

RESEARCH ARTICLE

Open Access

Gender wage discrimination by distribution of income in Mexico, 2005-2020

Miguel Ángel Mendoza González*

* Miguel Ángel Mendoza González is a tenured professor, researcher and supervisor of Regional and Urban Economics at the National Autonomous University of Mexico (UNAM) Faculty of Economics Postgraduate Division. Research profile ID: <http://orcid.org/0000-0001-7433-5194>. Corresponding author: mendozag@unam.mx.

Abstract

This paper aims to analyze the hourly gender wage gap between men and women in Mexico for the period 2005-2020. To this end, a number of variables is selected to reflect workers' human capital, household circumstances and workplace characteristics; then, a novel non-parametric method decomposes wage differentials between men and women into its composition and structure effects throughout the distribution of labor income. Results are consistent with the sticky-floor hypothesis, where male workers earn higher hourly wages than female workers at low income levels. However, differentials decrease in the upper part of the distribution and may even reverse, favoring women over men at the highest income levels.

JEL Classification: C14, J16, J24, J31, J71, O54

Key words: income distribution, gender wage gap, gender wage discrimination, non-parametric methods, sticky floors.

Introduction

According to the International Labor Organization (ILO, 2019), the wage gap between men and women remains high in Latin America and the Caribbean to this day, at around 17% per hour worked. This figure stands in contrast with ILO's reported progress towards closing the gender wage gap: the gender education gap has overturned—now working women are more educated than working men on average—, women's

participation into the labor market has increased, and the proportion of working women with higher human capital has also grown—largely as a result of lower fertility rates, the delay of motherhood, and more pressure to generate labor income coming from the growth of single-mother households.

The literature on gender inequality and salary discrimination by sex in Mexico has mainly focused on discussing explanatory factors for average trends and relied at large on the Blinder-Oaxaca methodology ([Del Razo, 2003](#); [Martínez-Jasso & Acevedo-Flores, 2002](#); [Meza-González, 2001](#); [Pagan & Ullibarri, 2000](#)). Although some question whether income inequality between men and women has increased or dwindled ([Mendoza-González, Carde-ro-García, & Ortiz-García, 2017](#)), most research concludes that the average gender wage gap widened during the 1990s ([Brown, Pagan, & Rodríguez-Oreggia, 1999](#); [Del Razo, 2003](#); [Popli, 2013](#)), and only started narrowing in 2001, led by the downturn of the American economy ([Arceo-Gomez & Campos-Vazquez, 2014](#); [Mendoza-González et al., 2017](#); [Rodríguez-Pérez & Castro-Lugo, 2014](#)).

The use of mean income in the estimation of the gender wage gap has also been criticized because the distribution of labor income tends to be skewed ([Machado & Mata, 2005](#); [Melly, 2005](#)). Instead, the median could serve as a better indicator, given that differentials are likely to vary along the income distribution; this is, among workers with low, medium, and high salaries ([DiNardo, Fortin, & Lemieux, 1996](#); [Machado & Mata, 2005](#); [Ñopo, 2008](#)).

Only a few papers have studied wage differentials using income distribution functions in Mexico. [Pagan and Ullibarri \(2000\)](#) estimated the Jenkins index to analyze the gender income gap by socio-economic subgroups, while [Popli \(2013\)](#) applied a nonparametric-distributional approach and found a reduction in average discrimination against women in the 1980s, but a growing discriminatory trend during the 1990s. In the same line, [Arceo-Gomez and Campos-Vazquez \(2014\)](#) identified a stable pattern in the lower part of the gender wage gap distribution, with a decreasing trend in the middle segment and an increasing trend in the upper segment. On their part, [Rodríguez-Pérez and Castro-Lugo \(2017\)](#) studied different hypotheses regarding the effects of trade liberalization on wage differentials.

In this context, this research aims to contribute to the literature on the hourly wage gap between men and women along the income distribution by studying the Mexican case from 2005 to 2020 from a gender discrimination approach and using a new non-parametric methodology proposed by [Chernozhukov, Fernández-Val and Melly \(2013\)](#), as well as quantile parametric models for comparison.

The rest of the paper is structured as follows. Section 2 will briefly explain the analytical approach and review the relevant literature on income differentials in Mexico. Section 3 will provide an outline of the methodology, while section 4 will describe the data and analyze the results. Finally, Section 5 will discuss the findings and suggest future lines of research.

Theoretical approach and discussion on gender income differentials in Mexico

The theory of human capital assumes that workers with greater endowments—of education and experience, for instance—will earn higher incomes. Presumably, men with higher human capital will rightly earn higher wages than women with lower human capital endowments ([Aigner & Cain, 1977](#); [Becker, 1957](#); [Becker, 1981](#)). Income differentials between men and women could then be explained by the difference between men and women's endowments ([Boeri & Van Ours, 2013](#); [Jacobsen, 2007](#)). Building on this idea, gender wage discrimination would arise when, for no apparent reason, men and women with the same endowment of human capital attain different labor incomes. In this sense, gender wage differentials can be understood as a non-monetary

source of income, where men tend to be valued differently from women in the labor market ([Becker, 1957](#); [Jacobsen, 2007](#)). There are other social, market and location factors that could favor male workers and disserve female workers, such as being the head of a family, having children under the age of six, having a job in a non-urban area or working for smaller companies ([Block & Walker, 1981](#); [Boeri & Van Ours, 2013](#)).

In the case of the Mexican economy, evidence points towards large wage differentials between men and women in the late 1980s and in the 1990s, that diminished significantly during the 2000s and 2010s ([Arceo-Gomez & Campos-Vazquez, 2014](#); [Mendoza-Cota & García-Bermúdez, 2009](#); [Rodríguez-Pérez & Castro-Lugo, 2014](#)). Most research has identified an average gender wage gap that fluctuates between 5 and 15 percent, depending on the measure of income and on the period analyzed ([Cardero, Mendoza, & Galán, 2015](#); [Martínez-Jasso & Acevedo-Flores, 2002](#); [Mendoza-Cota & García-Bermúdez, 2009](#)).

There is little research on the Mexican economy that uses a measure other than average income differentials and that focuses on the distribution of gender income gaps. [Pagan and Ullibarri \(2000\)](#) investigated the role of heterogeneity in socio-demographic subgroups on the gender earnings gap using microdata from the National Survey of Urban Employment (ENEU) and the Jenkins index to decompose the gap in its explained and unexplained components. They found a higher index –signaling discrimination– for people with lower levels of education, those with a bachelor’s degree, and for relatively older people with more work experience; the index was also inversely related to the size of the company and was higher in the private and informal sectors. In a similar fashion, [Popli \(2013\)](#) analyzed gender wage differentials for the years 1996 and 2006 with data from the National Household Income and Expenditure Survey (ENIGH), by applying non-parametric distributional methods, such as the Jenkins measure (Lorenz curve and concentration center), to the earnings distribution between men and women and compared them with parametric methods of the Blinder-Oaxaca type. She found that average discrimination against women decreased in the 1980s but showed a growing trend in the 1990s with a simultaneous decrease in gender wage gaps at the lower tail of the income distribution and without relevant changes in the upper tail. On their part, [Arceo-Gómez and Campos-Vázquez \(2014\)](#) worked with the population censuses of 1990, 2000 and 2010, and found stable patterns in the lower half of the gender wage-gap distribution but an increase in the upper tail. More recently, [Rodríguez-Pérez and Castro-Lugo \(2017\)](#) studied different hypotheses regarding the effect of trade liberalization on the wage gap with data from the National Occupation and Employment Survey (ENOE); most of their results indicated decreasing wage differentials against older women in the lower deciles of the labor income distribution.

