CEE Centro de Estudios Económicos
www.colmex.mx El Colegio de México, A.C.

Serie documentos de trabajo

Self-Selection Patterns among Return Migrants: Mexico 1990-2010

Raymundo M. Campos Vázquez y Jaime Lara Lara

El Colegio de México

DOCUMENTO DE TRABAJO
Núm. IX – 2011

Self-Selection Patterns among Return Migrants: Mexico 1990-2010

Raymundo M. Campos-Vazquez[♠]
El Colegio de México

Jaime Lara[♣]
El Colegio de México

Abstract

This paper analyzes the self-selection patterns among Mexican return migrants during the period from 1990 to 2010. Using census data, we can identify return migrants who have lived in the United States within the previous 5 years but who currently live in Mexico. To calculate the selection patterns, we nonparametrically estimate the counterfactual wages that the return migrants would have experienced had they never migrated by using the wage structure of nonmigrants. We find evidence that the selection patterns change over time toward negative selection. For example, in 1990, the wages that the male return migrants would have experienced had they not migrated was 6 percent larger than the wages of male nonmigrants. However, by 2010, the difference had declined to -14 percent. The increasing negativity of the degree of selection is robust to the analysis of specific subgroups: rural and urban, men and women, and states with high migration rates and low migration rates. Moreover, the negative selection results for the period from 2000 to 2010 are robust to the use of different surveys that define a return migrant by using distinct characteristics. Additionally, we observe that the wages of return migrants are larger than those that the migrants would have obtained had they not migrated. This finding shows that migration has a positive effect on the Mexican economy.

JEL Codes: F22; J61; O54.
Keywords:: Mexican Migration; Self-Selection; Return Migration; Wages.

This version: September 2011

[♠] Centro de Estudios Económicos, Camino al Ajusco 20, Col. Pedregal de Santa Teresa. 10740, México DF. Telephone: +52-55-54493000, 4153. http://raycampos.googlepages.com, Email: rmcampos@colmex.mx.
[♣] Centro de Estudios Económicos, Camino al Ajusco 20, Col. Pedregal de Santa Teresa. 10740, México DF. Telephone: +52-55-54493000, 3135. Email: jlara@colmex.mx.

1 Introduction

International migration is not always a permanent decision. Some migrants return to their countries of origin after staying for a period of time in the country of destination. Return migrants may bring skills or capital to the home economy and thereby contribute to the positive effects of migration in the source countries. Mexico has become the largest source of immigrants in the United States. Mexican immigrants accounted for 31.3 percent of the new arrivals in the 1990s (Chiquiar and Hanson, 2005). Their return migration rate is also high. The 2010 Mexican census shows that of the 994,869 individuals who left their country to live in the United States from 2005 to 2010, 307,783 returned to Mexico by 2010. In other words, 30.9 percent of the migrants returned home.[1] In this article, we investigate the self-selection patterns among the return migrants in Mexico.

Policymakers around the world are engaged in a broad debate on the implications of immigration and the optimal migration policy. One of the requirements for an informed discussion is accurately determining the skills of the migrant population. Unfortunately, no consensus exists regarding the self-selection patterns of Mexican immigrants. Chiquiar and Hanson (2005) conclude that Mexican immigrants are located in the middle of the Mexican wage distribution by using the Mexican and U.S. census data from 1990 and 2000. In contrast, Fernandez-Huertas (2011) uses the labor force survey for the period from 2000 to 2004. This dataset collects information on the migrants from Mexico before the individuals migrate. He finds strong evidence of negative selection in the fact that migrant wages are less than 20 percent of the earnings of the rest of the Mexican population.

Within the wide range of issues that have been examined by the recent literature on migration, we focus on two important questions regarding Mexican return migrants: 1) Are Mexican return migrants positively or negatively selected? 2) Do return migrants improve their productivity in the labor market of the source economy compared with their productivity had they not migrated? Given the scant empirical evidence regarding the return migration of Mexican migrants, our paper contributes to the understanding of the self-selection patterns among these migrants. The second question has implications for the

[1] Other data sources provide similar results. The National Survey of Demographic Dynamics 2006 (ENADID) presents a return migration rate of 33.72 percent in 2006 for those who left the country within the previous 5 years.

analysis of labor market effects on the economy of the source country. The high proportion of return migration among Mexican migrants in the U.S. suggests that the economic effects of return migration could be large. If the migrants acquire some skills as a result of their migration, then return migration has a positive effect on the source economy.

Only a few recent articles have used data on return migrants in the Mexican labor market. Lacuesta (2006) uses the wages of return migrants to calculate the wages that the migrants remaining in the United States would have obtained in Mexico. The information is from the 2000 census in Mexico and the United States. He shows that immigrants who stay in the United States come from the middle part of the distribution of human capital. However, his paper focuses on comparisons between return migrants and nonreturning migrants.

Gitter et al. (2008) analyze the effect of return migration on the probability of employment by using the Mexican Family Life Survey (MxFLS), which was conducted in 2002. They focus on return migrants who have been in the United States between 1997 and 2002. The researchers' results indicate that migration does not affect the probability of employment. Unfortunately, their paper does not address counterfactual wage distributions.

Ambrosini and Peri (2011) use the 2002 and 2005 rounds of MxFLS and define return migrants as workers who lived in Mexico in 2005 and have spent more than one year in the US between 2002 and 2005. Because of the sample size and the short time period, the number of return migrants identified is small (i.e., only 56). Thus, making strong inferences is difficult. However, the scholars find evidence that spending some time in the US enhances one's earning abilities and accounts for some mildly positive selection among the return migrants. The data from the Mexican census used in our estimates present an adequate sample of return migrants and allow us to observe changes over longer periods of time. However, the census does not include all of the variables that can be found in the MxFLS or the panel structure.

To determine the self-selection patterns, we calculate the counterfactual wage distributions of the return migrants had they stayed in Mexico. To calculate this counterfactual wage, we follow DiNardo et al.'s (1996) method and reweight the wage distributions of the nonmigrants such that the distribution of observable characteristics between the return migrants and the nonmigrants is as similar as possible. We use the 1990,

2000 and 2010 population censuses and focus on males and females separately. Using the census, we can identify individuals who have been in the United States within the previous 5 years. However, restricting the time frame to a short migration period may bias the results. Hence, we use different data sources that also include migration information to verify the sensitivity of our results.

The self-selection patterns among return migrants have changed over time. In 1990, the selection was slightly positive for both men and women. In 2000, the selection among the women stayed positive, but the men were drawn more from the middle of the wage distribution. The result changed in 2010. The selection among the men became negative, and the women were drawn more from the middle of the distribution, with negative selection on average. For example, in 1990, the wages that male return migrants would have experienced had they not migrated were 6 percent larger than the wages of nonmigrant males, but by 2010, this difference had declined to -14 percent. Moreover, migration allows those who return to obtain higher wages because the increase in human capital or savings can be applied to productive activities. In other words, there is a wage premium associated with migration and return. In 1990, the wage premiums for the men and women were 36 and 38 percent, respectively, whereas in 2010, the wage premiums were 5 and 7 percent, respectively. As in the case of selection, the wage premiums to migrate and return have worsened over time for both men and women.

When we analyze the geographical subgroups, we find that the self-selection patterns differ, but we find a tendency toward negative selection among every subgroup as time moves forward. Among the rural population, we observe positive selection patterns. However, among the urban population, the selection becomes less positive or even negative. If we divide the Mexican states between those that have historically shown a high migration rate and the rest of the country, we find that the degree of negative selection is higher in states with high migration rates but that the tendency toward negative selection is present in both groups.

To confront any concerns about our estimates, we perform different robustness tests. The results vary little when we restrict our sample to only the working population or young individuals. By using additional datasets that define return migrant differently from the census, we find results consistent with those of our basic specifications, except when we

concentrate on the short-term flow of return migrants. Here, the degree of selection is more negative, possibly because low-skilled individuals tend to do make more than one trip. In sum, we find robust evidence that the self-selection patterns among the current return migrants are negative.

According to Bratsberg and Borjas's (1996) model, in countries such as Mexico, where payments to human capital are more unequal than those in the United States, return migrants should be selected negatively with respect to the nonmigrant population. This hypothesis only holds clearly for the men in 2010 and is clearly rejected for the 1990 census. Our results are more consistent with the observation that low-skilled individuals may face costs that prevent them from migrating in the first place. The fact that the selection tends to become negative over time supports the hypothesis that migration networks may relax the costs faced by low-skilled individuals, as suggested by McKenzie and Rapoport (2010). However, further research is necessary to establish whether other factors, such as the enforcement of immigration laws or changes in the demand for immigrants, explain the change in the degree of selection among return migrants.

In the next section, we review the literature on selection and examine how return migration is related to selection and productivity improvements. Section 3 explains the identification strategy. Section 4 provides more details about Mexican return migration and describes the datasets employed in this study. In section 5, we discuss the results. Finally, in section 6, we conclude this paper.

2 Selection and Return Migration

Immigrants are not necessarily a representative sample of the population in the sending countries. Incentives to migrate differ among the various groups of the population depending on their observable and unobservable characteristics. Several authors have attempted to model the selection patterns of international migration.

On the one hand, Chiswick (1999) develops a model showing that immigrants are positively selected. On the other hand, Borjas (1987) shows that immigrants are selected from groups with lower qualifications when the returns to skills are more dispersed in the sending countries compared with the dispersion in the destination economies. In contrast,

immigrants tend to possess higher qualifications when the returns to skills in source economies are more egalitarian compared with the returns to skills in destination economies. Comparing theoretical predictions with US data, Borjas (1987) shows that positive selection is much more likely among immigrants from advanced countries, where the returns to skills are lower, whereas negative selection is more likely among developing countries, where the returns to skills are more unequal. Then, negative selection is expected for a country such as Mexico. A surge in empirical research has attempted to corroborate this claim.[2]

Previous studies regarding the case of Mexico have obtained mixed results. Using the US and Mexican population censuses of 1990 and 2000, Chiquiar and Hanson (2005) find evidence that Mexican immigrants in the US tend to be located in the middle of the wage distribution in Mexico. Orrenius and Zavodny (2005) developed a similar model and found empirical evidence consistent with Chiquiar and Hanson's (2005) results by using data from the Mexican Migration Project (MMP).

However, the US census may provide an incomplete picture of Mexican migrants.[3] Ibarraran and Lubotsky (2007) used the 2000 Mexican census to estimate the level of education of the Mexican migrants from 1995 to 2000. The researchers found that the Mexican migrants in the US are less educated by half a year than the remaining population in Mexico. The researchers argue that the missing migrants in the Mexican census (i.e., those who traveled with their entire families) cannot reverse the result of negative selection.

Fernandez-Huertas (2011) also challenges the empirical findings of intermediate selection. He uses the Labor Force Survey (ENE, a quarterly survey similar to the CPS in the US) from 2000 to 2004. In ENE, one can identify a Mexican emigrant to the United States before and after his or her departure. However, one can only identify short-term departures because ENE follows households for 5 consecutive quarters. Using the wages of these individuals before they migrate, Fernandez-Huertas (2011) finds that Mexican male

[2] Chiquiar and Hanson (2005), Fernandez-Huertas (2011), Ibarraran and Lubotsky (2007), Kaestner and Malamud (2010), Lacuesta (2006),McKenzie and Rapoport (2010) and Orrenius and Zavodny (2005). However, see also Grogger and Hanson (2011) for a model in which absolute wage differences, not relative differences, are the main determinant of migration.

[3] Ibarraran and Lubotsky (2007) point out that the migrants in the US census often overreport their education levels, possibly because of failures in translation or their inappropriate understanding of the survey options. Additionally, the researchers show that the US census underestimates the size of the illegal population, which is generally composed of low-skilled workers.

immigrants from 2000 to 2004 earn lower wages and have less education than individuals who remain in Mexico. This finding provides evidence of negative selection.

McKenzie and Rapoport (2010) point out that migration networks can partially reconcile some of the conflicting findings in the literature. They show that access to migration networks can theoretically alleviate some of the migration costs, with low-skilled migrants experiencing the greatest benefits. Hence, the researchers' model predicts positive self-selection in communities with weak migration networks and negative self-selection in communities with strong migration networks. Using the *Encuesta Nacional de la Dinamica Demografica* 1997 (ENADID), McKenzie and Rapoport find that their theoretical prediction holds for males between 15 and 49 years old in areas with populations less than 100,000 people. The probability of migration increases with education in communities with weak migrant networks and decreases with education in communities with strong migrant networks.

Recent papers have not reported consistent results. On the one hand, Kaestner and Malamud (2010) use the Mexican Family Life Survey (MxFLS) to find that male Mexican migrants are selected from the middle of the observed skill distribution. However, when the researchers control for migration costs, the evidence of intermediate selection diminishes. They also point out that no relationship exists between immigrant status and the distribution of the unexplained component of wages. On the other hand, Ambrosini and Peri (2011) use the same data source to find evidence of negative selection that is similar to the evidence found by Fernandez-Huertas (2011). Future research must explain why the same data source produces such different results.

