Análisis Económico, 41(106), 153–174, enero - abril de 2026, ISSN: 0185-3937, e-ISSN: 2448-6655

https://doi.org/10.24275/uam/azc/dcsh/ae/2026v41n106/Gallardo

Wage inequality of indigenous workers in Mexico

La inequidad salarial de los trabajadores indígenas en México

Roberto Gallardo del Ángel*, Mario Miguel Ojeda Ramírez**
y Julia Aurora Montano Rivas***

*Facultad de Economía, Universidad Veracruzana, México.
Correo electrónico: rogallardo@uv.mx. ORCID: http://orcid.org/0000-0003-4661-0675

**Facultad de Estadística e Informática, Universidad Veracruzana, México.
Correo electrónico: mojeda@uv.mx. ORCID: http://orcid.org/0000-0001-6161-3968

***Facultad de Estadística e Informática, Universidad Veracruzana, México.
Correo electrónico: julmontano@uv.mx. ORCID: http://orcid.org/0000-0002-0434-4113

ABSTRACT

This paper is a study of the wage distribution of indigenous workers and its changes over time. It offers an analysis of wage inequality for this type of worker compared to non-indigenous people by occupation type. The main objective is to provide some evidence that wages have become less unequal for both types of workers in recent years. Several statistical tools are used for measuring inequality and estimating the wage gap, such as the Gini index, Wasserstein distance, kernel densities, Oaxaca-Blinder decomposition and quantile regression. This study confirms that Indigenous workers receive lower wages compared with non-Indigenous workers, but the overall wage inequality is decreasing for all workers, with the larger decrease for indigenous people, although the improvement in wage equality is mainly reflected in the middle and upper parts of the wage distribution.

RESUMEN

Este documento es un estudio de la distribución salarial de los trabajadores indígenas y sus cambios a lo largo del tiempo. Ofrece un análisis de la desigualdad salarial para este tipo de trabajadores, en comparación con las personas no indígenas por tipo de ocupación. El objetivo principal es proporcionar alguna evidencia de que los salarios se han vuelto menos desiguales para ambos tipos de trabajadores en los últimos años. Se utilizaron varias herramientas estadísticas para medir la desigualdad y estimar la brecha salarial, como el índice de Gini, la distancia de Wasserstein, las densidades kernel, la descomposición de Oaxaca-Blinder y la regresión de cuantiles. Este estudio confirma que los trabajadores indígenas reciben salarios más bajos en comparación con los trabajadores no indígenas, pero la desigualdad salarial general está disminuyendo para todos los trabajadores, siendo la mayor disminución para las personas indígenas, aunque la mejora en la igualdad salarial se refleja principalmente en las partes media y alta de la distribución salarial.

Received: March/05/2024
Accepted: April/28/2025
Posted: January/12/2026

Keywords:
| Wage inequality |
| Oaxaca-Blinder decomposition |
| Quantile regression |

Palabras clave:
| Inequidad salarial |
| Descomposición Oaxaca-Blinder |
| Regresión de quantiles |

Clasificación JEL | JEL Classification |
J15, J71, C31

INTRODUCTION

The present research is a study about the changes in the distribution of labor income for Mexican workers, with a focus on those considered native or indigenous. The study aims to compare the wage income distributions at two points in time and observe for changes in the wage gap for different occupations to observe whether indigenous workers' wages have become more equal over time. The main scientific objective is to examine whether indigenous workers have improved in terms of income and job positions compared with non-indigenous people.

In recent years, there has been a renewed interest in studying social groups that are at an economic disadvantage, such as women and minorities. One of these groups is the original people on the American continent, wrongly and commonly referred to as Indians. Institutionalized discrimination against indigenous people originated during the conquest of México by Spaniards, with the establishment of a “Castas” system, which divided individuals and households according to the admixture of three main races: European, Indian or native, and African.

Unlike developed countries, where statistics based on ethnicity (race) are available, México has not yet developed consistent information systems based on ethnicity and cultural diversity. Flores and Telles (2012) mentioned that the little official interest in race or ethnicity comes from the hope that most indigenous populations would merge and accept a whitening process. But until now, there is still a large portion of the Mexican population that considers themselves as indigenous or descendants of the original population. For example, in 2010, the United Nations Development Program estimated 10.8 million indigenous people distributed nationwide, while the Mexican Population Census estimated only 6.7 million in the same year. This is because classifying this population is a difficult task and would depend on measurable physiological features. From the 2022 Mexican National Household Income and Expenditure Survey (ENIGH), the population that was able to speak an Indigenous language was about 7.6 million, and those that self-identified as indigenous were little more than 34.4 million, from a total population of almost 129 million Mexicans. The problem is that 70% of the indigenous population lives below the poverty line and has serious disadvantages when competing in the labor market with those considered non-indigenous.

As an example, in terms of education, only 3.4% of the total population that self-identify as indigenous completed tertiary education, and this percentage decreases to 0.3% for those that speak an indigenous language. Table 1 shows that the non-indigenous population usually has a higher level of education which is also associated with a higher mean income and paid work in subordinated positions.

As mentioned before, an important issue will be which concept to use to describe an indigenous person, either the self-identification or whether one speak an indigenous dialect. It is difficult to define a boundary to build a working category of an indigenous person because most Mexicans have, to some degree, Indigenous ancestry due a continuous process of admixture. Table 1 shows the differences in classification of Indigenous population. If only language is considered to define an indigenous person, then, there are about 7.6 million of those individuals (about 6% of the entire population) in México. If any Mexican can identify himself/herself as indigenous, then there are about 34.4 million of indigenous people (about 27% of the entire population).

Still, the question remains on defining the concept of “indigenous” that allows us to classify those who are native or descendants of the original population of México. For example, in Canada, an indigenous people are those original to the country together with their descendants, and this is a genetic definition that has important legal consequences. By contrast, as most of the Mexican population may claim Indigenous ancestry, the official definition is not a racial one but a cultural one. Currently, an indigenous person is someone who speaks an Indian dialect and lives in a predominantly Indian community. The term “Mestizo” is applied to the rest of the population that, despite having Indian ancestry, may have adopted European-style society and values. Perhaps the best survey to study discrimination against indigenous people is the 2022 Mexican National Survey of Discrimination (ENADIS), which provides detailed information on ethnicity, cultural diversity, religious affiliation, disabilities, etc. However, the ENIGH contains information that allows the classification on the individuals that speak an indigenous dialect, and those that self-identify themselves as Indigenous. Additionally, the ENIGH contains detailed information on the types of income sources for everyone, which is the focus of our analysis as this study about the distributions of labor income for both indigenous and non-indigenous populations.

Table 1
Indigenous population in México, 2022

Self-identified as Indigenous

	Population			Formal Education			
	Pop>3	*Men*	*Women*	*No education*	*Primary*	*Secondary*	*Tertiary*
Indigenous	34,434,036	16,487,413	17,946,623	2.19%	10.48%	11.61%	3.36%
No Indigenous	90,126,735	43,053,458	47,073,277	3.38%	21.65%	31.71%	15.62%
Total	124,560,771	59,540,871	65,019,900	5.57%	32.13%	43.31%	18.98%

	Marital Status				*Mean Age*	*Mean Income*	*Occupied & Subordinated*
	Pop>12	*Married*	*Single*	*Separated*			
Indigenous	28,928,754	53.61%	32.43%	13.96%	34.9	$20,698.78	81.40%
No Indigenous	28,928,754	53.61%	32.43%	13.96%	34.7	$30,206.92	70.47%

Speak an indigenous dialect

	Population			Formal Education			
	Pop>3	*Men*	*Women*	*No education*	*Primary*	*Secondary*	*Tertiary*
Indigenous	7,579,934	3,621,950	3,957,984	0.97%	2.76%	2.05%	0.31%
No Indigenous	116,980,837	55,918,921	61,061,916	4.61%	29.37%	41.27%	18.67%
Total	124,560,771	59,540,871	65,019,900	5.57%	32.13%	43.31%	18.98%

	Marital Status				*Mean Age*	*Mean Income*	*Occupied & Subordinated*
	Pop>12	*Married*	*Single*	*Separated*			
Indigenous	6,673,731	63.90%	21.60%	14.51%	39.1	$15,597.67	98.88%
No Indigenous	98,745,397	50.12%	35.61%	14.27%	34.5	$28,405.73	71.75%

Own estimation using the ENIGH 2022. The classification of indigenous is for those with more than 3 years of age. Percentages are in terms of the total population with more than three years of age or more than 12 years old. Mean income not adjusted by selection. Married category also includes those in common law. Separated category includes those divorced and widowed.

Source: own estimations.

