fedea

Fundación de
Estudios de
Economía Aplicada

Understanding Different Migrant Selection Patterns in Rural and Urban Mexico

by

Jesús Fernández-Huertas Moraga*

Documento de Trabajo 2013-02

January 2013

** FEDEA and IAE, CSIC.

Los Documentos de Trabajo se distribuyen gratuitamente a las Universidades e Instituciones de Investigación que lo solicitan. No obstante están disponibles en texto completo a través de Internet: *http://www.fedea.es.*
These Working Paper are distributed free of charge to University Department and other Research Centres. They are also available through Internet: *http://www.fedea.es.*
ISSN:1696-750X

Jorge Juan, 46
28001 Madrid -España
Tel.: +34 914 359 020
Fax: +34 915 779 575
infpub@fedea.es

N.I.F. G-78044393

Understanding Different Migrant Selection Patterns in Rural and Urban Mexico*

Jesús Fernández-Huertas Moraga†
FEDEA and IAE (CSIC)

January 9, 2013

Abstract

The productive characteristics of migrating individuals, emigrant selection, affect welfare. The empirical estimation of the degree of selection suffers from a lack of complete and nationally representative data. This paper uses a dataset that addresses both issues: the ENET (Mexican Labor Survey), which identifies emigrants right before they leave and allows a direct comparison to non-migrants. This dataset presents a relevant dichotomy: it shows negative selection for urban Mexican emigrants to the United States for the period 2000-2004 together with positive selection in Mexican emigration out of rural Mexico to the United States in the same period. Three theories that could explain this dichotomy are tested. Whereas higher skill prices in Mexico than in the US are enough to explain half of the negative selection result in urban

*An earlier version of this paper was circulated as "Wealth constraints, skill prices or networks: what determines emigrant selection?" I have received financial support from the ECO2008-04785 project, funded by the Spanish Ministry for Science and Innovation. I am thankful to Ronald Findlay, Eric Verhoogen and David Weinstein for their help and support. This paper has also benefited from useful comments and suggestions from Donald Davis, Timothy Hatton, Rosella Nicolini, Kiki Pop-Eleches, Nikos Theodoropoulos, two anonymous referees and seminar participants at Columbia University, the Second IZA Migration Meeting in Bonn, the 2008 Workshop on the Labor Market Effects of Immigration in Seville and the 2008 Migration and Development Conference in Lille. I would also like to thank Ognjen Obucina for his thorough research assistance. Of course, remaining errors are only mine.

†FEDEA, Jorge Juan, 46, 28001 Madrid, Spain. E-mail: jfernandezhuertas@fedea.es

Mexico, its combination with network effects and wealth constraints fully account for positive selection in rural Mexico.

Keywords: international migration, selection, wealth constraints, household survey
JEL Classification Numbers: F22, O15, J61, D33

1 Introduction

The goal of this paper is to explain why the pattern of emigrant selection varies in rural and urban Mexico. Fernández-Huertas Moraga (2011) shows that emigrants from Mexico to the United States earn an average wage before migrating lower than the average wage of those who decide to stay home. This is what Borjas (1999) defines as negative selection. However, Fernández-Huertas Moraga (2011) also shows that positive selection exists in rural Mexico, where rural Mexico is formed by those who live in localities with 2,500 inhabitants or less.[1]

The literature offers three main arguments that could explain these facts. This paper examines the relative merits of these three competing arguments. It must be noted though that they are neither exclusive nor exhaustive. Previous papers (see below) had already shown the qualitative validity of the three arguments in different frameworks and with distinct datasets. The contribution of this paper is to assess both their qualitative and their quantitative relevance in a common framework and with the same dataset: the Encuesta Nacional de Empleo Trimestral (ENET), Mexico's Labor Force Survey.[2]

The first argument is developed by Borjas (1987), who disregards the role of migration costs. If the return to skill were to be lower in rural Mexico than in the United States

[1]Whether positive or negative selection prevails in Mexico is not a settled question. Chiquiar and Hanson (2005), Lacuesta (2010) and Mishra (2007) argue for intermediate to positive selection in Mexico as a whole whereas Ibarrarán and Lubotsky (2007) report negative selection. Cuecuecha (2005) and Caponi (2010) obtain mixed results. McKenzie and Rapoport (2007) and Orrenius and Zavodny (2005) find positive selection in rural Mexico. See Hanson (2006) and Fernández-Huertas Moraga (2011) for a complete review of these results. More recent papers using the Mexican Family Life Survey, such as Ambrosini and Peri (2012) or Kaestner and Malamud (2012), obtain results in line with Fernández-Huertas Moraga (2011).

[2]This is the dataset Fernández-Huertas Moraga (2011) uses to study emigrant selection. He discusses its main advantages and disadvantages. A relevant concern is the attrition rate in the panel: 11 percent after one quarter and 26 percent after one year. Though large, these figures are comparable to the attrition rates of commonly used datasets, such as the US CPS, whose attrition rate is 20-30 percent after one year (Neumark and Kawaguchi, 2004).

whereas it were to be higher in urban Mexico, then we should expect positive selection out of rural Mexico and negative selection out of urban Mexico.

The second explanation comes from McKenzie and Rapoport (2010). They propose that the existence of different selection patterns in different migrant datasets can be reconciled by the existence of migration networks. Migration networks reduce migration costs so that emigrants out of areas with larger migration networks tend to be more negatively selected than emigrants out of areas with smaller migration networks. Thus, this could explain the different selection patterns in rural and urban Mexico if migration networks were more present in urban than in rural areas.

Finally, a third argument, also from McKenzie and Rapoport (2007) among others in a different setup, is related to the existence of wealth constraints affecting the migration decision. Even in the presence of higher returns to migration for low skill individuals relative to high skill individuals in rural Mexico, which would lead to negative selection, it could happen that these low skill individuals cannot cover migration costs by borrowing, thus resulting in positive selection of migrants.