In spite of the rising interest in the gender wage gap, to the best of our knowledge there are no studies updating and analyzing annual trends in gender wage discrimination by income level in Mexico that cover the period from 2005 to 2020, nor is there any research of this type using the non-parametric methodology proposed by [Chernozhukov, Fernández-Val and Melly \(2013\)](#). Such is the realm of this research, which uses workers’ individual characteristics (education, work experience, seniority and age), household circumstances (if the worker is the head of the household and has six-year-old children or younger), the size of the firm where he/she is employed (micro, small, medium, and large), and the location of the workplace (urban or rural). The non-parametric methodology of Chernozhukov-Fernández-Melly (CFM henceforth) will decompose the wage distribution differentials by sex into their composition and structure effects –or into their explained component and gender discrimination component. Results of applying this methodology are consistent with the sticky-floor hypothesis: male workers are paid higher hourly salaries than women at lower income levels. Results also show a decreasing trend in labor income differentials that begins in the lower half of the distribution and follows all the way to the median; then, in the upper segment of the income distribution the wage gap seems to revert: female workers get better returns relative to their human capital endowments at higher income levels.

The methodology of gender income differentials

The traditional Blinder-Oaxaca methodology (Blinder, 1973; Oaxaca, 1973) allows comparing the means of labor income by sex and decomposing labor income differentials in two parts (Fortin, Lemieux & Firpo, 2011). The first part is known as the explanatory component, or the endowment or composition effect, and is calculated as the weighted sum of the difference in the means of the explanatory variables between men and women (endowments of education, experience, seniority, etc.), where weights are the estimated parameters (returns) of women's labor income functions. The second part of the decomposition is the structural effect or the non-explanatory component, which is unexplained by variables in the model and can therefore be attributed to gender discrimination (Boeri & Van Ours, 2013; Lu, 2019). This component is defined as the weighted sum of the difference in returns between men and women, where weights correspond to the estimated coefficients of men's endowments.

This methodology has a problem in identifying 'wage discrimination' in the structural component, as Altonji and Blank (1999) have noted, because it is correlated with the variables included in the composition effect and may contain other individual characteristics of the workers, as well as other types of discrimination in the labor market and in other areas of the economy. Nonetheless, Ponthieux and Meurs (2015) defend the premise of wage discrimination in the structural component arguing the persistence of a wage gap between workers with similar productive characteristics but different gender and building on current approaches for various dimensions of economic inequality between men and women. Thus, they assert that it is still possible to identify gender wage discrimination distinctly from plain wage discrimination –unequal pay for equal work– in the case of the gender wage gap. Moreover, the authors show that the coefficients of the composition part in empirical studies on wage differentials by gender have a tendency to stabilize at values close to zero, while the structural part is increasingly relevant, since it has exhibited a constant growth regardless of the inclusion of a great number and a wide variety of explanatory variables, which could signal stronger evidence on gender wage discrimination.

The novelty of the CFM methodology relative to the Blinder-Oaxaca is in the incorporation of a non-linearity assumption in the wage differentials between men and women. The composition and structural effects are assumed to be non-linear with respect to the median, so they are expected to vary in the range of labor income distribution functions; hence, results can differ at low and high levels of labor income. This makes sense, given that workers at the bottom of the labor income distribution tend to have lower endowments of education and work experience, as well as diminishing returns (payments) to such endowments, whereas workers at the top of the distribution are characterized by greater knowledge accumulation and also greater returns (Fortin et al., 2011; Machado & Mata, 2005; Melly, 2005; Chen, Chernozhukov, Fernández-Val & Melly, 2017; Chernozhukov, et al., 2013).

Chen and co-authors (2017) postulate that, if we let 1 indicate the population of men and 0 the population of women, then F^{\leftarrow} could denote the quantile or left-inverse function of the distribution function F ; also, the variable Y_j could express wages for men and women and X_j their job-market relevant characteristics. The conditional distribution functions $F_{Y_1|X_1}(y|x)$ and $F_{Y_0|X_0}(y|x)$ would then describe the stochastic assignment of wages to workers' characteristics x for men and women. Let now $F_{Y(1|1)}$ and $F_{Y(0|0)}$ represent observed wage distribution functions for men and women, and $F_{Y(0|1)}$ represent the wage distribution function that would have prevailed for men had they faced the women's conditions $F_{Y_0|X_0}$. The distribution $F_{Y(0|1)}(y) := \int_{\mathcal{X}_1} F_{Y_0|X_0}(y|x) dF_{X_1}(x)$ is called counterfactual and is built by integrating the conditional distribution of women's wages with respect to the distribution of characteristics for men, as long as the support condition $\mathcal{X}_1 \subset \mathcal{X}_0$ holds, which guarantees the compatibility of men and women's characteristics. The difference in the quantile wage function between women and men can be broken down as shown in equation (1):

$$\frac{F_{Y(1|1)}^{\leftarrow} - F_{Y(0|0)}^{\leftarrow}}{\text{Total effect}} = \frac{[F_{Y(0|1)}^{\leftarrow} - F_{Y(0|0)}^{\leftarrow}]}{\text{Composition effect}} + \frac{F_{Y(1|1)}^{\leftarrow} - F_{Y(0|1)}^{\leftarrow}}{\text{Structure effect}} \quad (1)$$

(gender discrimination)

Quantile functions are formed by combining the conditional distribution in the population of women $F_{Y(0|1)}(y) := \int_{\mathcal{X}_1} F_{Y_0|X_0}(y|x) dF_{X_1}(x)$, where $y \in \mathcal{Y}_0$, with the covariate distribution in the population of men $Q_{Y(0|1)}(\tau) := F_{Y(0|1)}^{\leftarrow}(\tau)$, $\tau \in (0,1)$. Then, the quantile effect (QE) functions over a set of quantile indexes $\mathcal{T} \subset (0,1)$ define the composition effect $\Delta_C(\tau) = Q_{Y(0|1)}(\tau) - Q_{Y(0|0)}(\tau)$, the structure effect $\Delta_S(\tau) = Q_{Y(1|1)}(\tau) - Q_{Y(0|1)}(\tau)$, and the total effect $\Delta(\tau) = Q_{Y(1|1)}(\tau) - Q_{Y(0|0)}(\tau)$, $\tau \in (\mathcal{T})$.

To estimate the QE function, the conditional distribution $F_{Y_0|X_0}$ and the covariate distribution F_{X_1} had to be modeled and estimated. For the same reason as [Lu \(2019\)](#) –too many binary dummy variables in the model–, a *logit* link function was chosen to implement the distribution regression estimator of the conditional distribution, as in $\hat{F}_{Y_0|X_0}(y|x) = \Lambda(x'\hat{\beta}(y))$, where Λ is the *standard logistic distribution* function, and $\hat{\beta}(y)$ is the distribution regression estimator:

$$\hat{\beta}(y) = \arg \max_{b \in \mathbb{R}^{dx}} \sum_{i=1}^{n_j} [1\{Y_{ji} \leq y\} \log \Lambda(X'_{ij}b) + 1\{Y_{ji} > y\} \log \Lambda(-X'_{ij}b)] \quad (2)$$

The CFM methodology proposes the inference of several functional tests based on a two-test statistic, the Kolmogorov-Smirnov test and the Cramer-von-Misses-Smirnov test, which [Chen et al. \(2017\)](#) use to establish a set of null hypotheses with: 1) correct parametric specification of the model for conditional distribution; 2) zero quantile effects at all the quantile indexes of interest, this is $\Delta(\tau) = 0$ for all $\tau \in (\mathcal{T})$; 3) constant quantile effects at all the quantile indexes of interest, such that $\Delta(\tau) = \Delta(0.5)$ for all $\tau \in (\mathcal{T})$; 4) first-order stochastic dominance, $\Delta(\tau) \geq 0$ for all $\tau \in (\mathcal{T})$; and, 5) negative first-order stochastic dominance, or $\Delta(\tau) \leq 0$ for all $\tau \in (\mathcal{T})$.

The non-parametric methodology of the present analysis is the same applied by [Lu \(2019\)](#) and similar to that used by [Popli \(2013\)](#) and [Arceo and Campos \(2014\)](#). However, unlike these works, this study did infer the functional tests proposed by [Chernozhukov et al. \(2013\)](#), even though it does not consider possible self-selection biases into the labor market, which is probably its main limitation.

Data and results

This section describes the handling of the data for the construction of hourly wage indicators, as well as variables to gauge workers' human capital, their household circumstances, and the size and location of their workplace. Additionally, the section presents descriptive statistics and offers an analysis of the results of distribution-function estimates for gender income differentials and their disaggregation into the composition effect and the structure effect –the latter regarded as gender discrimination.

