

ORIGINAL ARTICLE

Open Access

# Self-selection patterns among return migrants: Mexico 1990-2010

Raymundo M Campos-Vazquez\* and Jaime Lara

\* Correspondence: [rmcampos@colmex.mx](mailto:rmcampos@colmex.mx)  
Centro de Estudios Económicos,  
Camino al Ajusco 20, Col. Pedregal  
de Santa Teresa, 10740, México, DF,  
México

**Abstract:** This paper analyzes the self-selection patterns among Mexican return migrants during the period 1990–2010. 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 stayers. We find evidence that the selection patterns in observable skills change over time from positive selection in 1990 toward negative selection in 2010. Additionally, we observe that the wages of return migrants are larger than those that the migrants would have obtained had they not migrated.

**Keywords:** Mexican Migration, Self-Selection, Return Migration, Wages

**JEL Codes:** F22, J01, J61, O54

## 1. Introduction

International migration is not always a permanent decision. Some migrants return to their countries of origin after a period 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.<sup>1</sup> In this article, we investigate the self-selection patterns among the return migrants in Mexico.

We focus on two important questions regarding Mexican return migrants: 1) Is selection stable over time? 2) Are there gender differences in selectivity patterns? This is the first paper focusing in changes in selection among return migrants over time. Hence, this paper contributes to the understanding of the self-selection patterns among return migrants in the long-run. Also, we show that women face different selection patterns than men. This is an important result that has not been addressed in previous literature on Mexican return migrants. 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 or if they overcome credit constraints, then return migration has a positive effect on the source economy.

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. (1996) method and reweight the wage distribution of the non-migrant population in Mexico (stayers) such that the distribution of observable characteristics between return migrants and stayers is as similar as possible.<sup>2</sup> This is a similar empirical strategy employed by Coulon and Piracha (2005) for return migrants in Albania.

Using the 1990, 2000 and 2010 population censuses, we find that the self-selection patterns among return migrants have changed over time. In 1990, average counterfactual wages of return migrants were higher than average wages of stayers for both men and women. In other words, 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 men became negative and the women were drawn more from the middle of the distribution. Both men and women show negative selection on average by 2010. Moreover, migration allows those who return to obtain higher wages because the increase in human capital or savings can be invested on productive activities. There is a wage premium associated with migration and return. As in the case of selection, the wage premiums to migrate and return have worsened over time for both men and women. Similar patterns hold when restricting to geographical subgroups and with other robustness tests. Our identification strategy only uncovers selection in observed variables (i.e. age, education, geographical variables). Following Reinhold and Thom (2012), we find suggestive evidence that this wage premium is in part due to the migration experience and not to positive selection in unobserved ability 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 present the results. We discuss some implications of our findings in section 6. Finally, in section 7, 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, as the case of Mexico and the United States. Due to the importance of Mexican immigration in the United States, a surge in empirical research has attempted to corroborate whether the implications of Borjas (1987) model hold.<sup>3</sup>

However, 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 lower 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.<sup>4</sup> One implication of this framework is that return migrants exhibit the same type of selection than migrants who settle in the destination country. Then, the type of selection of migrants can be assessed using information of return migrants.<sup>5</sup>

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. We closely follow the contribution of Coulon and Piracha (2005). Rooth and Saarela (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 performance of returning migrants in the Finnish market.

For the case of Mexican return migrants, there is a growing literature. Using the 2002 and 2005 rounds of the Mexican Family Life Survey (MxFLS), Ambrosini and Peri (2012) find mildly positive selection among the return migrants. Because of the sample size of return migrants (only 56 observations) and the short time period, making strong inferences is difficult. Orrenius and Zavodny (2005) find positive selection in education among return migrants using the Mexican Migration Project dataset (MMP). However, this sample is not nationally representative, and only provides insight of rural regions with high migration. Other contributions show that return migrants tend to be more positive selected than permanent migrants on unobservables, but they are similar in observable characteristics (Biavaschi 2012) and that return migration does not affect the probability of employment once return has occurred (Gitter et al. 2008).

The role of unobservables in the selection of return migrants with respect to population staying in Mexico has not reached a consensus in the literature. Ambrosini and Peri (2012) find no significant evidence of selection in unobservable characteristics. Lacuesta (2010) finds a wage premium for return migrants, which is due to positive selection in unobservable characteristics. However, Reinhold and Thom (2012) show that part of the wage premium of return migrants could be explained by migration experience and not by unobserved variables. Hence, future research should address whether unobservable characteristics play an important role for return migration results. In this paper, we focus only on selection on observable characteristics.

Our main contributions are three-fold. First, we show that the degree of selection in observable skills among return migrants has changed over time (i.e., from positive selection in 1990 to negative selection in 2010). Second, our results imply that women face different patterns of self-selection, less negative on average than men. The previous contributions are important given scant evidence of selection and gender differences over

time. Finally, we find that there is a wage premium for migrating from and returning to the country of origin. Additionally, our results indicate that the wage premium produced by migration has also evolved over time with a decreasing pattern from 1990 to 2010.

### 3. Empirical strategy

Most of the Mexican immigrant studies on self-selection patterns have attempted to assess the robustness of Chiquiar and Hanson (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 et al. (1996). The same methodology has also been applied in the return migration literature by Coulon and Piracha (2005).