The present work focuses on the study of wage distribution, although, as a secondary objective, there is an analysis of the wage gap between indigenous and non-indigenous workers. This study covers two points in time, twelve years apart, to allow comparison during times of similar macroeconomic conditions (DGP growth), therefore using the ENIGH surveys of 2010 and 2022, which also coincides with two different presidential midterms where the country was governed by right and leftist governments. This is important because two different governments may have different approaches to transfers and social welfare. The only exception is the Oaxaca-Blinder decomposition which uses the data from the ENIGH 2020, trying to figure out the main reasons for the wage gap for a year where the real GDP was decreasing, therefore revealing a possible larger and detectable labor discrimination.

In Figure 1 shows the Lorenz curve for Indigenous workers (dashed line) which is completely below the curve for non-Indigenous workers and therefore has a larger Gine index. But Figure 2 shows the Lorenz curves for the year 2022, with the curves for Indigenous and non-Indigenous crossing each other, therefore it is difficult to assess which distribution is more unequal. Our scientific objective will be to use income inequality measures to assess indigenous workers' wage distributions over time compared with non-indigenous workers.

Figure 1
Lorenz curves for Indigenous vs Non-Indigenous, 2010

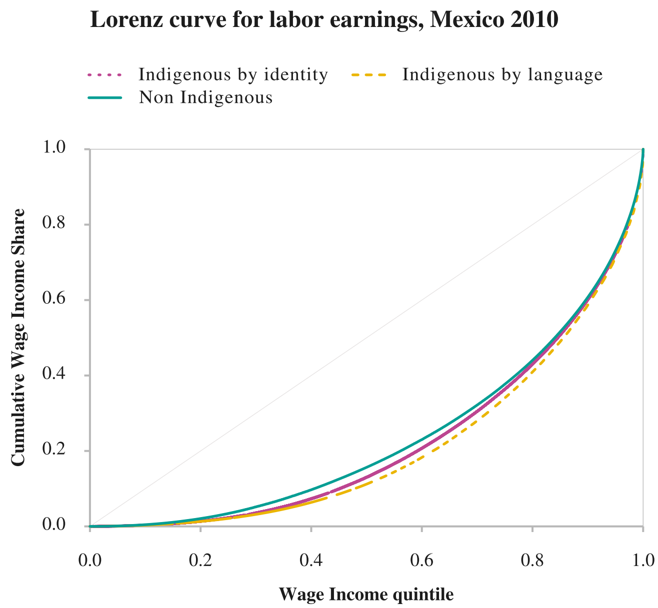
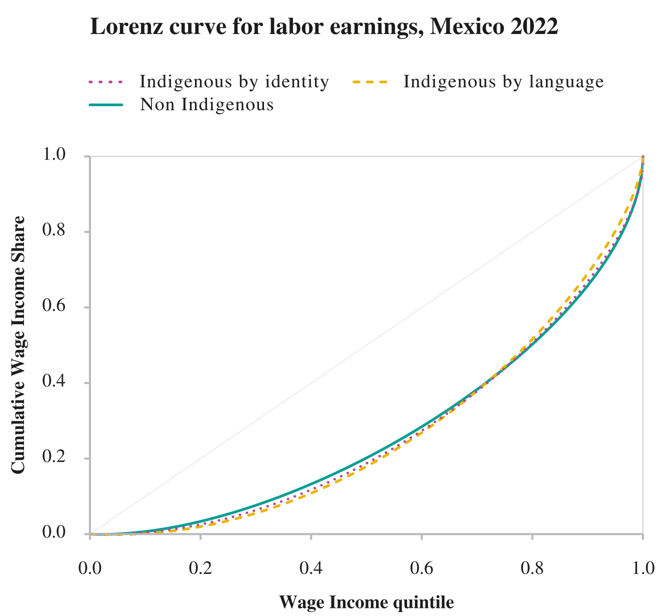


Figure 2
Lorenz curves for Indigenous vs Non-Indigenous, 2022



Source: own estimation using the data from Mexican National Survey of Household Income and Expenditure (ENIGH) 2010 and 2022.

Another objective in this paper is to analyze the wage distributions by occupations to observe for labor market segregation. Unfortunately, indigenous workers are associated with informal and elementary occupations, mainly in the agriculture sector of the economy, and also participate less in management and professional jobs. The difference in occupations also reflect the fact that there are relatively more indigenous workers who are informal, which deprives them of social security and benefits derived from meaningful and formal jobs. The scientific objective in this study is the wage distributions of Mexican workers based on ethnicity and by type of occupation.

Table 2 shows the percentage of workers by occupation type classified by the two available concepts of indigenous person. In each concept, indigenous workers are clearly underrepresented in most occupations, including managers, professionals, technicians and others, but better represented in jobs in agriculture and as manual laborers (general labor). This is more evident when using the concept of individuals that speak an indigenous dialect, as it highlights larger contrasts in occupations characterized by low pay and productivity.

This study focuses mainly on the concept indigenous person based on culture and language, although the self-identification as indigenous will also be used. The concept based on culture and language offers a clearer picture of people that have being traditionally discriminated against and left behind with fewer opportunities. That, when using the self-identification information, the labor income gap between indigenous and non-indigenous is smaller, as many workers that consider themselves part of the original people, which may be correct in terms of genetics, but difficult to assess as a more homogeneous category.

Table 2
Total workforce by occupations, 2022 (SINCO)

Occupation type	*% Indigenous (by dialect) vs Non-indigenous*		*% Indigenous (by Identity) vs Non-indigenous*	
	Indigenous	*Non-indigenous*	*Indigenous*	*Non-indigenous*
0. Migrant worker	0.002%	0.24%	0.02%	0.22%
1. Managers	0.07%	3.74%	0.57%	3.25%
2. Professionals	0.33%	15.09%	2.75%	12.67%
3. Technicians	0.10%	5.67%	1.12%	4.65%
4. Clerical	0.72%	13.16%	3.68%	10.20%
5. Services and sales	0.34%	7.75%	2.11%	5.99%
6. Agriculture	2.02%	5.68%	4.00%	3.71%
7. Trades	0.86%	8.52%	3.16%	6.22%
8. Operators	0.26%	9.39%	2.06%	7.59%
9. Manual laborers	2.58%	23.46%	9.28%	16.76%
Total	100%		100%	

SINCO = National System for Classification of Occupations.

Source: own estimations with ENIGH 2022 data.

This work begins with an introduction that includes information on the current situation of the indigenous population in México, and the scientific objectives to pursue. The next section contains a review of the literature about discrimination against Indigenous people in México. In the second section there is a brief explanation of the methodologies used in this analysis. The third contains the main findings and the last section contains the main conclusions.

I. LITERATURE REVIEW

The existing literature on the economic conditions of indigenous people in México is perhaps limited to development studies and focus primarily on poverty and social and economic disadvantages. Psacharopoulos & Patrinos (1994), in the survey on poverty in Latin America published by the World Bank, found that extreme poverty for the indigenous population is four times larger than that for those considered non-indigenous in México, a pattern that repeats itself in other Latin American countries. They conclude that differences in levels of poverty between indigenous and non-indigenous people came from differences in formal education and human capital.

More recently, Navarrete (2013) highlights the lower returns on education and fewer opportunities for indigenous people. This work applies an Oaxaca-Blinder decomposition to study poverty gap for indigenous people using the Mexican Family Life Survey (MxFLS) 2002 and 2006 and found that poverty is chronic in México and affecting more Indigenous people than non-Indigenous. People in Indigenous communities have less returns to education, lack access to health services and formal employment, among other problems that enhance poverty. A similar analysis by Servan *et al*., (2014) also found that the indigenous population has limited access to health and education as well as fewer employment opportunities. Cano & Mason (2016) mention that in México, although the Penal Code prohibits labor discrimination and the Federal Labor Law emphasizes that there should be no discrimination based on ethnic origin, discrimination against indigenous people persists nowadays, mainly when people self-identify as Indigenous, which has been a reason for discrimination at work and in schools. To find out the relationship between Indigenous self-identification and the salary received in the labor market, the authors take a sample of the 2010 census data considering the variables related to salary, some of which are age, years of education, self-employment, work hours per week, monthly income, and different types of jobs. They carried out a regression analysis, in which they discovered that people who declare themselves Indigenous receive a 9% discount on their salary, and if they mention that they speak an Indigenous language, their salary has an additional discount. Speaking Spanish at home is reflected in their income. They consider their way of acting and thinking, their cognitive ability, their employment status, and their community of origin as the factors that influence the wage gap.

Villarreal (2010) used the 2006 MIT México Panel Study, which includes information on the skin color of respondents as well as individual characteristics of occupation, poverty, and education attainment, and with a logistic regression explains how racial features, such as skin color, determine occupational opportunities, education attainment, and economic affluence. Comparatively, individuals with darker skin usually have less favorable economic conditions, fewer occupation opportunities, and less education. Following Villarreal's (2010) findings, Flores and Telles (2012) use the 2010 America's Barometer by the Latin American Public Opinion Project (LAPOP 2010) and use similar methodology to include parents' economic status. This study found that class (socio-economic status) determines mainly occupation, while race (skin color) determines education. Individuals with dark skin, with low-level occupations and low economic status, will have more chances to pass on these conditions to their children. The study found a money-whitening effect, where high-income individuals are classified to have white skin. This study is relevant because it introduces the concept of class or socio-economic status into the analysis of the indigenous population and shows how social mobility is also a problem, especially for marginalized indigenous households.