4.1. Data and descriptive statistics

The empirical part of this research worked with microdata from the National Survey of Occupation and Employment in Mexico (ENOE by its acronym in Spanish) in virtue of its many advantages, the most relevant being that it is intended to capture in-depth information on the characteristics of occupations and employment. Given ENOE's reliable regularity and timely publication, it can offer an overview of current trends in wage inequality. The period of study runs from 2005 to 2020 and the scope is in the urban, rural and cities sectors, along with their urban and rural complements.

The survey includes a sample of individuals in a rotating panel that lasts four quarters, but only the information from the first quarter of each year will be used. This responds to a couple of reasons; first, the questionnaire of the first quarter is longer than the questionnaires in the other three quarters, the first one is expanded with all the socioeconomic characteristics of individuals. Secondly, considering four databases for each of the 15 years of the period of study would bear practical issues without adding new information to the analysis. In light of current world events, it is important noting that 2020 data is also from the first quarter of the year; therefore, the results do not reflect changes arising from the COVID-19 pandemic.

Labor income per hour refers to the salaried working population between 15 and 65 years old and was calculated as the reported monthly income divided by reported hours worked per month. There are four variables for workers' individual characteristics or human capital: education, experience, seniority and age. The education variable was constructed as years of schooling based on the completion of 7 degrees: primary (6 years), secondary (9 years), technical secondary (9 years), tertiary (12 years), technical tertiary (12 years), bachelor's degree (15-17 years), and postgraduate degrees –master's or doctor's– (over 15 years). Work experience is defined as potential years worked and was calculated as the worker's age minus years of schooling, minus six (ages 0 to 6). Seniority corresponds to years worked in the current company and was computed as the year of the survey minus the year the worker stated that he or she began working for the current company. Age indicates how old workers are.

Two more variables reflect relevant household circumstances: if the worker is the head of the household and if he or she has children of six years old or younger at home. The last couple of variables reflect characteristics of the workplace: the size of the firm where the worker is employed (micro, small, medium or large) and if the company is located in an urban area or not. Household conditions and workplace characteristics are qualitative variables, where “yes” equals 1 and “no” is equal to 0. The explanatory variables and the descriptive statistics used the survey expansion factor w , while labor income per hour used an expansion factor weighted by total hours worked (Arceo-Gomez & Campos-Vazquez, 2014; Mendoza-González et al., 2017). Finally, the natural logarithm of the ratio of the indicators between men and women served to estimate the rate of men's labor income or endowments with respect to women's. Hence, the gap is $\Delta = \log \left(\frac{y_h}{y_m} \right)$.

Table 1 shows summary statistics of the variables by sex and decile of the labor income distribution [0.1,0.2, ..., 0.9], as well as mean and standard deviation (sd) for 2005-2020 pooled data. From the information presented in Table 1, the gender hourly-wage gaps along the income distribution can be observed, as well as the indicators that make up the endowments of male and female workers, and the composition effect factors in the quantile distribution, in accordance with the CFM methodology. The gender hourly wage gap is slightly higher than average at the lowest income levels (deciles 0.1 and 0.2), is close to the average in the middle segment of the distribution (deciles 0.3 to 0.6), and is overturned at medium-high levels of income (deciles from 0.7 upwards). In other words, women at the upper tail of the distribution earn a higher labor income than men.

Table 1. Summary statistics of labor market variables in Mexico, 2005-2020 (pooled data)

	Quantiles										mean	sd
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9			
Continuous variables												
Hourly wage (pesos)												
Women	9.8	13.0	16.5	19.4	22.9	27.1	33.3	44.4	65.4	33.1	44.1	
Men	10.9	14.5	17.4	20.4	24.1	27.9	33.3	42.0	62.5	33.8	50.1	
Log Men/Women	0.11	0.11	0.05	0.05	0.05	0.03	0.00	-0.06	-0.05	0.02	0.13	
Worked hours (weekly)												
Women	14.0	24.0	30.0	36.0	40.0	44.0	48.0	48.0	56.0	38.0	16.6	

Table 1 (continued). Summary statistics of labor market variables in Mexico, 2005-2020 (pooled data)

	Quantiles									mean	sd
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9		
Men	25.0	36.0	40.0	45.0	48.0	48.0	51.0	56.0	66.0	46.8	15.9
Log Men/Women	0.58	0.41	0.29	0.22	0.18	0.09	0.06	0.15	0.16	0.21	-0.04
Education (years)											
Women	5.0	6.0	9.0	9.0	9.0	12.0	12.0	14.0	17.0	10.0	4.5
Men	4.0	6.0	7.0	9.0	9.0	9.0	12.0	12.0	16.0	9.4	4.3
Log Men/Women	-0.22	0.00	-0.25	0.00	0.00	-0.29	0.00	-0.15	-0.06	-0.06	-0.03
Work experience (years)											
Women	2.0	5.0	10.0	14.0	18.0	22.0	26.0	31.0	38.0	19.1	13.4
Men	2.0	6.0	10.0	14.0	18.0	23.0	27.0	32.0	39.0	19.6	13.7
Log Men/Women	0.00	0.18	0.00	0.00	0.00	0.04	0.04	0.03	0.03	0.02	0.02
Seniority (years)											
Women	1.0	1.0	1.0	2.0	3.0	5.0	7.0	11.0	18.0	6.5	7.6
Men	1.0	1.0	2.0	3.0	4.0	6.0	10.0	15.0	21.0	8.1	9.1
Log Men/Women	0.00	0.00	0.69	0.41	0.29	0.18	0.36	0.31	0.15	0.22	0.18
Age (years)											
Women	21.0	25.0	29.0	33.0	36.0	40.0	44.0	48.0	54.0	36.9	12.0
Men	20.0	24.0	28.0	32.0	36.0	40.0	44.0	49.0	55.0	36.7	12.6
Log Men/Women	-0.05	-0.04	-0.04	-0.03	0.00	0.00	0.00	0.02	0.02	-0.01	0.05
Qualitative variables (yes=1)											
Head of household											
Women	0.24	0.22	0.23	0.23	0.23	0.23	0.24	0.24	0.28	0.24	0.02
Men	0.52	0.54	0.56	0.59	0.64	0.67	0.73	0.76	0.83	0.65	0.10
Log Men/Women	0.80	0.89	0.90	0.95	1.02	1.06	1.10	1.14	1.08	1.00	1.78
Children under six years old											
Women	0.06	0.05	0.05	0.05	0.06	0.05	0.05	0.05	0.05	0.05	0.00
Men	0.18	0.20	0.22	0.22	0.23	0.24	0.26	0.26	0.25	0.23	0.02
Log Men/Women	1.14	1.36	1.41	1.40	1.43	1.56	1.57	1.57	1.58	1.45	2.32
Micro firm											
Women	0.83	0.73	0.60	0.51	0.50	0.38	0.32	0.21	0.13	0.47	0.22
Men	0.61	0.57	0.49	0.45	0.44	0.40	0.35	0.27	0.13	0.41	0.14
Log Men/Women	-0.32	-0.25	-0.19	-0.13	-0.14	0.05	0.10	0.22	-0.06	-0.13	-0.44
Small firm											
Women	0.09	0.13	0.18	0.21	0.19	0.24	0.27	0.31	0.28	0.21	0.07
Men	0.18	0.19	0.21	0.22	0.22	0.23	0.25	0.26	0.24	0.22	0.02
Log Men/Women	0.75	0.39	0.15	0.04	0.18	-0.07	-0.08	-0.17	-0.15	0.05	-1.03
Medium firm											
Women	0.04	0.07	0.11	0.13	0.13	0.15	0.16	0.16	0.18	0.13	0.04
Men	0.11	0.12	0.14	0.15	0.15	0.15	0.16	0.17	0.20	0.15	0.03
Log Men/Women	1.03	0.52	0.23	0.12	0.14	0.02	0.04	0.06	0.11	0.18	-0.53
Large firm											
Women	0.04	0.07	0.11	0.15	0.18	0.22	0.26	0.31	0.40	0.19	0.11
Men	0.10	0.12	0.16	0.18	0.19	0.22	0.24	0.30	0.43	0.21	0.10
Log Men/Women	0.92	0.56	0.34	0.21	0.06	-0.03	-0.08	-0.05	0.07	0.10	-0.15
Urban workplace											
Women	0.82	0.87	0.88	0.90	0.90	0.93	0.94	0.95	0.96	0.91	0.04
Men	0.76	0.77	0.83	0.83	0.87	0.88	0.90	0.93	0.95	0.86	0.06
Log Men/Women	-0.08	-0.11	-0.06	-0.07	-0.04	-0.05	-0.04	-0.02	-0.01	-0.05	0.36

Source: Author's estimates using pooled data from ENOE 2005-2020.

