log wage at the p^{th} percentile of the male wage distribution, such that $\int\limits_0^{\log y_p^m} \widehat{h}(\log y|m) = p$. Similarly $\log y_p^w$ and $\log r_p^w$ are the log wages at the p^{th} percentile of the female actual wage distribution and the counterfactual distribution, respectively. The left hand side of equation (11) is the difference between the male and female wages, at each percentile, calculated from the estimated kernel distributions. This total difference is decomposed into the difference explained by the 'endowments' ($TermA$ of equation (11)) and the 'unexplained' ($TermB$ of equation (11)) difference. The unexplained difference is interpreted as the indirect measure of discrimination. The results of this exercise are reported in Table 5.

Just like in the BO decomposition we get the explained difference as negative - better endowments for women should, if left to market forces, yield higher returns; this is true not only on average but is observed at every percentile of the distribution. Discrimination is greater at the lower tail of the distribution, and decreases as one moves up the distribution though not monotonically.[15] Changes in the discrimination (both the decline in the first decade and the upward trend in the subsequent decade) come from changes in the 'unexplained' differences at the lower tail of the wage distribution, with the levels of discrimination at the upper tail remaining largely unchanged. Just as the Jenkins measure indicated, 1994 is an interesting case, while the discrimination at the lower tail of the distribution has decreased it has increased at the upper tail, relative to 1984.

3.3 Robustness checks

As mentioned above there is a debate in the literature about what should and should not be included in the control variables. Numerous specifications were tried to check the robustness of the findings. The results from this exercise are summarized here, detail results are available from the author on request.

One of the alternative specification that was checked and is of particular interest is one with no occupation dummies. The results for the summary statistics (D_{BO}, D_J, D_{NP}) from this specification are qualitatively the same as the ones reported here, but are quantitatively different, with the discrimination in this specification half of what is reported above. This is contrary to the findings for the US where evidence suggests that including occupational dummies reduces the estimated discrimination (Neumark, 1988; Blau and Kahn, 2006).

Two other variables that deserve a mention are the marital status of the individual and the number of young children in the family. Including both the variables in either of the above two specifications significantly reduces the estimated discrimination, without changing the results qualitatively.[16]

[15] This is contrary to the findings of Wood *et al.* (1993).

[16] This finding is consistent with that for the US (Neumark, 1988). The marriage premium is high and significant for married men and low and insignificant for women. If marriage is related to unobserved productive characteristics then it can explain away all discrimination (Neumark, 2004, chapter 1).

But these results should be intrepreted with caution, as these two variables are not measured accurately in the data. While we can identify how many children there are in the household, we cannot attach a child to an adult woman in the household. This is because of the complex joint family structure prevalent in Mexico. The question regarding marital status was asked only from the 1996 survey on. Before that it is possible only to identify whether or not the head of the household is married, and the spouse of the head of the household.

Including industry dummies, in either of the above specifications, does not significantly alter the results - qualitatively or quantitatively.

4 Concluding discussion

Establishing direct causality between the impact of structural reforms on gender wage differentials is not straightforward. What this paper does is to look at how the wage differentials, by gender, have changed over time given the changing macro conditions in the country over these two decades.

Average raw wage differentials between men and women in Mexico are not significant amongst the wage earners, who are predominantly in the formal sector of the economy. This finding in itself is surprising, especially when one looks closely at the endowments and the returns to these endowments for men and women. Wage earning women have higher educational achievement and higher unionization rates relative to men, the returns to these endowments are also higher for women relative to men. These two factors combined together should make the average wages for women higher than the average wages for men.

Factors that let women down, it would seem, are – their age, their region of residence, and their occupation. In the absence of information on actual labor market experience, age is used here as a proxy for experience. Though the returns to age are marginally lower for women than for men, what drives the results is that the women in the sample are on average two years younger than men (their actual experience could be even lower).

When we look at the discrimination measures, there does seem to be evidence of discrimination against women – all the measures indicate that women should be earning, on average, higher than what they are earning currently. The findings are consistent across all the three measures used in the paper; though each measure provides additional information. The summary statistics often hide the distributional experience of discrimination. For example we find that while the average discrimination is similar in 1984 and 1994, the distributions are very different: with the 1994 distribution showing lower discrimination at the lower tail and higher discrimination at the upper tail, relative to 1984. On the other hand the summary statistics for 2002 indicate much lower levels of discrimination relative to 1984, but the distributional experiences are almost similar - with discrimination falling as one moves up the distribution.

The average wage discrimination fell in the first decade of the period considered here, with most of the decline coming right after the 1994 currency crisis, in the second decade the discrimination started to increase. The distribution of discrimination has also changed over time. The average fall in discrimination over time has resulted largely from the fall in the estimated discrimination at the lower tail of the wage distribution, with little to no change at the upper tail.