Caicedo (2009) analyzes the wage difference between the native population and Latin American immigrants in the United States, establishing how much of this gap is due to differences in human capital and how much to the different treatment that certain workers receive in the labor market. The analysis includes Mexicans, Cubans, and Central Americans. The results highlight large differences in human capital among workers but also the unequal treatment of the market exercised mainly towards some groups of Latin American workers. Luz-Tovar & Samario-Zarate (2023) is an Oaxaca-Blinder decomposition analysis for Mexican Indigenous workers and found that the wage gap was about 36% using data from the ENIGH 2018.

Ahmad (2020) performed a study on labor market discrimination in Finland using fictitious job applications with Finnish, English, Russian, Iraqi, and Somali names with similar human capital levels and credentials that were sent to different Finnish businesses. The probability of a callback was estimated, and the results showed that discrimination was based on the ethnic background of these fictitious workers, all with similar credentials. A similar analysis is Di Stasio & Larsen (2020), which also implemented 19 thousand job applications in five countries: Britain, Germany, the Netherlands, Spain, and Norway. The job posting included positions such as cooks, hairdressers, payroll checkers, sales representatives, receptionists, software developers, store assistants, and other skilled trades. The applications contained fictitious backgrounds for white minorities, Black people, Asians, and Middle Eastern countries against the white majority of these five countries. These studies show that employers have some biases against foreign minorities.

Solis *et al*. (2019) researched social factors as well as discriminatory practices that influence social and ethnic/racial inequality. For this, they carry out three focus groups in each of three localities in the state of Yucatan in four cities of the country. Each group was made up of three men and three women, considering socioeconomic levels and people with parents who speak an Indigenous language. They conducted interviews to learn more about discrimination and social problems, registering 565 pieces of evidence on discrimination, and they classified it into 1) not allowing entry to certain public and private areas, that is, they are denied entry to the school, to the place where they could work, to shops, parks, and gardens, etc.; 2) Preferences and limitations: They receive services with restrictions; for example, they have a job but are not entitled to benefits, they receive limited health services, and the transport service is marginalized, and 3) psychological abuse such as insults and humiliation by bosses, employers, teachers, etc. They conclude that what triggers racial discrimination are characteristics such as skin color, speaking an Indigenous language, and the type of clothing they wear.

Jarvis *et al*. (2018) is a work that deals with the indigenous population and sustainability. They explain that in México, as in other countries, the indigenous population works protecting the flora and fauna, supported by some federal programs that aim to protect natural resources and, at the same time, offer social benefits and economic incentives to Indigenous people. They argue that the salaries of Indigenous and non-Indigenous communities differ due to the structure of their economies. The labor support that companies receive depends on the skill, knowledge, and experience of the employees. Therefore, the salary of Indigenous people is lower. Studying the Indigenous people's income increases and the wage gap changes, the authors concluded that these federal programs play a significant role in the country's economy, closing the income gap between the Indigenous and non-Indigenous populations.

Canedo (2019), is a decomposition analysis of wage differentials. This study highlights the factors related to the wage gap between the Indigenous and non-Indigenous population of México, as well as the policies that can influence this difference. In the reviewed bibliography, they mention in a relevant way that the Indigenous working population receives low wages and lives in poverty. She used the data of the economically active population from the National Household Income and Expenditure Survey (ENIGH) of 2016. Based on the descriptive analysis, the Indigenous population has on average 6.9 years of study, their workday is 6 hours a day, and they report more years of experience; however, their salary is much lower than that of the non-Indigenous population. She applies the decomposition of a regression model (Oaxaca-Blinder) to evaluate the difference in wages between both populations through the observable and unobservable variables, finding a wage gap of 77.9 %, of which 37.2 % refers to variables such as place of residence, occupation, type of spoken language, among others. The analysis shows that the gap would be smaller if Indigenous people lived in larger towns and if they also had a similar educational level to non-Indigenous people, and that the wage gap of the female population is wider due to the presence of discrimination reflected by penalties in their salary.

Another interesting work is Pérez & Martínez (2022), which is a decomposition of wage differentials for Mexican migrants and native workers in the US. Using the Ñopo (2008) and Ñopo *et al*. (2011) matching method, they found there is inequality between these types of workers, and the wage gap is mainly due to the unexplained conditions related to taste and statistical discrimination. They argue that, despite the accumulation of human capital, Mexican migrant workers are more likely to work in jobs that the natives no longer want. Stabridis & Salgado-Viveros (2023) also uses a matching method to analyses labor discrimination against agricultural workers (jornaleros) in the Northwest of México.

II. DATA AND METHODOLOGY

The statistical on wage distribution we used the data from the Mexican National Household Income and Expenditure Survey (ENIGH) of 2010 and 2022, which is periodically released by the Mexican National Institute for Statistics, Geography and Informatics (INEGI). This survey contains detailed information about income sources as well as labor and individual characteristics, including ethnicity. This survey is representative nationwide, with a total sample of 108 thousand individuals for the 2010 survey and 310 thousand for the 2022 survey. Income from work is aggregated and merged by each member of every household, and the pertinent explanatory variables are selected. For the analysis of Oaxaca-Blinder decomposition the ENIGH 2020 was used.

Before proceeding to estimate the wage distribution of indigenous and non-indigenous workers, the labor earning data was corrected for selection and used for imputation on the missing information on wages and other labor earnings. For this, a Heckman correction model was performed as depicted in Heckman (1979), in order to find the real distribution of log wages. A two-step regression was estimated, first a Probit regression on labor market participation and later a Heckman selection model. It is usual to assume that the error terms in both equations are distributed normal; then the expected value of the dependent variable in the main equation is:

$$E(w_i \mid x_i, l = 1) = x_i \beta + \rho\sigma\lambda(v_i \gamma) \quad (1)$$

Where λ is the Mills' ratio and ρ is a parameter that represents the correlation between the first term $x\beta$ and the Mill's ratio. If $\rho = 0$ then this equation can be estimated with OLS, but if $\rho \neq 0$ it must be adjusted because, not doing so, we may obtain a biased estimate β. After doing the correction, the predicted values were estimated, and a new and corrected dataset was obtained. With this procedure, about 2/3 of the missing information was predicted for the labor income variable.

Gini index and Wasserstein distance

For an analysis of wage distribution, the Gini index is used, which is an income distribution measure that can be defined as the area between the Lorenz curve and the equality line. The Gini is a very well-known and widely applied income distribution measure. It is used to make comparisons among populations and observe changes over time. But one of the main critiques is that the Gini index is not quite decomposable, which means that it performs poorly when used to compare different subgroups. The principle of transfers is not so strong, which means that the effect of a transfer in the distribution will depend on the rank of the ordering. In a few words, there may be the case that the Lorenz curves cross each other. In such a situation, an additional analysis is needed to check whether the change in the income distribution of one group is superior to the other. For this reason, the Wasserstein distance was used as a measure of total change (transport) of one distribution to another, which may give us a magnitude of the change in a distribution.

The Wasserstein distance is used to solve the optimal transport problem and measures the dissimilarities between two distributions. It is defined as the path root of the total cost function of moving a mass from one location to another, where the cost is conceived as adding the Euclidean distance to move every point. If the wage distribution was to be moved from one point in time to another, the Wasserstein distance is the minimum cost of moving the whole distribution. The idea is to apply the Gini coefficient to some subgroups to observe changes in income inequality within each subgroup and then use the Wasserstein distance for validation of such changes.

To complement the analysis of the distribution of indigenous wages, a graphical approach was tested with the use of Kernel distributions, which allow to visualize the real distributions of wages.

Quantile regressions and Oaxaca-Blinder decomposition

A simple Mincerian wage regression may be useful to understand the size of the wage gap between indigenous and non-indigenous workers, but a lineal regression may not represent properly the differences among different income groups. For example, the effect of the indigenous category on wages in México may be found using the simplest wage equation in the form:

$$logwages_i = \beta_0 + \beta_1 S_i + \beta_2 E_i + \beta_3 E_i^2 + \beta_4 I_i + \beta_5 O + \mu_i \tag{2}$$

Where S_i represents education level, E_i represents the age of each worker, I_i represents the category of indigenous person, and O_i are other individual and labor market variables. The use of a simple linear regression may not account for the different categories of workers, occupations and other features that account for much of the heteroskedasticity problems in the data. For this reason, a quantile regression was used in order to observe changes in wages due to the category of indigenous in different labor income quantiles. This is to say, the regression analysis of wages does not use the mean as the reference, but the median and other quantiles.

