Economía Teoría y Práctica [issn: 2448-7481] ▪ Nueva Época, año 31, número 58, enero-junio 2023, pp. 71-92, http://dx.doi.org/10.24275/etypuam/ne/582023/Gallardo

Wage Income Distribution in Mexico: A Nonparametric Approach*

Distribución del ingreso salarial en México: Un enfoque no paramétrico

*Roberto Gallardo Del Angel***

Abstract

This paper offers an analysis of wage income inequality for Mexico and offers some insights about welfare improvements for several categories of workers. We analyze real wage distributions at different points of time, using mainly nonparametric techniques. Kernel densities and smoothing techniques are used to analyze changes in the distribution of wages and labor supply for the first quarters of 2010 and 2020. We also use stochastic dominance analysis to observe welfare improvements for each category of workers and the Wasserstein distance to confirm changes in wage inequality. Our main results show that overall wage income inequality decreased, though the change is small and the categories that improved are those traditionally considered informal and low human capital workers, such as young people, workers with only elementary education and manufacturing or agricultural workers. The welfare of these groups also improved during the same period, yet welfare gains are negative for highly educated and experienced workers with a high level of human capital, including unionized and government or health sector workers. Intra-group wage distribution became more unequal for these workers. The results contradict the technological-bias change found during the initial years of free trade and market reforms in the 1980s and 1990s.

Keywords: Wage inequality, technological-bias change, welfare gains
jel classification: J31, J24 and O33

Resumen

Este documento contiene un análisis sobre la desigualdad de ingresos salariales en México y ofrece algunas ideas sobre las mejoras en el bienestar para algunas categorías de trabajadores. Analizamos las distribuciones de salarios reales en diferentes momentos utilizando principalmente técnicas no paramétricas. Se utilizaron densidades Kernel y técnicas de suavizamiento (Smoothing) para analizar cambios en la distribución de salarios y oferta laboral para los primeros trimestres de 2010 y 2020. También usamos análisis de dominancia estocástica para observar mejoras en el bienestar para cada categoría de trabajadores y la distancia de Wasserstein para confirmar cambios en desigualdad salarial. Nuestros principales resultados muestran que la desigualdad general de ingresos salariales disminuyó, aunque el cambio es pequeño y las categorías que mejoraron son aquellos tradicionalmente considerados trabajadores con bajo nivel de capital humano e informales, como podrían ser los trabajadores jóvenes, los que solo tienen educación primaria y los que trabajan en la industria manufacturera o la agricultura. Durante el mismo período, estos grupos también mejoraron en términos de bienestar. Por el contrario, las mejoras en el bienestar son negativas para los trabajadores altamente educados y experimentados con un alto nivel de capital humano, incluidos los sindicalizados y que trabajan en el sector público o los trabajadores de la salud. Para estos trabajadores, la distribución del salario intragrupo se hizo más desigual. Los resultados contradicen el cambio de sesgo tecnológico encontrado durante los años iniciales de las reformas de libre comercio y mercado de los años 80 y 90.

Palabras clave: Desigualdad salarial, cambio de sesgo tecnológico, ganancias de bienestar.
Clasificación jel: J31, J24 y O33.

* Fecha de recepción: 22/01/2021. Fecha de aceptación: 22/09/2022. Fecha de publicación: 30/01/2023.

** Facultad de Economía, Universidad Veracruzana. E-mail: rogallardo@uv.mx. orcid: 0000-0003-4661-0675.

INTRODUCTION

This work offers an alternative analysis of wage inequality, using nonparametric techniques with some insights on possible welfare changes during the ten-year period from 2010 to 2020. We compared changes in the distribution of real wages from the beginning of 2010 with 2020 and observed how real wages have changed over time in some economic sectors. We used stochastic dominance analysis to observe how real wages changed during both the end-of-year and ten-year periods, in order to detect possible welfare gains for certain categories of workers. The objective was to compare different groups of workers that may be affected by both trade liberalization and institutional changes (e.g. end-of-year *aguinaldo* bonus, minimum wage increase, etc.), and then compare the distribution of log wages.

A literature review on wage inequality in Mexico reveals general agreement that over the last three decades, wage inequality has increased and later decreased. Coincidentally, the period began with structural changes due to the implementation of major free-trade reforms. One accepted explanation for the initial increase in wage inequality is the technological-bias change that increased the demand for skilled workers at the expense of low-paid and low-skilled workers. Another important factor is the persistent loss in real value of wages due to post-1980s institutional arrangements. For example, a worker earning a minimum wage now can only obtain 40% of what (s)he could 30 years ago. Castro Lugo and Huesca Reynoso [12] offer a review and possible reasons behind the rise in wage inequality during the 1980s to mid 1990s. The same authors [12] mentioned three possible explanations for the increasing wage inequality during this period: 1. demand-side sources, 2. supply-side sources and 3. institutions. The first implies a possible technological-bias change: a separate equilibrium for skilled and unskilled workers, with higher wages for skilled and lower for unskilled. The second has to do with changes in demographics in the labor market, such as greater participation of young and female workers, and finally, institutional problems such as labor union bargaining power, minimum wage structure and public transfers, among others.

Wage inequality in Mexico can partially be explained by technological-bias change. Mexico began free-trade reforms in the mid 80s, first becoming a member of the General Agreement on Tariffs and Trade (GATT) in 1986 and culminating with the signing of the North American Free Trade Agreement (NAFTA) in 1994. A wave of privatizations was followed by an increase in foreign direct investment and new technology brought into production. This may explain the increase in income inequality during the 1980s and 1990s as shown by Castro Lugo and Huesca Reynoso [12]. Using firm level data from the Industrial Census, Hanson and Harrison [18] concluded that free trade policies affected firms hiring mainly low-skill workers. Similar conclusions can be found in Esquivel and Rodríguez-López [15], who found that recent wage inequality can be explained by the wage lag between skilled and unskilled workers caused by rapid technological changes and trade liberalization. Similarly, Airola and Juhn [3] explain this phenomenon on the side of the increasing demand for skilled labor. Acemoglu [1] provides a relevant theoretical work that explains the reasons behind the increasing wage inequality caused by technological-bias change. He builds a separate equilibrium model for skilled and unskilled workers produced by skill-biased technical change. His findings are that skilled workers will have their wages increased, while those of the unskilled will decrease and overall unemployment will increase. Such skill-biased technical change can be explained by higher returns to education, specialization and competition, although we may expect the skill premium to decrease over time and the wage spread to stop growing for those workers in the long run.