#### 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 stayers with the observable characteristics of the return migrant population. Then we compare this counterfactual distribution with the observed distribution of the stayers' 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  $\Omega$ ,  $f^s$  as the density function of the stayers ( $s$  denotes stayers), 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 counterfactual distribution of the wages that the return migrants would have experienced had they been paid according to the wage structure of the stayers can be expressed as:

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

where, by Bayes' rule,  $\psi(z) = \frac{f(z|I=m)}{f(z|I=s)} = \frac{f(I=m|z)f(I=s)}{f(I=s|z)f(I=m)}$ . Hence, we only need to know  $\psi(z)$  and reweight the wage distribution for the stayers to obtain the counterfactual distribution of the wages that the return migrants would have obtained had they never migrated.

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. Although previous literature has failed in reaching a consensus on the role of unobservable characteristics, in our results section we discuss why this bias could be small on average.

After obtaining the counterfactual distribution, we compare it with the wage distribution of the stayers and obtain the type of selection. In other words, we nonparametrically characterize the wage distributions to obtain the following:

$$f_m^s(w) - f^s(w) = \int_{z \in \Omega} (\psi(z) - 1) f^s(w|z) f(z|I=s) dz \quad (2)$$

A positive difference indicates that a greater proportion of return migrants rather than stayers at wage  $w$ . If the difference is negative, then the proportion of return

migrants is lower at wage  $w$ . For negative selection, we must note a positive difference in low wages and a negative difference in high wages. On the other hand, if the selection is positive, we must observe a negative difference in low wages and a positive difference in high wages.

#### 4. Data and descriptive statistics

We use the Mexican Population Census for the years 1990, 2000 and 2010.<sup>6</sup> 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 (this variable is available only in 2000 and 2010). We restrict our sample to the individuals born in Mexico who were between 20 and 59 years old. 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 stayers by design. In addition to the information about migration, the census includes important socio-demographic data. We use the following variables: sex, education, age, indigenous language, income from employment, hours worked and geographical location.

To estimate the wage distribution, we only use individuals who reported a positive hourly wage. However, it is important to emphasize that the reweighting procedure uses the full population of return migrants and stayers, not just the working population.<sup>7</sup> Additionally, we consider observations with more than 100 hours worked or with hourly wages less than 1 Mexican Peso (close to 0.09 USD) as missing wages.

##### 4.1. Descriptive statistics

Table 1<sup>8</sup> 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.

In order to construct our geographic variables, we classified six groups of states in accordance with their migration rates in 1950s.<sup>9</sup> 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, 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 1950s: 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, Sonora and Tamaulipas.

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

**Table 1 Descriptive statistics: Full population**

	1990		2000		2010	
	S	RM	S	RM	S	RM
N	3,433,584	6,868	4,536,879	38,320	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	0.75	0.47	0.76
Rural	0.25	0.28	0.22	0.32	0.20	0.36
High Migration Region	0.21	0.50	0.20	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
High Migration x Rural	0.06	0.20	0.05	0.19	0.05	0.15
Low Migration x Rural	0.06	0.00	0.06	0.01	0.06	0.05
North Region x Rural	0.02	0.03	0.02	0.02	0.02	0.02
Intermediate Migration x 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	7.90	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: S denotes stayers and RM return migrants. 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, Tlaxcala and Sinaloa; iv) north region: Baja California, Baja California Sur, Chihuahua, Coahuila, Sonora and Tamaulipas.. Proportions do not add 1 because Mexico City and Nuevo Leon are omitted. 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.

proportion of stayers 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 stayer population was stable across all regions in the three censuses.

Another characteristic of the geographical location patterns of the return migrants is 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 stayers decreased from 25 percent in 1990 to 20 percent in 2010.

Locations of the return migrants in the rural sector of different regions 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 fewer return migrants than stayers. However, the proportion of the indigenous population being return migrants increased from 2 to 4 percent,

whereas the proportion of indigenous people among the stayer population has decreased from 8 in 1990 to 7 percent in 2010.

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

In terms of educational groups descriptive statistics 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 stayers in the same level. In addition, the proportion of return migrants with incomplete primary schooling is similar to that of stayers. Nevertheless, whereas in 1990 the proportion of individuals in secondary and higher educational groups was similar for both return migrants and stayers, by 2010 the proportion of stayers 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

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 the existence of any gender differences. Table 2 shows the main characteristics with differences among men and women. 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 stayers. Conversely, the male return migrants exhibit less education than the stayer population. Positive selection in terms of education is disappearing among the women and becoming more negative in the case of the men.