These scholars have obtained their results while paying little or no attention to the following fact: an important proportion of migrants do not permanently reside in the country of destination. Bratsberg and Borjas (1996) developed theoretical implications of these migrants' existence. In their model, two types of individuals decide to return after migration: 1) individuals whose decisions to temporarily migrate are due to optimal decisions within their life cycles, and 2) individuals who return once they discover that their incomes in the country of destination are sufficiently worse than predicted. The model shows that return migration as an optimal life cycle decision occurs when the migrants discover that their returns to skills in their countries of origin are larger than those the

migrants would have obtained had they not moved temporarily.[4]

Bratsberg and Borjas (1996) also show that return migration accentuates the selection type in the original flow. If the original flow is characterized by low-skilled individuals, then the high-skilled migrants will have incentives to return, which accentuates the negative selection among the remaining migrants in the destination economy. If the original flow is characterized by high-skilled individuals, then the low-skilled migrants will have incentives to return, which accentuates the positive selection.

Within this theoretical framework, Coulon and Piracha (2005) analyze the migrants who have returned to Albania by using information from the source country. The researchers show that the decision to migrate may temporarily be an optimal decision because the wages are greater than what they would have been had the return migrants decided to permanently stay in Albania. With respect to the type of selection, the migrants who returned exhibited negative selection whereas those who never migrated would have gained higher wages had they been paid with the same returns to skills that the returning migrants received. Rooth and Sarela (2007) concentrate on Finnish immigrants in Sweden. Finland and Sweden have free mobility of labor between themselves. The returns to observable skills are higher in Finland than in Sweden. The researchers obtain the result predicted by the Bratsberg and Borjas model (i.e., negative selection) by using the data regarding the performances of returning migrants in the Finnish market.

The possibility of improving skills is not the only reason that the literature has explored to explain return migration. Dustmann (2003) shows that parents' concerns about their children can increase return migration. Dustmann and Weiss (2007) constructed a theoretical model in which higher preferences for consumption in the home country or high purchasing power of the host country's currency in the migrants' home country could lead to return migration. Gibson and McKenzie's (2011) empirical evidence for a group of highly skilled migrants shows that the decision to return is strongly linked to family and lifestyle reasons rather than to the income opportunities in different countries. Determining the importance of these explanations for return migration among Mexicans is beyond the

[4] Other models have similar implications. Dustmann and Kirchkamp (2002) construct a model in which return migration is also related to increases in wages after the migrants' return. Their model shows how the existence of different activities after the migrants' return can lead to different optimal time periods for the migration process.

scope of our article.

For the case of Mexico, only Lacuesta (2006) and Ambrosini and Peri (2011) have investigated the type of selection among return migrants. We discussed the differences and advantages of our approach earlier. We use information from Mexican censuses to show that the degree of selection among return migrants has changed over time (i.e., from positive selection in 1990 to negative selection in 2010). Additionally, our results indicate that the increase in wages produced by migration has decreased. The changes in selection are consistent with a decline in costs produced by migrant networks, as proposed by McKenzie and Rapoport (2010).

3 Empirical Strategy

Most of the Mexican immigrant studies on self-selection patterns have attempted to assess the robustness of Chiquiar and Hanson's (2005) finding of intermediate selection. To compare our results with those obtained in that article, we also construct counterfactual densities of wages, which is a methodology originally developed by DiNardo, Fortin and Lemiux (1996). This methodology also allows us to compare the results with those obtained by Fernandez-Huertas (2011), whose study reports the highest degree of negative selection in the literature.

3.1 Counterfactual densities

We aim to calculate the distribution of the wages that return migrants would have obtained had they never migrated. We can do so by combining the wage structure of nonmigrants with the observable characteristics of the return migrant population. Then we compare this counterfactual distribution with the observed distribution of the nonmigrants' earnings to establish the type of selection among the return migrants. We refer to w as wages, z as the observed characteristics of the individual in domain Ω, f^s as the density function of the nonmigrants (s denotes stayers), f^m as the density function of the return migrants, and f_m^s as the counterfactual density function of the wages that the return migrants would have earned had they never migrated. We define I as an indicator of

whether the individual is a stayer (*s*) or a return migrant (*m*).

The wage distribution for the nonmigrants is

$$f^{s}(w) = \int_{z \in \Omega} f^{s}(w \mid z) f(z \mid I = s) dz \tag{1}$$

and the wage distribution for the return migrants is

$$f^{m}(w) = \int_{z \in \Omega} f^{m}(w \mid z) f(z \mid I = m) dz \tag{2}$$

The counterfactual distribution of the wages that the return migrants would have experienced had they been paid according to the wage structure of the nonmigrants is

$$f_{m}^{s}(w) = \int_{z \in \Omega} f^{s}(w \mid z) f(z \mid I = m) dz \tag{3}$$

This expression represents the counterfactual density of the return migrants had they never migrated. We assume that the wage density function does not depend on the distribution of characteristics *z*. Therefore, we can use the characteristics of the return migrant population and integrate them over *z* in the function for the wage distribution of the nonmigrants. Instead of integrating, DiNardo, Fortin and Lemiux (1996) propose modifying equation (3) as follows:

$$f_{m}^{s}(w) = \int_{z \in \Omega} \psi(z) f^{s}(w \mid z) f(z \mid I = s) dz \tag{4}$$

where $\psi(z) = \dfrac{f(z \mid I = m)}{f(z \mid I = s)}$. Hence, we only need to know $\psi(z)$ and reweight the wage distribution for the nonmigrants to obtain the counterfactual distribution of the wages that the return migrants would have obtained had they never migrated. Using Bayes' Rule, the weight $\psi(z)$ can be rewritten as the following:

$$\psi(z) = \frac{f(I = m \mid z) f(I = s)}{f(I = s \mid z) f(I = m)} \tag{5}$$

The reweighting function $\psi(z)$ assigns higher weights to nonmigrants with values of *z* close to the characteristics of the return migrants and lower values to individuals with characteristics that are not so close to those of the return migrants. Thus, the reweighted population has values of *z* similar to those of the return migrants.

A possible bias in the methodology lies in the role of unobservable characteristics. For example, if the return migrants tend to have greater motivation, then our methodology will assign excessively low counterfactual wages. Conversely, if the migrants tend to be less motivated, then we will give them excessively high counterfactual wages. In a recent article, Kaestner and Malamud (2010) showed that there is little evidence of selection in the unobservables between migrants and nonmigrants in Mexico. This finding suggests that the potential bias caused by variables outside the vector *z* is small. Nevertheless, in the robustness test section, we implement a flexible estimation procedure by using different observable characteristics of *z*.

To estimate the weight $\psi(z)$, we notice that $f(I = s)$ and $f(I = m)$ are the proportions of the nonmigrants and the return migrants in the population. The other two elements, $f(I = s \mid z)$ and $f(I = m \mid z)$, are the conditional probabilities of being a nonmigrant (*s*) or a return migrant (*m*). We can easily estimate the conditional probabilities from the pooled population by using probit or logit conditioning on the set of characteristics *z*. After obtaining $\psi(z)$, we introduce it into the wage distributions for the nonmigrants, as in equation (4), to calculate the counterfactual wage distribution of the return migrants. We calculate the wage distributions by using nonparametric distributions.

Following Chiquiar and Hanson (2005) and Fernandez-Huertas (2011), we focus on the distribution of the wages that the return migrants would have obtained had they not migrated $f_m^s(w)$. After obtaining this counterfactual distribution, we compare it with the wage distribution of the nonmigrants and obtain the type of selection. In other words, we nonparametrically characterize the wage distributions to obtain the following:

$$f_m^s(w) - f^s(w) = \int_{z\in\Omega}(\psi(z)-1) f^s(w \mid z) f(z \mid I = s)dz \quad (6)$$

A positive difference indicates that a greater proportion of the migrants returned rather than stayed at the given level of wages. If the difference is negative, then the proportion of return migrants is lower. For negative selection, we must note a positive difference in low wages and a negative difference in high wages. However, if the selection is positive, then we must observe a negative difference in low wages and a positive difference in high wages.

An alternative way to characterize the counterfactual distribution of wages is to simply use the new factor $\psi(z)$ to compute the reweighted statistics in the distribution of the nonmigrants. These new statistics characterize the distribution of the wages that the return migrants would have obtained had they not migrated. We can compare these wages by utilizing the statistics arising from the observed distribution of the nonmigrants to determine the differences. We can calculate the standard errors of the statistics by bootstrapping the procedure.

The described methodology considers the full population of return migrants and nonmigrants regardless of whether the individuals work. However, the manner in which the characteristics in vector *z* affect the rate of labor market participation can differ between the return migrants and the nonmigrants. To eliminate any differences that might arise because of different labor market participation rates, in the robustness test section, we estimate the model by restricting the sample to only the individuals who are working and show that this procedure does not affect the main estimates.

4 Data and Descriptive Statistics

We use the Mexican Population Census for the years 1990, 2000 and 2010. One can obtain these data from the website of the National Institute of Statistics in Mexico.[5] The census includes a question about the place of residence 5 years before the survey takes place. Additionally, in the 2000 and 2010 censuses, we can identify the individuals who

[5] The 10 percent samples are available through the INEGI website (http://www.inegi.org.mx).

have migrated within the 5 years preceding the census but then returned during that period. We use these questions to identify the return migrants.

Two different types of people qualify as return migrants from the United States. First, we include individuals born in Mexico who lived in the United States 5 years prior to the census and resided in Mexico when the Census took place. Second, we include individuals born in Mexico who lived in Mexico 5 years prior to the census but migrated to the United States during that period and resided in Mexico when the census information was collected.

We restrict our sample to the individuals born in Mexico who were between 20 and 59 years old.[6] Unfortunately, the census does not allow us to identify international migratory activities beyond the 5 years prior to the census's survey date. Hence, individuals who may have migrated before that period are considered as nonmigrants by design. However, to counteract this possible bias, we use other data sources that do not restrict the time period of the return migration to verify the sensitivity of this result, as explained below.

In addition to the information about migration, the census includes important socio-demographic data. We use the following variables: sex, education, age, indigenous membership, income from employment, hours worked, type of activity, unemployment, geographical location and marital status.

To estimate the wage distribution, we only use individuals who reported a positive hourly wage.[7] However, it is important to emphasize that the reweighting procedure uses the full population of return migrants and nonmigrants, not just the working population. Additionally, we consider the individuals who reported more than 100 hours worked to have earned an invalid wage. Later, we analyze the sensitivity of our results by considering the differences in the participation rates between the return migrants and the nonmigrants.

[6] The 1990 census recorded 3,433,584 nonmigrants and 6,868 return migrants. The 2000 census recorded 4,535,926 nonmigrants and 38,112 return migrants. Of the return migrants in 2000, 17,235 are return migrants who lived sometime within the last 5 years in the US. In 2010, there were 5,521,552 nonmigrants and 108,691 return migrants. Of the return migrants in 2010, 21,978 were return migrants who lived sometime within the last 5 years in the US.

[7] In 1990, 1,581,113 nonmigrants and 3,083 return migrants had valid hourly wages; 2,154,906 nonmigrants and 17048 return migrants had valid hourly wages in 2000; and 2,429,803 nonmigrants and 54,235 return migrants had valid hourly wages in 2010.

4.1 Descriptive Statistics

Table 1[8] shows the main features of the return migrants (column *RM*) and the stayers (column *S*) among the three censuses. The return migrants tend to be younger than the rest of the population by 1 to 3 years. In addition, the return migrants are mostly composed of men; the proportion of men among the return migrants increased from 66 percent in 1990 to 76 percent in 2010.

We classified the six groups of states in accordance with their migration rates in 1950.[9] Using this classification, we try to identify the individuals' access to migration networks. High-migration states are Aguascalientes, Durango, Guanajuato, Jalisco, Michoacan, San Luis Potosi and Zacatecas. All of them are located at the center of Mexico. Low-migration states are Campeche, Chiapas, Quintana Roo, Tabasco and Veracruz and Yucatan. All of these states are located in southern Mexico. The third group is composed of states that exhibited an intermediate rate of migration in 1950: Colima, Mexico State, Guerrero, Hidalgo, Morelos, Nayarit Oaxaca, Puebla, Queretaro, Tlaxcala and Sinaloa. The fourth group consists of the states located in northern Mexico: Baja California, Baja California Sur, Chihuahua, Coahuila and Sonora Tamaulipas. Finally, we consider the state of Nuevo Leon (NL) and Mexico City as isolated regions because of their economic importance.

The table shows how the geographical location patterns of the return migrants have changed over time. The number of migrants returning to states with historically high migration rates has declined from 50 percent to only 35 percent, whereas the proportion of nonmigrants is only 21 percent. More migrants have returned to the states with low and intermediate migration rates with each census. Although the number of migrants returning to the north declined from 1990 to 2000, this number remained stable from 2000 to 2010. At the same time, the proportion of the nonmigrant population was stable across all regions in the three censuses. Mexico City and Nuevo Leon showed a low rate of return migration, and the importance of these states to the return migrant population has decreased over time.

Another characteristic of the geographical location patterns of the return migrants is

[8] We used the weights provided in each of the censuses in our descriptive statistics and estimates, except when we defined the size of the sample N.

[9] We follow the classification proposed by Hanson (2007).

the growing importance of the rural sector. Although the rural sector accounted for 28 percent of the return migrant population in 1990, by 2010 that proportion had risen to 36 percent. This change occurred even though the importance of the rural sector to the nonmigrants decreased from 25 percent in 1990 to 20 percent in 2010.

The locations of the return migrants in different regions of the rural sector have followed a pattern similar to that of the total population. The proportion of return migrants in the rural sector has decreased in the high-migration states but has increased in states with low and intermediate migration rates and has remained stable in the north.