To complement the analysis of the wage gap between indigenous and non-indigenous workers, the Oaxaca-Blinder decomposition was estimated, first proposed by Oaxaca (1973), to explain the differences in mean income between the two groups. The differences in labor earnings between these two groups of workers, let us say indigenous and non-indigenous can begin separating by group $Y_i^N = \beta_{ik}^N X_{ik}^N + \epsilon_i$ for non-indigenous and $Y_j^I = \beta_{ik}^I X_{ik}^I + \epsilon_j$ for indigenous.

Both are explained by a vector of K covariates, respectively. Each β is the vector of intercepts with $K-1$ slope coefficients of the regression for each group N, I. The error term ϵ is distributed normally with mean zero and variance σ^2. Then the difference in means between these two groups is just:

$$\Delta\bar{Y}=\bar{Y}^N-\bar{Y}^I=(\bar{X}^N-\bar{X}^I)'\hat{\beta}^I+\bar{X}^{I'}(\hat{\beta}^N-\hat{\beta}^I)+(\bar{X}^N-\bar{X}^I)'(\hat{\beta}^N-\hat{\beta}^I) \tag{3}$$

The first term on the right is group difference in covariates, the so-called endowment effect, which considers that both groups are different due to their characteristics. The second part explains the differences in coefficients, including the intercept, which, in part, can be attributed to discrimination. And the last part is the interaction term that explains that both differences, in covariates and coefficients, are present at the same time (interaction term). The above equation is known as the threefold decomposition, and it uses the Indigenous group as the reference for this decomposition.

II. ESTIMATIONS AND RESULTS

Figures 3 and 4 show the distribution of wages in México using kernel densities separated by quantiles for the years 2010 and 2022 and by occupations, as defined by the North American System of Occupations (SINCO), with numeration as shown in Table 2. From this graphical analysis it can be observed that, in the year 2010, the income distributions for agricultural workers (number 6 in the ordinate axis) are right skewed and their lowest quantile is small compared with other occupations. Other distributions are also skewed to the right and evidently truncated as well, like those for manual laborers and clerical workers, but relatively larger for the lowest quantiles.

Looking at the year 2022, most of the distributions of wages are moving to the left and becoming more normal. Comparing the kernel densities, the distribution for agricultural workers in 2022 becomes less skewed with a larger lowest 20% quantile and becomes bimodal, perhaps due to specialization and the effects of higher levels of education. The lowest quantiles in most distributions in the year 2022 are larger and much less skewed, so the mean and the median wages are much closer in 2022 than in 2010. These graphs show the distributions by occupation of the entire labor force, but the main interest is in the distributions of wages by ethnic background.

Figures 5 and 6 show the kernel densities for indigenous and non-indigenous workers for the years 2010 and 2022. Using these two categories, it can be observed that the kernel densities are skewed to the right, but this is because of the use of logarithms in the wage variable instead. The most important information is that the distribution for indigenous workers is always to the right of non-indigenous, which shows that the median and mean wages of non-indigenous workers are larger than those speaking an Indigenous dialect. Non-indigenous wage distribution dominates that of indigenous workers, but it is difficult to assess which one is more unequal just using a graphical analysis, especially for the year 2022.

Figure 3
Kernel distributions of Wages by Quantiles, 2010

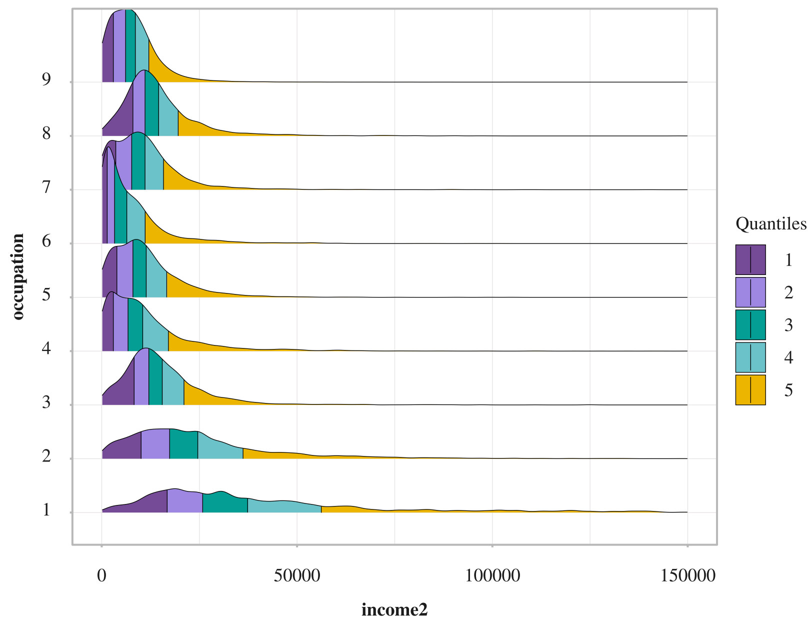
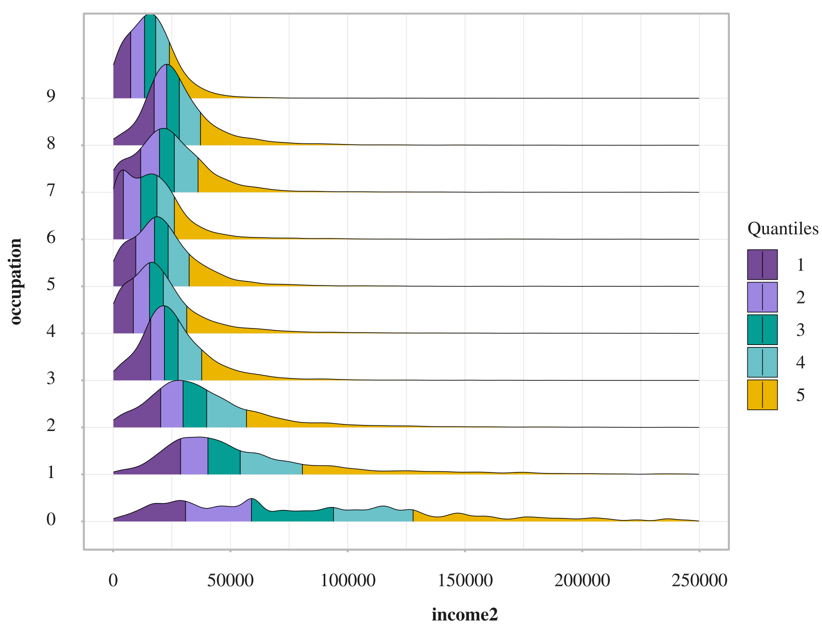


Figure 4
Kernel distributions of Wages by Quantiles, 2022



Note: the category ”0” in 2022 are those working in a foreign country (e.g. migrant workers). The R code used to estimate the quantile regression graph was from yuzaR Data Science (2022, Dec 22).

Source: own estimates using the data from Mexican National Survey of Household Income and Expenditure (ENIGH) 2010 and 2022.

Figure 5
Kernel densities Log Wages for Indigenous (Language) vs Non-Indigenous, 2010

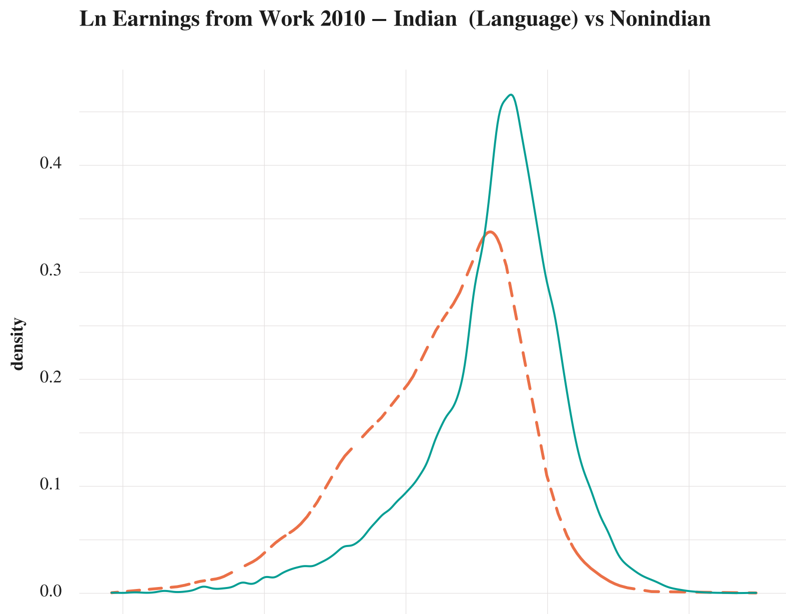
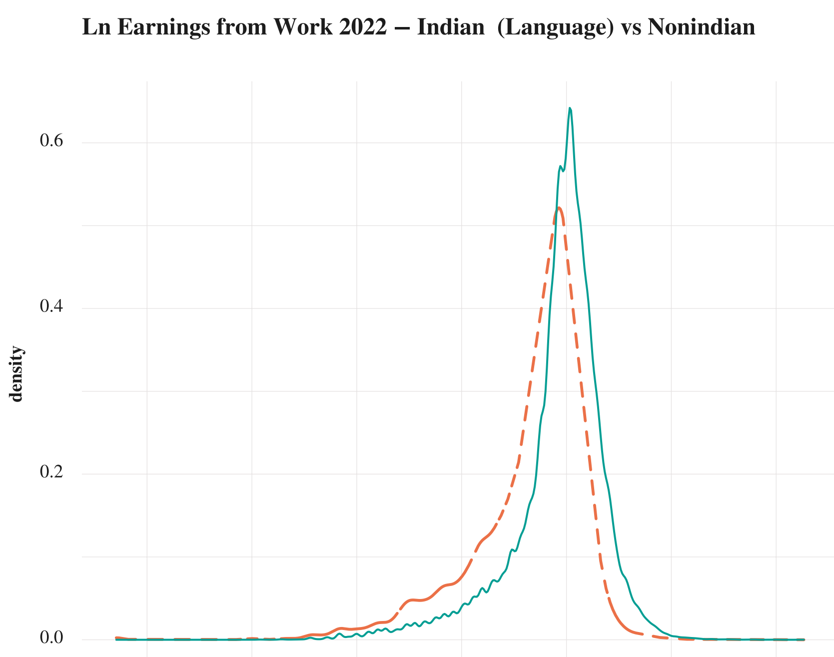