In sum, the descriptive statistics suggest that the selection in terms of education and wages are becoming more negative for both the men and the women. However, women

**Table 2 Descriptive statistics: Men and Women**

	1990		2000		2010	
	S	RM	S	RM	S	RM
<b>Men</b>						
N	1,643,304	4,537	2,143,446	29,165	2,584,619	85,208
Rural	0.26	0.32	0.22	0.35	0.20	0.39
Years of Schooling	6.87	6.76	8.30	7.65	9.46	8.12
Log hourly wage	3.18	3.62	3.02	3.15	3.20	3.11
<b>Women</b>						
N	1,790,280	2,331	2,393,433	9,155	2,936,933	23,483
Rural	0.25	0.20	0.21	0.24	0.21	0.27
Years of Schooling	5.90	7.27	7.70	8.63	9.09	9.37
Log hourly wage	3.22	3.72	3.03	3.27	3.20	3.25

Notes: S denotes stayers and RM return migrants. The sample is restricted to individuals who are 20–59 years old. 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.

tend to be more positive selected. In our analysis we focus on men and 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. Then we estimate selection patterns using equation (2). To estimate  $f(I = s|z)$  and  $f(I = m|z)$ , 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 dummy 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, northern states and the states of Mexico City and Nuevo Leon, rural, high-migration rural, low-migration rural, intermediate-migration rural and north rural. To include education, we used dummies for each aforementioned levels of education. In addition, we used an indigenous language dichotomous variable. Also, in order to absorb a possible correlation with the unobservable component of wages, we include interactions of years of schooling with regions and age groups, interactions of age with each region, and interactions of rural communities with each level of education. In total we use 46 variables in the estimation. Using the logit estimates, we obtained the weight  $\psi(z) = \frac{f(I=m|z)f(I=s)}{f(I=s|z)f(I=m)}$  and constructed the counterfactual distribution of the wages that the return migrants would have obtained had they never migrated, as indicated in equation (1), by using kernel methods.<sup>11</sup> Also, with this weight we can construct descriptive statistics of the counterfactual distribution.

### 5.1. Selection over time

Table 3 contains statistics of the difference between the counterfactual wage distributions of return migrants and the observed wage distribution for stayers each year (equation 2). For men, in 1990 the counterfactual wage of return migrants was higher than for the stayers by 8 percent approximately. Each of the percentiles also showed a positive difference indicating that the counterfactual distribution of return migrants was to the right of stayers, especially in low percentiles. The standard errors were obtained bootstrapping the procedure with 250 repetitions and imply that all differences are statistically significant. In 2000, average selection was negative in 5 percent, with the counterfactual distribution of return migrants at the right of stayers at lower percentiles, but at the left in higher percentiles. In 2010, the average selection had worsened approximately to minus 14 percent, with almost every percentile of counterfactual wages of return migrants at the left of stayers, except the lowest 5 percentile.

Table 4 shows observed wages and counterfactual wages for subgroups in the population. For men, we note that the counterfactual wages are lower than the observed wages after return (columns CF and RM in each year). This means that there is a wage premium for return migration compared to the option of never migrating. By 1990, the difference between observed wages after return and counterfactual wages was of 36 log points (43%), changing over time to 18 log points (20%) and to 5 log points in 2010. Then the wage premium of migrants has reduced over time.



**Table 3 Wage differences: Men and Women 1990–2010**

	Men			Women		
	1990	2000	2010	1990	2000	2010
Mean	0.078 [0.006]	-0.050 [0.004]	-0.142 [0.003]	0.125 [0.010]	0.087 [0.008]	-0.021 [0.006]
Var	-0.058 [0.005]	-0.129 [0.004]	-0.141 [0.003]	-0.031 [0.006]	-0.060 [0.005]	-0.055 [0.004]
5 per	0.206 [0.011]	0.089 [0.006]	0.000 [0.000]	0.182 [0.007]	0.221 [0.010]	0.059 [0.006]
10 per	0.143 [0.008]	0.040 [0.002]	-0.041 [0.001]	0.105 [0.003]	0.118 [0.010]	0.000 [0.011]
25 per	0.105 [0.000]	0.028 [0.001]	-0.049 [0.000]	0.121 [0.014]	0.107 [0.010]	0.000 [0.006]
50 per	0.069 [0.005]	-0.049 [0.004]	-0.128 [0.003]	0.121 [0.015]	0.080 [0.016]	-0.031 [0.010]
75 per	0.045 [0.007]	-0.142 [0.009]	-0.193 [0.010]	0.105 [0.014]	0.065 [0.015]	-0.041 [0.013]
90 per	0.036 [0.018]	-0.198 [0.015]	-0.357 [0.008]	0.095 [0.019]	0.054 [0.017]	-0.069 [0.007]
95 per	0.025 [0.009]	-0.187 [0.011]	-0.321 [0.015]	0.072 [0.017]	0.028 [0.014]	-0.049 [0.006]