The indigenous population has produced fewer return migrants than nonmigrants. However, the proportion of the indigenous population serving as return migrants increased from 2 to 4 percent, whereas the proportion of indigenous people among the nonmigrant population has decreased from 8 in 1990 to 7 percent in 2010.

With regard to the years of education, the difference between return migrants and nonmigrants has changed over time. Whereas in 1990, the return migrants had 0.57 more years of education than the nonmigrants, in 2010, the return migrants had 0.84 fewer years of education than the nonmigrants. The average education level has increased for both groups.[10]

In terms of educational groups, the results indicate that the selection of return migrants has evolved toward negative selection. Over the years, the proportion of return migrants with no formal education is lower compared with the proportion of nonmigrants in the same level. In addition, the proportion of return migrants with incomplete primary schooling is similar to that of nonmigrants. Nevertheless, whereas in 1990, the proportion of individuals in secondary and higher educational groups was similar for both return migrants and nonmigrants, by 2010, the proportion of nonmigrants in high school and college had become larger than the proportion of return migrants in high school and college.

4.2 Male and Female Differences

[10] We define the educational groups by six consecutive levels: No Education, Primary Incomplete, Primary, Secondary, High School and College. Primary Incomplete, Primary, Secondary, High School and College indicate that the individual completed 1-5 years, 6-8 years, 9-11 years, 12-16 years and 17 years or more of schooling, respectively. This classification reflects the structure of the Mexican educational system.

Most of the previous studies on selection and Mexican migration to the United States have focused on men. However, women represent an important proportion of return migrants (i.e., 34 percent in 1990 and 24 percent in 2010). Hence, it is important to investigate any possible gender differences. Table 2 shows the main characteristics of the male population, and Table 3 contains the characteristics of the women. Both tables include important labor market characteristics, such as wages, and labor market participation rates.

In both cases, the return migrants are younger than and exhibit almost the same propensity to be married as the nonmigrant population. However, the two populations differ in many other respects. For example, with regard to the size of locality, the female return migrants exhibit a lower tendency to reside in rural areas than the males. The female rural population represented less than 30 percent of the return migrants, whereas the male rural population constituted more than 30 percent of the return migrants in the three censuses.

In terms of education, the female return migrants have a higher level of education than the nonmigrants. Conversely, the male return migrants exhibit less education than the nonmigrant population. In 1990, the difference in education levels between the female return migrants and nonmigrants was 1.37 years, which decreased to 0.28 in 2010. In contrast, the difference between the male return migrants and nonmigrants was -0.11 years in 1990. By 2010, the negative difference had become -1.34 years. Positive selection in terms of education is disappearing among the women and becoming more negative in the case of the men.

Both the male and female return migrants exhibit larger rates of unemployment. We measure this rate as the proportion of the population who stated that they had sought work the week before the census. Additionally, in both cases, fewer return migrants are employed compared with the nonmigrant population.

The return migrants have higher wages than the nonmigrants, except for the men in 2010. As in the case of education, the wage differences for both the men and the women are more favorable for the return migrants in 1990 than in 2010. For the males in 1990, the log wage difference was 0.42, which decreased to 0.13 in 2000 and became a negative difference of -0.09 in 2010. For the women, the difference has always been positive in favor of the return migrants. It was 0.49 in 1990 but decreased to 0.25 in 2001 and to only

0.05 in 2010. Using our estimation methodology, we aim to explain the part of these differences that are attributable to the differences in human capital prior to the migration and to the migration process itself.

With respect to the number of hours worked per week, both the men and the women tend to work less if they are return migrants, although the difference is small. Interestingly, both the male and female return migrants exhibit a greater tendency to become entrepreneurs or to be self-employed. This trend remained stable during the period of the study. However, the return migrants are employed in a lower proportion of jobs that offer health insurance (i.e., formal sector jobs) as part of the employment benefits.

The descriptive statistics suggest that the selection in terms of education and wages are becoming more negative for both the men and the women. This pattern also holds regardless of whether we focus on the full population or the working population (i.e., the last row in tables 2 and 3). Hence, we can infer that excluding the women from the analysis can skew the results toward negative selection. We account for this bias in our analysis by estimating the selection among the men and the women as separate cases.

5 Results

Following the previous literature, we construct the counterfactual distribution of the wages that the return migrants would have obtained had they never migrated. As a first step, we look into the distributions of log hourly wages before estimating the counterfactuals. We consider the distributions of the men and women separately in each of the censuses.

To estimate $f(I = s \mid z)$ and $f(I = m \mid z)$, we used a logit model for the full sample, with a dependent variable indicating whether the individual was a return migrant. We divided age into 8 groups of 5 years, and we formed indicative variables for each group. To consider the high dependence of return migration on geographical variables, we used dummies for each of the following regions: high-migration states, low-migration states, intermediate-migration states, North, Mexico City, NL, rural, high-migration rural, low-migration rural, intermediate-migration rural and north rural. In addition, we used an indigenous membership variable. To include education, we used dummies for each

aforementioned level of education. Using the logit estimates, we obtained the weight $\psi(z) = \frac{f(I = m \mid z) f(I = s)}{f(I = s \mid z) f(I = m)}$ and constructed the counterfactual distribution of the wages that the return migrants would have obtained had they never migrated, as indicated in equation (4), by using kernel methods.

5.1 Selection by year

Figure 1[11] shows the results for the men in the 1990 census. Figure 1a shows the observed wage distributions for the return migrants and the nonmigrants. Figure 1c shows the difference between the two distributions. There is a vertical line at the median of the nonmigrants. We can see that the wage distribution of the return migrants is clearly to the right of the distribution of the nonmigrants. This finding implies positive selection.

Figure 1b shows the observed wage distribution of the nonmigrants and the counterfactual distribution of the wages that the return migrants would have obtained had they not migrated. Figure 1d shows the difference between these two distributions. Compared with the nonmigrants, a greater proportion of the return migrants are located in the middle and upper parts of the distribution, and fewer migrants are located in the lower part of the distribution. This evidence suggests positive selection among the male return migrants in 1990.

Table 4 shows another approach to observing these patterns. The table contains statistics on the observed wage distributions and the counterfactual wage distributions for the men and the women in each year. For the men in 1990, we can see that the average wage and the wage in each percentile of the return migrants' wage distribution are higher than those for the nonmigrants. On average, the return migrants earn 42 percent more than the nonmigrants. The same is observed for the counterfactual distributions. Each percentile of the return migrants' counterfactual distribution is located to the right of the percentiles of the nonmigrants' distribution, except for the 90th and 95th percentiles. In 1990, the return

[11] We constructed all of the figures by using an Epanechnikov kernel. We used two times Silverman's (1986) optimal bandwidth. As in the analysis of Chiquiar and Hanson (2005), we found that the appearance of densities resulting from the use of the optimal bandwidth presented problems for our analysis. Two times the optimal bandwidth resulted in figures similar to those reported previously in the literature. Furthermore, the bandwidth chosen has no effects on the counterfactual statistics that use the reweighting factor.

migrants would have earned, on average, 6 percent more than the nonmigrants had the return migrants never left the country (i.e., 3.17 *minus* 3.11). The difference between the observed wages that the return migrants received and the wages that they would have obtained equals 36 percent, which is a large positive effect. We call this effect the wage premium for migrating and returning.

One can note a similar pattern of positive self-selection in the case of the women in 1990 in Figure 2. The wage distribution of the return migrants has fewer individuals in the median of the nonmigrant wage distribution and a larger mass in the right tail of the wage distribution. After we estimate the counterfactual, we observe a decrease in the differences across both distributions, but evidence of positive selection still exists. Table 4 shows that the wages for the female return migrants are 49 percent higher than the nonmigrants' wages. If the female migrants had not migrated, then they would have obtained wages that are 11 percent higher. The wage premiums for migrating and returning equals 38 percent.

The distributions of the men in the 2000 census can be seen in Figure 3. The wage distributions show that fewer return migrants are in the lower part of distribution and that more return migrants are in the middle and upper parts of the distribution. After we estimated the counterfactual distribution, we found that the positive selection pattern observed in 1990 changes. The return migrants are drawn more from the middle of the distribution and less from the upper and lower parts of the distribution. In the graph, it is difficult to know whether the selection is positive or negative.

Table 4 also shows the statistics of the wage distributions in the year 2000. For the men, the percentiles of the wage distribution for the return migrants are still to the right of the percentiles of the distribution for the nonmigrants, but the average difference between the two groups has decreased. However, when we analyze the statistics in the counterfactual distribution, the return migrants are only located to the right of the nonmigrants in the 5th to 25th percentiles, whereas the nonmigrants tend to receive higher wages in the higher percentiles. This finding shows that the return migrants tend to be located disproportionally in the middle of the distribution. As a result, the wages that the return migrants would have obtained had they not migrated is slightly less than the nonmigrants' wages by 5 percent on average. Additionally, whereas the return migrants' wages were more than 36 percent higher than the nonmigrants' wages in 1990, this number

had decreased to 18 percent by 2000.

Figure 4 shows the data on the women in the 2000 census. The wage distribution of the return migrants is clearly to the right of the nonmigrants' wage distribution. The counterfactual distribution shows a pattern in which the return migrants are drawn more from the middle and the upper middle parts of the distribution and less from the lower part. This finding suggests a pattern of positive selection for the women in the 2000 census. The percentiles in Table 4 show that the selection pattern remains positive for the women. In both cases, the wages obtained after migrating are higher than the wages that the women would have obtained had they not migrated. Additionally, the percentiles of the return migrants' wage distributions are higher than the percentiles of the nonmigrants' distribution. On average, the return migrants' wages are 25 percent higher than the nonmigrants' wages, and the counterfactual wages of the return migrants are 8 percent larger than the nonmigrants' wages.

In 2010, the pattern of selection is negative for men, as shown in Figure 5. Using the counterfactual distribution, we can show that there are more return migrants in the lower-to-middle part of the wage distribution. If we analyze the statistics shown in Table 4 for the men in 2010, then the counterfactual distribution of the return migrants and the distribution of the nonmigrants only coincide in percentile 5. In the rest of the percentiles, the return migrants are to the left of the nonmigrants. This finding indicates that the return migrants have lower wages than those of the nonmigrants. The difference between the wages that the return migrants would have obtained had they not migrated and those of the nonmigrants is -14 percent. The wage premium of migrating and returning is equal to 5 percent. Hence, this wage premium has also decreased over time.

Figure 6 shows the self-selection patterns for the women in 2010. The return migrants are somewhat more concentrated near the median of the nonmigrants' wage distribution and less among the high and low wages. According to Table 4, the difference between the wages that the return migrants would have experienced had they not migrated and those of the nonmigrants is -2 percent. In the lower part of the distribution (i.e., in the 5th and 10th percentiles), the return migrants are located to the right of the nonmigrants' distribution, whereas in the higher wages, the return migrants are to the left of the nonmigrants. This finding indicates that the return migrants are selected from the middle of

the distribution.

To determine whether the differences between the nonmigrants' wage distribution and the counterfactual distribution of the wages that the return migrants would have earned had they not migrated are statistically significant, we calculate the standard error of the difference by using 250 bootstrap repetitions. Table 5 shows the results. Overall, the standard errors are small, and most of the differences are statistically significant, except for the women in 2010 in the 20th and 25th percentiles of the wage distribution.

5.2 Selection over time

The above results show that the self-selection among the return migrants tends to become negative over time. The same is true for both the men and the women. One concern about comparing changes in self-selection by using the estimates in each year is that doing so may confound the changes in the compositions of immigrant and resident populations with the changes in skill prices.[12] To address this concern, we must keep the skill returns constant. In the case of the return migrants in 1990, we estimate the following weighting function:

$$\psi_{m90}^{s10} = \frac{f(I = m, y = 1990 \mid z)}{f(I = s, y = 2010 \mid z)} \tag{7}$$

With this weight, we adjust the characteristics of the nonmigrants in 2010 according to the characteristics of the return migrants in 1990. Using this weight over the distribution of the nonmigrants in 2010 generates the counterfactual distribution of the wages that the return migrants would have earned in 1990 had they been paid as nonmigrants in 2010. For the nonmigrants in 1990, we calculate the following:

$$\psi_{s90}^{s10} = \frac{f(I = s, y = 1990 \mid z)}{f(I = s, y = 2010 \mid z)} \tag{8}$$

[12] Chiquiar and Hanson (2005) p. 264.

Applying this weight over the distribution of the nonmigrants in 2010 generates the counterfactual distribution of the wages that the nonmigrants would have earned in 1990 had they been paid as nonmigrants in 2010. Using equations (7) and (8), we can nonparametrically estimate the degree of selection for the return migrants in 1990 by evaluating this estimate in terms of the skill prices in 2010:

$$f_{m90}^{s10}(w) - f_{s90}^{s10}(w) = \int_{z\in\Omega} (\psi_{m90}^{s10} - \psi_{s90}^{s10}) f^{s10}(w,z)dz \tag{9}$$

Following a similar approach, we can estimate the degree of selection in 2000 by evaluating equation (9) in terms of the skill prices in 2010. For the 2010 census, we only use the 2010 estimates generated in the previous section.

Figures 7 and 8 show the difference between the return migrants' wage distribution and nonmigrants' wage distribution when their characteristics are priced as nonmigrants in 2010 for the men and the women, respectively. The vertical line shows the median log wages in 2010 for the nonmigrants. For the men, Figure 7 shows a pattern of positive selection in 1990, intermediate selection in 2000 and negative selection in 2010. For the women, Figure 8 shows positive selection in 1990 and 2000 but intermediate selection in 2010.