One of the reasons why men earn on average higher incomes from labor than women is the difference in the total amount of hours worked. Although both men and women at lower-income deciles tend to work less hours compared to higher-income workers, the amount of women's work hours is lower relative to men's across all deciles of the distribution.

Regarding the human capital factors that affect gender income differentials—years of schooling, potential work experience and years worked for current employer—the statistics show clearly that women tend to outsmart men, having more years of schooling or at least the same across all deciles. In general, practically all income levels present a positive relationship between education and labor income. Potential work experience is also positively related to higher income levels, with male workers usually recording more years active, although their advantage relative to the level of income is close to zero and constant. As to seniority, the number of years worked for the current employer has a positive relationship with higher levels of income. On average, 22% of men have worked for their current employer longer than working women, and this is the component most favorable to men, thus negatively affecting working women. Statistics indicate that at the lowest deciles (0.1 and 0.2) there is no particular gender advantage, but from the third decile onwards seniority starts accumulating and then diminishes in a non-asymptotic manner. In the case of age, years since birth can explain higher wages at higher income levels, but results do not flag this factor as affecting income differentials between men and women.

Household circumstances have mixed effects. Being the head of household is an important characteristic positively associated with male workers, but not an important component in the case of female workers; however, the interaction between age and being the household head tends to favor men over women. Female workers with children under the age of six are a small portion of working women (6%) and this factor does not seem to have a relationship with income levels. In contrast, around 23% of male workers have children younger than six years old, with this percentage growing at high-income levels.

Lastly, indicators for the characteristics of the workplace show a general inclination, whereby low-income workers (men and women) are largely concentrated in microenterprises, while medium- and high-income workers tend to be in small and large companies. Women with low wages tend to concentrate in micro and small enterprises in relation to their male peers. And although the majority of all workers are located in cities, both men and women at higher-income levels work mainly in urban areas, whereas women are more likely to work in non-urban areas than men at low-income levels.

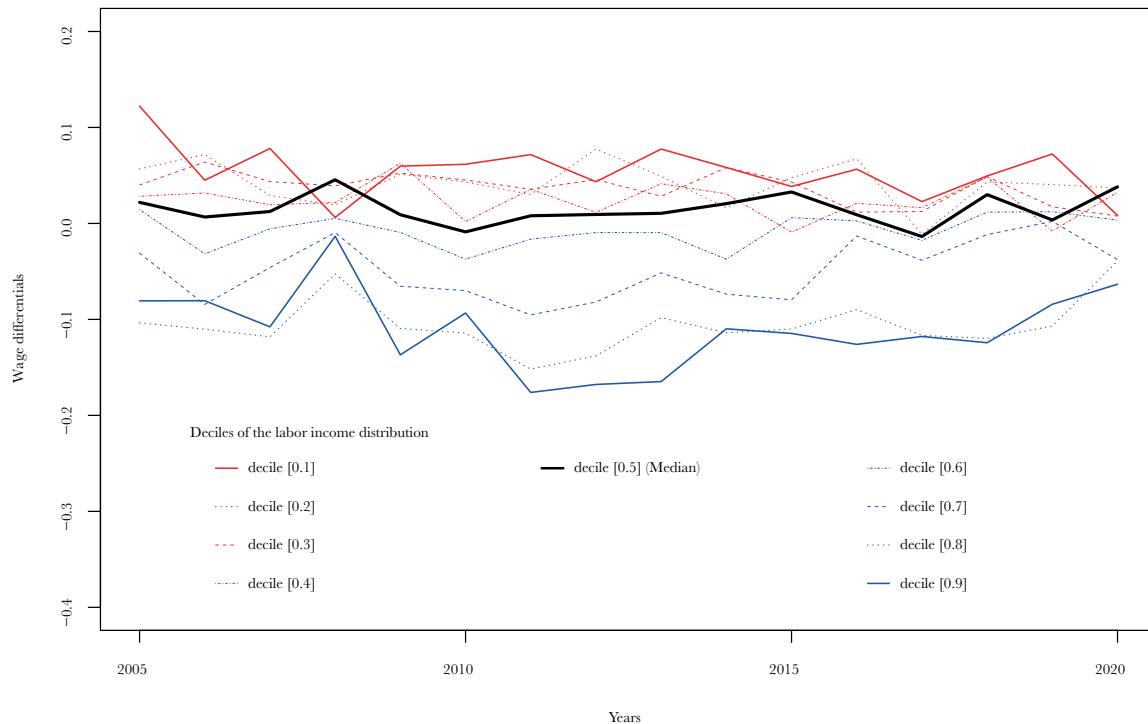
4.2. Analysis of results

Following the CFM non-parametric methodology, distribution functions for gender differentials in hourly wages were specified in order to compare and weigh the importance of endowments, contexts and returns in total effects, as well as in their constituent parts: the composition effects and, especially, the structure effects or gender discrimination. The distribution functions included workers' individual characteristics as endowments, while household conditions and workplace characteristics were included as context. Next, the logit method on the normal conditional distribution function was used to run the regression (Lu, 2019). The number of bootstrap simulations that were used to estimate the standard deviation in the tests and the confidence intervals for the selection of the subset of quantile indexes was 100. The expansion factor of the ENOE survey was used as a weight. The quantile tails were trimmed $\mathcal{T} = [0.1, \dots, 0.9]$ to eliminate extreme cases of low and high wages, as recommended by the CFM methodology. As mentioned, all variables were built from ENOE microdata for the period 2005-2020. The *counterfactual library* designed by Chen et al. (2017) in the R software was used to program the CFM methodology.

Figure 1 charts the gender wage differentials resulting from the CFM non-parametric model by deciles from 2005 to 2020. The lines show a stable behavior in gender hourly-wage differen-

tials after the economic crisis of 2009. This means that, on average, men receive a higher labor income than women and this has not changed in the last 15 years. Hourly wage differentials are characterized by being positive in deciles below the median and negative in deciles above the median from 2005 to 2020. Figure 1 also shows that hourly wage differentials in deciles below the median remained relatively stable, while in deciles above the median –which show an effect in favor of women– the differentials changed the course they had prior to the economic crisis of 2009 and fell towards 2020.

Figure 1. Trends in gender hourly wage differentials by deciles in Mexico, 2005-2020



Source: Author's estimation of the CFM methodology using pooled data from ENOE 2005-2020.

Notes: The logit method implements the regression estimator of the conditional distribution with normal link function distribution. Number of quantile regressions (n_{reg}) = 100; bootstrap = 100; weights = fac (survey expansion factor); and tail trimming $\mathcal{T} = [0.1, 0.9]$.

Next, the analysis examines whether these trends can be explained by a combination of composition and structure effects or if the structural effect (gender discrimination) predominates ([Mendoza-González, et al., 2017](#); [Lu, 2019](#)). Table 2 illustrates the quantile effects tests by gender for the differential in hourly wages in its total, composition, and structure effects in 2005, 2010, 2015 and 2020.² The statistical probabilities of the correct specification of the parametric model indicate that the conditioned distributions for the total, composition and structure effects act differently than the sample of observed data. [Chen et al. \(2017\)](#) contend that this test is not about linear probability in logit models and, therefore, is not a classic test of correct specification. Nevertheless, the results shown in Table 2 indicate that the counterfactual is not identical to the empirical distribution in the referenced population; thus, the weights of the marginal functions in the endowments and returns in the total, composition, and structure effects are more relevant.