From a policy perspective the important things to note are the evidence of a sticky floor and glass ceiling. There is evidence of a sticky floor with unexplained wage differentials being higher at the lower tail, though these have narrowed over time. The wage differentials at the upper tail, the glass ceiling, however show no signs of relenting.

Two findings from this study need further exploration. First is the dramatic decrease in discrimination observed between 1994 and 1996. This was the period of economic recession in Mexico. Recessions are often perceived to increase the gender wage gap, as economic downturns are expected to break the tenuous connection that women have with the labour market (Langton and Konrad, 1998). The empirical evidence in support of this hypothesis is however mixed with labour market structure being more important (Baden, 1993). For Mexico one possible explanation for the decrease in the gender wage gap following the recession could be the differential impact the recession had on the industrial and the services sector. Women are predominantly in the service sector, this sector saw a lower downturn compared with the industrial sector which mainly employs men.[17]

The second finding that needs further exploration is the impact of the occupational structure on the wage gap. The findings here are contrary to those in the literature.

The above results however carry some qualifications. First, the study looks only at the wage earners. There could be issues with selection bias, selection into the formal wage earning sector cannot necessarily be assumed random.[18] Second, as the robustness checks carried out in this paper suggest, there always is an issue of the control variables – what should and should not go into X (the vector of control variables). It might be the case that the finding that women do not earn higher than men, even if they have higher education, is due to unobserved, productivity related, variables which have not been controlled for. Third, there are issues with pre-labor market discrimination which need to be explored.

[17]Alaez Aller and Ullibarri Arce (2001) find results supporting this explanation for Spain.

[18]For e.g., there is a higher concentration of women in the informal sector in Mexico, may be women do not enter the formal sector due to other socioeconomic constraints (e.g. care responsibilities within the family). Sample selection could also explain why the results here are different from the findings of Brown *et al.* (1999) and Sanchez *et al.* (2001).

References

1. Aguayo-Tellez E., J. Airola, and C. Juhn, 2006. "Did trade liberalization help women? The case of Mexico in the 1990s," Working Paper, University of Houston.

2. Aláez Aller, R. and M. Ullíbarri Arce, 2001. "Explaining the reduction in the gender wage gap during the 1990–94 economic recession in Spain," *Atlantic Economic Journal*, vol. 29, no. 1, pp. 63-74.

3. Alarcon-Gonzalez, D. and T. McKinley, 1999. "The adverse effects of structural adjustments on working women in Mexico," *Latin American Perspectives*, vol. 26, no.3, pp. 103-117.

4. Anderson J.B. and D. Dimon, 1995. "The impact of opening markets on Mexican male/female wage and occupational differentials," *The Social Science Journal*, vol. 32, no. 4, pp. 309-326.

5. Artecona, R. and W. Cunningham, 2002. "Effects of trade liberalization on the gender wage gap in Mexico," Policy Research Report on Gender and Development, World Bank.

6. Baden, S., 1993. "The impact of recession and structural adjustment on women's work in selected developing countries," Report No. 15, BRIDGE (development – gender).

7. Banco de Mexico. The Mexican Economy 1999.

8. Barsky, R., J. Bound, K.K. Charles and J.P. Lupton, 2002. "Accounting for the black-white wealth gap: A nonparametric approach," *Journal of American Statistical Association*, vol. 97, no. 459, pp. 663-673.

9. Blau, F.D. and M. Ferber, 1987. "Discrimination: Empirical evidence from the United States," *American Economic Review*, vol. 77, no. 2, pp. 316-20.

10. Blau, F.D. and L.M. Kahn, 2006. "The US gender pay gap in the 1990s: Slowing convergence," *Industrial and Labor Relations Review*, vol. 60, no.1, pp. 45-66.

11. Blinder, A.S., 1973. "Wage discrimination: Reduced form and structural estimates," *Journal of Human Resources*, vol. 8, pp. 436-455.

12. Breunig, R. and S. Rospabe, 2005. "Parametric vs. semi-parametric estimation of the male-female wage gap: An application to France," *Social Science Research Network Paper* 844565.

13. Brown C.J., J.A. Pagan and E. Rodriguez-Oreggia, 1999. "Occupational attainment and gender earnings differential in Mexico," *Industrial and Labor Relations Review*, vol. 53, no.1, pp. 123-135.

14. del Rio, C., C. Gradin and O. Canto, 2006. "The measurement of gender wage discrimination: The distributional approach revisited," ECINEQ Working Paper 2006-25.

15. DiNardo, J., N.M. Fortin and T. Lemieux, 1996. "Labor market institutions and the distribution of wages, 1973-1993: A semi-parametric approach," *Econometrica*, vol. 64, pp. 1001-1044.

16. Gardeazabal, J. and A. Ugidos, 2005. "Gender wage discrimination at quantiles," *Journal of Population Economics*, vol. 18, no. 1, pp. 165-179.