On the side of institutional variables, Fairris [16] and Cortez [13] analyze wage inequality induced by changes in union bargaining power. The first study analyzes data from the Mexican National Household Income-Expenditure Survey (ENIGH, Spanish initials) to capture the power of unions on wage spread. Fairris [16] concludes that unions have an effect of decreasing wage dispersion. Cortez [13] also uses ENIGH data from different years to observe the returns on both education and unionization. He concludes that changes in labor market institutions are responsible for higher wage inequality, increasing the return on unionization and minimum wages. Bell [6] found that minimum wages are not binding for most manufacturing workers due to their low level and lack of compliance in many cases. Fairris et al. [17] present evidence that changes in real and minimum wages are important for changes in overall wage inequality. Maloney and Méndez [21] and Bosch and Manacorda [7] focus on analyzing distribution shape and the effect of minimum wages on real wage determination. The former work compares densities by groups of formal and informal workers and uses kernel density estimation for some Latin American countries. Then they use lineal regression analysis to estimate the effect of minimum wages on the real hourly salary. The latter includes an analysis of workers earning minimum wages, using spikes. They use longitudinal micro data from the Mexican National Urban Employment Survey (ENEU, Spanish initials), which only represents urban workers.

The main objective of this study is to confirm or reject the previous trend of increasing income inequality in groups affected by technological-bias change and debate the possible effects of institutions such as unionization, transfers and minimum wages. We compare changes in wage inequality by worker category so as to observe welfare changes in the last decade and try to find evidence of technological-bias change in those worker categories supposedly more affected by this. We also want to observe changes in wage income for those workers with different amounts of human capital (e.g. formal education) that are also affected by transfers and globalization policies. For example, Campos-Vázquez [9] found that the lower wage inequality in recent years is due to labor market effects, where return to higher education is decreasing. Campos-Vázquez et al. [11] and Campos-Vázquez et al. [8] also support the idea that market forces are behind this lower wage inequality and other institutional factors may not be so relevant.

The first part of the article is an introduction and brief discussion on the sources of wage inequality that may be affecting the labor market in Mexico. The second part explains the data and the main techniques used to estimate wage inequality and welfare change. The third part contains the main results and economic analysis, and we end with a short conclusion and final comments.

Data and methodology

The Mexican National Occupation and Employment Survey (enoe, Spanish initials) is an improved labor survey that began collecting longitudinal data in 2005. The survey is quarterly, and respondents stay in the sample for five continuous quarters, with quarterly attrition loss of about 1/5. This survey is representative of the whole Mexican population and contains detailed information on job conditions, including wages, salaries and other labor income, as well as hours of work, individual and household characteristics. We were able to construct a corrected sample of 92,000 salaried workers, and we use monthly labor income, which includes wages, salaries and fringe benefits from employment from the last quarters of years 2009 and 2019, and the first quarters of years 2010 and 2020. We converted to real wages using the price index estimated by the Bank of Mexico, with 2018 as the base year. We used some relevant individual characteristics and labor market variables for all wage earners. Neither business and self-employment income nor income from capital are included in the sample.

Before proceeding to our analysis, we decided to use a traditional parametric approach on wages due to the missing data in the wage variable. In order to obtain a corrected sample and to overcome the problem of selection in this type of data, a two-step estimation was carried out. First, we estimated the probability of labor force participation using a tobit regression and then performed a Heckman correction to obtain estimates for the wage regression. The tobit regression on labor participation included total family income, number of children, education level and experience for each individual, as well as other explanatory variables. The Heckman regression was performed on a traditional wage equation, which includes education, experience and other labor market characteristics. After estimation, imputation was performed to produce a new and corrected sample of wages.

Kernel density estimation

Kernel Density Estimation is a nonparametric technique that estimates the real distribution of a data set. The meaning of *real* is in the context of a model-free distribution as opposed to the parametric family of distributions. The idea is to find a distribution that follows the observed data rather than assuming a specific parametric model that may fit the data properly. Using kernel densities allows us to observe some interesting behavior in the sample, such as clusters or groups around a mode. Assumptions on the data are minimal and less rigid than with parametric methods.

A density estimation problem is about reconstructing a probability density function $p(x)$ from a given set of data points $X_1, X_2, \ldots, X_n$. Instead of assuming a model from any traditional parametric family density functions, we want to find a smooth function that fits the data better: the real distribution. With this in mind, the best approximation to the real distribution is:

$$\hat{p}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right)$$

Where $\hat{p}(x)$ is a better fit of the real distribution that depends on the smooth kernel function K. Here the $(X_i - x)$ is the distance of every point from a designed test point x divided by a smoothing parameter h. The smoothing parameter is the key for the best fit of the distribution around the points $(X_i - x)$, which also interact with the sample size. A simple way to set up the bandwidth h is using a Gaussian kernel density estimator, commonly known as Silverman's rule of thumb:

$$h = 0.9\, min\left(\sigma, \frac{IQR}{1.34}\right) n^{-\frac{1}{5}}$$

Where IQR stands for interquartile range and σ is the standard deviation of the chosen points. Using kernel density estimations, we are able to get a glimpse of real data distribution, finding modes, the spread and localization of the distributions that may have economic significance.

Gini index and Wasserstein distance

A traditional approach for measuring income distribution is the Gini index, defined as the area between the Lorenz curve and the equality diagonal line. A general formula can be constructed defining the Lorenz curve as Y = L(X):

$$G = 1 - 2\int_0^1 L(X)dX$$

Although the Gini index is a very well known measure, it does not work well when comparing subgroups, as the Lorenz curves may cross. In order to complement the wage distribution analysis, we make use of the Wasserstein distance to find out how different two distributions are at two points in time. The Wasserstein distance compares two measures and is used to solve the *transport* problem. It is defined as the p_{th} root of the total cost of transporting a mass from one place to another where the cost is defined as the Euclidean distance to move every element (point) of that mass. Let X and Y be two random variables with marginal distributions u and v, respectively X~u and $Y \sim v$. We want to move every point x to each y using minimum effort (distance) until all the mass u is moved to the new v, assuming we are in a norm vector space χ where $x, y \in \chi$. The Wasserstein distance of order p is defined as:

$$W_p(u,v) = \left(\inf_{\delta \in \Delta(x,y)} \int_{\chi \times \chi} \|x - y\|^p \, d\delta(x,y) \right)^p$$

Where $\Delta(u, v)$ is the set of probability measures δ that intuitively constitutes a transport plan. Each $\delta(x, y)$ informs us of the proportion of mass at point x that must be transported to point y in order to move the total mass u to the new mass v. In our context, we want to transport the real wage income distribution from one year to another and estimate the Wasserstein distance, which is the minimum (infimum) cost to move the whole distribution to another one. Using this measure, we are validating the changes in the distribution already described by the Gini index.