Tests of the no-effect hypothesis also indicate that the total, composition, and structure gaps for gender hourly wages are different from zero in the period of study. The statistical tests indicate a non-linear relationship in total and structural wage gaps, and a constant relationship for the composition effects. Further, the stochastic dominance tests confirm the double preva-

² The statistical tests for all 15 years in the period 2005-2020 were estimated; however, only four years are presented in the interest of simplicity and considering that the results are rather similar to each other. A full display of Table 2 can be requested from the author.

lence, positive and negative, of wage gaps in total effects from 2005 to 2020. Lastly, the negative stochastic dominance prevails in the composition effects and is positive in the structure effects throughout the years studied (see Table 2).

Table 2. Inferential statistics on counterfactual quantile processes for hourly wages: Kolmogorov-Smirnov (KS) and the Cramer-von-Misses-Smirnov (CMS) tests, 2005-2020

	2005		2010		2015		2020	
Null Hypothesis	KS	CMS	KS	CMS	KS	CMS	KS	CMS
Hourly wages: function estimate quantile effect (QE) with deciles {0.10,0.20,...,0.80,0.90}								
<i>Quantile Effects -- Total</i>								
Correct specification of the parametric model	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
No effect: $QE(\tau)=0$ for all taus	0.00	0.00	0.00	0.00	0.00	0.00	0.08	0.02
Constant effect: $QE(\tau)=QE(0.5)$ for all taus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Stochastic dominance: $QE(\tau)>0$ for all taus	0.00	0.00	0.03	0.00	0.01	0.01	0.04	0.07
Stochastic dominance: $QE(\tau)<0$ for all taus	0.00	0.00	0.00	0.00	0.00	0.00	0.09	0.03
<i>Quantile Effects -- Composition</i>								
Correct specification of the parametric model	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
No effect: $QE(\tau)=0$ for all taus	0.25	0.02	0.00	0.00	0.00	0.00	0.00	0.00
Constant effect: $QE(\tau)=QE(0.5)$ for all taus	0.85	0.77	0.10	0.01	0.00	0.01	0.51	0.09
Stochastic dominance: $QE(\tau)>0$ for all taus	0.22	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Stochastic dominance: $QE(\tau)<0$ for all taus	0.87	0.87	0.81	0.81	0.93	0.93	0.83	0.83
<i>Quantile Effects -- Structure</i>								
Correct specification of the parametric model	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
No effect: $QE(\tau)=0$ for all taus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Constant effect: $QE(\tau)=QE(0.5)$ for all taus	0.03	0.03	0.58	0.45	0.17	0.05	0.00	0.03
Stochastic dominance: $QE(\tau)>0$ for all taus	0.88	0.88	0.83	0.83	0.91	0.91	0.90	0.90
Stochastic dominance: $QE(\tau)<0$ for all taus	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: Results of the Chernozhukov, Fernández-Val and Melly (2013) methodology on ENOE data.

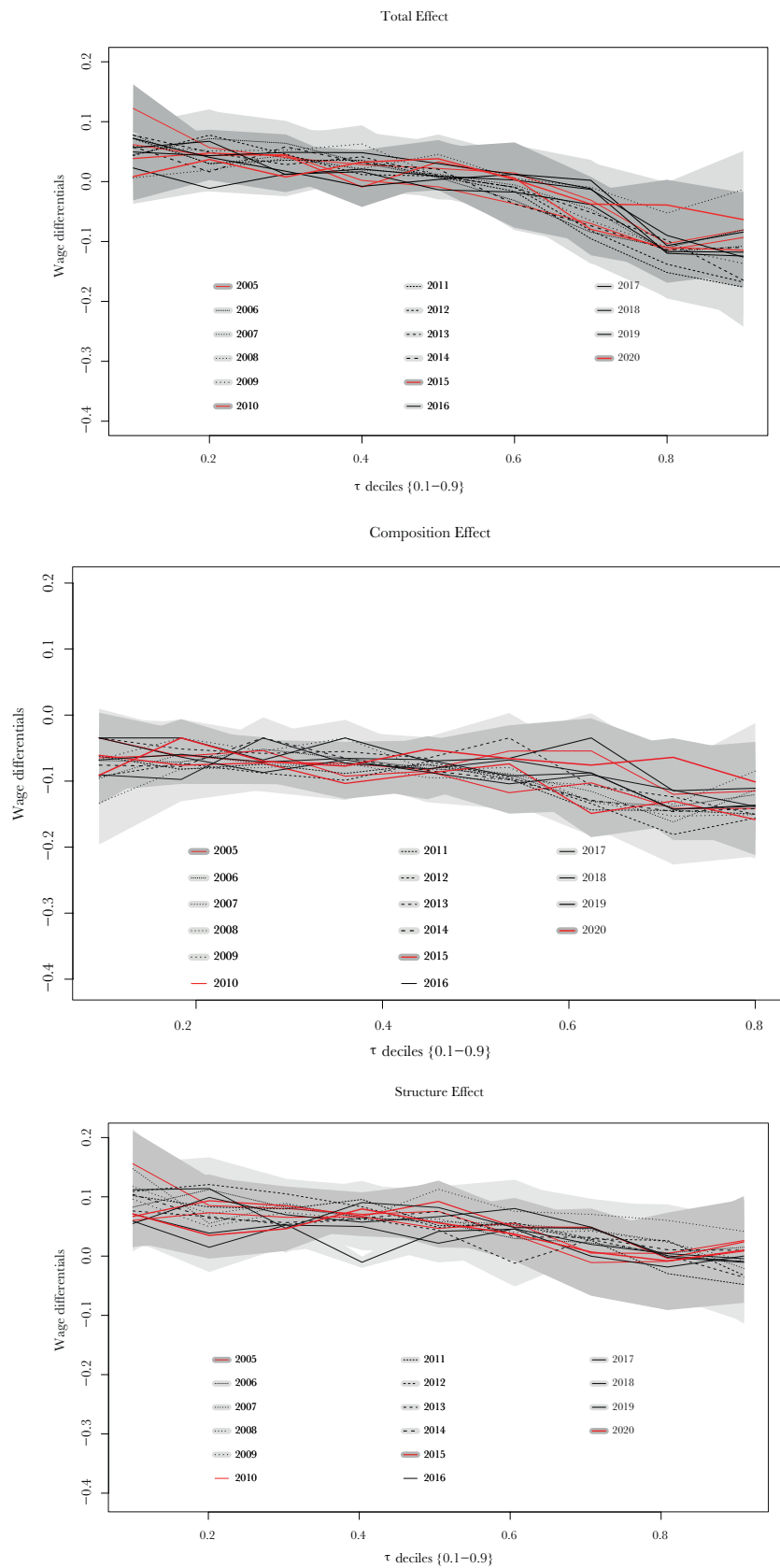
Notes: The *logit* method implements the distribution regression estimator of the conditional distribution with normal link function distribution; number of quantile regressions (nreg) = 100; bootstrap = 100; weights = fac (expansion factor); and tail trimming $\mathcal{T} = [0.1, 0.9]$.

Figure 2 depicts the gender hourly wage differentials and their 95% confidence intervals. Light gray color stains all years, except for 2005, 2010, 2015 and 2020, which are in gray to spot the possibility that the variability could affect the inferences. Withal, the ranges of the confidence intervals are very similar at all income levels, so there is no inference bias due to variability.

According to the results of the quantile effects in gender hourly-wage differentials in Figure 2, male workers secure hourly wages up to 12% higher than women at low segments of the income distribution. These differential decreases gradually until quantile 0.6, after which the hourly-wage gaps become increasingly negative until they reach 12%. This result is very interesting because it shows a gender hourly-wage gap against women at low labor-income levels and in their favor at the opposite end of the distribution for all years analyzed. In line with the

statistical tests, results for the composition effects show that hourly-wage gaps that are explained by the endowments are negative and increase at higher labor income levels.