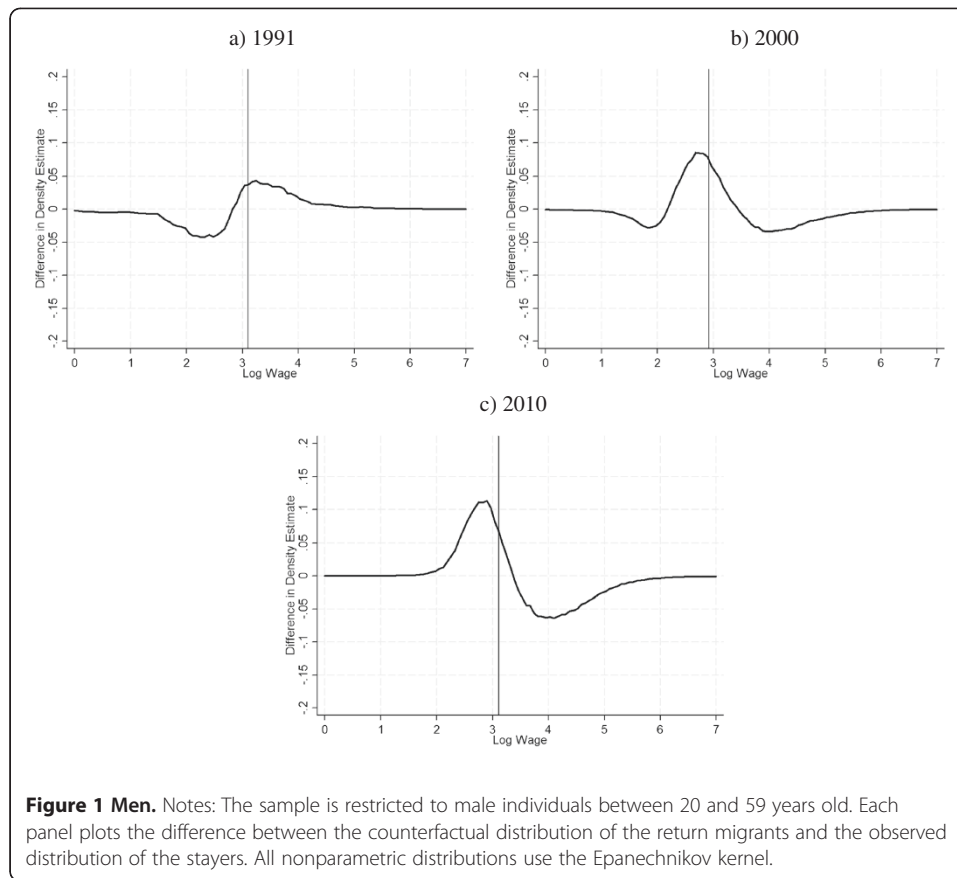
Notes: For each statistic, the table shows the difference between the log hourly wages of the return migrants' counterfactual distribution and the log hourly wages of the stayers (equation 2). We calculate the standard errors by using bootstrap with 250 repetitions. The standard errors are in brackets.

Figure 1 shows selection over time for men. The vertical line is the median of the distribution of the logarithm of wages for stayers. In 1990, the counterfactual distribution of return migrants shows a higher density in wages located to the right of the median wages of stayers. This is evidence of positive selection. However, by 2000 return migrants are located in greater proportion in wages close to the median of stayers, with

**Table 4 Wage distributions for subgroups**

	1990			2000			2010		
	S	RM	CF	S	RM	CF	S	RM	CF
<b>Men</b>	3.18	3.62	3.26	3.02	3.15	2.97	3.20	3.11	3.06
Rural	2.67	3.25	2.88	2.50	2.80	2.66	2.83	2.92	2.86
Urban	3.31	3.73	3.38	3.13	3.28	3.10	3.28	3.21	3.16
High Migration	3.23	3.40	3.16	3.07	3.03	2.93	3.24	3.07	3.07
Other States	3.17	3.81	3.35	3.01	3.24	3.01	3.20	3.14	3.06
<b>Women</b>	3.22	3.72	3.34	3.03	3.27	3.12	3.20	3.25	3.18
Rural	2.85	3.33	3.08	2.52	2.82	2.77	2.82	2.96	2.91
Urban	3.25	3.76	3.36	3.08	3.31	3.16	3.24	3.30	3.23
High Migration	3.17	3.28	3.19	3.02	3.01	2.99	3.18	3.15	3.13
Other States	3.23	3.92	3.43	3.03	3.39	3.18	3.20	3.29	3.20

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 stayers and return migrants, respectively. CF is the counterfactual distribution of the wages that the return migrants would have earned had they been paid as stayers. The counterfactual reweighting procedure uses the full population of stayers and return migrants.



less presence in low and high wages. By 2010, selection of return migrants is clearly negative. In sum, selection of male return migrants has worsened over time.