Stochastic dominance

We use stochastic dominance to observe whether any income distribution is superior to another. We want to compare real wage distribution during a period with low inflation, which may be difficult to observe. Using stochastic dominance analysis, we may be able to observe if the most recent real wage distribution dominates the older one in order to validate possible welfare gains. Stochastic dominance can be explained using a random variable X_1 which may dominate another X_2 if only the cumulative distribution function $F_1(X)$ is above the other $F_2(X)$. Strictly speaking, $F_1(X) \leq F_2(X)$ for any outcome X on the support $[a, b]$. If we use the definition of an increasing utility function $U(X)$, the expected utility may be defined as:

$$\int_a^b U(X)f(X)dX = u(b) - \int_a^b U'(X)F(X)dX$$

Where $F(X)$ and $f(X)$ are the cumulative distribution function and density function, respectively. We can compare two expected utilities given two different income distributions X_1 and X_2 in the form:

$$\int_a^b U(X)f_2(X)dX - U(b) - \int_a^b U(X)f_1(X)dX = \int_a^b U'(X)\big(F_2(X) - F_1(X)\big)dX$$

So if $U_1(X) > U_2(X)$ then the part $(F_2(X) > F_1(X))$ in the right will be positive for any point X. This is the definition of first-degree stochastic dominance we intend to apply in our comparative analysis. For a better understanding of the direction and magnitude of this dominance, we constructed a piece-wise function of the form:

$$D_i \begin{cases} \big(F_2(X) - F_1(X)\big) > 0 = 1 \\ \big(F_2(X) - F_1(X)\big) < 0 = -1 \end{cases}$$

With this function we decided to construct a stochastic domination index that shows the direction of the dominance as well as the intensity:

$$SDI = \frac{\sum_{i=1}^{N} D_i}{N}$$

This index ranges $-1 < SDI < 1$ and counts the amount of times there are more positive values than negatives. The positive sign means that $U_1(X) > U_2(X)$, and the negative shows the opposite. The closer to the absolute one $|1|$, the stronger the stochastic dominance is between the two distributions. A value close to zero means that there is no way to know if one distribution dominates the other.

Lowess smoothing

We also observe changes in labor supply, using per-hour wages and compare the supply curves over time. Using this information, we estimated a pseudo-labor supply using nonparametric techniques. We used the locally weighted scatter plot smoothing (lowess) to estimate and approach an empirical labor supply curve. Lowess smoothing uses traditional linear and nonlinear regression for a localized data sample. These localized subsets of data are constructed using the nearest neighbor algorithm, and a weighted function is used to give more weight to the closest points, usually a tri-cubic weight function of the $w(x) = (1-|d|^3)^3$ type, where d is the Euclidean distance. Linear and nonlinear regressions are used on these localized samples to find a linear or non-linear fit that is smoothed across the entire data set. The advantage of this method is that it does not demand strict underlying conditions and allows the data to speak for itself but requires a data set that is large enough to be effective.

In our analysis, lowess smoothing is implemented by plotting working hours supplied against the log of individual real wages. Smoothing is performed by averaging the nearest observations in the distribution and then performing regression analysis on reduced subsamples. The result is a pseudo-labor supply curve, which is defined by the data (as shown in the appendix). For example, figure 9 in the appendix shows an example of pseudo-labor supply for manufacturing workers in 2020 (blue line) plotted along with the 2010 supply curve (red line). For both years, the supply was elastic and then became inelastic at high wages, even bending backwards for very high wages. This is a common result in economics, predicted by theory. We also confirm that the lowess curve for manufacturing workers in the year 2010 dominates that of 2020. The interpretation is that any improvement in working conditions shows that the lowess curve for 2010 dominates that of 2020, which implies that fewer hours of work are needed to get the same real wage. Then, stochastic dominance can be applied to the lowess-supply curves to observe possible improvements in wage distribution.

Analysis

Kernel density estimations were performed for some worker categories in order to observe the spread and shape of log wages. We are interested in those groups of workers that might be more affected by both free-trade reforms and those prone to changes in institutional conditions. One example may be workers in the manufacturing sector, which may be more affected by the inflow of foreign direct investment and changes in labor conditions from international trade. On the other hand, unionized workers are more affected by changes in public policy and legal reforms.

Furthermore, labor market composition has also changed substantially in the last 30 years. The inclusion of younger and female workers with higher formal education may also have an impact on wage dispersion. We compare the kernel density estimations over time for several categories of workers according to their labor market and individual characteristics.

The kernel distributions were constructed using information on the logarithm of monthly real wage income reported by each worker in the first quarters of 2010 and 2020. The red line shows the Kernel estimation for 2010 and the blue line for 2020. Three dotted lines show minimum wages in 2010, and the two separate dotted lines to the right show 2020 minimum wages. The minimum wage lines for 2020 (general minimum wage and the border zone minimum wage on the far right) are closer to the mean and median wage and binding (the minimum wage is in a mode) for all groups, as the most recent increases are relatively large (4% in 2010 compared with 15% in 2020).

Figures 1 to 8 in the appendix show the kernel densities for different categories of workers. In each graph we include different kernel estimations for two different points in time (2010 and 2020) and vertical dashed lines to show the real minimum wage in those years. We observe that the unionized distribution is positively skewed while it is negative for non-unionized workers. Furthermore, there are fewer modes for unionized than non-unionized, meaning that there are more clusters or subgroups for workers that do not belong to a union. We also observe that unionized workers are further from the minimum wage lines and the left part of the kernel has no modes, which means that minimum wage cannot be associated or is not binding to these kinds of workers.