Figure 2. Quantile effects in gender hourly-wage differentials in Mexico, 2005-2020



Source: Results of the Chernozhukov, Fernández-Val and Melly (2013) methodology on ENOE data.
 Notes: The *logit* method implements the distribution regression estimator of the conditional distribution with normal link function distribution; number of quantile regressions (nreg) = 100; bootstrap = 100; weights = fac (expansion factor); tail trimming $\mathcal{T} = [0.1, 0.9]$; and 95% confidence interval. All years are colored in light gray, except for 2005, 2010, 2015 y 2020, which are in gray.

The results for the accumulation of human capital suggests that women should gain higher wages per hour than men, around the magnitude of 6% at low income levels and of 14% at high income levels. In addition, although the statistical tests show that there is a positive stochastic dominance in the structure effects or gender discrimination against women, Figure 2 shows that this is true only below quantile 0.7, after which it fluctuates between 2 and -4%. The conclusion is that there is enough evidence to claim gender discrimination in hourly wages at low and middle-income levels, but discrimination seems to disappear at high levels of labor income (see Table 2 and Figure 2).

To support the conclusion of gender discrimination in hourly wages at low- and middle-income levels, Table 3 displays the trends in relative importance of composition and structure quantile effects in total hourly gender wage differences in Mexico for selected years in the period of study. The first feature to highlight is the increase in the relative importance of the average structure effect in explaining total gender hourly wage differences, from 38.6% in 2005 to 45.2% in 2020. Secondly, it stands out that the relative importance of the structure effect tends to increase in the range of 50 to 75% at low income levels (deciles 0.1 to 0.4). Third and last, notice the lower relative importance of the structure effect in explaining the total gender hourly wage differences at the highest income levels, even though there is an increasing trend going from 1.1% in 2005 to 15.5% in 2020 in the middle deciles (0.6 to 0.9). It is important to note that the interpretation of the relative importance of the composition effect presented in Table 3 must be combined with the negative sign shown in Figure 2.

Table 3. Relative importance of composition and structure quantile effects in gender hourly wage gaps in Mexico, 2005-2020

Relative effects	Quantiles: decile ranges			mean
	[0.1 - 0.4]	[0.5=median]	[0.6 - 0.9]	
<u>Year: 2005</u>				
<i>Quantile Effects</i>				
-- Total	100	100	100	100
-- Composition	28.8	41.7	98.9	61.4
-- Structure	71.2	58.3	1.1	38.6
<u>Year: 2010</u>				
<i>Quantile Effects</i>				
-- Total	100	100	100	100
-- Composition	25.9	54.2	82.4	54.1
-- Structure	74.1	45.8	17.6	45.9
<u>Year: 2015</u>				
<i>Quantile Effects</i>				
-- Total	100	100	100	100
-- Composition	37.5	39.4	76.1	54.9
-- Structure	62.5	60.6	23.9	45.1
<u>Year: 2020</u>				
<i>Quantile Effects</i>				
-- Total	100	100	100	100
-- Composition	32.5	25.3	84.5	54.8
-- Structure	67.5	74.7	15.5	45.2

Table 4 (continued). Hourly-wage quantile models for women and men in Mexico, 2005-2020 (pooled estimate)

Quantiles		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Medium	Coefficient	0.202	0.167	0.140	0.118	0.098	0.075	0.053	0.026	-0.004
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.33
Large	Coefficient	0.306	0.262	0.229	0.202	0.178	0.154	0.131	0.098	0.067
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Urban workplace	Coefficient	0.129	0.096	0.082	0.070	0.064	0.059	0.051	0.047	0.038
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Men										
Education	Coefficient	0.049	0.052	0.055	0.057	0.060	0.062	0.064	0.067	0.070
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Work experience	Coefficient	-0.002	-0.003	-0.004	-0.004	-0.005	-0.006	-0.006	-0.007	-0.009
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Seniority	Coefficient	0.007	0.007	0.007	0.008	0.008	0.009	0.009	0.009	0.009
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Age	Coefficient	0.006	0.007	0.009	0.010	0.011	0.013	0.014	0.016	0.018
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Head of household	Coefficient	0.103	0.097	0.094	0.093	0.094	0.095	0.097	0.098	0.100
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Children aged six and under	Coefficient	-0.002	0.004	0.008	0.009	0.010	0.011	0.011	0.010	0.007
	p-value	0.45	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Firm size (factor)										
Micro (base)		0.368	0.426	0.463	0.492	0.518	0.534	0.555	0.577	0.611
Small	Coefficient	0.165	0.154	0.147	0.141	0.137	0.135	0.131	0.126	0.117
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Medium	Coefficient	0.180	0.160	0.150	0.140	0.130	0.124	0.118	0.110	0.093
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Large	Coefficient	0.287	0.260	0.240	0.227	0.215	0.206	0.195	0.187	0.178
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Urban workplace	Coefficient	0.108	0.104	0.105	0.104	0.103	0.104	0.103	0.100	0.102
	p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Men-Women coefficient ratios										
Education		0.75	0.77	0.79	0.81	0.82	0.83	0.84	0.85	0.87
Work experience		0.92	0.65	0.68	0.69	0.72	0.73	0.76	0.80	0.82
Seniority		0.51	0.49	0.49	0.50	0.52	0.55	0.57	0.60	0.66
Age		0.54	0.56	0.61	0.65	0.67	0.69	0.71	0.75	0.77
Head of household		11.23	17.91	27.23	150.31	-64.44	-40.86	162.42	-174.46	26.20
Children aged six and under		-0.03	0.09	0.16	0.16	0.19	0.20	0.18	0.16	0.11

Table 4 (continued). Hourly-wage quantile models for women and men in Mexico, 2005-2020 (pooled estimate)

Quantiles	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Firm size (factor)									
Micro (base)	1.27	1.07	0.96	0.90	0.85	0.80	0.76	0.71	0.68
Small	0.81	0.90	1.00	1.09	1.18	1.36	1.58	2.13	3.66
Medium	0.89	0.96	1.07	1.18	1.33	1.65	2.24	4.25	-22.23
Large	0.94	0.99	1.05	1.12	1.21	1.34	1.50	1.91	2.68
Urban workplace	0.84	1.08	1.28	1.48	1.62	1.76	2.01	2.12	2.66

Notes: Number of quantile regressions (nreg) = 100; bootstrap = 100; weights = fac (expansion factor); tail trimming $\mathcal{T} = [0.1, 0.9]$.

In the parametric version of the structure-effects component of the CFM methodology, the Coefficients for wage quantile models are compared by sex at each quantile $\mathcal{T} = [0.1, \dots, 0.9]$; these are called men-women coefficient ratios at the end of Table 4. Here, values above the unit indicate that the rate of men's returns is greater than the rate of women's, while for values below 1, women have the upper hand over men's returns.

In terms of human capital, the results show that education, work experience, and seniority are statistically different from zero in all estimates. Years of schooling is the most important factor and has a positive effect; the same happens with years worked for current employer, while potential years of work experience tend to reduce both women and men's wages. From the point of view of gender discrimination, male workers are expected to get higher returns than females, in a case of overvaluation of men with respect to women (Boeri & Van Ours, 2013; Jacobsen, 2007). However, the results of the estimates and the men-women coefficient ratios calculated for the human capital variables show that, in the three cases discussed here, women actually achieve higher returns than men, with the main advantage for women being their seniority. By inspecting these three variables, it turns out that the returns to women's human capital are higher compared to men's at low-income levels and lower at high-income levels. Together, these results suggest that human capital factors are not the main source of wage differentials against women in Mexico. Others have made this point previously by critiquing the theory of human capital and proposing alternative approaches (Mendoza-González et al., 2017; Andrés et al., 2019; Garza-Acevedo & Quintana-Romero, 2014).

From the perspective of discrimination and life expectancy theory (Modigliani, 1966; Sedgley & Elmslie, 2018), the expectations of higher incomes as one grows older depend on labor-market conditions; this is the reason why salary returns could be expected to increase with age at high income levels. Yet workers' age can be an element of wage discrimination itself when the income expected by age is determined differently for men and women. Wage discrimination exists in quantile models when coefficients estimated for returns related to age are higher for men than for women. In the case of the present research, the results of the estimates show there is no discrimination against women due to age, and everything seems to indicate that women have better salary conditions than men do at any income level.