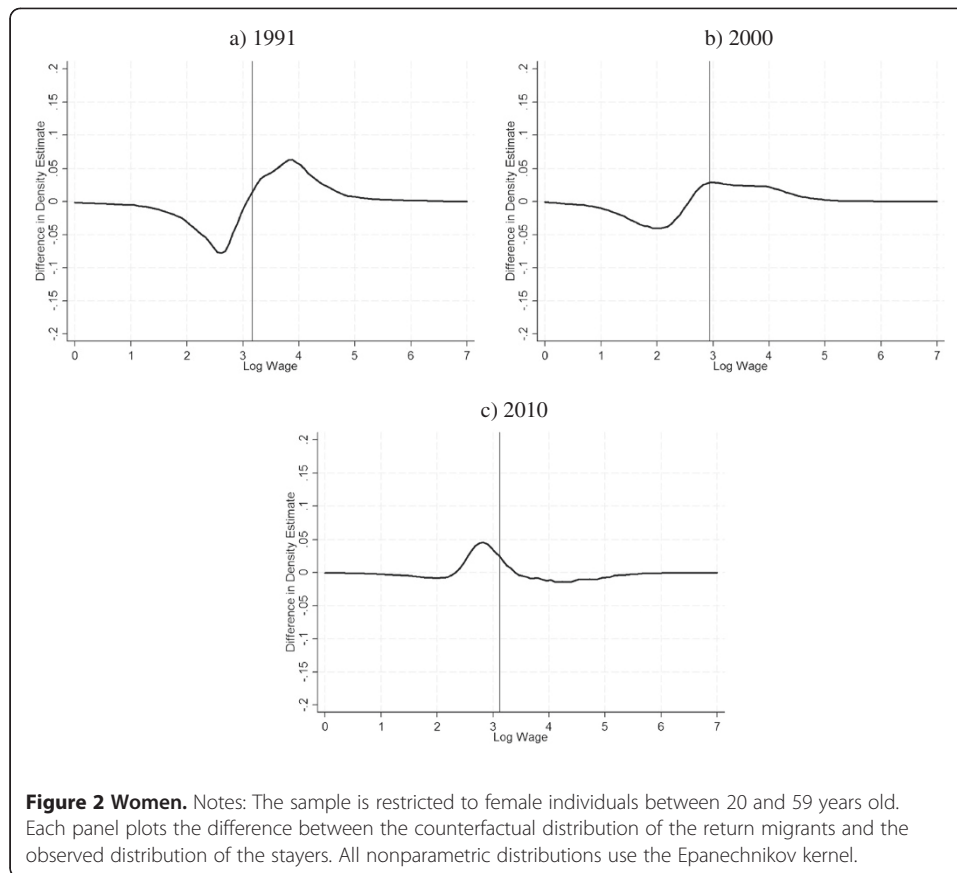
### 5.2. Differences between men and women

Table 3 shows that average selection was more positive for women than for men. In 1990, positive selection of women was approximately 12 percent on average, while men were 8 percent. In 2000, average selection for women remains positive and men were negatively selected on average. By 2010, the wage that female return migrants would have earned had they not migrated was higher than the wage of female stayers. For men, that was only true at lower wages. In 2010, when men exhibit a high degree of negative selection, women are only slightly negatively selected in approximately 2 percent. These differences in the average degree of selection are statistically significant given the low standard errors.

Figure 2 shows the selection results for women. Differences between men and women do not preclude some similarities. Both men and women show a worsening of selection over time. Nonetheless, selection is more negative for men than for women. Also, although the magnitude of the wage premium is similar for both men and women, both have a wage premium that is decreasing over time as shown in Table 4.

### 5.3. Extensions

Previous studies on self-selection among permanent migrants have shown that the type and degree of selection tends to differ when 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 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 return migrants over time.

Table 4 shows the results of the wage differences between the return migrants and the stayers in the urban and rural sectors. 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. As in the general case, women tend to be more positively selected than men in rural and urban sectors. 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 wage premium has fallen over time in the rural and urban sectors.

To investigate the effect of migration networks, we divide the Mexican states between those with high migration, as defined previously, and the rest. 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 (2010) proposal is correct, then we should

find more evidence of negative selection in states with a tradition of high migration rates. Table 4 shows the results.

The states with high migration rates show a greater degree of negative selection each year for both men and women. This finding is consistent with McKenzie and Rapoport (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, which is consistent with a higher level of negative selection.

#### **5.4. The role of unobservable characteristics**

A possible bias in the estimation is the role of unobserved characteristics. Then, the difference between wages after return and counterfactual wages would be a combination of the impact of return migration and the impact of selection on unobserved variables. Our methodology does not allow us to know the possible bias caused by this weakness. Moreover, return migrants face a double selection process.<sup>12</sup> First, they self-select from the population together with permanent migrants. In a second stage, return migrants self-select from all Mexican migrants in the United States. In both moments variables not included in vector  $z$  could be playing a role.

The evidence of selection on unobservable characteristics among return migrants is ambiguous. Ambrosini and Peri (2012) show that negative selection on unobservable variables among permanent migrants decreases and is close to zero in the case of return migrants.<sup>13</sup> For selection among return migrants, Lacuesta (2010) uses a measure of the duration in the United States to assess whether the wage premium comes mainly from the high value of migration experience and the possibility of overcoming liquidity constraints or if it is due to selection on unobservables. Using the 2000 Census, Lacuesta (2010) finds that the time spent in the US does not increase wages of return migrants and interprets this finding as evidence of positive selection on unobservables.

However, Reinhold and Thom (2012) point out that the causal interpretation of the duration parameter in a simple OLS wage equation is only possible if unobserved ability is not correlated with the optimal length in the United States. A likely bias occurs when the optimal duration is lower for migrants with higher unobserved ability. Hence, the coefficient on duration in a simple OLS wage equation is a lower bound of the true effect. Then, including variables correlated with unobserved ability should move the coefficient of duration in the right direction. Using the Mexican Migration Project dataset, Reinhold and Thom (2012) find that the measure of time spent in the United States positively affects wages and that becomes slightly more positive adding variables correlated with unobserved ability. They claim little role of unobservable characteristics.