In terms of stochastic dominance, Table 1 shows that in the short term (final quarters of 2009 and 2019), there is a welfare gain for unionized and non-unionized workers, but in the long term, the wage distribution of the first quarter of 2010 dominates the fourth quarter of 2019, which means that there is no long-term welfare gain. The minimum wage is binding for some non-unionized workers, as the vertical lines cut the kernel distributions in a mode. In terms of income distribution, inequality is larger for the non-unionized category, but intra-group inequality also increased in a ten-year period for unionized workers (see Table 2). Per-hour wages increased for non-unionized workers, while unionized workers saw their hourly wage decrease in a ten-year period, though unionized workers enjoy fairly higher wages (see Table 3). One possible reason is perhaps the reduction in wages and fringe benefits for unionized public workers, which has been a policy under the current federal administration, though a more detailed analysis is needed to support this hypothesis.

Young workers (29 years old and younger) and experienced workers (30 years old and older) also have multi-mode distributions, and the 2020 minimum wage seems to be binding for some subgroups. Young and non-unionized workers have seen their real mean wage increase, while experienced and unionized workers have seen their mean wages decrease. Also, young workers have a real welfare gain as their 2019Q4 distribution dominates that of 2010Q1. But for old workers there is not any clear gain at all. In terms of income inequality, both young and old categories have their intra-group inequality decreased by little. Young workers had their hourly wages increased (less labor supply per wage unit) in the ten-year period, while old workers have seen the opposite trend. This result contradicts the technological-bias change hypothesis. Perhaps institutional change is the source of these distribution changes (e.g. recent federal government-sponsored programs for unemployed young people).

We also observe that workers with elementary education have a negative skewed distribution with many modes in the left part, while those with tertiary education have a positive skewed distribution in the year 2020 and many clusters (modes) in the right part of their distribution. The new minimum salary seems to be binding for workers with elementary education, but not for workers with higher education. Looking at the stochastic dominance index, workers with tertiary education have a larger real wage than those with elementary education, but their long-term gain seems to be negative, while those with just elementary education made real improvements in welfare in the last decade. Intra-group income inequality has decreased for less educated workers, while it increased for highly educated workers. In terms of labor supply, Table 3 shows that younger workers with only elementary education provide less work for the same wage, while the opposite is true for highly-educated people. These findings support the idea of lower returns for higher education found by Campos-Vázquez [9].

Observing kernel estimations by economic sector, the distribution for agriculture and for manufacturing workers are located to the left of those workers in the government and health services in 2010 (lower mean wages). But in the year 2020, all four distributions are closer to each other. Through a careful examination of the stochastic dominance index in Table 1, we observe that from the first quarter of 2010 to the last quarter of 2019 both agriculture and manufacturing made important welfare improvements (2019-Q4 dominated the wage income distribution of 2010-Q1). The opposite results were found for those in the government sector and health services who experienced a welfare loss in terms of wage income, closing their wage gap with agriculture and manufacturing workers. Intra-group wage inequality has increased for health and government workers and decreased for workers in agriculture and manufacturing. In terms of labor supply, the hourly wage decreased for health and government workers (same wage for more work) in the ten-year period, while workers in manufacturing and agriculture had the opposite effect (same wage for less work) as shown in Table 3.

Table 1 reveals a positive value for the stochastic dominance index, which shows that the latter quarter dominates the previous one. A positive stochastic dominance index and close to one in the middle column shows that the kernel estimation of the last quarter of 2019 dominated the distribution of the first quarter of 2010. This long-term improvement in welfare was only possible for workers with supposedly low productivity, those in agriculture and manufacturing, and mainly young workers and those with little formal education.

The Gini index and Wasserstein distance in Table 2 shows that overall income inequality decreased from 2010 to 2020. But the groups that contributed to this decrease are those traditionally associated with low productivity, such as the young and those with only elementary education, as well as workers in the agriculture and manufacturing sectors. Those workers in sectors that require higher specialization, such as in the health and government sectors, unionized workers and those with tertiary education, have seen their wage distribution becoming more unequal.

Table 1: *Stochastic Dominance Index for Wage Income in Mexico*

Group	2009Q4→2010Q1	2010Q1→2019Q4	2019Q4→2020Q1
Union	0.999	−1	1
Non union	0.999	−1	1
Agriculture	−0.593	1	−0.587
Manufacture	−0.983	1	−0.985
Government	0.471	−0.854	−0.975
Health	0.130	−0.605	0.145
Elementary	−0.964	0.934	0.005
Tertiary	0.474	−1	−1
Young	−0.968	1	−0.995
Old	−0.993	0.273	0.889

Notes: This is the Stochastic Dominance Index with value from −1 to 1. The arrows show the dominance direction. If a value is positive, then later distribution dominates the previous and the direction of the arrow holds. If negative, then the direction of dominance reverses.

Table 2: *Gini Index and Wasserstein Distance for Wages in Mexico*

Group	Gini 2010	Gini 2020	Wasserstein
Union	0.411	0.463	0.031
Non union	0.451	0.483	0.022
Agriculture	0.652	0.591	0.016
Health	0.512	0.582	0.018
Manufacture	0.493	0.484	0.023
Government	0.459	0.512	0.031
Elementary ed	0.617	0.595	0.040
Tertiary ed	0.603	0.633	0.029
Young	0.623	0.606	0.002
Old	0.608	0.598	0.025
Total	0.622	0.606	0.018

Table 3: *Pseudo-Labor Supply for Mexico*

Group	Wage per hour of work		Dominance 2010→2020	
	2010	2020	Wasserstein	Stoch. Domin
Union	261.52	237.07	0.045	−0.476
Non union	137.02	141.31	0.071	1
Agriculture	94.22	99.24	0.157	0.947
Health	248.02	237.53	0.038	−0.876
Manufacture	129.78	142.20	0.158	1
Government	227.61	221.95	0.020	0.625
Elementary ed	109.33	114.32	0.063	1
Tertiary ed	269.58	234.11	0.119	−1
Young	125.11	130.50	0.057	1
Old	167.54	162.99	0.042	1
Total	149.68	150.71	0.059	1

Overall wage income per hour of work barely increased from 2010 to 2020, though the groups that improved their position (fewer hours of work for the same wage) are workers in agriculture and manufacturing, non-unionized workers, young workers and those with only elementary education. Unionized workers, workers in the health and government sectors and workers with tertiary education saw the same toil for less wage income in this ten-year period.

Stochastic dominance analysis on the lowess supply curve shows a negative index for workers whose 2020 labor curve dominated their 2010 curve, which implies that they are supplying more labor for the same real wage. Workers traditionally associated with low productivity are supplying less labor for the same real wage, such as those in the manufacturing and agricultural sectors, as well as those with only elementary education, young and non-unionized workers.