Out of these three arguments, the first one is independent from the structure of migration costs since Borjas (1987) considers them constant across skill groups. On the contrary, the networks and wealth constraints arguments are fundamentally based in the structure of migration costs. The true relationship between migration costs and skill levels is not only relevant to study why migrant selectivity evolves in one way or another but also to understand the consequences of different migration policies.[3]

One reason why migration costs can be decreasing in skills is through the positive relationship between these skills and wealth (McKenzie and Rapoport, 2007), which can then be combined with the existence of wealth constraints in migration. This paper tackles this theory by regressing, using semi-parametric analysis to account for non-linearities, the decision to migrate on a household wealth index extracted from the ENET. The results indicate that the probability of emigration is increasing in wealth for low wealth individuals and decreasing in wealth for high wealth individuals in rural Mexico (individuals living in localities with less than 2,500 inhabitants), consistent with the existence of wealth constraints and with the findings in McKenzie and Rapoport (2007) for the Mexican Migration Project[4] database.

[3]Borger (2010) provides an excellent example.

[4]The Mexican Migration Project, developed by Princeton University and the University of Guadalajara, surveys communities in Mexico. For more information, see http://mmp.opr.princeton.edu/. Also, see

However, the result for urban Mexico is that there is no relationship between wealth and the emigration probability. This could explain why there is positive selection in rural Mexico whereas there is negative selection of emigrants from Mexican urban areas.

As for the ability of skill prices to account for the different selection patterns, simple Mincer regressions are used first to show that the return to education in rural Mexico does not seem to be low enough to generate positive selection of emigrants to the United States. This finding is confirmed by the fact that observable skills account for as much of the observed degree of selection in urban Mexico as in rural Mexico. In order to estimate wages based on observable skills, the counterfactual wage density estimation procedure developed by DiNardo, Fortin, and Lemieux (1996) and applied by Chiquiar and Hanson (2005) is used.

Finally, network effects, as defined by McKenzie and Rapoport (2010), are shown to be more relevant in shaping migration decisions in rural Mexico. When networks are added as an additional observable variable to the DiNardo, Fortin, and Lemieux (1996) counterfactual wage estimation, all of the observed degree of positive selection in rural Mexico can be accounted for. When networks and wealth are jointly considered, much more than the observed degree of positive selection in rural Mexico is accounted for, implying a degree of negative selection in unobservables similar to that in urban Mexico.

In a cross-country setting, Belot and Hatton (2012) similarly show that a combination of the Roy model (Roy, 1951) in log utility terms, as in Borjas (1987), with poverty constraints can explain selection patterns to 29 OECD countries in 2000-2001. However, Grogger and Hanson (2011) and Rosenzweig (2007) question the usefulness of the Borjas (1987) log utility interpretation of the Roy model and argue instead for using a linear utility model to study selection. The contribution of this paper to this ongoing debate is to show a case where both models can be distinguished. The existence of positive selection in rural Mexico is coherent with both models once the log utility model is corrected to allow for wealth constraints but the existence of negative selection in urban Mexico is only compatible with the log utility model.[5]

The structure of the paper follows. First, the simple theory underlying this study is sketched. Second, a description of the ENET dataset and several stylized facts are presented.

Fernández-Huertas Moraga (2011) for a comparison of the ENET and MMP datasets.

[5] To be fair, the linear utility model could also be modified in its structure of migration costs in order to accommodate the possibility of negative selection. However, one would need to find something in the structure of migration costs that differs between urban and rural Mexico.

The following section explores how well different theories are able to explain the opposed selection patterns in rural and urban Mexico. Finally, the main conclusions of the paper are drawn.

2 Emigrant Selection Theory

This section reviews three simple variations to the classical selection framework derived by Borjas (1987) from the combination of the Roy (1951) selection model and the Sjaastad (1962) idea that migration is an investment decision in which individuals make the utility maximizing choice out of a set of alternatives. These variations offer explanations to the fact that emigrant selection patterns differ in rural and urban Mexico.

Following Borjas (1999), positive selection is defined as a situation in which:[6]

$$E\left(\log w_0 | \text{emigration}\right) > E \log\left(w_0 | \text{no emigration}\right)$$

where w_0 represents the wage level in the original location (rural or urban Mexico in this case).

Positive selection implies that emigrants are on average more productive (as reflected on their wage) than non-migrants. The above inequality can be easily computed from the ENET data for the Mexico-US case since both the wages of non-migrants and migrants right before migration can be observed. In addition, the difference between the two expectations can be interpreted as the degree of selection (DS):

$$DS \equiv E\left(\log w_0 | \text{emigration}\right) - E\left(\log w_0 | \text{no emigration}\right)$$

2.1 The differential returns to skill explanation

First, following Borjas (1987) and his simpler exposition in Borjas (1999), consider the case where migration costs, in time equivalent units are constant across skill levels so that emigrant selection is determined by the differences in returns to skills among competing destinations. Suppose that individuals maximize utility on a period by period basis and

[6] The definition in Borjas (1999) also includes that the earnings of immigrants will be higher than those of natives in the host country as long as the base average wage both groups have access to is the same.

that their decisions for each period do not affect their outcome in subsequent periods.[7] Utility consists of their log wage income net of time equivalent migration costs. Of course, migration costs are not incurred if the individual decides to stay home. Otherwise, there are three alternative destinations: rural Mexico ($0R$), urban Mexico ($0U$) and the United States (1). The structure of wages in each of these places is given by:

$$\log w_i = \mu_i + \delta_i x; \quad i = \{0R, 0U, 1\}$$

Individuals performance in the labor market depends on a vector of observable and unobservable characteristics summarized in the variable $x \geq 0$, whose density function over the population is $f(x)$. It can be assumed that base wages are ordered $\mu_1 > \mu_{0U} > \mu_{0R} > 0$ whereas no assumption will be made by now with respect to the returns to skill coefficients δ_1, δ_{0U} and δ_{0R}.