We performed a similar test to Reinhold and Thom (2012) with different datasets.<sup>14</sup> Using the 2000 and 2010 Census, we find similar results to Lacuesta (2010). However, experience in the United States is poorly measured in the Census because it only covers migration within the last five years. A database with a better measurement of duration in the United States is the 2006 Social Mobility Survey (EMS). That survey asked individuals if they had traveled to the United States to work for a month or more at least once in their lifetime.<sup>15</sup> Using this dataset we find results close to Reinhold and Thom (2012) in high migration states. Then, although we cannot rule out completely the role

of unobservable characteristics, part of our wage premium is due to the migration experience itself,<sup>16,17</sup>

## 6. Discussion

Self-selection among the return migrants tends to become negative over time. The same is true for both men and 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.<sup>18</sup> 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:

$$\psi_{m90}^{s10} = \frac{f(I = m, y = 1990|z)f(I = s, y = 2010)}{f(I = s, y = 2010|z)f(I = m, y = 1990)} \quad (3)$$

With this weight, we adjust the characteristics of the stayers in 2010 according to the characteristics of the return migrants in 1990.<sup>19</sup> Using this weight over the distribution of the stayers in 2010 generates the counterfactual distribution of the wages that the return migrants in 1990 would have earned had they been paid as stayers in 2010. For the stayers in 1990, we calculate the following:

$$\psi_{s90}^{s10} = \frac{f(I = s, y = 1990|z)f(I = s, y = 2010)}{f(I = s, y = 2010|z)f(I = s, y = 1990)} \quad (4)$$

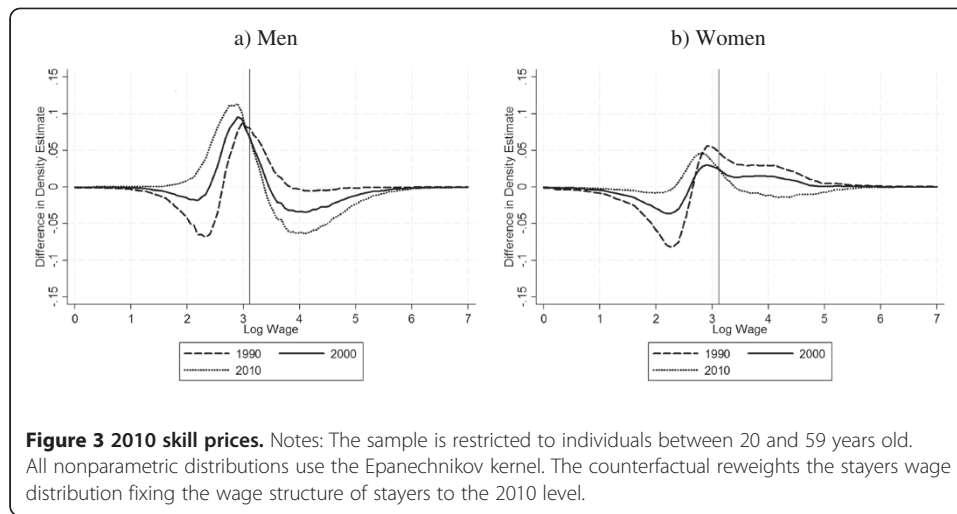
Applying this weight over the distribution of the stayers in 2010 generates the counterfactual distribution of the wages that the stayers in 1990 would have earned had they been paid as stayers in 2010. Using equations (3) and (4), we can nonparametrically estimate the degree of selection for the return migrants in 1990 by evaluating this estimate in terms of the skill prices in 2010:

$$f_{m90}^{s10}(w) - f_{s90}^{s10}(w) = \int_{z \in \Omega} (\psi_{m90}^{s10} - \psi_{s90}^{s10}) f^{s10}(w, z) dz \quad (5)$$

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

Figure 3 shows the difference between the return migrants' wage distribution and the stayers' wage distribution when their characteristics are priced as stayers in 2010 for the men and the women, respectively. The vertical line shows the median log wages in 2010 for the stayers. For men, Figure 3 shows a pattern of positive selection in 1990, intermediate selection in 2000 and negative selection in 2010. For women, Figure 3 shows positive selection in 1990 and 2000 but intermediate selection in 2010. These are the same patterns of selection obtained in the previous section. Then, the growing degree of negative selection is not due to a change in the returns of stayers characteristics.

The previous literature on migration between Mexico and the United States has pointed out the importance of the structure of costs and restrictions on credit to understand patterns of selection among Mexican migrants (Chiquiar and Hanson 2005; Orrenius and Zavodny 2005). If costs to migrate represent a higher proportion of income for the low skilled population or if this population cannot access to credit, the theoretical prediction of negative selection among Mexican migrants in the United States can be modified. Our



findings 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). In states with a long migration tradition, the self-selection patterns are negative except in 1990 for women, 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. Also, negative selection tends to increase as time goes on, which suggest that the development of migration networks produces a structure of costs close to models in Borjas (1987) and Bratsberg and Borjas (1996) in the Mexican case.

However, both economies have been exposed to shocks that could shape the self-selection patterns. These shocks may affect the degree of selection. In 1995, Mexican economy suffered a drastic fall in the GDP, also affecting the exchange rate. In 2008, the United States economy began a recession with negative impacts in Mexico. In the period covered by our data the enforcement of the United States migratory laws has changed from opportunities of legalization under Immigration Reform and Control Act of 1986 (IRCA) to a tightening in the border control in the last years. Our contribution is to show that self-selection patterns are moving towards higher negative selection. Hence, incorporating economic shocks and immigrant regulation in both countries is needed in future research.