17. Jenkins, S., 1994. "Earnings discrimination measurement: A distributional approach," *Journal of Econometrics*, vol. 16, no.1, pp. 81-102.

18. Langton, N. and A.M. Konrad, 1998. "The impact of labor market structure on sex differences in earnings," *Gender, Work and Organization*, vol. 5, no. 2, pp. 83-101.

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

20. Lustig, N., 2001. "Life is not easy: Mexico's quest for stability and growth." *Journal of Economic Perspectives*, vol. 15, pp. 85-106.

21. Neumark, D., 1988. "Employers' discriminatory behaviour and the estimation of wage discrimination," *The Journal of Human Resourses*, vol. 23, no. 3, pp. 279-295.

22. Neumark, D., 2004. *Sex Differences in Labor Markets*, Routledge.

23. Oaxaca, R., 1973. "Male-female wage differential in urban labor markets," *International Economic Review*, vol. 14, no. 3, pp. 693-709.

24. Pagan, J.A. and M. Ullibarri, 2000. "Group heterogeneity and the gender earnings gap in Mexico," *Economia Mexicana Nueva Epoca*, vol. IX, num. 1, pp. 23-40.

25. Pagan, J.A. and S.M. Sanchez, 2000. "Gender Differences in Labor Market Decisions: Evidence from Rural Mexico," *Economic Development and Cultural Change*, vol. 48, no. 3, pp. 619-637.

26. Racine, J.S. and C. Green, 2004. "Robust estimation of union and gender wage gaps," Manuscript, McMaster University.

27. Sanchez, S.M. and J.A.Pagan, 2001. "Explaining gender differences in earnings in micro-enterprise sector," in M. Correria and E. Katz (eds) *The Economics of Gender in Mexico: Work, Family, State and Market* (World Bank).

28. Wood, R.G., M.E. Corcoran and P. Courant, 1993. "Pay Differences among the highly paid: The male-female earnings gap in lawyers," *Journal of Labor Economics*, vol. 11, no. 3, pp. 417-441.

Table 1a: Average Sample Characteristics[1] - Women

	1984	1994	1996	2002
Personal Characteristics				
Log real wage (hourly)[2]	2.63 (0.85)	2.81 (0.85)	2.43 (0.83)	2.66 (0.84)
Age (in years)	31.46 (11.51)	30.34 (10.78)	31.96 (10.75)	33.77 (11.73)
Hours worked per week	40.11 (12.96)	41.25 (13.63)	41.62 (13.51)	41.19 (13.80)
Union member (1=yes)	0.31 (0.46)	0.23 (0.42)	0.23 (0.42)	0.19 (0.39)
Head of the household (1=yes)	0.19 (0.39)	0.14 (0.35)	0.16 (0.36)	0.18 (0.38)
Education Level (proportion of the sample)				
Primary incomplete	0.18 (0.38)	0.12 (0.32)	0.12 (0.33)	0.11 (0.31)
Primary complete	0.27 (0.44)	0.16 (0.37)	0.18 (0.38)	0.16 (0.37)
Junior high incomplete	0.04 (0.19)	0.04 (0.20)	0.03 (0.18)	0.04 (0.19)
Junior high complete	0.23 (0.42)	0.32 (0.47)	0.30 (0.46)	0.27 (0.45)
High school incomplete	0.03 (0.16)	0.04 (0.20)	0.04 (0.19)	0.06 (0.24)
High school complete	0.09 (0.28)	0.12 (0.33)	0.11 (0.31)	0.12 (0.33)
Some college	0.04 (0.19)	0.05 (0.22)	0.07 (0.25)	0.09 (0.28)
College complete	0.06 (0.23)	0.09 (0.29)	0.10 (0.31)	0.10 (0.30)
More than college	-	0.01 (0.08)	0.00 (0.07)	0.01 (0.08)
Region[3] of Residence (proportion of the sample)				
North	0.28 (0.45)	0.29 (0.45)	0.28 (0.45)	0.28(0.45)
Center	0.37 (0.48)	0.40 (0.49)	0.29 (0.45)	0.43(0.50)
Capital	0.10 (0.29)	0.16 (0.37)	0.14 (0.35)	0.16(0.37)
South	0.26 (0.44)	0.14 (0.35)	0.18 (0.39)	0.13(0.33)
Observations	1136	3417	3482	5625

1. Standard deviation in parentheses. Sample weights are used in all calculations.
2. In 2002 Mexican Pesos.
3. 32 states of Mexico are divided into four regions. **Northern states**: Baja California, Baja California Sur, Chihuahua, Coahuila, Durango, Nayarit, Nuevo Leon, San Luis Potosi, Sinaloa, Sonora, Tamaulipas, Zacatecas; **Central states**: Aquascalientes, Colima, Jalisco, Guanajuato, Hidalgo, Mexico, Michoacan, Morelos, Puebla, Queretaro, Tlaxcala; **Southern states**: Campeche, Chiapas, Guerrero, Oaxaca, Quintana Roo, Tabasco, Veracruz, Yucatan; **Capital**: Federal District.