An income maximizing individual will migrate whenever the wage in the destination j net of migration costs ($C_{ij} > 0$) exceeds the wage at her original location i or other possible destinations. This can be expressed with the following function:

$$I^{ij}(x) \equiv \log\left(\frac{w_j}{w_i + C_{ij}}\right) \simeq \log w_j - \log w_i - \pi_{ij}$$

where $\pi_{ij} = \frac{C_{ij}}{w_i}$ are migration costs in time-equivalent units. As a result, emigrants from rural Mexico to the US will be characterized by $I^{0R1}(x) > 0$ and $I^{0R1}(x) > I^{0R0U}(x)$, and emigrants from urban Mexico to the US will satisfy $I^{0U1}(x) > 0$ and $I^{0U1}(x) > I^{0U0R}(x)$.

Suppose π_{ij} are considered constant across characteristics and also that $\pi_{ij} = \pi \ \forall i \neq j$, then the existence of positive selection in emigration from rural Mexico to the United States would imply $\delta_{0R} < \delta_1$, whereas negative selection in emigration from urban Mexico to the United States would require $\delta_{0U} > \delta_1$. Thus, the expression to be tested with the ENET dataset is:

$$\delta_{0U} > \delta_1 > \delta_{0R} \tag{1}$$

If inequality (1) is true, an additional implication is that internal migration from rural to

[7] Alternatively, think of a Mincerian world (Mincer, 1958) where wages are constant over time or, in a more sophisticated yet still simple version, where the best prediction about future wages can be obtained from current wages.

urban Mexico should be positively selected. Also, emigrants to the US should be negatively sorted with respect to internal migrants between rural and urban Mexico.[8]

2.2 The networks effect explanation

A second reason why different patterns of selection arise in rural and urban Mexico can be found in the existence of migration networks. Munshi (2003) showed that the existence of Mexican migrant networks improves the economic opportunities of Mexican migrants in the United States, thus increasing the return to emigration. On the other hand, Carrington, Detragiache, and Vishwanath (1996) or McKenzie and Rapoport (2007) among others showed that migrant networks also help reducing the costs of the migratory move. Both phenomena can be modeled as a negative relationship between network size and migration costs: $\pi(n, x)$, with $\frac{\partial \pi}{\partial n} < 0$ and $\frac{\partial \pi}{\partial x} < 0$, where n is the network size. Under these conditions and assuming also that $\delta_{0U} = \delta_{0R} = \delta_0 > \delta_1$, McKenzie and Rapoport (2010) prove two propositions:

Proposition 1 *Larger migrant networks increase migration incentives (i) at all productive characteristics (x) levels, and (ii) more so at low x levels.*

Proposition 2 *With intermediate self-selection, where the support of x is $[0, \bar{x}]$ and $x_L > 0$, $x_U < \bar{x}$, where x_L and x_U represent the minimum and maximum level of productive characteristics x at which people emigrate, (a) An increase in the migration network increases the range of lower x levels that wants to migrate more than it increases the range of higher x levels that wants to migrate. (b) Providing that $f(x)$ is not increasing in x, larger migration networks reduce average levels of x among migrants (and increase average levels of x among non-migrants), therefore increasing the likelihood and/or degree of migrants' negative self-selection.*

Again, the implications are testable with the ENET dataset. If the existence of different migrant network structures in rural and urban Mexico were to explain their different selection patterns, it should be the case that migrant networks are more present in urban than in rural

[8] The definition of sorting comes from Grogger and Hanson (2011). A migration flow to a particular destination is positively sorted when their average wage is larger than the average wage of the migration flow to an alternative destination. Here I am comparing urban Mexico and the US as alternative destinations for rural Mexico inhabitants.

Mexico. In addition, ceteris paribus, higher levels of migration networks should be correlated with higher degrees of negative selection.

With respect to internal migration, the fact that migration costs are likely to be lower than for international migration would imply that networks should play a less relevant role. This less relevant role would be translated into a lower degree of negative selection for internal migrants and to a negative sorting of migrants to the US with respect to internal migrants out of rural Mexico.

2.3 The wealth constraints explanation

Finally, a third reason why selection patterns could be so different between urban and rural Mexico is the possible existence of wealth constraints affecting the migration decision in rural but not in urban Mexico. An individual is constrained in wealth when she would be willing to migrate given her expected return to migration ($I^{ij}(x) > 0$) but she cannot afford the trip. If credit markets worked efficiently, this individual should be able to borrow in order to undertake migration. Assuming that the credit market is not very developed or simply that collateral is required in order to obtain a loan, Hanson (2006) suggests an easy way to incorporate a wealth constraint to the migration decision:

$$\gamma_i C_{ij} \leq Y$$

where γ_i represents the fraction of the loan that must be collateralized and Y denotes the wealth level of the individual. It can be assumed that this wealth level is positively related to the productive characteristics of the individual:

$$Y = \rho + \sigma x$$

where $\rho > 0$ stands for the part of wealth which is unrelated to productive characteristics and $\sigma > 0$ reflects the positive relationship between productive characteristics and wealth.

Assume again that $\delta_{0U} = \delta_{0R} = \delta_0 > \delta_1$ and further that $C_{0R1} = C_{0U1} = C$. Given this additional constraint, individuals will decide to migrate from i to j whenever the following inequalities are satisfied at the same time:

$$
\begin{aligned}
I^{ij}(x) &> 0, I^{ij}(x) \geq I^{ih}(x); \forall i \neq j, h \\
x &\geq \frac{\gamma_i C - \rho}{\sigma} \equiv x_i^{CC}
\end{aligned}
$$

Under these conditions, the degree of selection will only depend on the value of γ_i. In fact, the degree of positive selection will be increasing in γ_i since higher levels of wealth constraints imply that the minimum level of skills at which individuals start to emigrate is higher. Thus, if differential levels of wealth constraints were to explain the different patterns of emigrant selection between urban and rural Mexico, it should be the case that:

$$
\gamma_{0R} > \gamma_{0U} \tag{2}
$$

so that the degree of positive selection is higher in rural than in urban Mexico. This is another test that can be performed in the ENET.