The results of these estimates also appear in Table 6. For the men, we find a positive selection of 5 percent in 1990, a negative selection of 5 percent in 2000 and a negative selection of 14 percent in 2010. Similar to the pattern obtained in the previous section, the results show that the degree of selection becomes more negative. For the women, the results also show that the degree of selection becomes more negative. In 1990, the women show an average positive selection of 12 percent. They show a positive selection of 5 percent in 2000 and a negative selection of 2 percent in 2010. In the column DIF, we show the differences in the statistics and the bootstrapped standard errors of the difference. Again, the differences in the distributions are statistically significant at the mean, variance and almost every percentile. The previous results show that the growing negativity in the degree of selection is not due to a change in the returns of the nonmigrants' characteristics.

5.3 Extensions

Previous studies on self-selection among permanent migrants have shown that the type and degree of selection tends to differ if we focus on different groups. Fernandez-Huertas (2011) finds that in the rural sector, selection is positive, whereas in the urban sector, selection is negative. McKenzie and Rapoport (2010) show that the degree of selection depends on the people's access to migration networks. In areas with highly developed migration networks, selection will tend to be negative, whereas in areas with underdeveloped networks, selection could be positive. We determine whether these patterns hold among the male return migrants.

Table 7 shows the results of the wage differences (only males) between the return migrants and the nonmigrants in the urban and rural sectors. The type of selection has always been positive in the rural areas. In 1990, the male return migrants would have obtained wages that are 21 percent higher than those of the nonmigrants if the migrants had never left the country. In 2000, this difference was equal to 17 percent, and in 2010, this difference declined to 4 percent. The 10th, 50th and 90th percentiles are higher in the return migrants' counterfactual distribution than in the nonmigrants' distribution in the rural sector in 1990 and 2000. In 2010, only the 10th percentile is higher in the counterfactual distribution of the return migrants, whereas the 50th percentile is the same for both distributions, and the 90th percentile is higher in the nonmigrants' distribution. This finding indicates that a small number of return migrants are low-wage workers and that this effect dominates the fact that few migrants with high wages exist in the rural sector in 2010.

In the urban sector, the type of selection changes from positive in 1990 to negative in 2000 and 2010. In other words, the selection is becoming more negative over time. In 1990, the percentiles of the return migrants' counterfactual wage distribution are larger than those of the nonmigrants' distribution, although the difference is only 2 log points in the 90th percentile. This finding suggests that, although the average difference is positive, the difference is due to the greater presence of return migrants in the middle of the distribution and not at the top. Although in 2000, the wages of the return migrants and the nonmigrants in the 50th percentile are equal, the return migrants' wages are greater than the nonmigrants in the 10th percentile and lower in the 90th percentile. In 2010, the percentiles of the return migrants' counterfactual wage distribution are lower than those of the nonmigrants'

distribution. In sum, the evidence from the urban sector indicates a change toward negative selection.

Overall, the rural sector exhibits positive selection, and the urban sector changes from positive to negative selection. However, the pattern is becoming more negative in both sectors. Thus, the change in the type of selection from 1990 to 2010 is not a result of the increased number of rural workers becoming return migrants. In both rural and urban areas, the return migrants receive higher wages than the wages they would have obtained had they not migrated (i.e., the observed wage of the return migrant minus the counterfactual wage). This difference has fallen over time in both groups. In 1990, the male return migrants in the rural sector obtained hourly wages that were 35 percent higher than the wages they would have obtained had they not migrated. In the urban sector, this difference was 34 percent. In 2000, the increase in wages dropped to 14 percent within the rural population and to 19 percent within the urban population. In 2010, the difference was 7 and 3 percent in the rural and urban sectors, respectively. Thus, the wage premium for migrating to the US and returning back to Mexico has fallen over time.

To investigate the effect of migration networks, we divide the Mexican states between those with a long migratory tradition and the rest.[13] We expect that the migration networks are more developed in states in which the rate of migration has been historically high. If McKenzie and Rapoport's (2010) proposal is true, then we should find more evidence of negative selection in states with a tradition of high migration rates.

Table 7 shows the results. The difference between the nonmigrants' wages and the wages that the return migrants would have earned had they never migrated increased from 8 percent in 1990 to 17 percent in 2010 in states with a history of high immigration rates. The 10th, 50th and 90th percentiles of the counterfactual wage distribution for the return migrants are always to the left of the percentiles of the nonmigrants' distribution. This finding provides evidence of negative selection. In the rest of the states, the pattern changed from positive selection of 17 percent in 1990 to negative selection of 13 percent in 2010. The percentiles in the rest of states show that in 1990, the counterfactual wage distribution of the return migrants lay to the right of the nonmigrants' wage distribution. However, the

[13] The states with the highest rates of migration in 1950 were Aguascalientes, Durango, Guanajuato, Jalisco, Michoacan, San Luis Potosi and Zacatecas.

selection pattern has become more negative over time.

The states with high migration rates show a greater degree of negative selection each year. This finding is consistent with McKenzie and Rapoport's (2010) hypothesis. Moreover, the type of selection has become more negative over time in both types of states. Interestingly, the wage premium associated with migration differs between the states with high migration rates and the rest. In the states with high migration rates, the wage premium is lower than that in the rest of the states. In 1990, return migration was associated with a wage premium of 23 percent in the states with high migration rates and 45 percent in the rest of the country. In 2000, the wage premium was 10 percent in states with high migration rates and 24 percent in the rest of the country. In 2010, the premium was null in states with high migration rates and equal to 7 percent in the rest of the states.

5.4 Robustness

To establish the validity of the results shown in Figures 1 to 8 and Table 4, we perform some robustness tests. One possible critique of our study is that the unobserved components may bias the results. However, Kaestner and Malamud (2010) use a different survey and find little evidence of selection bias from unobservables. Hence, if the unobservables are correlated with the nonlinear functions of the observable characteristics, then adding these nonlinear functions should mitigate the possible bias. To do so, we estimate $f(I = s \mid z)$ and $f(I = m \mid z)$ by using the interactions between the respondents' years of education and each of the six geographical areas, the interactions between the years of education and each of the age groups, the interactions between the age and the six geographical areas, and the interactions between the rural sector and each educational level. The counterfactual wage distributions changed only slightly. The average counterfactual wage changed by no more than 2 log points. Table 8 shows the results under the column ALL. The first column (OR) in the table for each year shows the counterfactual estimates of Table 4.

Additionally, we obtained the wage distributions of both the return migrants and the nonmigrants from the individuals with valid wages, but we constructed the counterfactual distribution while considering the characteristics of all of the migrants, as if they had

participated in the labor market in the same manner as the nonmigrants. Doing so may also have biased our results. To determine whether this decision has important effects on our estimates, we restrict the sample to only those with valid wages. The second column (REST) in Table 8 shows the estimates generated by using the restricted sample. Once again, the results show no major changes, and the counterfactual average wage does not differ by more than 2 log points from the results in Table 4. The type of selection still becomes more negative over time, and the wage premia still decline because of migration.

Another concern is that the results could be due to the differences in the age structures between the migrant population and the nonmigrants. To determine if the age structure changes our results, we restricted the male sample to only the migrants between 20 and 35 years old. Table 9 shows the results. In terms of selection, the results are similar to those provided in Table 4. There is a positive selection of 4 percent in 1990, a negative selection of only 3 percent in 2000 and a negative selection of 10 percent in 2010. The immigration wage premium also follows the same pattern as that shown in Table 4.

The census only captures the migrants who returned within five years before the data were collected. Thus, it is difficult to know whether the migrants have temporarily or permanently returned to Mexico. A second weakness in the census is that we are only aware of the international migratory activities over a period of 5 years. Many individuals who have migrated in the past are considered nonmigrants in the census. The census does not contain information that allows us to resolve the potential bias caused by these characteristics of the census. However, we use a survey that may counteract this possible bias. The 2006 Social Mobility Survey (EMS) asked individuals if they had traveled to the United States to work for a month or more at least once in their life.[14] Using this survey, we can completely separate the population that has migrated at least once to work in the United States from those who have not.

We apply the same methodology by using the EMS as a robustness test. We concentrate on men because the survey was designed to interview mostly men. Table 10 shows the results. In 2006, the wages that the return migrants would have obtained had they not migrated is 3 percent lower than the wages of the nonmigrants. The immigration wage

[14] The survey was designed by Centro de Estudios Espinosa Yglesias, a civil association funded by Fundacion Espinosa Rugarcia (http://www.ceey.org.mx).

premium is 7 percent. The evidence suggests that the return migrants are more concentrated to the left of the median of the nonmigrants' wage distribution given that the 10th percentile is larger for the migrants, but the 90th percentile is higher for the nonmigrants. This result is similar to our observation in the 2000 census for men. Additionally, the fact that the average selection is less negative in the EMS does not eliminate the possibility of a slightly positive selection in the earlier years, as we estimated from the 1990 census. Hence, the census's restriction of immigrant status to the previous 5 years does not affect the main results and, in fact, the EMS supports the evidence provided by the 1990 and 2000 censuses.

Recent articles about the self-selection among Mexican migrants (e.g., Fernandez-Huertas 2011, Kaestner and Malamud 2010, and Ambrosini and Peri 2011) focus on the period from 2000 to 2005. We also use a survey conducted in 2006. This survey captures information about the return migrants in that period by using the same methodology employed in the 2000 census. The National Survey of Demographic Dynamics (ENADID) asked if any of the members of the respondents' households traveled to the United States with the objective of living in that country since 2001 and whether the same member had returned within the same period. The survey also asked for the locations in which the members of the household had lived five years before the census took place.[15] Table 10 provides the results of this survey for the men. The difference between the wages that the return migrants would have earned had they not migrated and the wages of the nonmigrants is 4 percent. The immigration wage premium is equal to 11 percent. Although ENADID uses the same methodology as the census, the results are similar to those found by the EMS. Nevertheless, both surveys do not show the large negative difference found in the 2010 census. This finding suggests that the self-selection pattern among the return migrants became more negative after 2006. The shocks suffered by the US economy after 2006 may explain part of the increasingly negative self-selection pattern.

We can make an additional comparison by using the information regarding the return migrants in a survey that employs a methodology similar to that of the labor force survey used by Fernandez-Huertas's (2011). This survey reports the greatest negative selection regarding Mexican migrants in the literature. The National Survey of Employment

[15] The survey is available at INEGI (http://www.inegi.org.mx).

and Occupation (ENOE) is a Mexican survey designed to capture the evolution of employment and unemployment since 2005. Each household is interviewed for 5 consecutive quarters. From the second interview on, the households report whether a new member is immigrating or whether the household has lost a member because of migration. We use the same methodology employed for the men in Table 4 for the data extracted from the third quarter of 2005 to the second quarter of 2007. However, before taking the logarithm of the wages, we estimate the wages relative to the average wages of the quarter to avoid seasonal effects, as proposed by Fernandez-Huertas (2011).

Table 10 provides the results under the column ENOE. The wages that the return migrants would have earned had they not migrated is 19 percent less than the wages of the nonmigrants. Contrary to all previous estimates, the wages obtained by the return migrants is 12 percent less than the counterfactual wages. In sum, after the migrants return, the return migrants earn a wage that is 31 percent less than that of the nonmigrants. This finding suggests that the return migrants become more negatively selected if we analyze the flow and not the stock, possibly because of the existence of low-skilled migrants who tend to make more trips. Nonetheless, three different surveys of the census show negative selection for the return migrants.

5.5 Implications

An important aspect to consider is whether the return migrants exhibit different skills from the permanent migrants. According to the discussion in section 2, these skills should be different in the Bratsberg and Borjas (1997) framework. However, with our datasets, we cannot directly test this prediction. Nonetheless, our results are similar to the findings regarding Mexican permanent migrants that were reported in the recent literature. For example, using the 2000 Census, Ibarraran and Lubotsky (2007) find that male Mexican migrants have 0.56 fewer years of schooling than the males in nonmigrant households. This finding almost coincides with our statistics, which indicate that the male return migrants have 0.64 fewer years of schooling than the nonmigrants. With respect to Chiquiar and Hanson's (2005) results, we also find evidence of positive selection among both the men and the women in 1990 as well as evidence of positive selection among the

women and intermediate selection among the men in 2000. This finding suggests that the differences in skills between the return migrants and the permanent migrants who settle in the United States could be small.

The evidence suggests that migration networks play an important role in alleviating the migration costs experienced by low-skilled individuals, as suggested by McKenzie and Rapoport (2010). The tendency to find negative selection increases as time goes on, regardless of the subgroups being considered. Moreover, in states with a long migration tradition, the self-selection patterns are negative, as predicted by Bratsberg and Borjas (1996), whereas in states with weak migration networks, the self-selection patterns tend to be positive or less negative. Future research should tackle the importance of alternative hypotheses to explain why the selection among the Mexican return migrants becomes more negative over time. Factors such as changes in the demand for labor because of shocks to both economies or the effect of changes to the enforcement of migration laws may play an important role.

There are important differences in the selection patterns between the men and the women among the return migrants. Not considering the women in the analysis biases the results toward finding negative selection. If these differences also exist among permanent migrants, as Chiquiar and Hanson's (2005) results suggest, then omitting women from the analysis of the selection patterns may produce misleading conclusions.