Source: Author's calculations from ENIGH dataset for various years.

Table 1b: Average Sample Characteristics - Men

	1984	1994	1996	2002
Personal Characteristics				
Log real wage (hourly)	2.65 (0.79)	2.79 (0.87)	2.43 (0.82)	2.66 (0.79)
Age (in years)	33.77 (12.81)	33.07 (12.73)	33.50 (12.00)	35.00 (13.16)
Hours worked per week	46.52 (11.67)	48.96 (13.33)	49.52 (13.63)	48.94 (13.26)
Union member (1=yes)	0.22 (0.41)	0.16 (0.36)	0.15 (0.36)	0.14 (0.35)
Head of the household (1=yes)	0.68 (0.47)	0.64 (0.48)	0.67 (0.47)	0.62 (0.49)
Education Level (proportion of the sample)				
Primary incomplete	0.27 (0.45)	0.18 (0.39)	0.16 (0.37)	0.14 (0.35)
Primary complete	0.23 (0.42)	0.22 (0.41)	0.21 (0.40)	0.19 (0.39)
Junior high incomplete	0.06 (0.24)	0.07 (0.25)	0.05 (0.22)	0.05 (0.22)
Junior high complete	0.15 (0.36)	0.21 (0.41)	0.24 (0.43)	0.26 (0.44)
High school incomplete	0.03 (0.17)	0.04 (0.20)	0.06 (0.23)	0.06 (0.24)
High school complete	0.05 (0.21)	0.07 (0.25)	0.09 (0.29)	0.10 (0.30)
Some college	0.04 (0.20)	0.06 (0.23)	0.06 (0.22)	0.07 (0.25)
College complete	0.06 (0.24)	0.08 (0.26)	0.07 (0.26)	0.08 (0.26)
More than college	0.00 (0.06)	0.01 (0.07)	0.01 (0.09)	0.01 (0.10)
Region of Residence (proportion of the sample)				
North	0.28 (0.45)	0.27 (0.44)	0.29 (0.45)	0.28 (0.45)
Center	0.40 (0.49)	0.43 (0.49)	0.28 (0.45)	0.41 (0.49)
Capital	0.13 (0.34)	0.19 (0.39)	0.16 (0.37)	0.20 (0.40)
South	0.18 (0.38)	0.12 (0.32)	0.14 (0.35)	0.11 (0.31)
Observations	3011	7770	8004	10394

For sources and definitions refer to notes at the end of Table 1a.

Table 1c: Occupation – proportion of the sample

	1984		1994		1996		2002	
	Men	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Personal service workers -domestic	0.01 (0.08)	0.11 (0.31)	0.01 (0.08)	0.15 (0.35)	0.01 (0.11)	0.13 (0.34)	0.01 (0.12)	0.14 (0.35)
Agricultural workers	0.19 (0.39)	0.05 (0.21)	0.13 (0.33)	0.02 (0.12)	0.11 (0.31)	0.02 (0.13)	0.11 (0.31)	0.01 (0.12)
Vendors, peddlers- with no business representation	0.00 (0.05)	0.01 (0.08)	0.01 (0.09)	0.01 (0.11)	0.01 (0.09)	0.01 (0.09)	0.01 (0.11)	0.02 (0.13)
Less skilled workers- industrial production	0.10 (0.30)	0.01 (0.10)	0.14 (0.35)	0.03 (0.17)	0.11 (0.32)	0.02 (0.15)	0.13 (0.33)	0.03 (0.17)
Police and armed forces	0.03 (0.16)	0.00 (0.03)	0.04 (0.19)	0.00 (0.06)	0.05 (0.23)	0.01 (0.08)	0.05 (0.21)	0.01 (0.08)
Personal service workers - establishments	0.05 (0.22)	0.13 (0.34)	0.07 (0.25)	0.07 (0.26)	0.07 (0.26)	0.08 (0.27)	0.07 (0.25)	0.11 (0.31)
Workers, operators – industrial production	0.25 (0.43)	0.13 (0.33)	0.22 (0.41)	0.12 (0.33)	0.21 (0.41)	0.16 (0.37)	0.21 (0.41)	0.14 (0.35)
Salespersons	0.07 (0.25)	0.09 (0.28)	0.08 (0.27)	0.12 (0.32)	0.08 (0.26)	0.11 (0.31)	0.09 (0.29)	0.14 (0.35)
Transport workers	0.07 (0.25)	-	0.08 (0.28)	-	0.10 (0.30)	-	0.09 (0.29)	0.00 (0.04)
Administrative workers –industrial production	0.09 (0.29)	0.28 (0.45)	0.08 (0.27)	0.23 (0.42)	0.10 (0.29)	0.23 (0.42)	0.10 (0.29)	0.19 (0.40)
Technicians	0.03 (0.17)	0.07 (0.25)	0.04 (0.20)	0.10 (0.29)	0.04 (0.20)	0.06 (0.24)	0.03 (0.17)	0.05 (0.23)
Supervisors- industrial production	0.03 (0.18)	0.01 (0.09)	0.02 (0.15)	0.01 (0.11)	0.03 (0.17)	0.02 (0.12)	0.03 (0.17)	0.02 (0.12)
Workers in arts, entertainment and sports	0.01 (0.10)	0.00 (0.03)	0.01 (0.10)	0.01 (0.08)	0.01 (0.11)	0.01 (0.07)	0.01 (0.08)	0.00 (0.06)
Education workers	0.03 (0.16)	0.10 (0.30)	0.02 (0.16)	0.09 (0.29)	0.02 (0.15)	0.10 (0.30)	0.03 (0.16)	0.08 (0.27)
Professional	0.02 (0.14)	0.01 (0.12)	0.02 (0.15)	0.03 (0.16)	0.02 (0.15)	0.04 (0.20)	0.03 (0.17)	0.04 (0.19)
Senior directors, Administrators in public and private sector	0.03 (0.17)	0.01 (0.09)	0.03 (0.16)	0.02 (0.14)	0.03 (0.17)	0.02 (0.13)	0.02 (0.14)	0.02 (0.13)