The implication for internal migration patterns between rural and urban Mexico is again that wealth constraints should play a less relevant role there, considering that migration costs are lower than for international migration. Thus, it can be expected that the selection of internal migrants out of rural Mexico would be less positive than the selection of international migrants so that migrants to the US would be positively sorted with respect to internal migrants.

The following section reviews the ENET dataset and describes the different selection patterns found in rural and urban Mexico.

3 The ENET Dataset

The Encuesta Nacional de Empleo Trimestral (ENET) is a nationally representative household survey that was carried out quarterly by the Mexican Instituto Nacional de Geografía y Estadística (INEGI, 2005) between the second quarter of 2000 and the last quarter of 2004. This labor force survey is similar to the American CPS and it has been used in a number of different studies.[9]

The ENET has a panel structure that follows Mexican households for five consecutive quarters.[10] Every quarter, one fifth of the sample is renewed with an average attrition

[9] Robertson (2000) or Fernández-Huertas Moraga (2011) are two examples.

[10] Households are followed by going back to the same dwelling but movers are not tracked.

rate of 11 percent.[11] For the remaining four fifths, a person who is present in the quarter in which her household is observed but moves to the United States (or elsewhere) in the following quarter is considered an emigrant.[12] The characteristics of future emigrants can be compared directly to the characteristics of future non-migrants at the same point in time. Fernández-Huertas Moraga (2011) discusses some possible sources of bias in the ENET with respect to comparable data sources on Mexican migration to the US and shows that its main inconveniences, such as the omission of whole households migrating together or the inability to differentiate between first-time and repeated migrants, do not greatly affect the magnitude of the selection results. In particular, he shows his analysis can be replicated using similarly selected MMP samples. The large attrition rate in the ENET[13] does not seem to create large problems either. The selection results of the ENET have been confirmed by subsequently available datasets, such as the Mexican Family Life Survey, studied by Ambrosini and Peri (2012) or Kaestner and Malamud (2012) among others.[14]

Table 1 presents some characteristics for migrants to the US, internal migrants (defined as individuals who move to a different state within Mexico)[15] and non-migrants first in Mexico as a whole and then disaggregated for both rural and urban areas.[16]

(Table 1)

For Mexico as a whole, the table reproduces the negative selection result reflected in

[11]Attrition rates in the sample are detailed in the appendix. The results are robust to the inclusion or exclusion of quarters in which the attrition level is high. In addition, the observations that disappear from the sample are not statistically different from the observations that remain in the sample in the main observable characteristics.

[12]See the data appendix for ENET total migration numbers.

[13]Though large, it is comparable with the 20-30 percent attrition rate over a year in the CPS (Neumark and Kawaguchi, 2004).

[14]Ambrosini and Peri (2012) calculate a degree of selection of -0.23 for the Mexican Family Life Survey that can be compared with the -0.26 reported by Fernández-Huertas Moraga (2011) for the ENET.

[15]This is just a proxy for the real internal migration flow. Unfortunately, the ENET only reports the destination state for internal migrants. Thus, I am excluding individuals who may have migrated from rural to urban areas within a state and inappropriately including individuals who may have migrated to a rural area in a different state. The reason for the former exclusion is the risk of pulling together long-distance migrants with individuals who just move to a nearby town. Internal migrants across states represent 64.4 percent of overall internal migrants in the ENET.

[16]The distinction between rural and urban Mexico follows an ENET convention. The dichotomy is interesting because of the different selection patterns that characterize both populations. Appendix table A2 further disaggregates urban Mexico by locality size in as fine a division as allowed by the dataset.

Fernández-Huertas Moraga (2011). Concentrating on the working age population, Mexican male emigrants to the United States earned an average wage of 1.4 2006 US dollars per hour the quarter before they emigrated, lower than the average wage of 2.1 dollars earned by non-migrants. The same negative selection result is obtained for women. However, dividing the overall population between urban and rural Mexico, where rural Mexico refers to people living in localities with less than 2,500 inhabitants, it can be observed that the negative selection result is not homogeneous throughout the country. Rural Mexico represents 22 percent of the overall Mexican population but rural Mexican emigrants to the United States account for 45 percent of male migrants and for one third of female migrants. Thus, rural emigrants are over-represented in the total emigration flow to the US. They are also over-represented in the internal migration flow but not by as much (37 percent).

Positive selection characterizes migration flows out of rural Mexico whereas negative selection is obtained if we only look at urban Mexico. Male emigrants out of rural Mexico earn an average wage of 1.1 dollars per hour, higher than the 1 dollar per hour wage of those who do not emigrate out of rural areas. In contrast, male emigrants out of urban Mexico earn 1.6 dollars per hour, much less than the 2.3 dollars per hour usual wage obtained by those who remain behind.

Male internal migrants are in between non-migrants and US-bound emigrants with respect to wages for Mexico as a whole and in urban Mexico. However, they are noticeably below both non-migrants and US-bound emigrants in rural Mexico, where they earn the lowest average wage out of the three groups thus being negatively selected also there. For females, that is the case both in rural and urban Mexico.

In terms of other observable characteristics presented in table 1, emigrants to the US are shown to be younger than non-migrants both in rural and in urban Mexico (29 versus 35 years old) whereas the education levels are in line with the selection result in terms of wages. Whereas male emigrants out of urban Mexico tend to have 1.3 less years of education than non-migrants, male emigrants out of rural Mexico present 0.7 more years of education than non-migrants. Notice that this is not the case for internal migrants. These tend to be the youngest of the three groups (28 years old on average) but they are similarly educated to non-migrants in urban Mexico and more highly educated than both non-migrants and migrants to the US in rural Mexico. Thus, for rural internal migrants, there coexists a negative selection result in terms of wages with a positive selection result in terms of education.

Working-age women behave differently from men in Mexico as a whole and do not present relevant differences (except in levels) between rural and urban Mexico. Female emigrants to the US are negatively selected in terms of wages both in rural and in urban Mexico but they are positively selected in terms of education both in rural and urban Mexico. Female internal migrants are more positively selected than US emigrants in terms of education and more negatively selected in terms of wages. The explanation might be found in the fact that many women are tied-movers, that is, they accompany family members or travel to join them instead of moving for economic reasons so that there is a small percentage of female emigrants that actually work and earn a wage relative to men. This is the reason why what follows will focus on the behavior of male emigrants.