Conclusion and final comments

The objective of the present analysis is to open the debate on the possible sources of wage inequality in Mexico in recent years. We opted for nonparametric techniques to analyze short-and long-term changes in real wages for several categories of workers and also to observe important trends. One of our major research results shows that workers in groups with traditionally high levels of human capital are not experiencing improvements in their welfare in the long term, and their intra-group wage inequality is increasing. The stochastic dominance analysis also shows that short-term improvements are also becoming difficult to attain. These workers are receiving much lower wages for the same hour of work.

On the other hand, workers considered to have low human capital, such as young workers with only elementary education, as well as those in agriculture and manufacturing, are improving in intra-group income inequality as well as welfare in the ten-year period of analysis. Using stochastic dominance, we analyzed possible short-term changes in welfare during the end-of-year changes (bonuses and minimum wage increase) in 2009 and 2019, as well as the ten-year gap from the first quarter of 2010 to the fourth quarter of 2019, using real wage income. We observed that workers in traditionally low specialized sectors, such as young workers, workers in the agricultural and manufacturing sectors and those with only elementary education, are not getting short-term welfare improvement due to changes in their real wages at the end of the year. The end-of-year changes might be due to yearly bonuses (*aguinaldo*) and institutional changes such as the minimum wage. However, these categories are improving their welfare in the ten-year period from 2010Q1 to 2019Q4.

Workers traditionally associated with low specializations and low human capital improved in their labor supply, receiving relatively higher wages for the same labor, while the opposite was true for highly specialized workers and those with high human capital. Non-unionized, agricultural workers and workers in manufacturing, as well as those with only elementary education, increased their product per hour worked. The stochastic dominance and Wasserstein distances of lowess labor supply show possible improvements in productivity for these categories of low specialization and low human capital.

The above trends are difficult to explain using the framework of technological-bias change and separating equilibrium for low-skilled and high-skilled workers, as observed in the first decades of the 1980s and 1990s. As explained by Castro Lugo and Huesca Reynoso [12], technical-bias change was a possible reason for the increasing wage inequality during that period. But the current trend seems to be reversed, as many workers with high productivity and higher education have experienced increased intra-group inequality and long-term welfare loss.

References

[1] Acemoglu, Daron (1999), "Changes in Unemployment and Wage Inequality: An Alternative Theory and Some Evidence", *American Economic Review*, 89 (5), pp. 1259-1278.

[2] Aguilera Fernández, Albany and Castro Lugo, David (2018), "NAFTA and Wage Inequality in Mexico: An Analysis for Border Cities, 1992-2016", *Frontera Norte*, 30 (60), pp. 85-110.

[3] Airola, Jim and Juhn, Chinhui (2005), "Wage Inequality in Post-Reform Mexico", IZA Discussion Paper Series, No. 1525.

[4] Allegretto, Silvia A.; Dube, Arindrajit; Reich, Michael and Zipperer, Ben (2013), "Credible Research Designs for Minimum Wage Studies".

[5] Anderson, Gordon (1996), "Nonparametric Tests of Stochastic Dominance in Income Distributions", *Econometrica: Journal of the Econometric Society*, pp. 1183-1193.

[6] Bell, Linda A. (1997), "The Impact of Minimum Wages in Mexico and Colombia", *Journal of Labor Economics*, 15 (S3), pp. S102–S135.

[7] Bosch, Mariano and Manacorda, Marco (2010), "Minimum Wages and Earnings Inequality in Urban Mexico", *American Economic Journal: Applied Economics*, pp. 128-149.

[8] Campos-Vázquez, Raimundo; Esquivel, Gerardo and Lustig, Nora (2014), "The Rise and Fall of Income Inequality in Mexico, 1989-2010", *Falling Inequality in Latin America: Policy Changes and Lessons*, 140.

[9] Campos-Vázquez, Raimundo Miguel (2013), "Why Did Wage Inequality Decrease in Mexico after NAFTA?", *Economia Mexicana*, Nueva Época, 22 (2) pp. 245-278.

[10] Campos-Vázquez, Raimundo Miguel and Rodas Milán, James Alexis (2020), "El efecto faro del salario mínimo en la estructura salarial: evidencias para México", *El Trimestre Económico*, 87 (345), pp. 51-97.

[11] Campos-Vázquez, Raimundo Miguel; López-Calva, Luis F. and Lustig, Nora (2016), "Declining Wages for College-Educated Workers in Mexico: Are Younger or Older Cohorts Hurt the Most?", World Bank Policy Research Working Paper, (7546).

[12] Castro Lugo, David and Huesca Reynoso, Luis (2007), "Wage inequality in Mexico: a review", *Papeles de Población*, 13 (54), pp. 225-264.

[13] Cortez, Willy W. (2001), "What is Behind Increasing Wage Inequality in Mexico?", *World Development*, 29 (11), pp. 1905-1922.

[14] Esquivel, Gerardo and Cruces, Guillermo (2011), "The Dynamics of Income Inequality in Mexico since NAFTA" [with comment], *Economia*, 12 (1), pp. 155-188.

[15] Esquivel, Gerardo and Rodríguez-López, José Antonio (2003), "Technology, Trade, and Wage Inequality in Mexico before and after NAFTA", *Journal of Development Economics*, 72 (2), pp. 543-565.

[16] Fairris, David (2003), "Unions and Wage Inequality in Mexico", *Industrial & Labor Relations Review*, 56 (3), pp. 481-497.

[17] Fairris, David; Popli, Gurleen and Zepeda, Eduardo (2008), "Minimum Wages and the Wage Structure in Mexico", *Review of Social Economy*, 66 (2), pp. 181-208.

[18] Hanson, Gordon H. and Harrison, Ann (1999), "Trade Liberalization and Wage Inequality in Mexico", *Industrial & Labor Relations Review*, 52 (2), pp. 271-288.