Occupations are ranked from lowest to highest paying. The ranking is based on the 1984 average wages for the whole sample in each occupational category. The ranking does not change much across years.
Standard deviation in parentheses. Sample weights are used in all calculations.
Source: Author’s calculations from ENIGH dataset for various years.

Table 2: Estimated Coefficients of Earnings Equation

Dependent Variable: Log real wage (hourly); Method of estimation: Least Squares

	1984		1994		1996		2002	
Independent Variables	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**	**Men**	**Women**
Age	0.06*	0.05*	0.06*	0.06*	0.06*	0.04*	0.05*	0.05*
Age Square (x 10^{-3})	-0.71*	-0.55*	-0.57*	-0.55*	-0.68*	-0.44*	-0.52*	-0.59*
Union	0.24*	0.27*	0.10*	0.13*	0.19*	0.21*	0.19*	0.33*
Education dummies								
Primary incomplete	0.27*	0.50*	0.16*	0.22*	0.12*	0.22*	0.17*	0.31*
Primary complete	0.46*	0.76*	0.30*	0.46*	0.30*	0.35*	0.26*	0.41*
Junior high incomplete	0.54*	0.94*	0.40*	0.52*	0.36*	0.53*	0.39*	0.50*
Junior high complete	0.56*	0.99*	0.44*	0.61*	0.44*	0.56*	0.38*	0.56*
High school incomplete	0.67*	0.86*	0.52*	0.72*	0.57*	0.64*	0.51*	0.67*
High school complete	0.77*	1.00*	0.68*	0.96*	0.65*	0.78*	0.55*	0.77*
Some college	0.79*	1.02*	0.86*	0.89*	0.90*	0.91*	0.82*	0.94*
College complete	0.96*	1.06*	1.22*	1.14*	1.08*	1.06*	1.07*	1.10*
More than college	1.41*	-	1.51*	1.50*	1.35*	1.47*	1.39*	1.36*
Region dummies								
North	0.13*	0.24*	0.25*	0.17*	0.18*	0.19*	0.37*	0.37*
Centre	0.09*	0.07	0.24*	0.22*	0.04**	0.05***	0.27*	0.28*
Capital	0.21*	0.35*	0.34*	0.34*	0.19*	0.26*	0.25*	0.39*
Occupation dummies								
Personal service workers -domestic	0.04	-0.12	0.39*	0.34*	0.11***	-0.05	0.26*	0.26*
Vendors, peddlers- with no business representation	0.17	0.51**	0.41*	0.60*	-0.03	0.1	0.12**	0.22*
Less skilled workers- industrial production	0.26*	-0.12	0.30*	0.30*	0.14*	-0.09	0.26*	0.12
Polish and armed forces	0.23*	0.24	0.31*	0.36**	0.28*	0.44*	0.32*	0.49*
Personal service workers - establishments	0.46*	0.35*	0.50*	0.34*	0.31*	0.08	0.31*	0.35*
Workers, operators – industrial production	0.39*	0.38*	0.53*	0.33*	0.38*	-0.07	0.44*	0.22*
Salespersons	0.47*	0.18***	0.54*	0.40*	0.36*	-0.01	0.49*	0.33*
Transport workers	0.50*	-	0.53*	-	0.38*	-	0.39*	0.35***
Administrative workers –industrial production	0.69*	0.62*	0.81*	0.78*	0.66*	0.35*	0.67*	0.66*
Technicians	0.70*	0.69*	0.92*	0.84*	0.58*	0.43*	0.64*	0.73*
Supervisors- industrial production	0.80*	0.39**	0.77*	0.67*	0.73*	0.30*	0.78*	0.66*
Workers in arts, entertainment and sports	0.79*	0.86***	0.98*	1.09*	0.91*	0.73*	0.77*	1.73*
Education workers	0.94*	0.95*	1.07*	1.10*	0.79*	0.72*	0.83*	0.89*