More striking are differences in the selection patterns across gender. One possible explanation is that enforcement in migratory laws or border tightening has different effects on migration costs of men and women. Previous findings show that women migrate at older ages than men, and this is consistent with family reunification purposes (Campos-Vazquez and Sobarzo 2012). Border tightening may in fact decrease the age of first migration, increasing the degree of negative selection in the migrant population in the US and in the return migrant population in Mexico (due to less schooling acquired in the source country). How these factors affect self-selection differences between men and women is a question that must be addressed.

## 7. Conclusions

In this article, we analyzed the self-selection patterns of return migrants in Mexico using the censuses for the years 1990, 2000 and 2010. In particular, we followed

DiNardo et al. (1996) and Coulon and Piracha (2005) 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) but has not been utilized to analyze the selection of return migration into Mexico. We presented evidence suggesting that the self-selection patterns in observable characteristics among the Mexican return migrants have changed over time (i.e., from positive selection in 1990 to negative selection in 2010). The wages that the male return migrants would have earned had they not migrated is 8 percent larger than the wages of the male stayers. However, by 2010, this difference had declined to -14 percent. Among females, average selection changed from 13 percent approximately in 1990 to -2 percent in 2010. The growing negativity of the degree of selection is robust to the analysis of specific subgroups: rural and urban, and states with high migration rates and low migration rates.

Important differences exist among the different subgroups. Women tend to show more positive selection than men. For both men and women 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 on observables (Chiquiar and Hanson 2005). 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 observable skills between return migrants and permanent migrants could be small.

According to Bratsberg and Borjas (1996) model, in countries such as Mexico, where payments to observable human capital are more unequal than those in the United States, return migrants should be selected negatively with respect to the stayer population. This hypothesis only holds clearly for the men in 2010 and is clearly rejected for the 1990 census. The changes in selection are consistent with a decline in costs produced by migrant networks, as proposed by McKenzie and Rapoport (2010). However, other possibilities as changes in enforcement of US migration laws or shocks in both economies have to be assessed. 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 for migrating from and returning to the country of origin. Suggestive evidence indicates that this premium comes in part from experience in the United States and is not due to unobservable characteristics. 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. Also, this effect has diminished over time. Hence, further research is necessary to understand this effect and 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 and also improve the gains of return migrants in the source economy.

### 7.1. Endnotes

<sup>1</sup>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.

<sup>2</sup>We use the term “stayer” to refer to the population residing in Mexico with no migration experience to the United States in the previous 5 years.

<sup>3</sup>Chiquiar and Hanson (2005), Ibararan and Lubotsky (2007), Fernandez-Huertas (2011), McKenzie and Rapoport (2010), Kaestner and Malamud (2010) and Ambrosini and Peri (2012)

<sup>4</sup>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.

<sup>5</sup>For other reasons of return migration see Dustmann (2003), Dustmann and Weiss (2007) and Gibson and McKenzie (2011).

<sup>6</sup>The 10 percent samples are available through the Mexican Statistical Office website (<http://www.inegi.org.mx/>).

<sup>7</sup>The estimation calculates counterfactual wages as if return migrants participate in the labor market as stayers. Using the working population instead of the full population provides similar results.

<sup>8</sup>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.

<sup>9</sup>We follow a classification close to Hanson (2007).

<sup>10</sup>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.

<sup>11</sup>We use an Epanechnikov kernel. For the bandwidth, we employ Silverman (1986) optimal bandwidth multiplied by two.

<sup>12</sup>We thank an anonymous referee for this remark.

<sup>13</sup>This is consistent with Biavaschi (2012) who shows that return migrants are positively selected on unobservables from Mexican migrants in the United States.

<sup>14</sup>Results are available upon request.

<sup>15</sup>The survey was designed by Centro de Estudios Espinosa Yglesias, a civil association funded by Fundacion Espinosa Rugarcia (<http://www.ceey.org.mx>).

<sup>16</sup>Other sources of possible bias were analyzed. Geographic variables of the state of birth instead of state of residence causes minor changes in counterfactuals densities. Including only population with valid wages or restricting the age to only population between 20 and 34 years old do not affect the main results.

<sup>17</sup>An important question to ask is whether the main results hold when we use different datasets. When we look at different datasets with information about return migrants and apply the same methodology used in Census the findings are consistent with our previous results in selection. The EMS dataset, where we can fully separate return migrants from stayers beyond five years and keep only migrants having finished their education before migration, shows results similar to 1990 and 2000 Census. The



2006 National Survey of Demographic Dynamics (ENADID) covers information of return migrants in the same way that 2000 Census for return migrants between 2001 and 2006; using this dataset we find the same self-selection patterns than 2000 Census. Only when we use the National Survey of Employment and Occupation (ENOE), covering information about return migrants between 2005 and 2007, we find evidence of negative selection in higher degree than 2010 Census. Another interesting result using ENOE is that the wage premium of migration does not exist in the quarters following the date of return. This last survey is better suited to measure the characteristics of the flow of migrants who return to an existing household. These results are available upon request.