On the other hand, the socio-economic approach to gender wage discrimination (Boeri & Van Ours, 2013; Jacobsen, 2007) posits that women who are the head of household and have children under the age of six (single mothers) face important time restrictions, which leads to an overvaluation of working men with the same circumstances. According to the results presented in Table 4, this statement does not hold true when comparing men and women's hourly wages in Mexico between 2005 and 2020. It is worth remembering that the proportion of working women who are heads of family (23%) is much lower than the proportion of men (64%), as well as the proportion of female workers with six-year-old children and younger (6%) relative to men (23%) (see Table 1).

The estimated coefficients of the quantile models show in general that being the head of a family and/or having children aged six or under is related to higher wage returns and is linear or relatively constant at any income level (see Tables 4). From a comparison of the coefficients of the hourly-wage models, an overvaluation of men is inferred with respect to women when they are the head of the household and have children less than or equal to six years of age. The results show that being a female head of household does not cause a salary differentiation with respect to those who are not heads, while for men it is important to be the head of a family, although the effect is linear at any income level (see Table 4). With hourly wages, the condition of having children under or equal to six years of age does not seem to be an element of discrimination against women in general (see Table 4).

The relationship between firm size and wages has been the subject of different approaches and discussions ([Lallemand & Rycx, 2007](#); [Mellow, 1982](#); [Pagan & Ullibarri, 2000](#)), but in general it can be argued that the larger the company, the better the chances to earn a decent income. However, if conditions favor men, then there is a greater likelihood of gender discrimination against women. Results from the quantile models provide enough evidence to claim that there is wage discrimination against women in small, medium, and large firms, which is exacerbated at medium and high income levels (see Table 4).

In terms of location ([Block & Walker, 1981](#)), higher returns are expected in jobs that take place in metropolitan areas rather than in non-metropolitan areas, or urban versus non-urban settings. Results demonstrate that, when the coefficients of hourly-wage quantile models are compared, there is a wage premium for working in urban areas that is higher at the lower income levels and tends to diminish at higher levels of income (see Table 4).

Discussion and conclusions

This work has analyzed the trends of gender wage differentials in Mexico along the income distribution through the use of non-parametric methods and quantile regressions. Results confirm that men earn higher salaries than women in general. However, and in contrast with other studies, evidence shows a relative stability of median hourly-wage differentials from 2005 to 2009 and a slight but consistent increase from 2010 to 2020.

Results also show a non-linear relationship between gender wage differentials and labor income levels that is consistent with the sticky floor hypothesis –which posits that male workers earn higher hourly salaries than female workers at lower income levels– but departs from the idea of a non-linear U-shaped relationship, where greater gender wage differentials are observed at the low and high ends of the labor-income distribution at the same time –which would be in line with both the sticky-floor and the glass-ceiling hypotheses ([Arceo-Gomez & Campos-Vazquez, 2014](#))–, even though other studies have documented this trend for Mexico ([Pagan & Ullibarri, 2000](#); [Popli, 2013](#); [Rodríguez-Pérez & Castro-Lugo, 2017](#)). Findings here suggest that women earn higher wages than men per hour at higher levels of income, but men's hourly salaries are up to 15% larger than women's at low-income levels, which implies that women have clear disadvantages when the number of worked hours is taken into account to estimate labor income.

The human capital endowment did not turn out to be relevant in the models here estimated, since the composition effect was negative; on the other hand, the structure effect increased its relative importance (45%) in explaining why men earn higher wages than women, which accrues as more solid evidence on gender wage discrimination in Mexico. This conclusion is in accordance with findings from most studies on the subject, showing that the coefficients of the composition part in empirical studies on gender wage differentials have a tendency to stabilize at values close to zero, while coefficients of the structural part are increasingly more relevant, and that this is consistent regardless of the inclusion of a large number and a wide variety of

explanatory variables ([However, Ponthieux and Meurs, 2015](#)). Instead, endowment valuation –or returns to human capital factors– largely explains the gender wage gap in Mexico, with the structural effect –or gender discrimination– emerging as the main explanatory component.

Based on the outcomes from the parametric estimations of the wage quantile models, the conclusion is that education, work experience and seniority (three out of four factors of human capital education) are not the main sources of wage discrimination against women. In turn, age (the fourth factor) is a cause for wage discrimination against women when hourly wages are considered, especially at low-income levels. In particular, being the head of household and being the mother of children aged six or younger are two important factors of discrimination against women; men can earn almost twice what women make because of these two circumstances alone. Evidence was also found for discrimination against women when hourly wages are considered in small, medium, and large firms, which worsens at medium-high levels of income.

In summary, results indicate that the greatest wage differentials and wage discrimination against women occur in the working population with the lowest levels of hourly income. This begs for great efforts in international organizations and national institutions in Mexico towards public policies that reform cultural, social and labor market conditions to reduce the wage gap and wage discrimination against women of scarce resources.

In terms of public policy, the European Union Action Plan stands out as an example, because it specifies five areas to stimulate these changes, especially the second point proposes as an objective “equal pay for equal work and work of equal value” ([López Díaz and Santos del Cerro, 2015](#)). In this context, the current research joins to highlighting the need for mechanisms that focus on gender wage discrimination at low income levels and aspire to the ideal of a genderless remuneration of work.

Finally, the following lines for research are suggested to overcome some of the limitations in this work. The main restriction of the non-parametric CFM methodology could be that it does not consider possible self-selection bias into the labor market; this may affect the results if women of higher unobservable ability are overrepresented at the top of the hourly wage distribution, relative to what it is observed in the distribution of men. Hence, a crucial line of future research would look into the self-selection issue. On the other hand, as it currently is, the CFM methodology may be expanded into a research line that analyzes the effects of explanatory factors on wage discrimination against women and to simulate scenarios of wage-differential distributions by percentiles, instead of by deciles as done in this paper.

Another research proposal would be undertaking a long-term analysis that integrates other authors’ results for the 1980s and the 1990s decades, involving the challenging task of harmonizing labor-market surveys. In the same spirit of expanding present results, another suggestion is to perform a comparative analysis using more of the available data sources, such as the labor-market surveys (ENOE), the income-expenditure surveys (ENIGH), and the population census samples, to identify and contrast the sticky-floor and glass-ceiling hypotheses. It would be equally interesting to incorporate other variables for different types of occupations and industries. Lastly, the increase in the age coefficient across all deciles of the income distribution observed in the results of the parametric model could also be further investigated from a discrimination and life expectancy theory perspective.

Abbreviations

ILO: International Labor Organization; CFM: Chernozhukov, Fernández-Val and Melly (2013); ENEU: National Survey of Urban Employment in Mexico (*Encuesta Nacional de Empleo Urbano*); ENIGH: National Household Income and Expenditure Survey in Mexico (*Encuesta Na-*

cional de Ingreso y Gasto de los Hogares); ENOE: National Occupation and Employment Survey in Mexico (*Encuesta Nacional de Ocupación y Empleo*); QE: Quantile effect; KS: Kolmogorov-Smirnov test; CMS: Cramer-von-Misses-Smirnov test.

Acknowledgements

I would like to thank Professor María Elena Cardero at the National Autonomous University of Mexico (UNAM) for previous discussions on concepts and methodologies that motivated this research. Alejandra Ortiz-García, lecturer at UNAM Acatlán and PhD Candidate at the University of Manchester, provided helpful suggestions to the different versions of this paper, as well as language and proofreading revisions to the final draft. Recommendations from the journal's referees improved the main messages of the research and were greatly appreciated.

Funding

This research was supported by two projects financed by UNAM: “PAPIIT-IN308721 *Public policies for urban economic reactivation and restructuring in Mexico in the face of economic and social impacts of COVID-19*” and “PAPIIT-IN306513 *Employment in Mexico of men and women. A regional and spatial analysis of segregation, segmentation and discrimination*”.

Availability of data and materials

Upon request from the author.

Competing interests

The author declares no competing interests.