Table 2 (Continued)

Professional	0.79*	1.04*	0.95*	1.04*	0.66*	0.49*	0.93*	1.00*
Senior directors, Administrators in public and private sector	1.18*	1.05*	1.42*	1.11*	1.21*	0.85*	1.27*	1.17*
Constant	0.36*	0.23	0.40*	0.20***	0.22*	0.56*	0.48*	0.15
R-squared	0.54	0.60	0.56	0.57	0.50	0.53	0.51	0.54
Observations	3011	1136	7770	3417	8004	3482	10394	5625

For sources and definitions refer to notes at the end of Table 1a.

* significant at 1% level; ** significant at 5% level; *** significant at 10% level.

1. Nine education dummies are included in estimation, base category is – ‘no education’.
2. Three regional dummies are included in estimation, base category is – ‘south’
3. Fifteen occupational dummies are included in estimation, base category is – ‘agricultural workers’

Table 3: Gender Wage Differentials

	1984	1994	1996	2002
Raw log wage differential[1] (standard error)	0.02 (0.03)	-0.02 (0.02)	-0.00 (0.02)	0.01 (0.01)
Conditional log wage differential[2] (standard error)	0.19* (0.02)	0.20* (0.01)	0.16* (0.01)	0.16* (0.01)
Blinder-Oaxaca Decomposition				
Component attributable to:				
Endowments	-0.17	-0.20	-0.14	-0.14
Discrimination (t-stat)	0.19 (7.41)	0.18 (10.12)	0.14 (9.02)	0.15 (12.47)
Discrimination index (D_{BO})	20.75	20.05	14.76	15.89
Jenkins Measure				
Average discrimination index (D_J)	21.74	20.88	16.38	18.53
Nonparametric Measure				
Discrimination index (D_{NP})	24.68	19.56	13.17	15.36

* significant at 1% level; ** significant at 5% level; *** significant at 10% level.

1. The raw log wage differential is simply the difference between the average log wages for men and women.
2. The conditional log wage differential is obtained by running an OLS regression on the pooled sample of men and women. The dependent variable in the regression is log wages and the independent variables are: age and age squared, dummy variables for gender, education, union status, region of residence, and occupation. The conditional log wage difference is the coefficient on the 'gender' dummy.

Source: Author's calculations from ENIGH dataset for various years.

Table 4: Jenkins Measure – summary statistics

	1984	1994	1996	2002
Mean predicted reference wage for women ($\bar{r}^w$)	19.09	23.71	15.20	19.31
Mean predicted wage for women ($\bar{y}^w$)	16.36	19.84	13.38	17.11
$\bar{r}^w - \bar{y}^w$	2.73	3.87	1.82	2.20
Cumulative sample share (%)	Generalized Lorenz ordinates for $\lvert\bar{r}^w - \bar{y}^w\rvert$			
10	0.22	0.15	0.04	0.15
20	0.48	0.37	0.12	0.33
30	0.71	0.63	0.24	0.55
40	0.90	0.91	0.38	0.76
50	1.12	1.19	0.55	0.99
60	1.31	1.47	0.73	1.23
70	1.48	1.72	0.92	1.42
80	1.74	1.99	1.13	1.57
90	2.06	2.41	1.39	1.72
100	2.73	3.87	1.82	2.20