[19] Huesca-Reynoso, Luis; Castro-Lugo, David and Rodriguez-Pérez, Reyna E. (2010), "Cambio tecnológico y sus efectos en el mercado de trabajo: una revisión analitica", *Economia, Sociedad y Territorio*, 10 (34), pp. 749-779.
[20] Li, Qi and Racine, Jeffrey S. (2007), *Nonparametric Econometrics: Theory and Practice*, Princeton University Press.
[21] Maloney, William and Mendez, Jairo (2004), "Measuring the Impact of Minimum Wages. Evidence from Latin America". in *Law and Employment: Lessons from Latin America and the Caribbean*, pp. 109-130, University of Chicago Press.
[22] Negrete Prieto, Rodrigo and Luna Ramirez, Lilia Guadalupe (2016), "¿Cuál es el monto de trabajadores en Mexico cuyas remuneraciones son un reflejo del salario minimo? Realidad, datos y espacio", *Revista Internacional de Estadística y Geografía*, 7 (1), pp. 76-111.
[23] Newmark, David; Schweitzer, Mark and Wascher, William (2000), "The Effects of Minimum Wages Throughout the Wage Distribution", NBER Working Paper, (7519).
[24] Robertson, Raymond (2004), "Relative Prices and Wage Inequality: Evidence from Mexico", *Journal of International Economics*, 64 (2), pp. 387-409.
[25] Rodríguez Pérez, Reyna Elizabeth; Huesca Reynoso, Luis and Camberos Castro, Mario (2011), "Mercado laboral, inequidad salarial y cambio tecnológico regional", *Frontera Norte*, 23 (45), pp. 7-33.
[26] Wolfstetter, Elmar; Dulleck, Uwe; Inderst, Roman; Kuhbier, Peter and M. Lands-Berger, Michael (1993), *Stochastic Dominance: Theory and Applications*. Citeseer.