References

- Aigner, D. J., and Cain, G. G. (1977). Statistical theories of discrimination in labor markets. *ILR Review*, 30(2), 175-187. doi: 10.1177/001979397703000204.
- Altonji, J. and Blank, R. (1999). Chapter 48 Race and gender in the labor market. In Ashenfelter, O. and Card, D. (Eds.) *Handbook of labor economics Volume 3C*, (pp. 3143-3259). Elsevier. doi: 10.1016/S1573-4463(99)30039-0.
- Andrés-Rosales, R., Czarnecki, L. and Mendoza, M. A. (2019). A spatial analysis of precariousness and the gender wage gap in Mexico, 2005–2018. *Journal of Chinese Sociology*, 6(13), 1-21. doi: 10.1186/s40711-019-0104-2.
- Arceo-Gomez, E., y Campos-Vazquez, R. (2014). Evolución de la brecha salarial de género en México. *El Trimestre Económico*, 81(323), 619-653.
- Becker, G. S. (1957). *The economics of discrimination*. Chicago: University of Chicago Press.
- Becker, G. S. (1981). An economic theory of discrimination. In W. E. Block y M. A. Walker (Eds.) *Discrimination, affirmative action, and equal opportunity* (pp. 129–146). Vancouver: The Fraser Institute.
- Blinder, A. S. (1973). Wage discrimination: Reduced form and structural estimates. *The Journal of Human Resources*, 8(4), 436-455. doi:10.2307/144855.

- Block, W., and Walker, M. (1981). Introduction. In W. Block y M. Walker (Eds.), *Discrimination, affirmative action, and equal opportunity* (pp. 5-36). Vancouver: The Fraser Institute.
- Boeri, T., and Van Ours, J. (2013). *The economics of imperfect labor markets* (2nd ed.) New Jersey: Princeton University Press.
- Brown, C. J., Pagan, J. A., and Rodríguez-Oreggia, E. (1999). Occupational attainment and gender earnings differentials in Mexico. *ILR Review*, 53(1), 123-135. doi:10.1177/001979399905300107.
- Cardero, M. E., Mendoza, M. A., and Galán, P. (2015). The employment of women in the manufacturing industry after NAFTA. Discrimination and segregation. *Global Journal of Human-Social Science: Arts and Humanities - Psychology*, 15(2), 52-69.
- Chen, M., Chernozhukov, V., Fernández-Val, I., and Melly, B. (2017). Counterfactual analysis in R: A vignette. *CEMMAP Working paper CWP64/17*. Retrieved from: <https://www.cemmap.ac.uk/wp-content/uploads/2020/08/CWP6417.pdf>
- Chernozhukov, V., Fernández-Val, I., and Melly, B. (2013). Inference on counterfactual distributions. *Econometrica*, 81(6), 2205-2268. doi:10.3982/ECTA10582.
- Del Razo, L. M. (2003). Estudio de la brecha salarial entre hombres y mujeres en México (1994-2001). Documentos de Investigación de la Secretaría de Desarrollo Social 16776.
- DiNardo, J., Fortin, N. M., y Lemieux, T. (1996). Labor market institutions and the distribution of wages, 1973-1992: A semiparametric approach. *Econometrica*, 64(5), 1001-1044.
- Fortin, N., Lemieux, T. and Firpo, S. (2011). Chapter 1 Decomposition methods in economics. In Ashenfelter, O. and Card, D. (Eds.) *Handbook of labor economics Vol. 4A* (pp. 1-102). Elsevier.
- Garza-Acevedo, B. E., and Quintana-Romero, L. (2014). Determinantes de la desigualdad salarial en las regiones de México: 2005-2010. Una visión alternativa a la teoría del capital humano. *Paradigma Económico*, 6(1), 33-48.
- ILO (2019) Women in the World of Work. Pending Challenges for Achieving Effective Equality in Latin America and the Caribbean (Thematic Labour Overview 5). Lima: International Labor Organization Regional Office for Latin America and the Caribbean, 188. Retrieved from: https://www.ilo.org/wcmsp5/groups/public/---americas/---ro-lima/documents/publication/wcms_736930.pdf
- Jacobsen, J. (2007). *The economics of gender* (3rd ed.). Oxford: Blackwell Publishing.
- Koenker, R. (2018). Quantile regression in R: A vignette. Retrieved from: <https://cran.r-project.org/web/packages/quantreg/vignettes/rq.pdf>
- Lallemand, T., and Rycx, F. (2007). Employer size and the structure of wages: A critical survey. *Reflets et perspectives de la vie économique*, 46(2-3), 75-87.
- López Díaz, E., and Santos del Cerro, J. (2015). Discriminación de género en el mercado laboral español y políticas de regulación públicas. *Revista de Evaluación de Programas y Políticas Públicas*, 5, 63-82, doi: 10.5944/reppp.5.2015.15130.
- Lu, H. (2019). Gender Wage Gap in Canada: An Analysis using Counterfactual Distributions Regression. *Economics - Research Papers University Ottawa*. Retrieved from: <https://ruor.uottawa.ca/handle/10393/39202>.
- Machado, J. A. F., and Mata, J. (2005). Counterfactual decomposition of changes in wage distributions using quantile regression. *Journal of Applied Econometrics*, 20(4), 445-465. doi:10.1002/jae.788.
- Martínez-Jasso, I., y Acevedo-Flores, G. J. (2002). La brecha salarial en México con enfoque de género, capital humano, discriminación y selección muestral. *Ciencia UANL*, 21(2), 43-80.

- Mellow, W. (1982). Employer size and wages. *Review of Economics and Statistics*, 64(3), 495-501. doi: 10.2307/1925949.
- Melly, B. (2005). Decomposition of differences in distribution using quantile regression. *Labour Economics*, 12(4), 577-590. doi:10.1016/j.labeco.2005.05.006.
- Mendoza-Cota, J. E., y García-Bermúdez, K. J. (2009). Discriminación salarial por género en México. *Problemas del Desarrollo*, 40(156), 77-99. doi: 10.22201/iiec.20078951e.2009.156.7754.
- Mendoza-González, M. Á., Cardero-García, M. E., y Ortiz-García, A. S. (2017). Algunos hechos estilizados y explicativos sobre el diferencial y la discriminación salarial por sexo en México, 1987-2015. *Investigación Económica*, 76(301), 103-135. doi:10.1016/j.inveco.2017.12.004
- Meza-González, L. (2001). Wage inequality and the gender wage gap in Mexico. *Economía Mexicana Nueva Época*, 10(2), 291-323.
- Modigliani, F. (1966). The life cycle hypothesis of savings, the demand for wealth and the supply of capital. *Social Research*, 33(2), 160-217.
- Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *Review of Economics and Statistics*, 90(2), 290-299.
- Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, 14(3), 693-709. doi: 10.2307/2525981.
- Pagan, J. A., and Ullibarri, M. (2000). Group heterogeneity and the gender earnings gap in Mexico. *Economía Mexicana Nueva Época*, 9(1), 23-40.
- Ponthieux, S. and Meurs, D. (2015). Chapter 12 Gender Inequality. In Atkinson, A and Bourguignon, F. *Handbook of Income Distribution Volume 2A* (pp. 981-1146). doi: 10.1016/B978-0-444-59428-0.00013-8.
- Popli, G.K. (2013). Gender wage differentials in Mexico: A distributional approach. *Journal of the Royal Statistical Society*, 176(2), 295-319. doi:10.1111/j.1467-985X.2011.01031.x.
- Rodríguez-Pérez, R. E., y Castro-Lugo, D. (2014). Discriminación salarial de la mujer en el mercado laboral de México y sus regiones. *Economía, Sociedad y Territorio*, 14(46), 655-686.
- Rodríguez-Pérez, R. E., and Castro-Lugo, D. (2017). Trade liberalization and gender wage gap in Mexico. *Modern Economy*, 8(10), 1167-1185. doi: 10.4236/me.2017.810081.
- Sedgley, N., and Elmslie, B. (2018). Discrimination, human capital, and life expectancy in a model of economic development. *International Journal of Economic Theory*, 14(3), 211-232. doi: 10.1111/ijet.12155.