In addition to differences at the mean, figure 1 shows how the wage distribution of male emigrants and non-migrants reflects the negative selection result for urban Mexico and the positive selection result for rural Mexico. The wage distribution is calculated as the kernel density estimate[17] of the distribution of the logarithm of real hourly wages in 2006 dollars relative to their quarter average (to avoid time trend effects) registered for the group of migrant and non-migrant men aged 16 to 65 years old in the period going from the second quarter of 2000 to the third quarter of 2004. The wage distribution is calculated both for rural and urban Mexico. In the case of urban Mexico, it can be seen that the wage distribution of migrants lies to the left of the wage distribution of non-migrants, evidencing the existence of negative selection. The distance between the averages of both wage distributions, previously defined as the degree of selection, is -0.29 (0.02 is the standard error). For rural Mexico, both wage distribution are displaced to the left of the urban wage distributions but this time most migrant wages lie to the right of non-migrant wages, suggesting the existence of positive selection out of rural Mexico. The computed degree of positive selection is 0.18 (0.03 is the standard error).

(Figure 1)

For completeness, figure 1 also represents the wage distribution of internal migrants both out of rural and out of urban Mexico, confirming in both cases the pattern found in table 1.

[17] The estimated density is $\hat{g}(w) = \frac{1}{hN}\sum_{i=1}^{N} K\left(\frac{w-w_i}{h}\right)$ where N is the number of observations. $K(u) = \frac{3}{4}(1-u^2)$ for $-1 < u < 1$ and $K(u) = 0$ otherwise is the Epanechnikov kernel, where $u = \frac{w-w_i}{h}$. The optimal bandwidth (Silverman (1986)) is $h = 0.9\hat{\sigma}N^{-\frac{1}{5}}$ with $\hat{\sigma} = \min\{S, \frac{IQR}{1.349}\}$ where S is the sample standard deviation and IQR is the inter-quartile range. To prevent over-smoothing and following Leibbrandt, Levinsohn, and McCrary (2005), I use a bandwidth which is 0.75 times this optimal level.

Wages for internal migrants out of urban areas are in between non-migrants and US-bound emigrants whereas the wages of internal migrants out of rural areas are the lowest of all the groups. This implies that emigrants to the US are negatively sorted with respect to internal migrants out of urban Mexico but very positively sorted with respect to internal migrants out of rural Mexico.

The following section addresses the differences between the urban Mexico and rural Mexico patterns.

4 Assessing three migrant selection theories

This section explores which of the three theories summarized in section 2 could better accommodate the existence of positive selection in rural Mexico together with negative selection in urban Mexico in the period 2000-2004: skill prices, network effects or wealth constraints.

4.1 Skill prices

The expression to test is inequality (1) in section 2. If skill prices were higher in urban Mexico than in the United States and, in addition, higher in the United States than in rural Mexico, then that could explain why positive selection prevails in rural Mexico while there is negative selection in urban Mexico and this would confirm Borjas (1987) classical theory.

The main problem with such a test is to determine the concept of skill prices that would be relevant to the migration decision. One way to test the theory without specifying the concept is to look directly at the selection patterns of US-bound versus internal migrants. If inequality (1) is true, this should entail negative sorting of US migrants out of rural Mexico with respect to migrants between rural and urban Mexico. Table 1 and figure 1 tell the exact opposite story: migrants to the US are very positively sorted in terms of observed wages. However, the rejection of the theory is not conclusive for two reasons. First, the measure of rural-urban migration is just an approximation to the theoretical concept and, second, the result on education levels coincides with the theory, as US migrants are negatively sorted with respect to internal migrants in rural Mexico in terms of schooling years.

An alternative test is to follow most of the literature identifying the theoretical δ with the return to education.[18] Under this identification, running simple Mincer regressions on

[18] Cragg and Epelbaum (1996), Hanson (2006) and Ibarrarán and Lubotsky (2007) are just some examples.

rural and urban Mexico and on Mexican immigrants in the United States and comparing the coefficients on the return to schooling can be done as an approximation to the test. Table 2 presents the results from this exercise. The data for Mexican immigrants in the United States come from the American Community Survey (ACS)[19] and, for comparability purposes, it refers to recent Mexican immigrants in the United States, defining them as those who arrived there a year before the survey takes place.[20]

(Table 2)

Concentrating on the coefficient of schooling years, table 2 shows that the market price of an additional year of education is slightly higher in urban than in rural Mexico and also higher in both cases (0.09) than in the United States (0.03). These estimates are in line with the findings reported in Hanson (2006) and would imply negative selection of Mexican emigrants to the United States both out of rural and of urban Mexico and negative sorting with respect to internal migrants out of rural Mexico. The only contribution to the previous literature is the calculation of the schooling coefficient both for urban and for rural Mexico,[21] which turns out to be of similar magnitude although significantly higher (at a 95 percent confidence level) in urban Mexico than in rural Mexico. These results would suggest that the Borjas (1987) hypothesis, summarized by equation (1), can be rejected. However, table 2 also presents the calculation of Mincer regressions only for the population of future working-age Mexican emigrants to the United States and for future internal migrants. Confining our attention to these samples, which are the ones that ultimately emigrate, it can be observed that the return to an additional year of schooling is still higher in urban Mexico (0.06) than in rural Mexico (0.04) and for both higher than in the United States (0.03). The returns to education are in both cases significantly lower for emigrants than for the rest of the population. In the case of rural Mexico, although the point estimate suggests otherwise, it is no longer possible to reject the hypothesis that the return to an additional year of schooling is lower in rural Mexico than in the United States so that equation (1) could still be true.

[19] See Ruggles, Sobek, Alexander, Fitch, Goeken, Hall, King, and Ronnander (2004).