6 Conclusions

In this article, we analyzed the self-selection patterns of return migrants in Mexico by using the censuses for the years 1990, 2000 and 2010. In particular, we followed DiNardo et al.'s (1996) methodology to calculate the counterfactual wages that the return migrants would have earned had they not migrated. This methodology has been used to analyze the selection of Mexican immigrants into the United States (Chiquiar and Hanson 2005; Fernandez-Huertas 2011) but has not been utilized to analyze the selection of return migration into Mexico. We presented evidence suggesting that the self-selection patterns among the Mexican return migrants have changed over time (i.e., from positive selection in 1990 to negative selection in 2010). For example, the wages that the male return migrants

would have earned had they not migrated is 6 percent larger than the wages of the male nonmigrants. However, by 2010, this difference had declined to -14 percent. The growing negativity of the degree of positive selection is robust to the analysis of specific subgroups: rural and urban, men and women, and states with high migration rates and low migration rates. Furthermore, the negative selection results are robust with respect to the dataset used. We employed different datasets that measure the flow and stock of return migrants in different ways. The negative selection result is stronger in the sample that measures the flow of the return migrants.

Important differences exist among the different subgroups. Women tend to show more positive selection than men. For men in the rural sector, selection has been positive since 1990. However, states with high migration rates have shown negative selection since 1990. This last result is consistent with the role of migration networks in alleviating migration costs.

In general, the self-selection patterns tend to coincide with the results found in the literature on Mexican migrants living in the United States. For example, previous studies indicate that women are positively selected and that men show intermediate selection. We also find similar results for the censuses taken in the years 1990 and 2000. However, we find that the selection of the return migrants became more negative from 1990 to 2010. The similarity between our results and those of previous studies on Mexican migrants in the United States suggests that the differences in skills between return migrants and permanent migrants could be small.

An interesting result is that the observed wages of the return migrants are higher than the wages that they would have earned had they not migrated. In other words, there is a wage premium to migrate and return. This premium shows that migration has a positive effect on the Mexican economy. The previous literature on Mexican migration to the United States has neglected to study this effect. Hence, further research is necessary to understand the factors driving migrants to return to Mexico. Such research would help policymakers design return migration policies that may reduce the concern of a massive permanent migration wave to the United States.

References

- Ambrosini W, Peri, G (2011). The Determinants and the Selection of Mexico-US Migrations. Manuscript UC Davis.
- Borjas, G (1987). Self-Selection and the Earnings of Immigrants. American Economic Review. 77(4): 531-553.
- Bratsberg B, Borjas G (1996). Who Leaves? The Outmigration of the Foreign-Born. Review of Economics and Statistics. 78(1): 165-176.
- Chiquiar D, Hanson G (2005). International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States. Journal of Political Economy. 113(2): 239-281.
- Coulon A, Piracha M (2005). Self-selection and the Performance of Return Migrants: The Source Country Perspective. Journal of Population Economics. 18(4): 779-807.
- Dustmann C (2003). Children and Return Migration. Journal of Population Economics. 16 (4): 815-830.
- Dustmann C, Kirchkamp O (2002). The optimal migration duration and activity choice after re-migration. Journal of Development Economics. 67(2):351-372.
- Dustmann C, Weiss Y (2007). Return Migration: Theory and Empirical Evidence for the UK . British Journal of Industrial Economics. 45(2): 236-256.
- Fernández-Huertas J (2011). New Evidence on Emigrant Selection. Review of Economics and Statistics. 93(1): 72-96.
- Gibson, J., McKenzie, D. (2011). The economic determinants of emigration and return migration of the best and the brightest: Evidence from the Pacific. Journal of Development Economics. 95(1): 18-29.
- Grogger, J, Hanson G (2011). Income Maximization and the selection and sorting of international migrants. Journal of Development Economics. 95 (1): 42-57.
- Hanson G (2007). Emigration, Labor Supply, and Earnings in Mexico, in Mexican Migration to the United States, ed. G. Borjas, University of Chicago Press and National Bureau of Economic Research, Chapter 9, pp. 289-328.
- Gitter S, Gitter R, J. Southgate D (2008). The Impact of Return Migration to Mexico. Estudios Económicos. 23(1): 3-23.
- Ibarraran P, Lubotsky D (2007). Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census, in Mexican Migration to the United States, ed. G. Borjas, University of Chicago Press and National Bureau of Economic Research, Chapter 5, pp. 159-191.
- Kaestner R, Malamud O (2010). Self-Selection and International Migration: New Evidence from Mexico. NBER Working Paper 15765.
- Lacuesta A (2006). Emigration and Human Capital: Who Leaves, Who Comes Back and What Difference Does it Make? Banco de España Working Papers 0620.
- McKenzie D, Rapoport H (2010). Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. The Review of Economics and Statistics. 92(4): 811-821.
- Orrenius P, and Zavodny, M (2005). Self-selection among undocumented immigrants from Mexico. Journal of Development Economics, 2005. 78(1):215-240.

- Rooth D-O, Saarela J (2007). Selection in Migration and Return migration: Evidence from Micro Data. Economic Letters. 94(1): 90-95.
- Silverman, B. W (1986). Density Estimation for Statistics and Data Analysis. London: Chapman and Hall.

Table 1. Descriptive statistics: Full population.

	1990		2000		2010	
	S	RM	S	RM	S	RM
N	3,433,584	6,868	4,535,926	38,067	5,521,552	108,691
Age	34.4	33.3	35.1	32.5	36.6	34.3
Male	0.48	0.66	0.47	0.73	0.47	0.76
Rural	0.25	0.28	0.22	0.32	0.20	0.36
High Migration Region	0.21	0.5	0.2	0.45	0.21	0.35
Low Migration Region	0.15	0.01	0.15	0.04	0.15	0.09
Intermediate Migration Region	0.34	0.15	0.36	0.28	0.37	0.36
North Region	0.14	0.25	0.14	0.16	0.14	0.16
Mexico City	0.12	0.05	0.10	0.04	0.09	0.02
NL	0.04	0.03	0.04	0.03	0.04	0.02
High Migration: Rural	0.06	0.20	0.05	0.19	0.05	0.15
Low Migration: Rural	0.06	0.00	0.06	0.01	0.06	0.05
North Region: Rural	0.02	0.03	0.02	0.02	0.02	0.02
Intermediate Migration: Rural	0.10	0.05	0.09	0.10	0.08	0.14
Indigenous	0.08	0.02	0.07	0.03	0.07	0.04
Years of Schooling	6.36	6.93	7.98	7.91	9.26	8.42
No Education	0.15	0.08	0.07	0.03	0.05	0.02
Primary Incomplete	0.25	0.24	0.19	0.19	0.11	0.12
Primary	0.26	0.32	0.25	0.30	0.20	0.28
Secondary	0.20	0.21	0.24	0.27	0.28	0.35
High School	0.10	0.11	0.21	0.18	0.28	0.20
College	0.04	0.04	0.05	0.03	0.07	0.03

Notes: The sample is restricted to individuals who are 20-59 years old. Indigenous is a dichotomic variable representing the population that speaks an indigenous language. The states were divided into the following groups. i) High migration: Aguascalientes, Durango, Guanajuato, Jalisco, Michoacán, San Luis Potosí and Zacatecas; ii) low migration: Campeche, Chiapas, Quintana Roo, Tabasco and Veracruz; iii) intermediate migration: Colima, Estado de México, Guerrero, Hidalgo, Morelos, Nayarit Oaxaca, Puebla, Querétaro and Tlaxcala y Sinaloa; iv) north region: Baja California, Baja California Sur, Chihuahua, Coahuila and Sonora Tamaulipas. Nuevo Leon (NL) and Mexico City are considered separately. Rural represents the population living in areas with 2500 inhabitants or fewer. Years of schooling includes only the completed years. Primary incomplete, Primary, Secondary, High School and College indicate 1-5 years of schooling, 6-8 years of schooling, 9-11 years of schooling. 12-16 years of schooling and 17 years of schooling or more, respectively.

Table 2. Descriptive statistics: Men

	1990		2000		2010	
	S	RM	S	RM	S	RM
N	1,643,304	4,537	2,142,705	28,962	2,584,619	85,208
Age	34.5	33.4	35.2	32.5	36.5	34.3
Married	0.74	0.73	0.73	0.73	0.70	0.72
Rural	0.26	0.32	0.22	0.35	0.2	0.39
Years of Schooling	6.87	6.76	8.3	7.66	9.46	8.12
No Education	0.12	0.09	0.06	0.04	0.04	0.03
Primary Incomplete	0.24	0.26	0.18	0.20	0.11	0.13
Primary	0.26	0.32	0.24	0.31	0.20	0.30
Secondary	0.20	0.19	0.25	0.28	0.29	0.36
High School	0.12	0.10	0.21	0.15	0.28	0.17
College	0.05	0.04	0.06	0.03	0.08	0.02
Unemployed	0.021	0.031	0.012	0.020	0.045	0.073
Employed	0.73	0.59	0.74	0.59	0.71	0.65
Log hourly wage	3.11	3.53	3.02	3.15	3.20	3.11
Hours worked	46.3	45.4	48.9	47.3	49.2	48.0
Self employment	0.24	0.29	0.21	0.26	0.21	0.24
Entrepeneur activities	0.03	0.05	0.03	0.04	0.03	0.04
Health insurance			0.58	0.37	0.57	0.32
Schooling if working	7.15	7.31	8.62	8.13	9.65	8.31

Notes: The sample is restricted to individuals who are 20-59 years old. Married includes marriages without civil contracts. Persons are unemployed if they have searched for a job within the last week. Our calculation of wages excludes unknown or invalid wages. Hourly wages are in constant pesos as of June 2010 according to the Consumer Price Index of Banco de Mexico. Health Insurance is a dummy variable indicating whether the worker's current job provides him with health insurance.

Table 3. Descriptive statistics: Women

	1990		2000		2010	
	S	RM	S	RM	S	RM
N	1,790,280	2,331	2,393,221	9,105	2,936,933	23,483
Age	34.3	33	35	32.4	36.6	34.2
Married	0.73	0.77	0.70	0.71	0.67	0.73
Rural	0.25	0.2	0.21	0.24	0.21	0.27
Years of Schooling	5.9	7.27	7.7	8.64	9.09	9.37
No Education	0.17	0.07	0.08	0.03	0.06	0.02
Primary Incomplete	0.25	0.20	0.19	0.14	0.12	0.09
Primary	0.27	0.33	0.25	0.28	0.21	0.24
Secondary	0.20	0.24	0.23	0.27	0.28	0.33
High School	0.09	0.12	0.21	0.24	0.28	0.28
College	0.02	0.03	0.04	0.04	0.06	0.04
Unemployed	0.004	0.005	0.003	0.004	0.011	0.018
Employed	0.22	0.18	0.33	0.26	0.37	0.31
Log hourly wage	3.17	3.66	3.02	3.27	3.19	3.24
Hours worked	40.8	39.9	40.2	40.2	40.5	39.9
Self employment	0.13	0.19	0.20	0.25	0.23	0.31
Entrepeneur Activities	0.02	0.04	0.02	0.04	0.02	0.04
Health insurance			0.64	0.54	0.62	0.49
Schooling if Working	8.57	9.24	9.50	10.24	10.53	10.35

Notes: The sample is restricted to individuals who are 20-59 years old. Married includes marriages without civil contracts. Persons are unemployed if they have searched for a job within the last week. Our calculation of wages excludes unknown or invalid wages. Hourly wages are in constant pesos as of June 2010 according to the Consumer Price Index of Banco de Mexico. Health Insurance is a dummy variable indicating whether the worker's current job provides him with health insurance.

Table 4. Wage distributions statistics

	1990			2000			2010		
	S	RM	CF	S	RM	CF	S	RM	CF
MALE									
N	1194679	2659		1477186	15095		1599986	48671	
Mean	3.11	3.53	3.17	3.02	3.15	2.97	3.20	3.11	3.06
Var	1.10	1.45	1.05	0.78	0.83	0.64	0.65	0.59	0.51
5 per	1.41	1.75	1.56	1.78	2.04	1.89	2.11	2.11	2.11
10 per	2.09	2.41	2.21	2.07	2.24	2.11	2.33	2.33	2.29
25 per	2.56	2.88	2.66	2.45	2.58	2.47	2.69	2.66	2.64
50 per	3.10	3.45	3.12	2.92	2.98	2.87	3.11	3.02	2.98
75 per	3.66	4.20	3.69	3.49	3.56	3.35	3.61	3.43	3.42
90 per	4.34	4.98	4.34	4.18	4.44	3.97	4.25	4.05	3.91
95 per	4.82	5.42	4.82	4.59	4.96	4.37	4.65	4.59	4.34
WOMEN									
N	386434	424		677720	1973		829817	5564	
Mean	3.17	3.66	3.28	3.02	3.27	3.10	3.19	3.24	3.17
Var	0.90	1.39	0.86	0.83	1.02	0.75	0.74	0.88	0.67
5 per	1.82	2.18	2.01	1.66	1.91	1.85	1.95	1.88	2.00
10 per	2.23	2.41	2.38	1.99	2.17	2.11	2.27	2.27	2.29
25 per	2.64	2.92	2.74	2.42	2.58	2.51	2.66	2.66	2.66
50 per	3.14	3.50	3.25	2.94	3.12	3.01	3.12	3.13	3.09
75 per	3.72	4.42	3.79	3.61	3.86	3.63	3.71	3.77	3.67
90 per	4.21	5.05	4.28	4.20	4.52	4.22	4.30	4.52	4.23
95 per	4.64	5.67	4.68	4.52	4.93	4.52	4.64	5.00	4.59

Notes: N corresponds to individuals with valid hourly wages each year. The sample is restricted to individuals who are 20-59 years old. S and RM represent columns for the observed wage distributions of nonmigrants and migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as nonmigrants. The counterfactual reweighting procedure uses the full population of nonmigrants and return migrants. To estimate the reweighting factor, we use the following variables: i) eight groups of age; ii) regional dummies for North, Mexico City, NL rural, high-migration rural, low-migration rural, intermediate-migration rural and north rural; iii) dummies for each of the following levels of education: no education, primary incomplete, primary, secondary, high school and college; and iv) an indicator of an indigenous condition.