Figure 6
Kernel densities Log Wages for Indigenous (Language) vs Non-Indigenous, 2022



Note: the red dashed line is the distribution of indigenous workers' wages. Heckman regression was used to correct for selection in the sample.

Source: own estimates using the data from the Mexican National Survey of Household Income and Expenditure (ENIGH) 2010 and 2022.

As explained before, our objective is to find out whether the distribution of wages for indigenous people has improved over recent time (becoming less unequal). One way to observe this is to estimate the Gini index at two points in time and pay attention to subgroups that are representative of the population under analysis. Usually, the Indigenous population is confined to rural communities, with many individuals working as manual laborers and in the agriculture sector. Table 3 shows the Gini index by occupations for both indigenous and non-indigenous workers. There is a decrease in wage inequality for both types of workers and most occupations, with the only exceptions being managers and indigenous technical workers. The Wasserstein distance was estimated for all occupations to confirm the magnitude of the change, with the indigenous workers experiencing the largest changes in labor income inequality compared to non-indigenous workers. For example, indigenous workers in agriculture jobs experienced the largest decrease in income inequality, followed by those in professional jobs. But the largest change (increase) was for indigenous workers in management positions, but only few indigenous workers hold positions of responsibility in México. Therefore, most of the improvement in the distribution of wages for indigenous people may come from those working in the agriculture sector.

But overall, all workers in México have seen a decrease in wage inequality from 2010 to 2022 and an increase in real wage income. For example, in the year 2010, the real mean wage income was $19,306, at 2018 prices, and in 2022, the mean wage was about $21,941 pesos. The real median income was $12,947 in 2010 and twelve years later became $16,741, which means that the gap is more than a thousand pesos less in 2022.

Table 3
Gini and Wasserstein distance for Indigenous and Non-indigenous workers, 2010 and 2022

	Indigenous			*Non-indigenous*		
	Gini 2010	*Gini 2022*	*W. distance*	*Gini 2010*	*Gini 2022*	*W. distance*
Managers	0.40	0.44	0.15	0.46	0.45	0.04
Professionals	0.42	0.32	0.10	0.42	0.37	0.02
Technician	0.28	0.33	0.06	0.36	0.32	0.06
Clerical	0.52	0.42	0.02	0.51	0.44	0.01
Sales	0.41	0.33	0.04	0.42	0.38	0.03
Agriculture	0.63	0.40	0.11	0.55	0.45	0.02
Trades	0.52	0.37	0.02	0.41	0.35	0.02
Operators	0.35	0.31	0.05	0.34	0.31	0.05
Labor work	0.43	0.39	0.01	0.40	0.36	0.03
Total	0.51	0.39	0.04	0.48	0.40	0.03

Source: own estimates using the data from Mexican National Survey of Household Income and Expenditure (ENIGH) 2010 and 2022.

Overall, wage inequality is decreasing from the year 2010 to 2022, with major changes for indigenous workers. The wage distributions of both types of workers are getting similar, and indigenous workers are increasing their participation in occupations that require higher education and training. Wage income is now better distributed, but the gap between non-Indigenous and Indigenous people is still important. For example, using the ENIGH of 2020, the mean wage was $19,038.05 Mexican pesos for the non-Indigenous workers and $12,879.42 for the Indigenous workers, and the wage gap was $6,158.63 pesos. Table 4 shows the three-fold Oaxaca-Blinder decomposition, showing that $1,559.12 can be explained by the different individual endowments and $3,772.44 can only be explained by other factors such as discrimination, among other things.

Table 4
Mean Wages and Wages gap using three-fold Oaxaca decomposition, 2020

Group	*Mean Income*	*3-Fold decomposition due to:*	*Mean Income*
Non-Indigenous	$19,038.05	Endowments	$1,559.12 (134.58)
Indigenous	$12,879.42	Coefficients	$3,772.44 (252.12)
Total wage Gap	$6,158.63	Interaction	$827.07 (188.28)
		Total Wage Gap	**$6,158.63**
The R package “Oaxaca” was used to estimate the decomposition by Hlavac (2014).			

Source: own estimates using the data from Mexican National Survey of Household Income and Expenditure (ENIGH) 2020.

Most of this unexplained part of the wage differential between indigenous and non-indigenous can be attributed to experience (age) and the condition of the employee, while the explained part can be attributed mainly to university education. Figures 7 and 8 show the main features used for estimating the wage gap between indigenous and non-indigenous workers in México using the Oaxaca-Binder decomposition. The covariate that better describes the unexplained part of the income gap is age as a proxy of job market experience and seniority.

Figure 7
Three-fold decomposition of labor earnings Indigenous vs Non-Indigenous, 2020

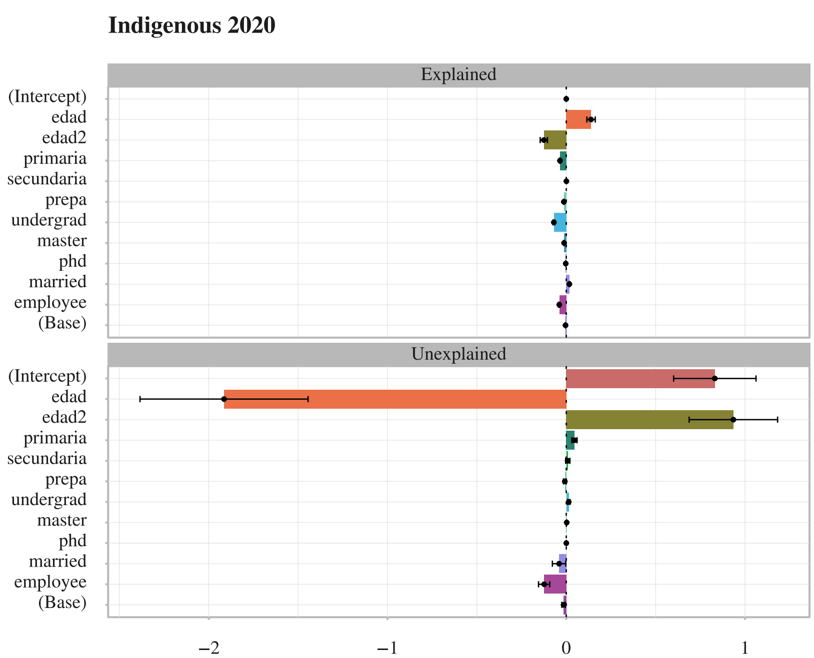
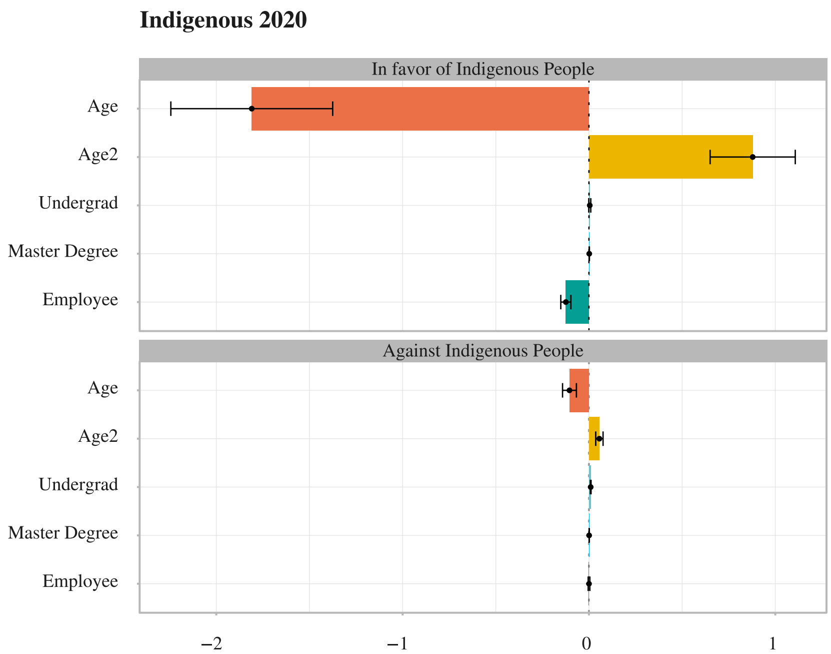


Figure 8
Factors against and in favor of Indigenous workers, 2020



Source: this methodology Reimers (1983) and packges R of Oaxaca Hlavac (2014).