[20] Summary statistics for the ACS are provided in appendix B. Alternative definitions of Mexican immigrants in the US do not alter the results. The ACS is preferred to other sources, like the Current Population Survey in the United States, because it enumerates more immigrants than the latter. Still, the ACS is likely to under-count Mexican immigrants in the US, especially if they are undocumented. See Hanson (2006) and Fernández-Huertas Moraga (2011) for details.

[21] This is not meant to imply that rural and urban Mexico are different labor markets. The exercise is purely descriptive.

For internal migrants out of rural Mexico, the coefficient is 0.06 but it is not significantly different either from that of future US emigrants or from that of Mexicans already in the US.

Heckman, Lochner, and Todd (2003) question the appropriateness of using traditional Mincer regressions to compute the return to education. Lacking the desired data, addressing all of the concerns that they stress is out of the scope of this paper. Still, they argue that one of the quantitatively more important biases in the calculation of rates of return arises from the assumptions of linearity in education and from the separability between schooling and work experience. Relaxing the assumption of linearity does not alter the conclusions from table 2, as it can be observed in figure 2.

(Figure 2)

Figure 2 graphs the coefficients from regressing log wages on the same variables as in table 2, but this time substituting the schooling years variable for several schooling categories. The first graph shows that the structure of schooling returns is similar in urban and in rural Mexico and clearly above the returns to schooling for Mexican immigrants in the US. The second repeats the exercise just for Mexican emigrants. Although the graphed point estimates suggest that returns to schooling are higher for Mexican emigrants out of rural Mexico at low schooling levels and higher for emigrants out of urban Mexico at high schooling levels, the fact is that none of these results is statistically significant at a 95 percent confidence level.

The conclusion is thus that Borjas (1987) theory seems roughly to fit the selection of emigrants out of urban Mexico but it has more problems predicting the selection pattern out of rural Mexico despite the fact that the validity of the theory cannot be clearly rejected.[22]

Since Mexican emigrants to the United States are younger than non-migrants, one could think that the negative selection result in terms of wages results from a seniority effect. Older individuals have more experience in the labor market and are thus able to obtain higher wages. In general, it is interesting to understand which part of the selection result

[22]Skill prices can also be computed using a linear utility framework, as suggested by Grogger and Hanson (2011) and Rosenzweig (2007). In that case, ignoring migration costs (that could be varying by skill levels), college graduates have much more incentives to emigrate from Mexico than any other group in the population, which would lead to positive selection both out of rural and out of urban Mexico. Given that selection is negative out of urban Mexico, the linear utility model does not seem able to explain the differing selection patterns out of urban and rural Mexico to the United States. See figure B2 in appendix B for more details.

is due to differing observable characteristics of emigrants and how they are rewarded and which part of the result is due to unobservable characteristics. One way of performing this calculation non-parametrically is to use DiNardo, Fortin, and Lemieux (1996) reweighing procedure, following Chiquiar and Hanson (2005) and Fernández-Huertas Moraga (2011), both for urban and rural Mexico.

If the information on emigrant wages in the ENET is ignored and their wage distribution is inferred only from their observable characteristics, as suggested by DiNardo, Fortin, and Lemieux (1996) and Chiquiar and Hanson (2005), the actual wage distribution of non-migrants computed in figure 1 can be compared now not to the actual wage distribution of emigrants but to a counterfactual wage distribution. The counterfactual reweighs the non-migrant wage distribution by the observable characteristics of migrants. The reweighing factor is computed as the conditional odds of migrating (from a logit model of the probability of emigration).[23] This is what Chiquiar and Hanson (2005) define as the "appropriate weight" (see page 262) since it conditions migrants participating in the labor market (in this case, reporting wages) in the same way as non-migrants, thus abstracting from the differences in labor market participation and wage reporting observed in table 1. The result can be viewed in figure 3.

(Figure 3)

Figure 3 shows the kernel density estimate of the non-migrant wage distribution (solid line) already calculated in figure 1 together with the counterfactual density (dashed lines) corresponding to the wage emigrants should be earning according to their observable characteristics. As a result, the difference between the two densities reflects the part of selection that is due only to observable characteristics of the migrants. The rest of the difference with the actual wage distribution of the emigrants can be considered as the effect of unobservables in selection.

The difference between the graphs in figure 3 and figure 1 can be summarized in terms of averages. The degree of selection on observables can be computed as the difference between the average of the counterfactual migrant wage distribution and the average of the actual non-migrant wage distribution. This degree of selection on observables is -0.15 (0.01 is the

[23]The logit regresses the migration dummy from the ENET on the following variables (used in Chiquiar and Hanson (2005)): schooling groups, age, age squared, marital status and interactions of these variables with the schooling groups. The results of this auxiliary regression are available from the author upon request.

standard error)[24] for urban Mexico and 0.09 (0.01 is the standard error) for rural Mexico in the case of emigrants to the US. This means that the degree of selection on observables for Mexican emigrants to the US coincides in sign with the actual degree of selection: positive for rural and negative for urban Mexico. Observable characteristics account for 51 percent (s.e.=5 percentage points) of the observed negative selection in urban Mexico and for 48 percent (s.e.=15 percentage points) of the observed degree of positive selection in rural Mexico. The most striking result is that of internal migrants out of rural Mexico. As it could be expected from their summary statistics in table 1, they are positively selected in observables: 0.13 (s.e.=0.01) while their overall degree of selection (figure 1) was -0.06 (s.e.=0.04). This implies a degree of negative selection in unobservables (-0.20; s.e.=0.05) similar to that of US-bound migrants out of urban Mexico (-0.14; s.e.=0.03) coupled with a degree of positive selection in observables (0.13) higher than that of US-bound emigrants in rural Mexico (0.09), which implies negative sorting of the latter in observables (-0.04), as predicted by the theory, but positive in unobservables (0.29).