Table 5. Wage differences. Men and Women 1990-2010.

	Men			Women		
	1990	2000	2010	1990	2000	2010
Mean	0.058	-0.051	-0.134	0.108	0.076	-0.017
	[0.005]	[0.003]	[0.002]	[0.009]	[0.007]	[0.006]
Var	-0.051	-0.142	-0.143	-0.038	-0.084	-0.068
	[0.005]	[0.003]	[0.003]	[0.005]	[0.003]	[0.004]
5 per	0.152	0.105	0.000	0.194	0.186	0.049
	[0.023]	[0.000]	[0.005]	[0.016]	[0.010]	[0.006]
10 per	0.116	0.041	-0.041	0.151	0.118	0.017
	[0.013]	[0.002]	[0.000]	[0.007]	[0.006]	[0.011]
25 per	0.097	0.028	-0.049	0.107	0.097	0.009
	[0.006]	[0.000]	[0.000]	[0.008]	[0.010]	[0.006]
50 per	0.023	-0.049	-0.126	0.110	0.067	-0.031
	[0.012]	[0.003]	[0.007]	[0.008]	[0.014]	[0.010]
75 per	0.036	-0.145	-0.187	0.072	0.017	-0.037
	[0.010]	[0.007]	[0.003]	[0.005]	[0.016]	[0.013]
90 per	0.004	-0.219	-0.342	0.069	0.013	-0.069
	[0.013]	[0.009]	[0.016]	[0.012]	[0.012]	[0.008]
95 per	0.000	-0.223	-0.310	0.041	0.000	-0.049
	[0.008]	[0.011]	[0.008]	[0.019]	[0.008]	[0.006]

Notes: For each statistic, the table shows the difference between the log hourly wages of the nonmigrants and the log hourly wages of the return migrants' counterfactual distribution in Table 4. We calculate the standard errors by using bootstrap with 250 repetitions. The standard errors are in brackets.

Table 6. Return migrants and non-migrants paid as non-migrants in 2010.

	1990			2000			2010		
	S	RM	DIF	S	RM	DIF	S	RM	DIF
MEN									
Mean	3.02	3.07	0.046	3.11	3.06	-0.050	3.20	3.06	-0.134
			[0.005]			[0.002]			[0.002]
Var	0.57	0.48	-0.086	0.6	0.49	-0.118	0.65	0.51	-0.143
			[0.005]			[0.003]			[0.003]
5 per	1.97	2.11	0.134	2.04	2.11	0.069	2.11	2.11	0.000
			[0.006]			[0.002]			[0.005]
10 per	2.26	2.33	0.071	2.29	2.33	0.041	2.33	2.29	-0.041
			[0.009]			[0.000]			[0.000]
25 per	2.58	2.66	0.089	2.64	2.66	0.021	2.69	2.64	-0.049
			[0.000]			[0.002]			[0.000]
50 per	2.98	2.98	0.000	3.02	2.98	-0.041	3.11	2.98	-0.126
			[0.009]			[0.001]			[0.007]
75 per	3.39	3.39	0.000	3.49	3.39	-0.105	3.61	3.42	-0.187
			[0.005]			[0.002]			[0.003]
90 per	3.93	3.90	-0.036	4.12	3.9	-0.219	4.25	3.91	-0.342
			[0.006]			[0.008]			[0.016]
95 per	4.40	4.34	-0.065	4.52	4.3	-0.219	4.65	4.34	-0.310
			[0.018]			[0.011]			[0.008]
WOMEN									
Mean	2.93	3.05	0.122	3.07	3.12	0.048	3.19	3.17	-0.017
			[0.008]			[0.005]			[0.006]
Var	0.64	0.59	-0.053	0.69	0.63	-0.058	0.74	0.67	-0.068
			[0.007]			[0.004]			[0.004]
5 per	1.75	1.95	0.204	1.87	2	0.129	1.95	2.00	0.049
			[0.013]			[0.011]			[0.006]
10 per	2.10	2.29	0.183	2.18	2.29	0.111	2.27	2.29	0.017
			[0.010]			[0.006]			[0.011]
25 per	2.49	2.60	0.108	2.58	2.64	0.069	2.66	2.66	0.009
			[0.011]			[0.005]			[0.006]
50 per	2.85	2.97	0.123	2.98	3.02	0.041	3.12	3.09	-0.031
			[0.012]			[0.002]			[0.010]
75 per	3.34	3.43	0.090	3.54	3.56	0.020	3.71	3.67	-0.037
			[0.018]			[0.014]			[0.013]
90 per	3.90	4.05	0.153	4.16	4.18	0.020	4.30	4.23	-0.069
			[0.015]			[0.011]			[0.008]
95 per	4.34	4.44	0.092	4.52	4.52	0.000	4.64	4.59	-0.049
			[0.020]			[0.000]			[0.006]

Notes: The sample is restricted to individuals who are 20-59 years old. S and RM represent columns for the observed wage distributions of nonmigrants and migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as nonmigrants. The counterfactual reweighting procedure uses the full population of nonmigrants and return migrants. Standard errors are calculated using bootstrap with 250 repetitions. Standard errors in brackets.

Table 7. Wage distributions for subgroups. Men.

	1990			2000			2010		
	S	RM	CF	S	RM	CF	S	RM	CF
RURAL									
N	255636	597		421374	5842		563870	23748	
Mean	2.55	3.11	2.76	2.48	2.79	2.65	2.81	2.92	2.85
10 per	1.17	1.71	1.49	1.62	2.07	1.89	2.06	2.24	2.15
50 per	2.59	3.10	2.81	2.42	2.76	2.58	2.80	2.86	2.80
90 per	3.75	4.53	3.87	3.43	3.68	3.45	3.67	3.71	3.61
URBAN									
N	939043	2062		1055812	9253		1036116	24923	
Mean	3.26	3.65	3.31	3.13	3.28	3.09	3.27	3.20	3.17
10 per	2.34	2.56	2.41	2.19	2.31	2.26	2.41	2.39	2.39
50 per	3.21	3.51	3.26	2.98	3.09	2.98	3.20	3.09	3.09
90 per	4.44	5.05	4.46	4.30	4.59	4.12	4.34	4.05	4.05
HIGH									
N	232892	1209		284071	6944		318680	17362	
Mean	3.15	3.30	3.07	3.06	3.03	2.93	3.24	3.07	3.07
10 per	2.18	2.29	2.13	2.21	2.24	2.16	2.44	2.38	2.37
50 per	3.10	3.25	3.08	2.96	2.90	2.83	3.13	2.98	3.02
90 per	4.34	4.59	4.20	4.15	4.03	3.83	4.23	3.90	3.87
OTHER									
N	961787	1450		1193115	8151		1281306	31309	
Mean	3.10	3.72	3.27	3.01	3.24	3.00	3.19	3.13	3.06
10 per	2.07	2.56	2.32	2.04	2.22	2.07	2.29	2.29	2.29
50 per	3.07	3.61	3.24	2.92	3.03	2.92	3.09	3.02	2.98
90 per	4.34	5.15	4.47	4.19	4.66	4.08	4.25	4.12	3.96

Notes: The sample is restricted to male individuals who are 20-59 years old. S and RM represent columns for the observed wage distributions of nonmigrants and migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as nonmigrants. The counterfactual reweighting procedure uses the full male population of nonmigrants and return migrants.

Table 8. Robustness tests in counterfactual distribution for return migrants.

	1990			2000			2010		
	OR	REST	ALL	OR	REST	ALL	OR	REST	ALL
MEN									
Mean	3.17	3.19	3.18	2.97	2.98	2.97	3.06	3.06	3.06
Var	1.05	1.04	1.06	0.64	0.65	0.64	0.51	0.51	0.51
5 per	1.56	1.64	1.57	1.89	1.89	1.89	2.11	2.08	2.10
10 per	2.21	2.23	2.22	2.11	2.11	2.11	2.29	2.29	2.29
25 per	2.66	2.69	2.67	2.47	2.47	2.47	2.64	2.64	2.64
50 per	3.12	3.16	3.14	2.87	2.89	2.87	2.98	2.98	2.98
75 per	3.69	3.72	3.72	3.35	3.39	3.35	3.42	3.42	3.41
90 per	4.34	4.35	4.36	3.97	4.01	3.99	3.91	3.91	3.90
95 per	4.82	4.84	4.86	4.37	4.42	4.41	4.34	4.34	4.32
WOMEN									
Mean	3.28	3.26	3.29	3.10	3.11	3.12	3.17	3.16	3.17
Var	0.86	0.91	0.88	0.75	0.76	0.77	0.67	0.67	0.67
5 per	2.01	1.93	2.00	1.85	1.86	1.86	2.00	1.99	2.00
10 per	2.38	2.34	2.37	2.11	2.11	2.11	2.29	2.29	2.29
25 per	2.74	2.72	2.76	2.52	2.54	2.53	2.66	2.66	2.66
50 per	3.25	3.25	3.28	3.01	3.02	3.03	3.09	3.05	3.08
75 per	3.79	3.79	3.82	3.63	3.65	3.68	3.67	3.65	3.67
90 per	4.28	4.31	4.31	4.22	4.25	4.26	4.23	4.23	4.23
95 per	4.68	4.71	4.71	4.52	4.55	4.55	4.59	4.59	4.59

Notes: OR represents the counterfactual estimates for return migrants in Table 4. RES follows the same procedure except that restricts the reweighting procedure to individuals with valid hourly wage. In ALL we include more variables when we estimate the reweighting factor in the full sample: interactions between years of schooling and each of the six geographical areas, interactions between the years of schooling and each of the age groups, interactions between the age and the six geographical areas, and interactions between the rural sector and each educational level.

Table 9. Robustness test: Men 20-34.

	1990			2000			2010		
	S	RM	CF	S	RM	CF	S	RM	CF
N	656020	1698		770578	10039		747004	28845	
Mean	3.08	3.46	3.12	2.94	3.09	2.91	3.10	3.05	3.00
10 per	2.17	2.41	2.23	2.07	2.22	2.11	2.33	2.33	2.29
50 per	3.06	3.40	3.10	2.86	2.92	2.82	3.02	2.98	2.96
90 per	4.21	4.81	4.21	3.96	4.25	3.83	4.04	3.90	3.78

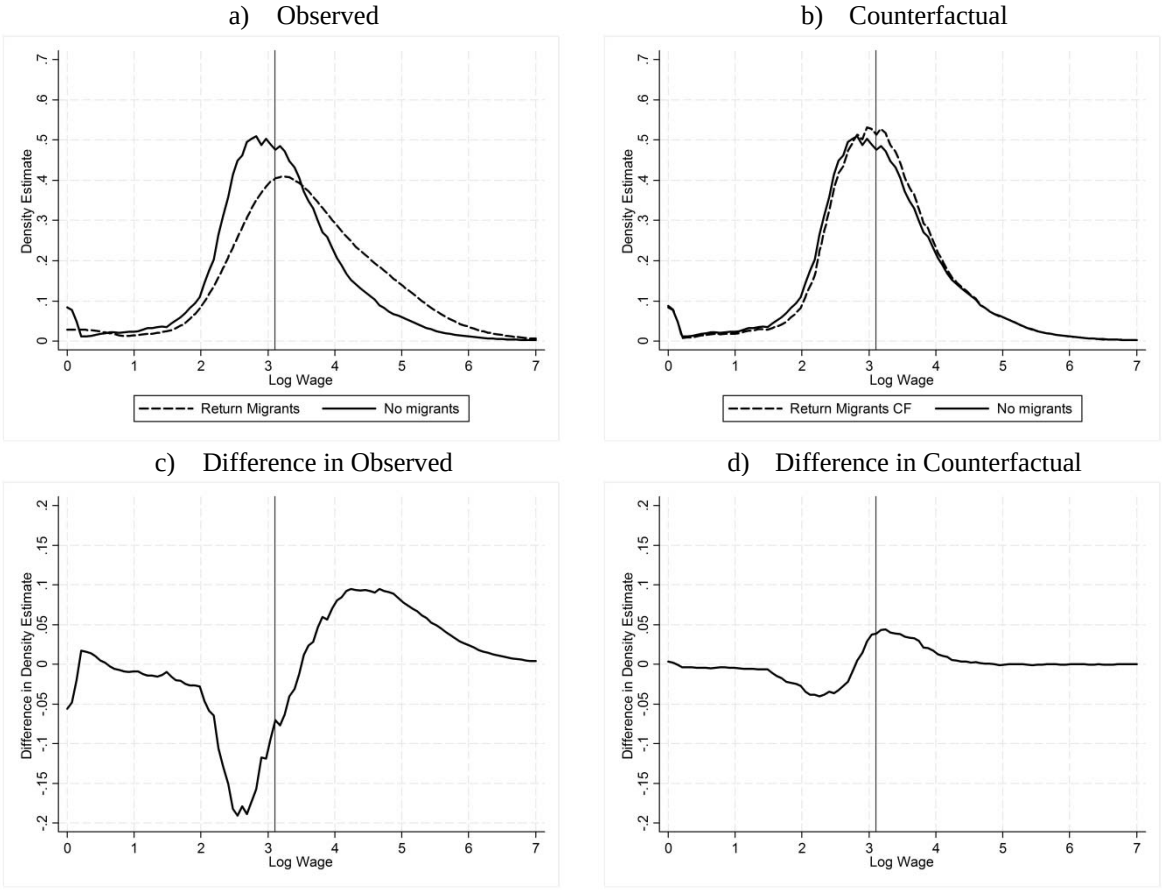
Notes: The sample is restricted to male individuals who are 20-34 years old. S and RM represent columns for the observed wage distributions of nonmigrants and migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as nonmigrants. The counterfactual reweighting procedure uses the male population of nonmigrants and return migrants who are 20-34 years old.