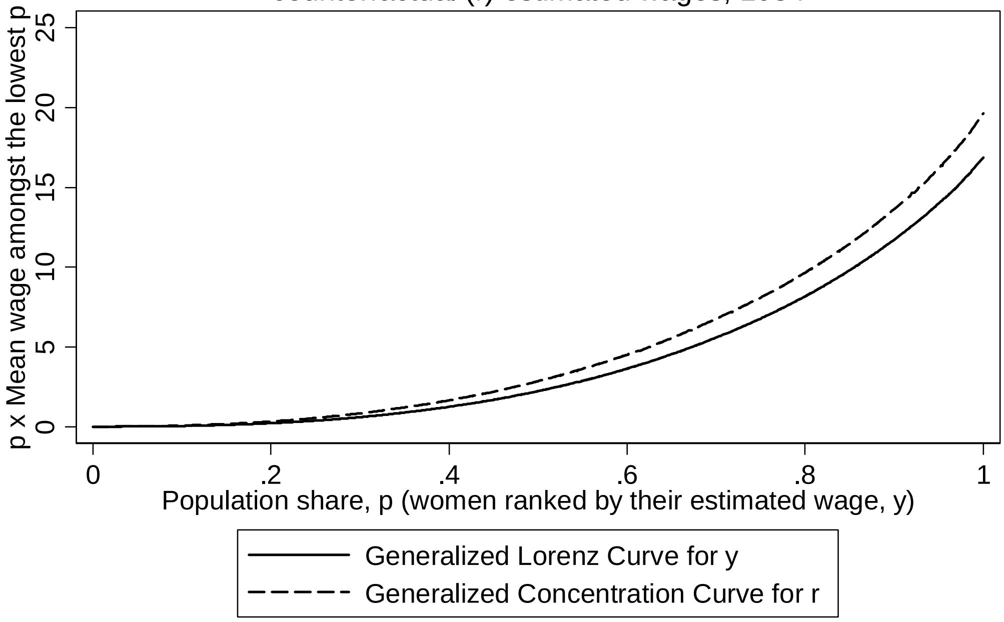
Source: Author's calculations from ENIGH dataset for various years.

Table 5: Nonparametric Measure – summary statistics[1]

	1984		1994		1996		2002	
	Term A	***Term B***	***Term A***	***Term B***	***Term A***	***Term B***	***Term A***	***Term B***
Wage percentiles (%)								
5	-0.30	0.62	-0.29	0.34	-0.13	0.23	-0.16	0.27
10	-0.30	0.49	-0.25	0.28	-0.10	0.16	-0.15	0.22
25	-0.25	0.27	-0.20	0.20	-0.09	0.12	-0.15	0.19
50	-0.20	0.14	-0.19	0.14	-0.14	0.12	-0.13	0.14
75	-0.21	0.11	-0.23	0.07	-0.19	0.09	-0.14	0.07
90	-0.09	0.08	-0.15	0.14	-0.12	0.09	-0.13	0.02
95	-0.03	0.08	-0.02	0.21	-0.03	0.09	-0.05	0.07

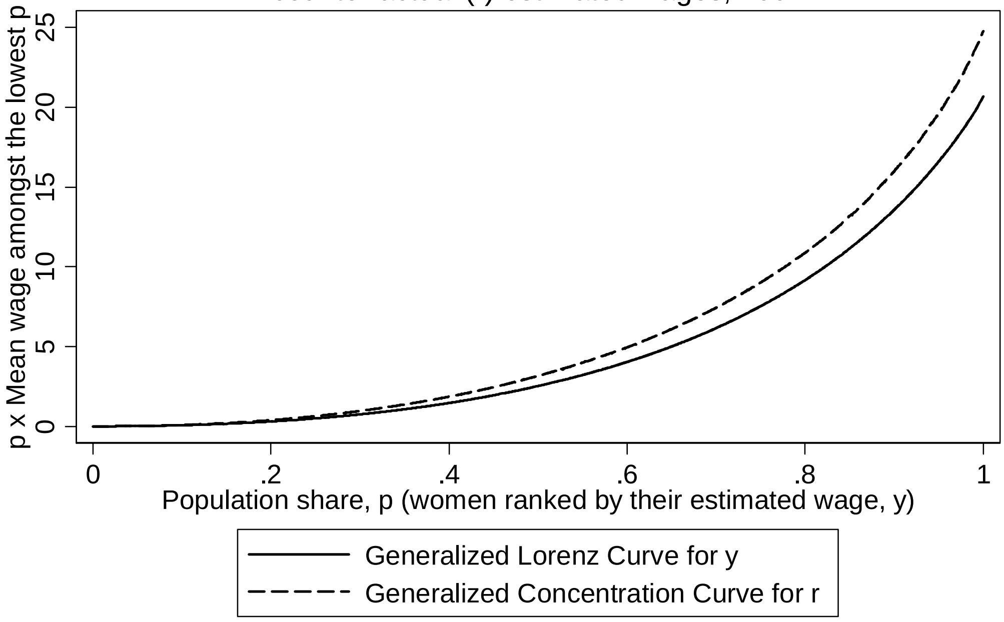
1. *Term A* is the wage difference explained by 'endowments'; *Term B* is the unexplained wage difference, interpreted as the indirect measure of discrimination. Refer to equation (11) in the main text on how these are calculated.
Source: Author's calculations from ENIGH dataset for various years.