APPENDIX

Figure 1: *Kernel Density for Unionized Workers 2010 and 2020*

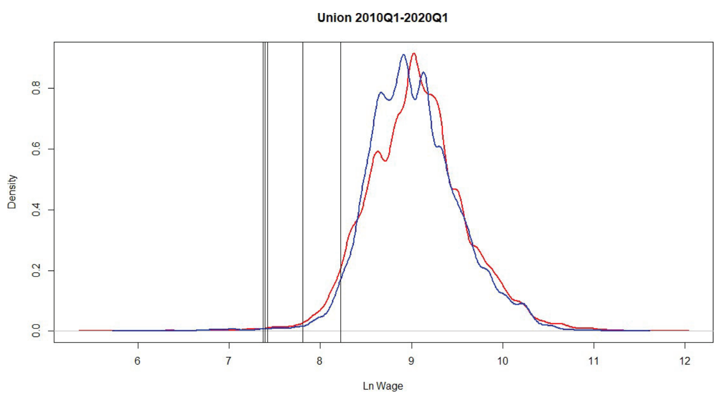

Figure 2: *Kernel Density for Non-Unionized Workers 2010 and 2020*

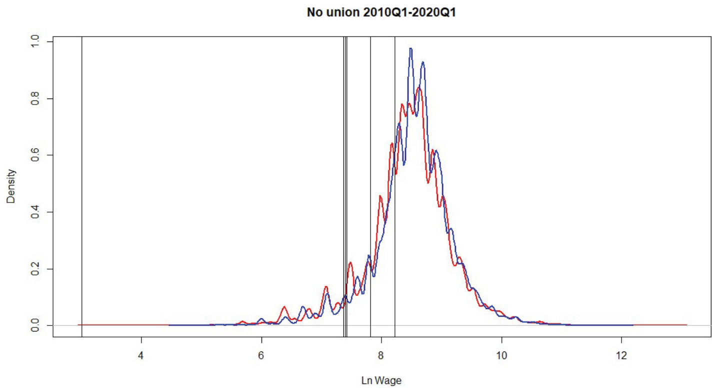

Figure 3: *Kernel Density for Young Workers 2010 and 2020*

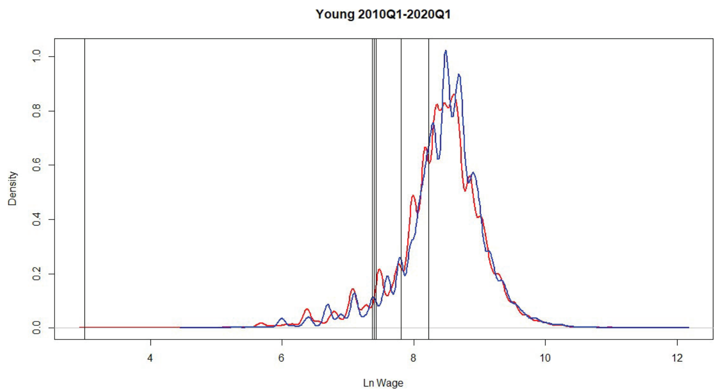

Figure 4: *Kernel Density for Experienced Workers 2010 and 2020*

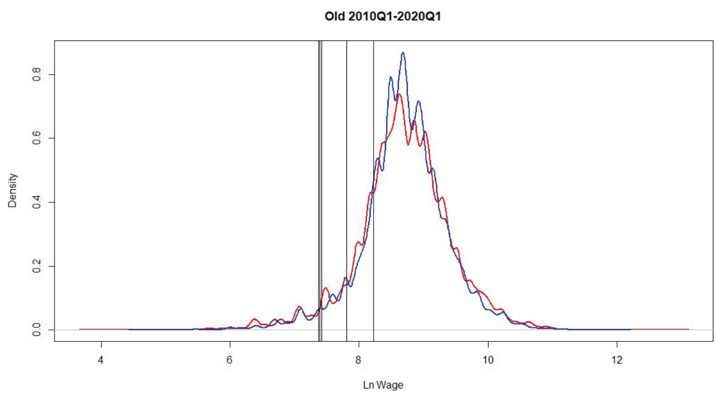

Figure 5: *Kernel Density for Workers with Elementary Education 2010 and 2020*

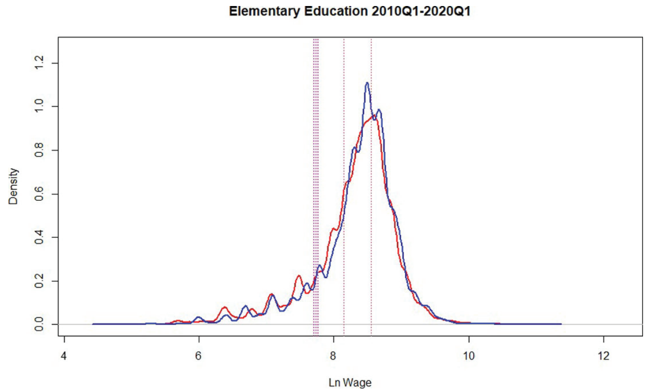

Figure 6: *Kernel Density for Workers with Tertiary Education 2010 and 2020*

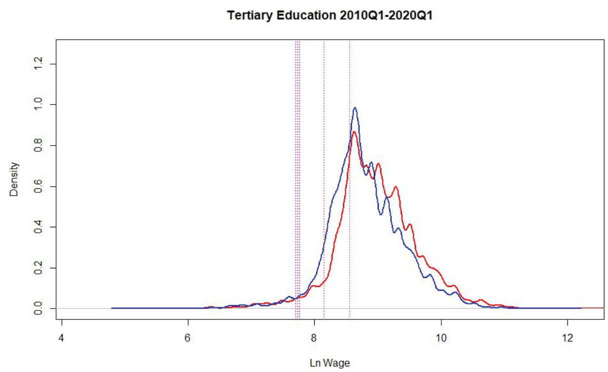

Figure 7: *Kernel Density by Economic Sector 2010*

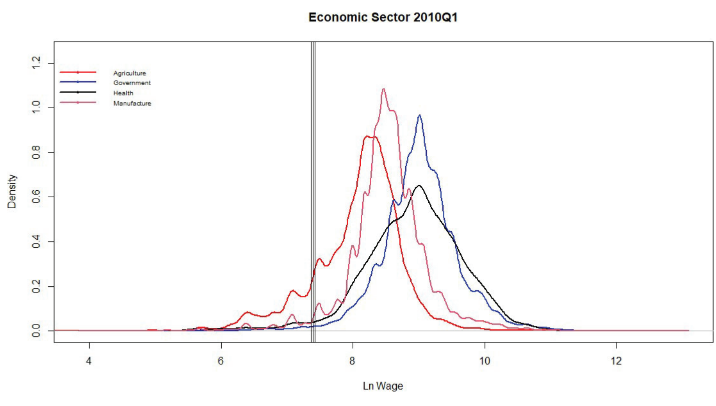

Figure 8: *Kernel Density by Economic Sector 2020*

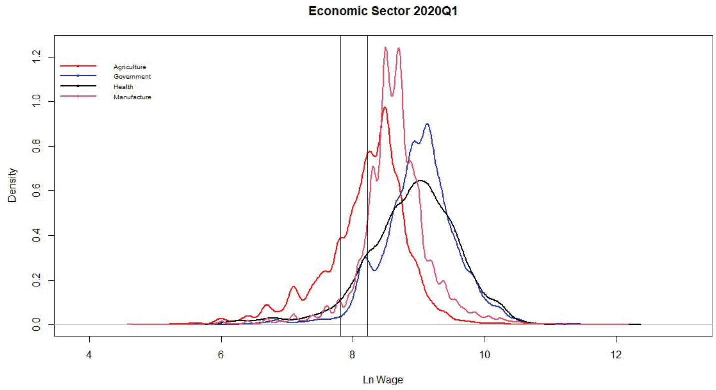

Figure 9: *Lowess Labor Supply for Manufacturing Workers 2020*

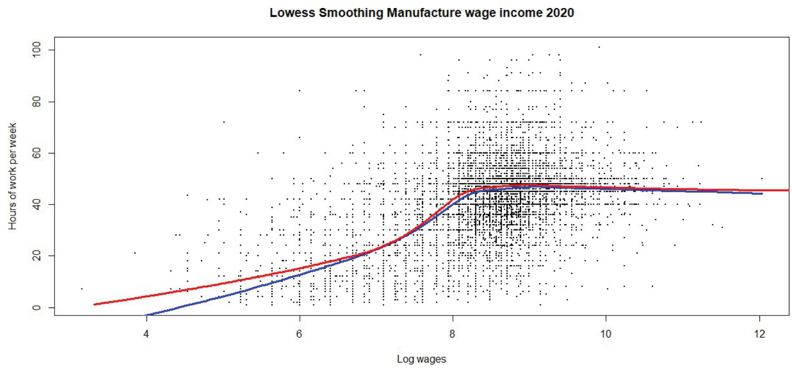

Figure 10: *Lowess Labor Supply for Agricultural Workers 2020*

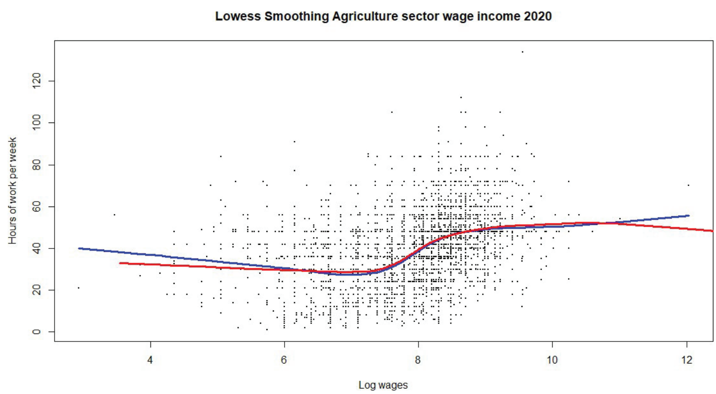

Figure 11: *Lowess Labor Supply for Health Workers 2020*

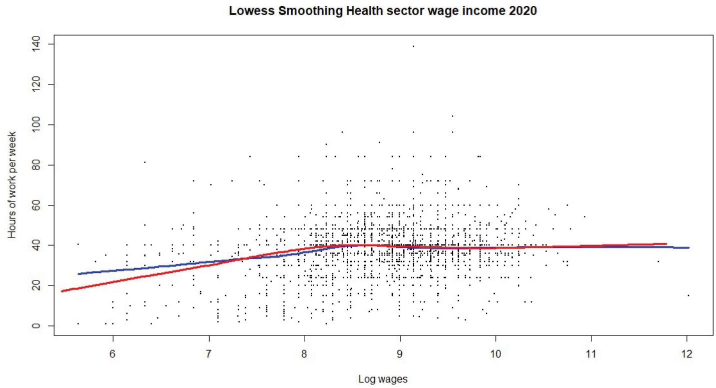

Figure 12: *Lowess Labor Supply for Government Workers 2020*

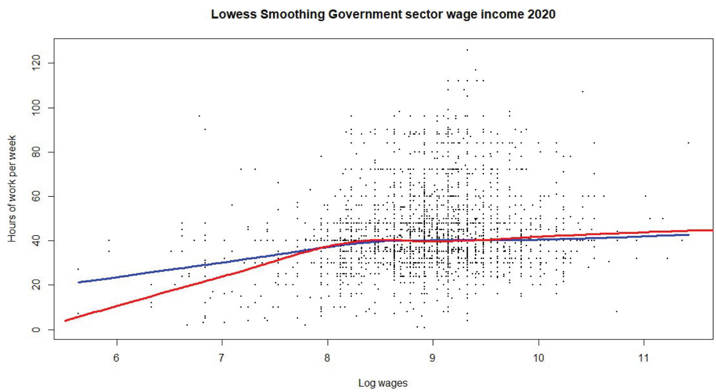

Figure 13: *Lowess Labor Supply for Workers with Only Elementary Education 2020*

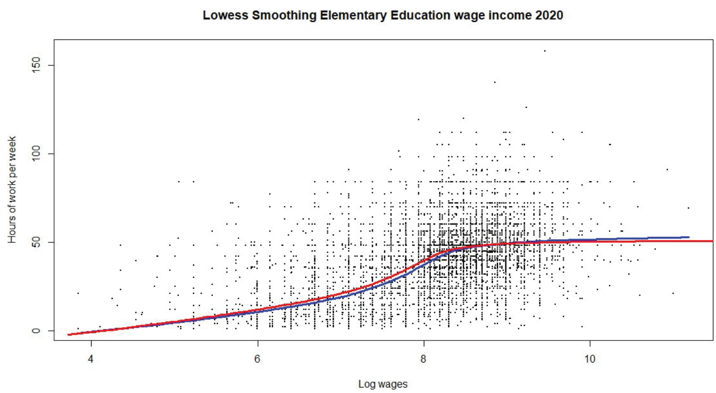

Figure 14: *Lowess Labor Supply for Workers with Tertiary Education 2020*

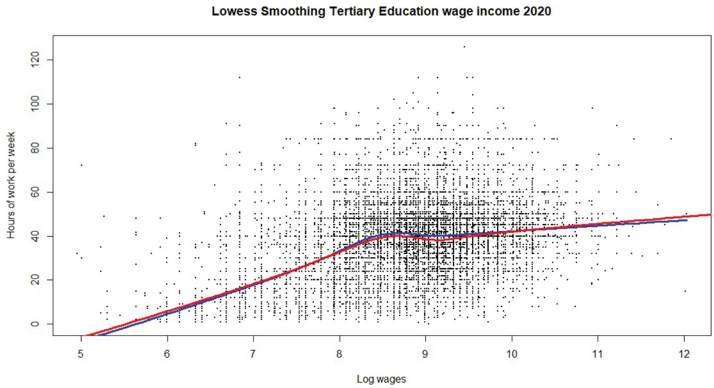

Figure 15: *Lowess Labor Supply for Unionized Workers 2020*

Figure 16: *Lowess Labor Supply for Non-Unionized Workers 2020*