<sup>18</sup> Chiquiar and Hanson (2005), p. 264.

<sup>19</sup> Derivation of  $\psi_{m90}^{s10}$  is similar to  $\psi(z)$ . In the notation, instead of the wage distribution of stayers ( $I = s$ ) we use the wage distribution of stayers in 2010 ( $I = s, y = 2010$ ) and instead of the wage distribution of return migrants ( $I = m$ ) we use the wage distribution of return migrants in 1990 ( $I = m, y = 1990$ ). In both cases, since  $f(I = m)/f(I = s)$  and  $f(I = m, y = 1990)/f(I = s, y = 2010)$  are constants, assuming that are equal to 1 does not affect the reweighting procedure.

#### Competing interests

The "IZA Journal of Migration" is committed to the IZA Guiding Principles of Research Integrity. The authors declare that they have observed these principles.

#### Acknowledgements

We thank three anonymous referees and the editor for valuable comments. All remaining errors are our own.

Responsible editor: Corrado Giuliatti

Received: 27 September 2012 Accepted: 10 November 2012

Published: 24 December 2012

#### References

- Ambrosini W, Peri G (2012) The Determinants and the Selection of Mexico-US Migrants. *The World Economy* 35(2):111–151
- Biasvaschi C (2012) Recovering the Counterfactual Distribution with Selective Return Migration, IZA DP No. 6795
- Borjas G (1987) Self-Selection and the Earnings of Immigrants. *Am Econ Rev* 77(4):531–553
- Bratsberg B, Borjas G (1996) Who Leaves? The Outmigration of the Foreign-Born. *Rev Econ Stat* 78(1):165–176
- Campos-Vazquez R, Sobarzo H (2012) The Development and Fiscal Effects of Emigration on Mexico. Migration Policy Institute. April, Washington DC
- Chiquiar D, Hanson G (2005) International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States. *J Political Econ* 113(2):239–281
- Chiswick B (1999) Are Immigrants Favorably Self-Selected? *Am Econ Rev* 89(2):181–185
- Coulon A, Piracha M (2005) Self-selection and the Performance of Return Migrants: The Source Country Perspective. *J Population Econ* 18(4):779–807
- DiNardo J, Fortin M, Lemieux T (1996) Labor Market Institutions and the Distribution of Wages. *Econometrica* 64(5):1001–1044
- Dustmann C (2003) Children and Return Migration. *J Population Econ* 16(4):815–830
- Dustmann C, Kirchkamp O (2002) The optimal migration duration and activity choice after re-migration. *J Develop Econ* 67(2):351–372
- Dustmann C, Weiss Y (2007) Return Migration: Theory and Empirical Evidence for the UK. *British J Industrial Econ* 45(2):236–256
- Fernandez-Huertas J (2011) New Evidence on Emigrant Selection. *Rev Econ Stat* 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. *J Development Econ* 95(1):18–29
- Gitter S, Gitter R, Southgate D (2008) The Impact of Return Migration to Mexico. *Estudios Económicos* 23(1):3–23
- Hanson G (2007) Emigration, Labor Supply, and Earnings in Mexico. In: Borjas G (ed) *Mexican Migration to the United States*, Chapter 9th edn. University of Chicago Press and National Bureau of Economic Research, pp 289–328, <http://ideas.repec.org/h/nbr/nberch/0097.html>
- Ibarraran P, Lubotsky D (2007) Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census. In: Borjas G (ed) *Mexican Migration to the United States*, Chapter 5th edn. University of Chicago Press and National Bureau of Economic Research, pp 159–191, <http://www.nber.org/books/borj06-1>, <http://www.nber.org/chapters/c0103>
- Kaestner R, Malamud O (2010) Self-Selection and International Migration: New Evidence from Mexico. In: NBER Working Paper 15765

- Lacuesta A (2010) A Revision of the Self-selection of Migrants Using Returning Migrant's Earnings. *Annals of Econ Stat* 97/98:235–259
- McKenzie D, Rapoport H (2010) Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks. *Rev Econ Stat* 92(4):811–821
- Orrenius P, Zavodny M (2005) Self-selection among undocumented immigrants from Mexico. *J Dev Econ* 78(1):215–240
- Reinhold S, Thom K (2012) Migration Experience and Earnings in the Mexican Labor Market. Working Paper, <https://files.nyu.edu/kt44/public/MigExperienceAndEarnings.pdf>
- Rooth D-O, Saarela J (2007) Selection in Migration and Return migration: Evidence from Micro Data. *Econ Lett* 94(1):90–95
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London

doi:10.1186/2193-9039-1-8

**Cite this article as:** Campos-Vazquez and Lara: Self-selection patterns among return migrants: Mexico 1990-2010. *IZA Journal of Migration* 2012 **1**:8.

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- ▶ Convenient online submission
- ▶ Rigorous peer review
- ▶ Immediate publication on acceptance
- ▶ Open access: articles freely available online
- ▶ High visibility within the field
- ▶ Retaining the copyright to your article

---

Submit your next manuscript at ▶ [springeropen.com](http://springeropen.com)

---