There are a multiplicity of factors that could be related to the unobservable component of the degree of selection. The negative selection on unobservables in urban Mexico could in principle be related to the existence of an Ashenfelter dip that reduces wages right before migration but Fernández-Huertas Moraga (2011) shows this is not the case. In addition, rural Mexico presents the opposite result so this is an unlikely explanation. Another explanation could be the existence of low unobserved ability in the case of urban Mexico emigrants (Borjas, 1987) together with high unobserved ability for emigrants out of rural Mexico.[25] McKenzie and Rapoport (2011) have shown that living in a Mexican migrant household decreases the probability of high school completion by 13-14 percent on average, with most of the effect coming from young males migrating before completion. In the absence of emigration, these individuals would have become more educated, moving the counterfactual emigrant wage distribution in figure 3 to the right, and thus reducing the degree of negative selection explained by observable components in urban Mexico while increasing the degree of positive selection explained by observables in rural Mexico.

The following two subsections review two additional explanations proposed by the liter-

[24]These and all of the following on degrees of selection are bootstrapped standard errors obtained by randomly sampling with replacement half of the observations.

[25]Chiswick (1978, 1999) explains a variety of reasons why emigrants could be positively selected in unobservables.

ature.

4.2 Network Effects

McKenzie and Rapoport (2010) proved propositions 1 and 2, rewritten in subsection 2.2 of this paper. These propositions suggest that larger migration networks should be correlated with more negative selection of emigrants. The reason is that migration networks reduce costs (or increase benefits) from migration relatively more for individuals at the low end of the skill distribution. However, the fact that this assertion is true does not say anything about its ability to disentangle the differences in selection between urban and rural Mexico. In a sense, McKenzie and Rapoport (2010) showed the qualitative validity of propositions 1 and 2 whereas what will be assessed in this section is its quantitative relevance in explaining differing selection patterns.

McKenzie and Rapoport (2010) perform their exercise in a different survey: the Encuesta Nacional de la Dinámica Demográfica (ENADID) for 1997.[26] Their results suggest that the effect of migration networks on the probability of emigrating for the first time to the US in the period 1996-1997 is 29 percent lower in localities with more than 100,000 inhabitants but they do not compute directly the effect of the locality size on the degree of selection.[27] They measure their migration network variable as the proportion of individuals aged 15 and over in a given community (municipality) who have ever migrated to the US. Unfortunately, this information is not present in the ENET.[28] For comparability purposes, I use their migration network variable calculated from the ENADID in what follows.[29]

[26]The ENADID is a nationally representative household survey that INEGI carried out in 1992, 1997 and 2009. For more information on the ENADID, see McKenzie and Rapoport (2007, 2010).

[27]Their coefficient on the effect of the interaction between education and the migration network on the probability of emigrating becomes less negative (implying less negative selection) when they take localities larger than 100,000 inhabitants out of the sample. Although this difference is not significant, this would go against the fact that negative selection prevails in urban Mexico whereas positive selection prevails in rural Mexico.

[28]In unreported results, I construct a municipal network variable from the ENET as the average municipal emigration rate to the United States of individuals aged 16 and 65 in the 2000-2004 period. The correlation coefficient between this variable and the ENADID network variable is 0.73. Using this alternative variable as a measure of networks for what follows does not change the results.

[29]I match the ENADID and ENET on municipality codes. Only 7 percent of the observations remain unmatched. If I follow the restriction in McKenzie and Rapoport (2010) by dropping municipalities where less than 50 households were interviewed, 32 percent of the observations are unmatched but the results are

First, table 3 presents some preliminary evidence by reproducing the summary statistics computed in table 1 but this time dividing the sample between municipalities with high community migration prevalence (high migration network) and low community migration prevalence (low migration network), with the cutoff value determined by the median at the national level (5.4 percent).

(Table 3)

From table 3, it can be seen that the migration rate to the US is substantially larger in high network areas than in low network areas whereas the opposite is true for the internal migration rate, suggesting that migration to urban Mexico and to the US might be seen as substitutes for rural Mexico males. As it could be expected, individuals who migrate to the US tend to come in all areas from municipalities with a higher network prevalence. In terms of selection, the summary statistics show a clear relationship between negative selection and network prevalence both in rural and urban Mexico, as McKenzie and Rapoport (2010) proved. However, even if higher network prevalence leads to more negative selection, it does not seem likely that networks can explain the different selection patterns in urban and rural Mexico. The reason is that network prevalence is higher (11 percent on average) in rural Mexico, where selection is positive, than in urban Mexico (8 percent), where selection is negative.

A final exercise that can be performed to assess the impact of networks on the computed degree of selection is to redo the calculation in subsection 4.1. Figure 3 represented counterfactual wage distributions for Mexican emigrants to the United States and internal migrants based just on their observable characteristics (schooling, age and marital status). Coming back to DiNardo, Fortin, and Lemieux (1996) reweighing procedure, assume that the network prevalence variable constitutes another observable characteristic that can be used when computing counterfactual wages. The result from including migration networks in the computation of counterfactual wage densities for emigrants in rural and urban Mexico[30] can be observed in figure 4.

(Figure 4)

Figure 4 appears almost identical to figure 3. Adding the network variable moves the

not altered significantly.

[30]See subsection 4.1 for an explanation of the computation of the wage counterfactual. The weights are calculated as in footnote 23 and adding the network variable and its interaction with the schooling categories. Results from the auxiliary logit regression are available from the author upon request.

density for rural Mexico migrants to the US slightly to the right, meaning that observables should now imply more positive selection. For internal migrants out of rural Mexico, the density moves to the right, meaning that selection on observables for this group is less positive. This is confirmed by looking at the averages. The average degree of selection in figure 4 for migrants to the US out of urban Mexico is -0.12 (s.e.=0.01). This is 43 percent (s.e.=5 pp) of the actual degree of negative selection. Thus, adding networks has no significant effect on the explanatory power of observables for the case of urban Mexico. The result is very different for rural Mexico, though. The average degree of selection stemming for figure 4 is 0.22 (s.e.=0.01) for emigrants to the US out of rural Mexico. This means that observables more than explain the 0.18 actual degree of positive selection in rural Mexico. This translates into a statistically zero degree of selection in unobservables: -0.03 (s.e.=0.05). For internal migrants out of rural Mexico, the degree of selection in observables becomes 0.06 (s.e.=0.02), less positive than when networks where excluded in subsection 4.1 (0.13) but still implying negative selection in the remaining unobservables (-0.13; s.e.=0.05). This is enough to generate positive sorting of migrants to the US with respect to internal migrants out of rural Mexico both on observables (0.15) and on unobservables (0.09).