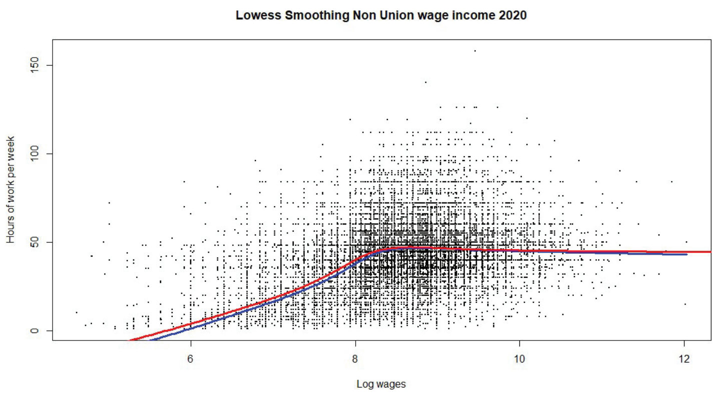

Table 4: *Standard Statistics for Wage Income 2010 and 2020 by Group*

Group	Obs	Mean	SD	Median	Min	Max	Range	Skew	Kurtosis
Union 2010Q1	11,964	9.001	0.524	8.981	5.562	11.828	6.266	0.386	1.300
Union 2020Q1	12,815	8.901	0.505	8.841	5.925	11.717	5.792	0.457	1.202
Non union 2010Q1	65,316	8.477	0.666	8.548	3.088	12.930	9.842	−0.616	2.847
Non union 2020Q1	79,426	8.517	0.598	8.559	4.606	12.034	7.428	−0.586	3.432
Agriculture 2010Q1	4,536	8.096	0.665	8.171	3.088	10.912	7.824	−0.827	2.266
Agriculture 2020Q1	4,405	8.218	0.628	8.300	4.801	11.008	6.207	−0.982	2.254
Manufacture 2010Q1	12,482	8.540	0.565	8.576	4.004	12.930	8.926	−0.494	5.371
Manufacture 2020Q1	17,526	8.609	0.502	8.612	4.793	12.034	7.241	−0.352	6.150
Government 2010Q1	6,662	8.946	0.570	8.908	5.103	11.828	6.725	−0.024	3.063
Government 2020Q1	6,652	8.858	0.560	8.815	5.486	11.635	6.149	0.003	2.008
Health 2010Q1	3,599	8.874	0.688	8.832	5.562	11.787	6.225	−0.493	2.290
Health 2020Q1	4,124	8.783	0.636	8.724	5.819	12.020	6.200	−0.320	2.025
Elementary 2010Q1	13,219	8.290	0.647	8.433	4.769	10.931	6.162	−1.185	2.680
Elementary 2020Q1	10,774	8.346	0.614	8.482	4.606	11.190	6.584	−1.285	3.281
Tertiary 2010Q1	19,765	8.974	0.641	8.944	5.562	12.930	7.368	−0.069	1.856
Tertiary 2020Q1	28,613	8.815	0.610	8.735	4.957	12.034	7.077	0.046	2.019
Young 2010Q1	36,418	8.407	0.642	8.471	3.088	11.657	8.569	−0.982	3.188
Young 2020Q1	38,381	8.449	0.590	8.482	4.606	12.020	7.413	−0.964	3.756
Old 2010Q1	41,678	8.692	0.668	8.681	3.841	12.930	9.089	−0.322	2.274
Old 2020Q1	55,338	8.656	0.590	8.633	4.606	12.034	7.428	−0.214	2.813