Table 10. Robustness test: Different datasets.

	ENADID			EMS			ENOE		
	S	RM	CF	S	RM	CF	S	RM	CF
MALE									
N	23231	415		3767	585		221648	1048	
Mean	2.97	3.04	2.93	2.66	2.70	2.63	-0.25	-0.56	-0.44
Var	0.62	0.80	0.52	0.69	0.59	0.59	0.59	0.51	0.51
10 per	2.09	2.15	2.12	1.60	1.73	1.68	-1.04	-1.32	-1.17
50 per	2.89	3.00	2.87	2.71	2.66	2.64	-0.29	-0.54	-0.43
90 per	3.98	4.14	3.83	3.63	3.69	3.52	0.70	0.22	0.33

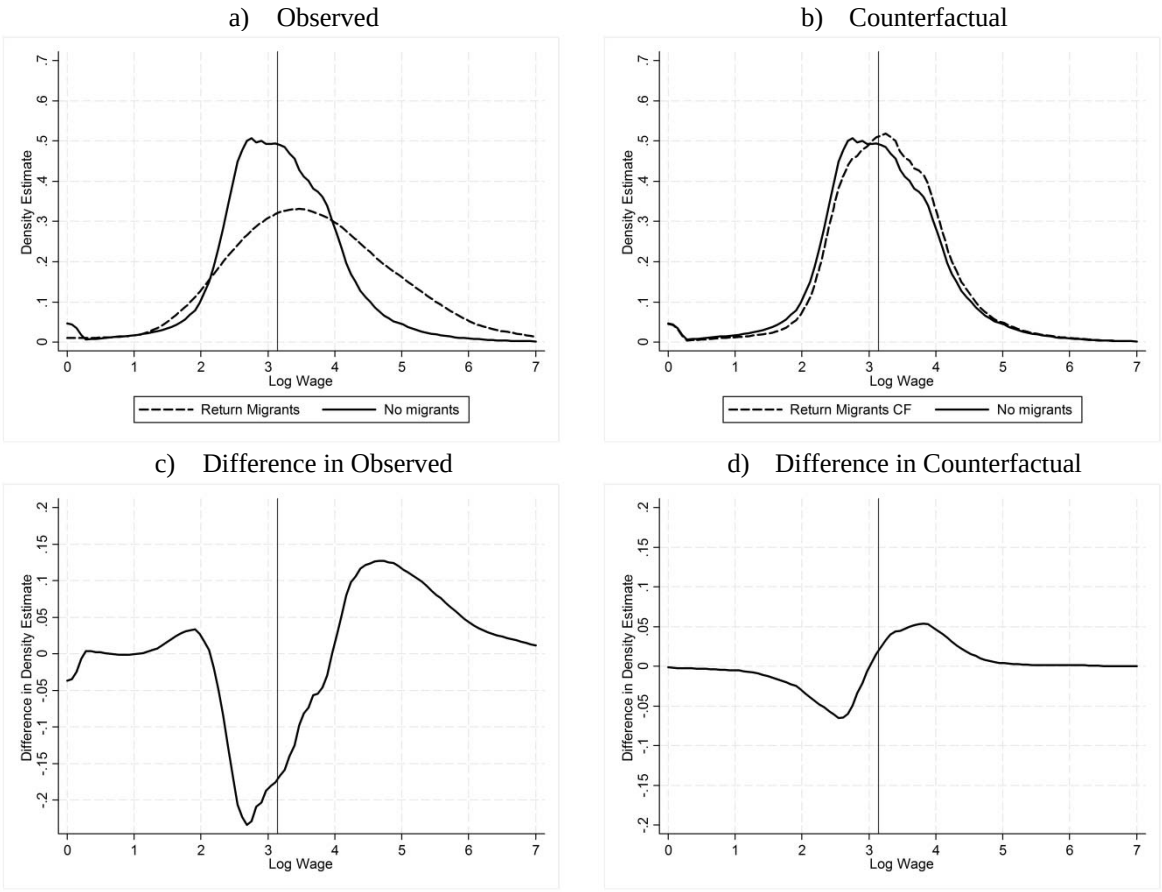
Notes: ENADID: National Survey of Demographic Dynamics, conducted in 2006. EMS: Social Mobility Survey, conducted in 2006. ENOE: National Survey of Employment and Occupation. We use the surveys from the third quarter of 2005 to the second quarter of 2007. We restrict the sample to men between 20 and 59 years old, except in the given EMS for which data were only available for men between 25 and 59 years old. S and RM represent the columns for the observed wage distributions of the nonmigrants and the migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as nonmigrants. The counterfactual reweighting procedure uses the male population of the nonmigrants and the return migrants. The variables used to calculate the reweighting factor are the same as those in Table 4, except in the case of ENOE in which the indigenous variable is not available.

Figure 1. 1990 men.



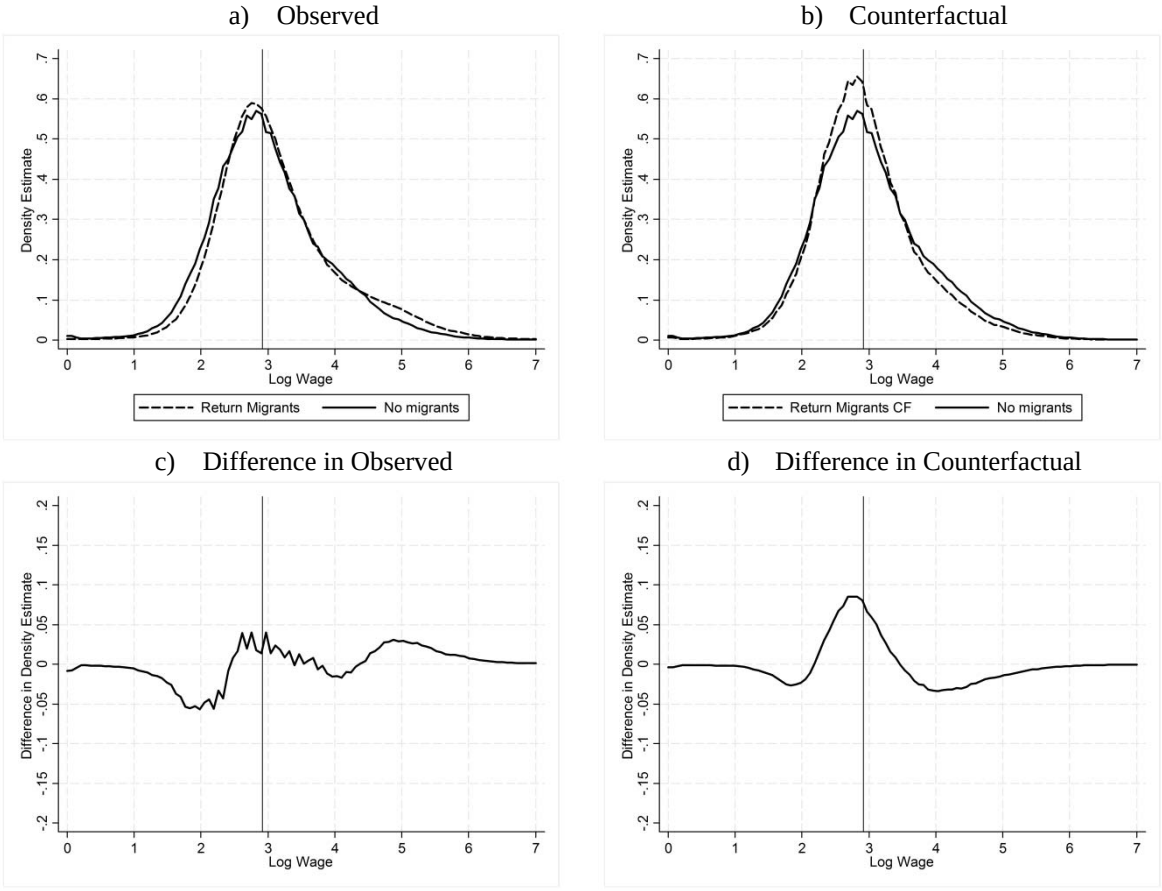
Notes: The sample is restricted to male individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

Figure 2. 1990 women.



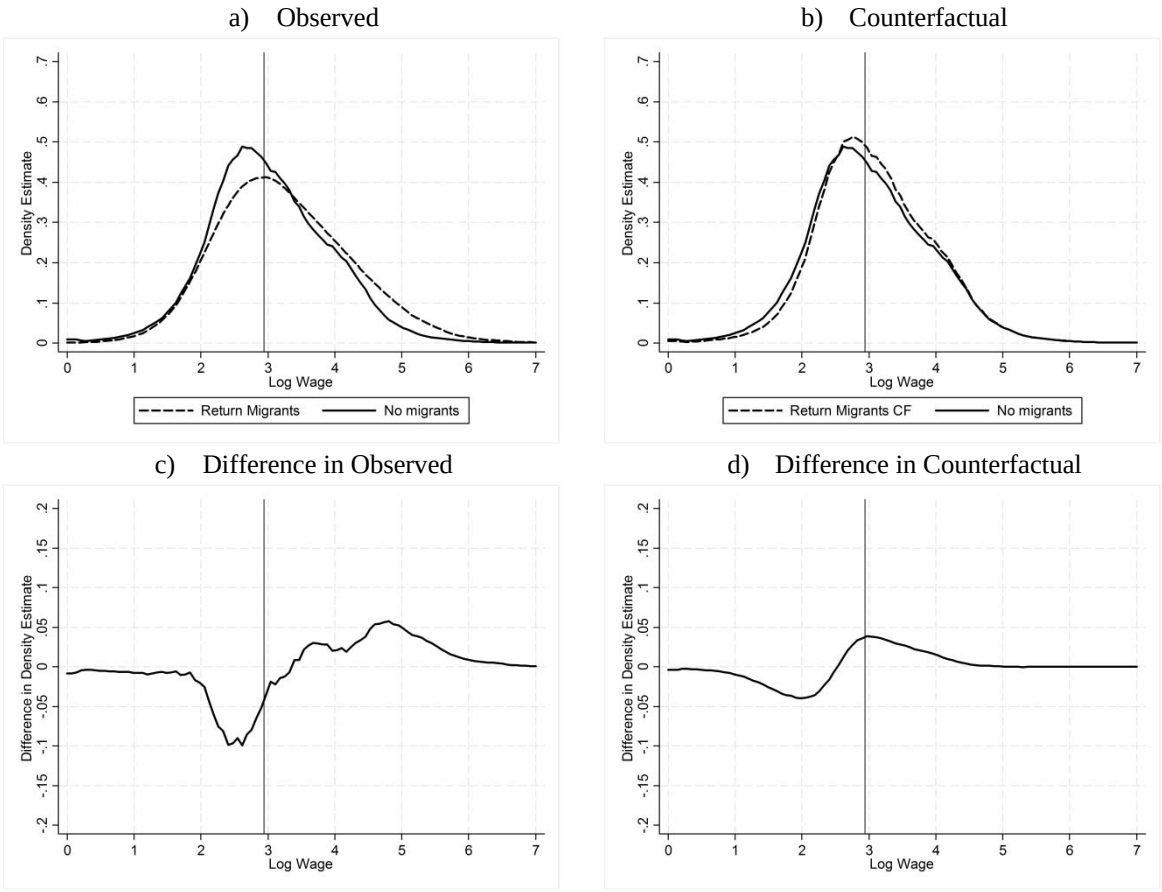
Notes: The sample is restricted to female individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

Figure 3. 2000 men.