Figure 1a. Joint distribution of actual (y) and counterfactual (r) estimated wages, 1984



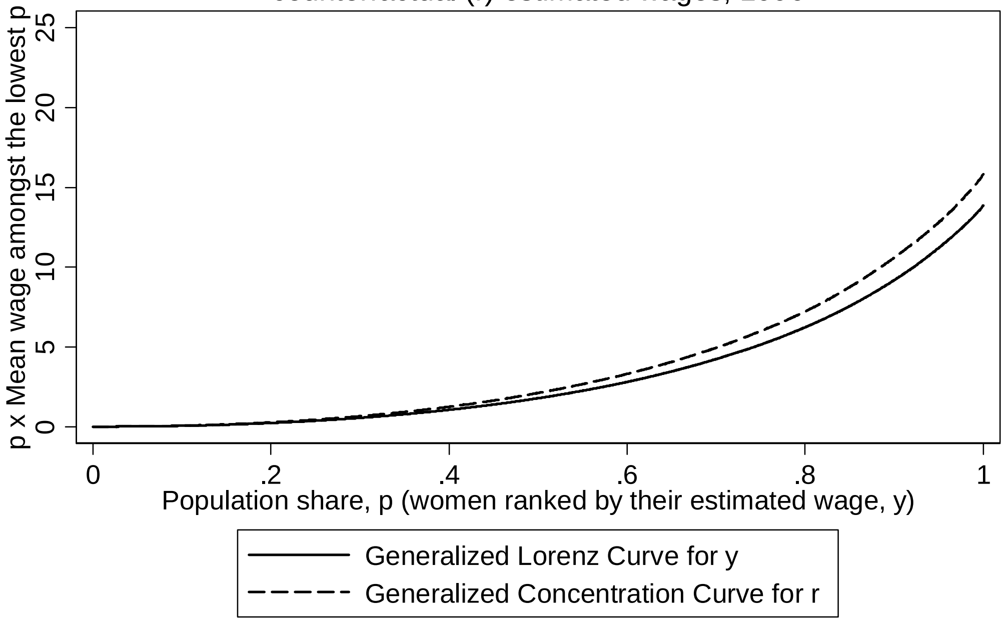
Source: Authors calculations from ENIGH 1984.

Figure 1b. Joint distribution of actual (y) and counterfactual (r) estimated wages, 1994



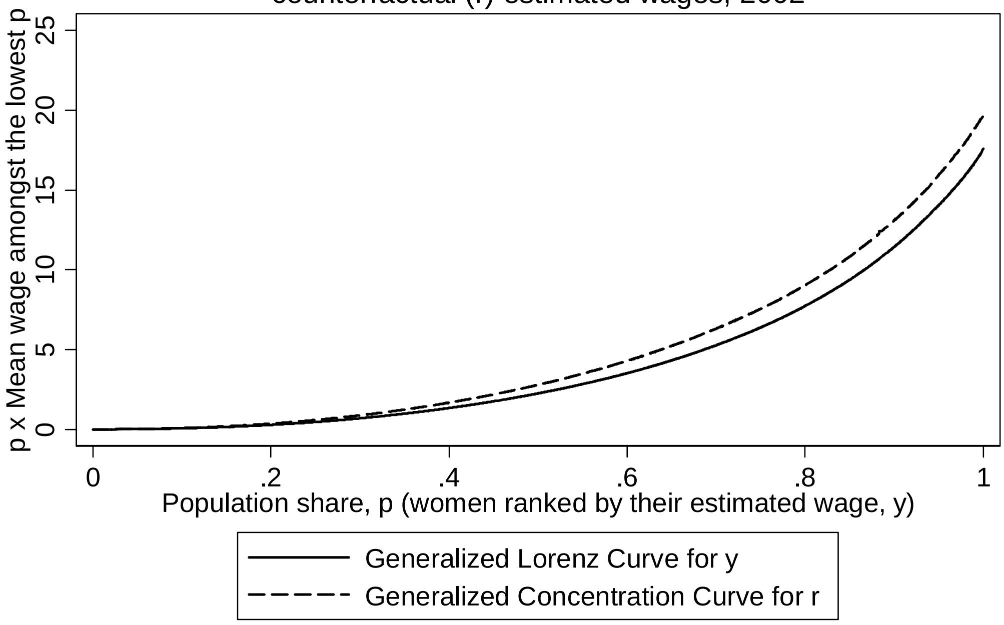
Source: Authors calculations from ENIGH 1994.

Figure 1c. Joint distribution of actual (y) and counterfactual (r) estimated wages, 1996



Source: Authors calculations from ENIGH 1996.

Figure 1d. Joint distribution of actual (y) and counterfactual (r) estimated wages, 2002



Source: Authors calculations from ENIGH 2002.

Figure 2a. Density of log wages, 1984

Source: Authors calculations from ENIGH 1984.

Figure 2b. Density of log wages, 1994

Source: Authors calculations from ENIGH 1994.

Figure 2c. Density of log wages, 1996

Source: Authors calculations from ENIGH 1996.

Figure 2d. Density of log wages, 2002

Source: Authors calculations from ENIGH 2002.