The Oaxaca-Blinder decomposition uses the mean wage of both groups to calculate the wage gap, using linear regression for estimation. However, using the mean wage as a reference may not be the better choice, as wage data has many clusters and groups of workers with different characteristics. This data has serious problems of heteroscedasticity, and the wage gap may be overestimated for some workers and underestimated for others. Things are complicated because the income data is usually not normal and become more skewed when analyzing some subsets or groups in the data.

Following this argument, quantile regressions were performed to observe the effect of wages between indigenous and non-indigenous workers without considering the mean wage as the reference. For example, a median regression (quantile=0.5) on wages proves to be a better model than the OLS (lower AIC or less loss of information). Figures 9 and 10 show the median regression estimates for years 2010 and 2022, with changes in real wages at 2018 prices.

From Figures 9 and 10 it can be observed that differences in estimates are important, especially in terms of occupations. Using the median model, it can be observed that the drop in wages for being an indigenous worker was $2,300 real pesos in 2010 and $1,603 in 2022, compared to the non-Indigenous, everything else constant. If the linear model is used, it will appear that the indigenous wage gap is increasing instead of decreasing, which is not correct. Figures 11 and 12 in the appendix show the confidence intervals (CI) for the indigenous variable. The CI for the 2010 data shows that the linear model may explain in the middle and upper wages in the distribution, but the CI for the 2022 data shows that the quantile regression is a better fit, as wages cannot be explained by the linear model.

Figure 9
Linear vs Quantile Wage regression, 2010

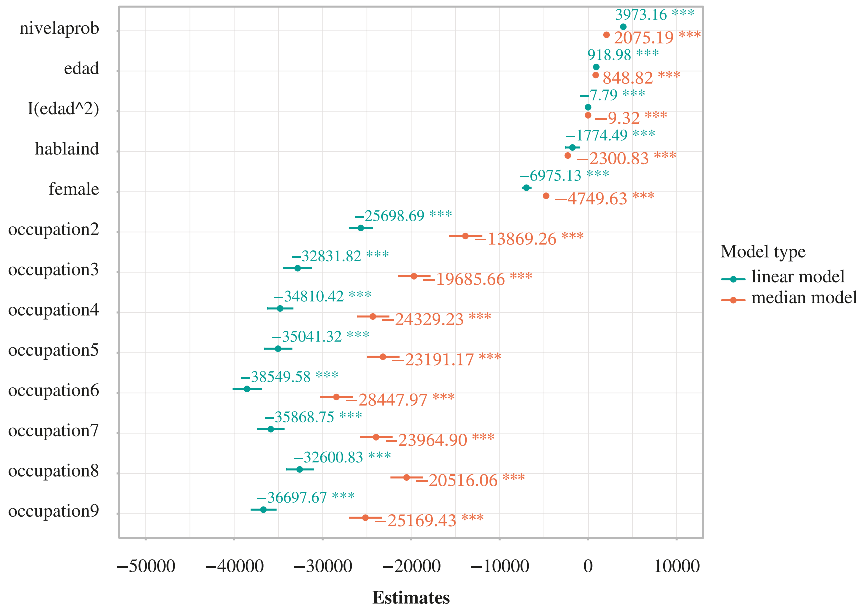
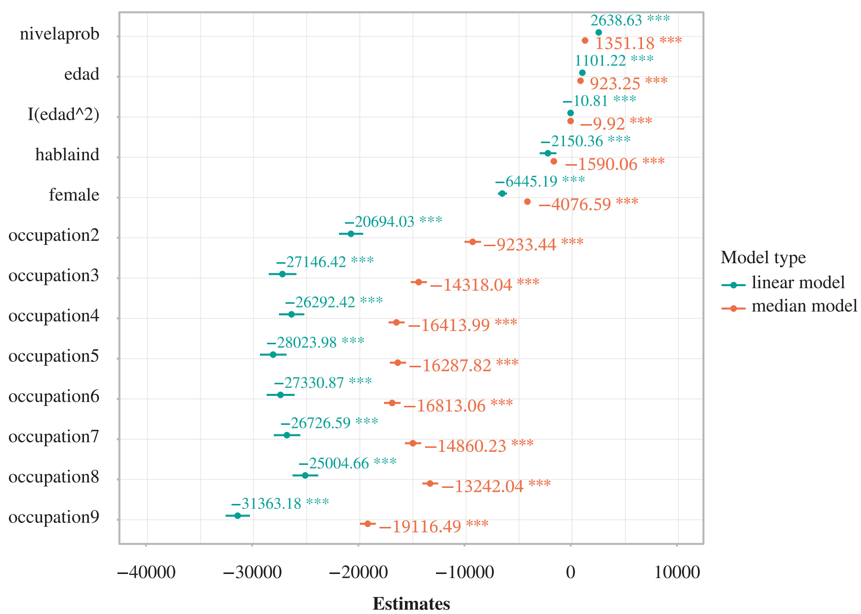


Figure 10
Linear vs Quantile Wage regression, 2022



Note: Estimates in blue are from the OLS model and in red are from a quantile regression. Estimates are changes in real wages of 2018. The estimates by occupations are negative because the benchmark is Managers = occupation1. The R code used to estimate the quantile regression was from yuzaR Data Science (2022, Dec 22).

Source: own estimates using the data from the Mexican National Household Income and Expenditure Survey (ENIGH) 2010 and 2022.

But the OLS and quantile regression estimates only show a partial view of reality, taking as reference some benchmark workers, because the effect on wages for being a native or indigenous is different along the distribution. For example, in previous Figure 6 it can be observed that the right tail in the kernel for indigenous workers is above that of non-indigenous workers and int the Lorenz curves in Figure 2 only indigenous workers on the highest quantiles cross the Lorenz curve of non-indigenous workers. Therefore, running a regression for the lower 20% quantile for the year 2022, which represents the indigenous workers with the lowest wages, the finding is a higher earnings loss due to their condition as indigenous people. The estimates for 50% and 80% quantiles in Table 5 and Table 6 in the appendix show that the effect on wages due to the indigenous variable is less in the year 2022, which implies that mostly indigenous workers in the upper part of the distributions improved their labor income.

CONCLUSIONS

Being an indigenous worker in México means that wages will generally be less than those of non-indigenous workers. The typical models of labor discrimination also confirm that ethnicity negatively affects wages in a similar way gender (female) does. The major difference is that, in recent years, there have been some policies, though inconsistent in their instrumentation, directed to improve labor conditions for female workers, such as demanding the same level of political representation, but nothing has been delivered to indigenous people in terms of better access to job market opportunities. One reason might be the practice of "blending," "admixture," or "assimilation," in which the indigenous population is slowly becoming less and less indigenous, adopting urban and mestizo style values and standards of living.

Non-indigenous workers' wage distribution dominates entirely and completely the indigenous one. However, the distribution of wages for both types of workers is similar in recent years. Still, the Oaxaca-Blinder decomposition shows that the wage differential is driven by the unexplained part and, although it cannot be blamed only on ethnic discrimination, although there is support in the literature that ethnic discrimination is one of the main forces behind lower wages and fewer opportunities for indigenous workers (e.g. Canedo, 2019).

Overall wage inequality is decreasing in México, and the largest decrease comes from indigenous workers compared to non-indigenous workers. But not all indigenous workers benefit from these improvements, as the major gains are for those in the middle and left tail of the distribution, which are the "better off". Perhaps a good public policy may be to increase transfers to low-income indigenous workers and households.

This analysis also found that the unexplained part of the wage gap between indigenous and non-indigenous workers is due to experience (age) and, to a lesser degree, undergraduate education. Perhaps a good public policy might be to offer job opportunities and access to certain public goods, such as education, based on minority clauses (being an indigenous person). Although it seems that the indigenous population may become assimilated with the rest of the population within the normal process of urban and economic development.