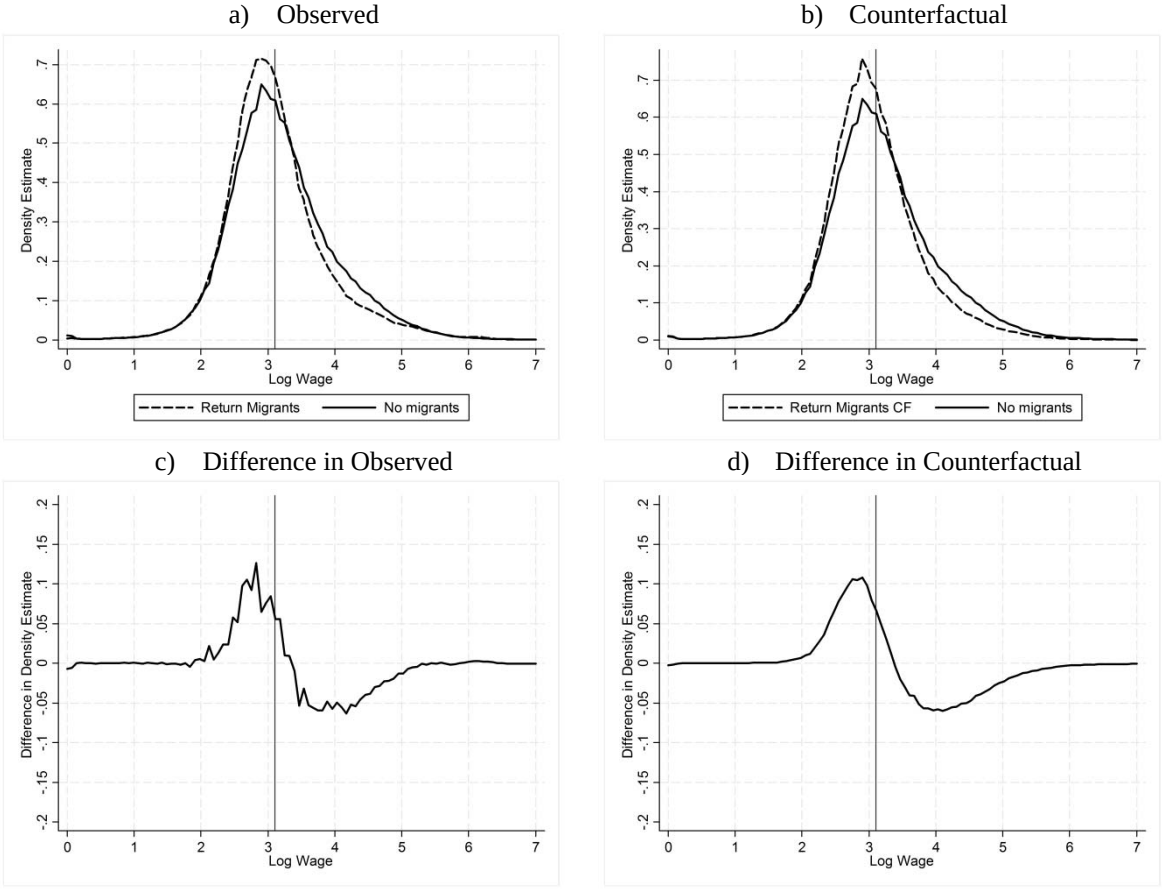
Notes: The sample is restricted to male individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

Figure 4. 2000 women.



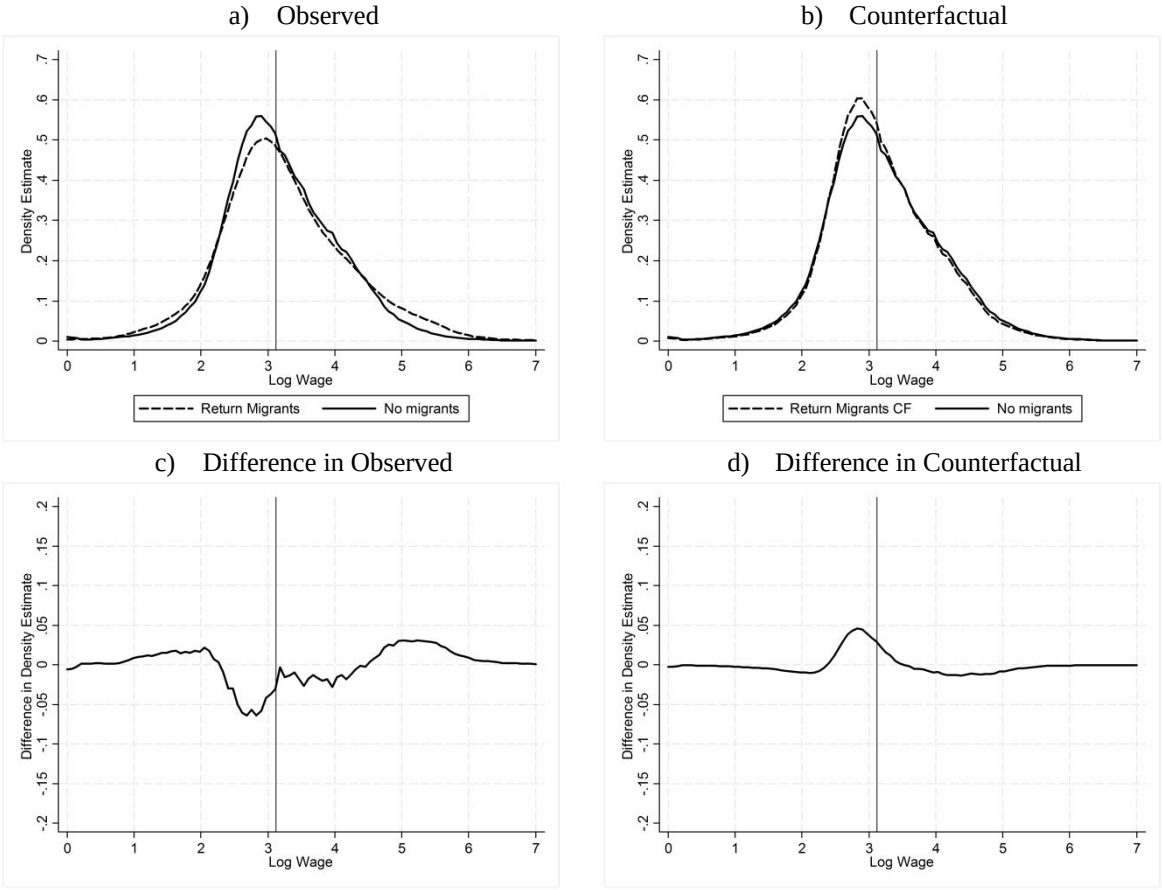
Notes: The sample is restricted to female individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

Figure 5. 2010 men.



Notes: The sample is restricted to male individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

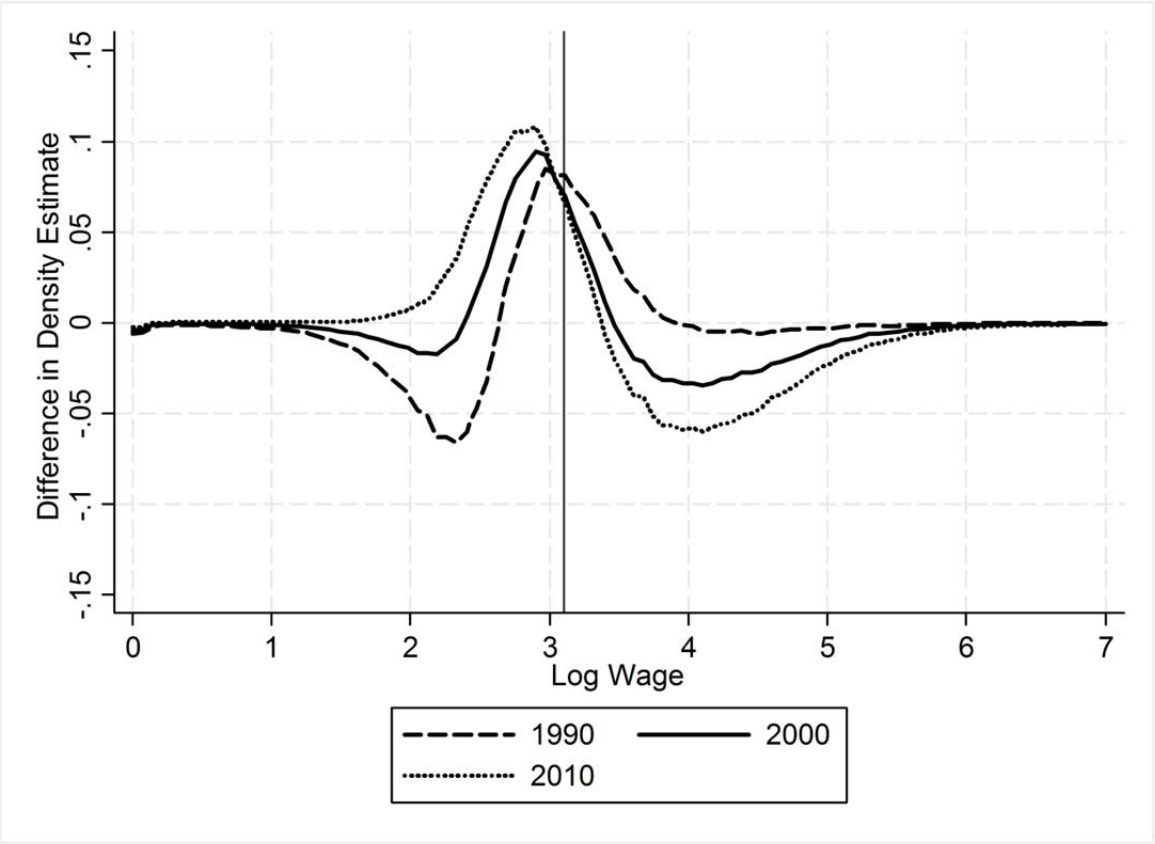
Figure 6. 2010 women



Notes: The sample is restricted to female individuals between 20 and 59 years old. Panel a plots the observed wage distributions; Panel b plots the counterfactual distribution of wages that the return migrants would have earned had they never migrated; Panel c plots the difference in the observed wage distributions; and Panel d plots the difference between the counterfactual distribution and the observed distribution of the nonmigrants. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the nonmigrants' distribution for the characteristics of the return migrants.

Figure 7. Men

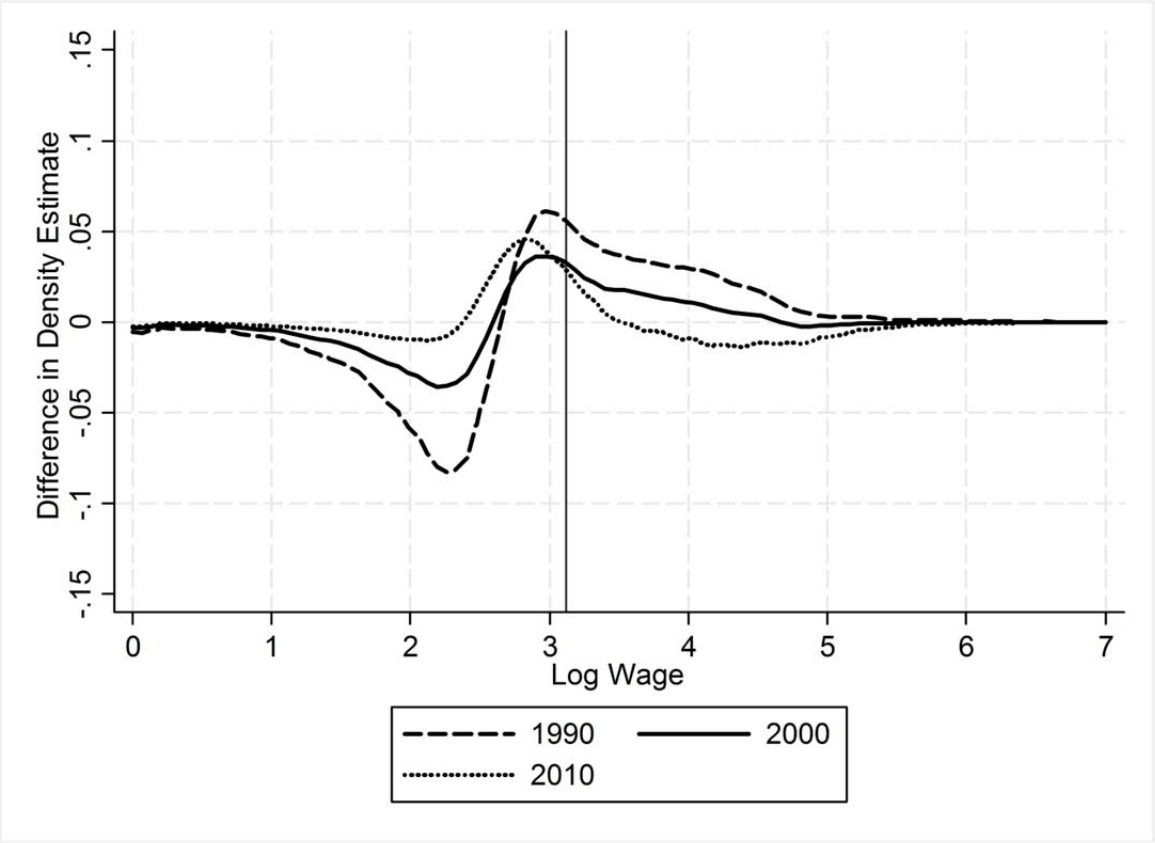
Differences between return migrants counterfactual and non-migrants wage distributions with 2010 skill prices.



Notes: The sample is restricted to male individuals between 20 and 59 years old. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the non-migrants wage distribution fixing the wage structure of non-migrants to the 2010 level with the characteristics of return migrants.

Figure 8. Women

Differences between return migrants counterfactual and non-migrants wage distributions with 2010 skill prices.



Notes: The sample is restricted to female individuals between 20 and 59 years old. All nonparametric distributions use the Epanechnikov kernel. The counterfactual reweights the non-migrants wage distribution fixing the wage structure of non-migrants to the 2010 level with the characteristics of return migrants.