REFERENCES

Ahmad, A. (2020). When the name matters: An experimental investigation of ethnic discrimination in the Finnish labor market. *Sociological Inquiry*, *90*(3), 468–496. https://doi.org/10.1111/soin.12276

Caicedo, M. (2009). Desigualdad salarial en el mercado laboral estadounidense: La situación de los inmigrantes mexicanos, cubanos y centroamericanos. *Gaceta Laboral*, *15*(2), 5–31.

Canedo, A. (2019). Labor market discrimination against indigenous peoples in México: A decomposition analysis of wage differentials. *Iberoamericana-Nordic Journal of Latin American and Caribbean Studies*, *48*(1), 12–27. https://doi.org/10.16993/iberoamericana.433

Cano, J. & Mason, P. (2016). Acculturation and the labor market in México. *IZA Journal of Labor Policy*, *5*(1), 1–29. https://doi.org/10.1186/s40173-016-0077-6

Di Stasio, V. & Larsen, E. (2020). The racialized and gendered workplace: applying an intersectional lens to a field experiment on hiring discrimination in five European labor markets. *Social Psychology Quarterly*, *83*(3), 229–250. https://doi.org/10.1177/0190272520902994

Flores, R. & Telles, E. (2012). Social stratification in México: Disentangling color, ethnicity, and class. *American Sociological Review*, *77*(3), 486-494. https://doi.org/10.1177/0003122412444720

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*(1) 153-161. https://doi.org/10.2307/1912352

Hlavac, M. (2014). *Oaxaca: Blinder-Oaxaca decomposition in R*. http://dx.doi.org/10.2139/ssrn.2528391

INEGI. (2010). *Encuesta Nacional de Ingresos y Gastos de los Hogares 2020: ENIGH: serie anterior, nueva construcción*. México: INEGI. https://www.inegi.org.mx/programas/enigh/nc/2010/#microdatos

INEGI. (2020). *Encuesta Nacional de Ingresos y Gastos de los Hogares 2020: ENIGH: nueva serie*. México: INEGI. https://www.inegi.org.mx/programas/enigh/nc/2020/

INEGI. (2022). *Encuesta Nacional de Ingresos y Gastos de los Hogares 2020: ENIGH: nueva serie*. México: INEGI. https://www.inegi.org.mx/programas/enigh/nc/2022/

Jarvis, D., Stoeckl, N., Hill, R. & Pert, P. (2018). Indigenous land and sea management programs: Can they promote regional development and help "close the (income) gap"? *Australian Journal of Social Issues*, *53*(3), 283–303. https://doi.org/10.1002/ajs4.44

Luz-Tovar, C. D. L. & Samario-Zarate, J. (2023). The effect of ethnicity on the mexican labor market: an analysis of wage discrimination in the indigenous population. *Contaduría y administración*, *68*(2). 1-27. http://dx.doi.org/10.22201/fca.24488410e.2023.4649

Navarrete, J. C. (2013). Poverty Dynamics in México, 2002-2005. An Ethnicity Approach. *Brazilian Review of Econometrics*, *33*(1), 69–89. https://doi.org/10.12660/bre.v33n12013.8944

Ñopo, H. (2008). Matching as a tool to decompose wage gaps. *The review of economics and statistics*, *90*(2), 290-299. https://doi.org/10.1162/rest.90.2.290

Ñopo, H., Daza, N. & Ramos, J. (2011). Gender earnings gaps in the world (5736). *IZA Discussion Papers*.

Oaxaca, R. (1973). Male-female wage differentials in urban labor markets. *International Economic Review*, *14*(3), 693-709. https://doi.org/10.2307/2525981

Pérez, R. E. R. & Martínez, D. V. (2022). Desigualdad salarial por tipo de calificación laboral de los mexicanos inmigrantes en Estados Unidos. *Lecturas de Economía*, (97), 217-254. https://doi.org/10.17533/udea.le.n97a345715

Psacharopoulos, G. & Patrinos, H. A. (1994). Indigenous people and poverty in Latin America. *Finance & Development*, *31*(1). 41-43.

Reimers, C. W. (1983). Labor market discrimination against Hispanic and black men. *The review of economics and statistics*, *65*(4), 570–579. https://doi.org/10.2307/1935925

Servan, E., Torres, P., Orozco, E. & Sosa, S. (2014). An explanatory analysis of economic and health inequality changes among Mexican indigenous people, 2000-2010. *International Journal for Equity in Health*, *13*(1), 1–8. https://doi.org/10.1186/1475-9276-13-21

Solis, P., Krozer, A., Arroyo C. & Güémez B. (2019). *Ethnic/racial discrimination in México: A taxonomy of discrimination practices*. Centro de estudios Sociológicos. El Colegio de México.

Stabridis, O. & Salgado-Viveros, C. (2023). Effects of Gender and ethnicity on the Wage Gap Among Farmworkers in Northwestern México. *Frontera Norte*, 35, 1-28. https://doi.org/10.33679/rfn.v1i1.2339

Villarreal, A. (2010). Stratification by skin color in contemporary México. *American Sociological Review*, *75*(5), 652–678. http://www.jstor.org/stable/20799484

yuzaR Data Science (2022). *Quantile regression as the most useful alternative for ordinary linear regression* [Video] https://www.youtube.com/watch?v=Gtz8ca_4hVg&t=600s

APPENDIX

Figure 11
Confidence interval Linear vs Quantile Wage regression, 2010

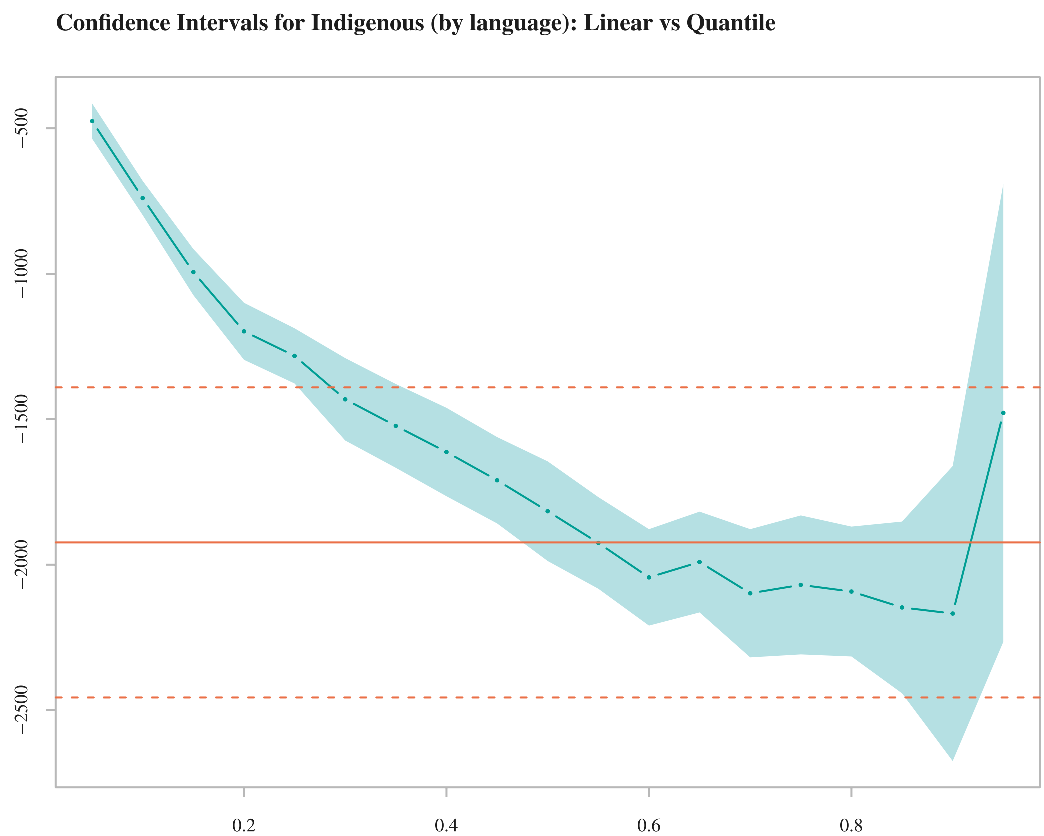
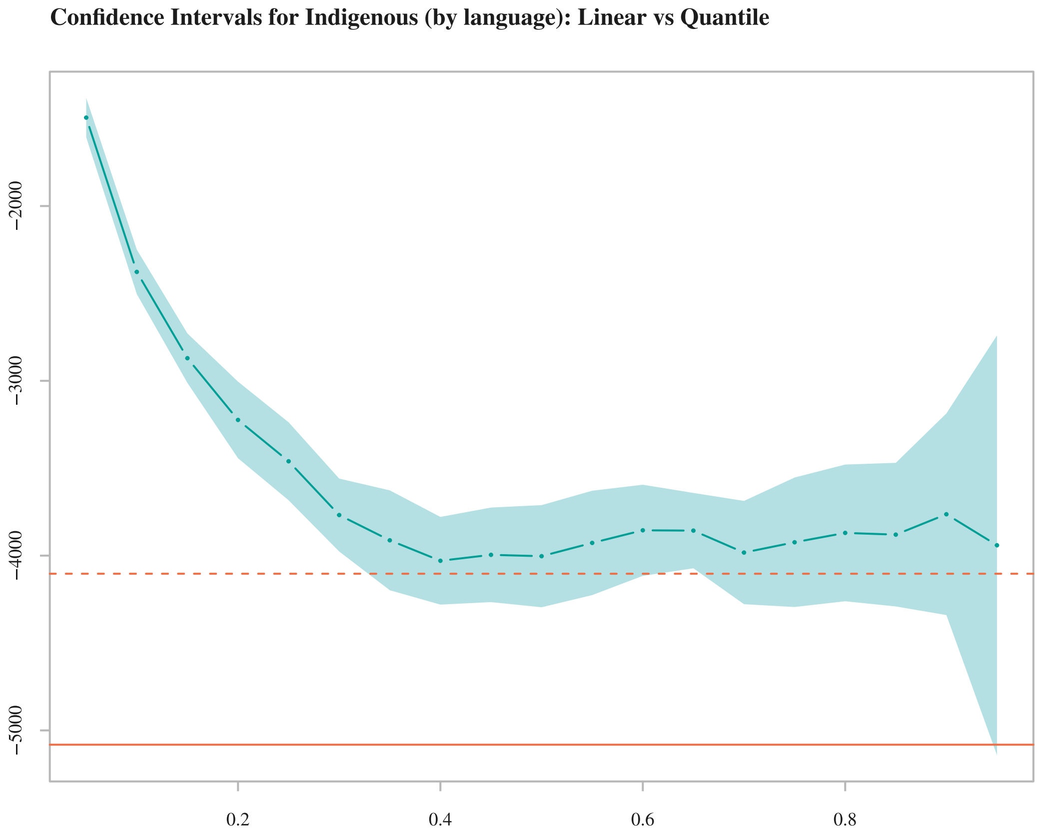


Figure 12
Confidence interval Linear vs Quantile Wage regression, 2022



Note: estimates in blue are from the OLS model and in red are from a Quantile regression. Estimates are changes in real wages of 2018.

Source: own estimates using the data from the Mexican National Household Income and Expenditure Survey (ENIGH) 2010 and 2022.

Table 5
Estimates for Linear vs 20%, 50% and 80% Quantile Wage regressions, 2010

Characteristic	*OLS*			*QR 20%*			*QR 50%*			*QR 80%*		
	Beta	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*
Education	3,973	3,809, 4,138	<0.001	1,110	1,044, 1,175	<0.001	2,075	1,986, 2,164	<0.001	3,849	3,719, 3,978	<0.001
Age	919	799, 1,039	<0.001	657	626, 688	<0.001	849	801, 896	<0.001	920	839, 1,001	<0.001
Age^2	-7.8	-9.3, -6.2	<0.001	-7.4	-7.8, -7.0	<0.001	-9.3	-10, -8.7	<0.001	-8.8	-9.9, -7.7	<0.001
Indigenous	-1,774	-2,601, -948	<0.001	-1,556	-1,731, -1,380	<0.001	-2,301	-2,570, -2,032	<0.001	-1,910	-2,316, -1,505	<0.001
Female	-6,975	-7,505, -6,445	<0.001	-3,491	-3,652, -3,331	<0.001	-4,750	-4,983, -4,517	<0.001	-6,744	-7,092, -6,395	<0.001
Occupation:												
1. Managers	—	—	—	—	—	—	—	—	—	—	—	—
2. Professionals	-25,699	-27,055, -24,343	<0.001	-8,350	-9,723, -6,977	<0.001	-13,869	-15,720, -12,018	<0.001	-32,622	-37,441, -27,804	<0.001
3. Technicians	-32,832	-34,427, -31,237	<0.001	-9,364	-10,704, -8,023	<0.001	19,686	-21,510, -17,861	<0.001	-44,984	-49,735, -40,233	<0.001
4. Clerical	-34,810	-36,254, -33,367	<0.001	-14,716	-16,022, -13,410	<0.001	-24,329	-26,140, -22,519	<0.001	-47,122	-51,853, -42,391	<0.001
5. Sales	-35,041	-36,062, -33,481	<0.001	-13,327	-14,676, -11,978	<0.001	-23,191	-25,016, -21,366	<0.001	-46,856	-51,600, -42,112	<0.001
6. Agriculture	-38,550	-40,171, -36,296	<0.001	-17,691	-19,005, -16,376	<0.001	-28,448	-30,275, -26,621	<0.001	-51,828	-56,579, -47,077	<0.001
7. Trades	-35,869	-37,377, -34,360	<0.001	-14,198	-15,523, -12,872	<0.001	-23,965	-25,779, -22,151	<0.001	-48,233	-52,958, -43,508	<0.001
8. Operators	-32,601	-34,142, -31,059	<0.001	-9,775	-11,099, -8,451	<0.001	-20,516	-22,327, -18,705	<0.001	-44,859	-49,596, -40,123	<0.001
9. General Labor	-36,698	-38,134, -35,262	<0.001	-14,597	-15,904, -13,291	<0.001	-25,169	-26,970, -23,369	<0.001	-49,737	-54,461, -45,013	<0.001

CI is Confidence Interval.

The negative coefficients in each occupation type represent how much less is paid in pesos of 2018 taking as benchmark the “Managers”.

Source: own estimations.

Table 6
Estimates for Linear vs 20%, 50% and 80% Quantile Wage regressions, 2022

Characteristic	*OLS*			*QR 20%*			*QR 50%*			*QR 80%*		
	Beta	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*	*Beta*	*95% CI*	*p-value*
Education	2,634	2,506, 2,762	<0.001	793	746, 840	<0.001	1,346	1,297, 1,395	<0.001	2,578	2,493, 2,662	<0.001
Age	1,107	1,015, 1,199	<0.001	875	847, 902	<0.001	916	887, 945	<0.001	961	913, 1,009	<0.001
Age^2	-11	-12, -9.7	<0.001	-9.7	-10, -9.3	<0.001	-9.9	-10, -9.5	<0.001	-9.4	-10, -8.8	<0.001
Indigenous	-2,068	-2,826, -1,310	<0.001	-2,023	-2,289, -1,756	<0.001	-1,603	-1,856,-1,351	<0.001	-902	-1,342, -463	<0.001
Female	-6,445	-6,845, -6,045	<0.001	-3,425	-3,558, -3,291	<0.001	-4,102	-4,232, -3,971	<0.001	-5,741	-5,963, -5,519	<0.001
Occupation:												
1. Managers	—	—	—	—	—	—	—	—	—	—	—	—
2. Professionals	-20,668	-21,771, -19,566	<0.001	-6,062	-6,710, -5,414	<0.001	-9,261	-10,018, -8,504	<0.001	-21,235	-23,454, -19,016	<0.001
3. Technicians	-27,139	-28,416, -25,862	<0.001	-7,792	-8,435, -7,148	<0.001	-14,361	-15,106, -13,616	<0.001	-30,898	-33,102, -28,694	<0.001
4. Clerical	-26,273	-27,434, -25,111	<0.001	-10,911	-11,552, -10,270	<0.001	-16,645	-17,379, -15,910	<0.001	-31,166	-33,364, -28,967	<0.001
5. Sales	-28,024	-29,247, -26,801	<0.001	-10,720	-11,376, -10,065	<0.001	-16,438	-17,182, -15,694	<0.001	-31,540	-33,744, -29,336	<0.001
6. Agriculture	-27,330	-28,617, -26,043	<0.001	-13,053	-13,739, -12,367	<0.001	-17,100	-17,860, -16,341	<0.001	-31,773	-33,990, -29,556	<0.001
7. Trades	-26,698	-27,920, -25,476	<0.001	-9,865	-10,519, -9,212	<0.001	-15,052	-15,804, -14,300	<0.001	-29,784	-31,992, -27,576	<0.001
8. Operators	-25,006	-26,188, -23,823	<0.001	-7,079	-7,717, -6,442	<0.001	-13,318	-14,053, -12,583	<0.001	-29,291	-31,483, -27,100	<0.001
9. General Labor	-31,359	-32,477, -30,242	<0.001	-12,977	-13,604, -12,351	<0.001	-19,204	-19,929, -18,479	<0.001	-35,965	-38,146, -33,784	<0.001

CI is Confidence Interval.

The negative coefficients in each occupation type represent how much less is paid in pesos of 2018 taking as benchmark the "Managers".

Source: own estimations.