In principle, the network variable could be capturing any municipality-specific component affecting the migration decision since it is the only variable that changes at a municipal level. However, McKenzie and Rapoport (2010) showed that the migration network variable effects did not disappear or change their magnitude even if they added municipality dummies to their main regression. In unreported results, I also use municipality dummies and keep the network-schooling interactions when building figure 4. This also takes care of the possible non-linear effect of networks on the probability of emigration.[31] The results do not change for US-bound emigrants both out of urban and of rural Mexico. However, the selection on observables for internal migrants out of rural Mexico becomes negative. One possible explanation might be that we are omitting for internal migrants an analogous definition of networks to that employed for US migrants. This is the effect that municipality dummies might be capturing for internal migrants.

The conclusion from this subsection is that network variables seem unlikely to be able to explain by themselves why there is negative selection in urban Mexico. Propositions 1 and 2 would suggest that network effects on the degree of selection should be more pronounced in

[31]See Bauer, Epstein, and Gang (2007, 2009)

urban than in rural Mexico but the ENET survey does not seem to support this view. Still, networks seem to affect selection in the expected direction and are able to fully deal with the selection result for US-bound emigrants in rural Mexico. The next subsection explores a third possible explanation to the different selection patterns.

4.3 Wealth constraints

The fact that migration is generally a profitable investment does not mean that every person who could obtain this profit will actually emigrate. It could happen that low-income individuals willing to emigrate cannot do so because they lack the financial resources to cover the migration costs. They could borrow to start the trip but sometimes they do not have the possibility of borrowing, either because the financial sector in the area in which they live is not specially developed or because they do not have access to a network (family or friends) that can lend them the money. If this is the case, this individual will be considered wealth constrained. Wealth constraints could be able to explain why emigrant selection is positive in rural Mexico and negative in urban Mexico. The reason is that even when low-skill individuals have relatively more incentives to migrate in both areas they could be constrained in rural Mexico and not in urban Mexico. This is the issue that this section will address.

The existence of wealth constraints, in addition to be able to sort out the rural-urban emigrant selection difference, is key to understanding the consequences of migration policies on the selection of emigrants. Borjas (1987) simplest model of negative selection, presented in section 2, suggests that any increase in migration costs, such as tougher enforcement at the border, will lead to an increase in negative selection. However, if migration costs are decreasing in the productive characteristics of the emigrant-sending country population, this does not need to be the case and the selection of emigrants could actually become positive or, at least, less negative.[32]

Observable migration costs seem too small to justify the fact that the real hourly wage is between four and five times larger in the United States for Mexican immigrants than in Mexico (ACS data for 2000-2004, see appendix B). This leads Hanson (2006) to consider that the real puzzle is why more people do not migrate. Wealth constraints could provide an explanation and their relevance can be tested with the ENET data. The reasoning is that individuals whose expected utility is higher in the United States decide to remain in Mexico

[32] See Borger (2010) for an application.

not necessarily because they cannot afford their trip but because they need to provide a buffer of savings for their family for the time it will take them to start sending remittances, for the probability of not being successful in crossing the border in the case of undocumented migrants, etc. For these individuals, the probability of emigrating should be increasing in a measure of their wealth, independently from their wage level.

Once a year, in the second quarter, the ENET surveys the characteristics of the building where the household lives, which enables the construction of a household wealth index. Filmer and Pritchett (2001) suggest that a principal components analysis can be used to this end. If this index is a good proxy for wealth (McKenzie, 2005) and there exist wealth constraints in migration, the decision to migrate should be positively related to the index, controlling for other observables and wage income, for low wealth individuals and have no relation to migration for high wealth individuals. McKenzie (2005) constructs one general wealth index out of thirty asset characteristics present in the ENIGH. In addition, he shows that three other indices made out of subgroups of characteristics also provide a good measure of wealth: a housing characteristics index, a utilities index and a durables index. Unfortunately, the ENET does not provide information on durables ownership but it has some other indicators not present in the ENIGH, such as kitchen equipment. Given these considerations, six wealth indices are constructed from the ENET data. On the one hand, one that replicates the utilities and housing characteristics index from McKenzie (2005) for Mexico as a whole, urban Mexico and rural Mexico. On the other hand, one that uses all the available information in the ENET with its thirty-six components again for Mexico as a whole, urban Mexico and rural Mexico. Their construction is detailed in Appendix A. Both of them are very similar although the McKenzie (2005) index does a better job at explaining the overall variance in the first principal component in the three cases and this is why it is preferred. An additional choice is between the overall Mexico index and those particular to urban and rural Mexico. The latter are preferred due to the greater ability of the urban one to discriminate wealth in urban Mexico, which can be seen, for example by comparing standard deviations. The standard deviation of the urban Mexico overall index in urban Mexico is 1.6 while that of the urban index is 1.9. Given that they are all highly correlated, the use of either for the analysis is immaterial for the results.

The assumption to be tested is whether wealth is positively related to the decision to migrate, controlling for other factors (especially productive characteristics reflected in the wage), for low wealth individuals. In contrast, wealth should not be relevant in the migration decision of wealthier individuals. Given that there is positive selection among the rural population and negative for the urban population, one should expect that wealth constraints would be easier to identify and more relevant in a context of positive selection like rural Mexico.

To this end, the binary variable mig_{it} taking the value 0 if the individual remains in Mexico in the quarter following the one in which the observation takes place and 1 if the individual emigrates to the United States in the following quarter is considered as the dependent variable.[33] The regressors that should be correlated with this variable are the log of the hourly wage ($\log w_{it}$), which should have a negative effect on emigration, and all other observable characteristics of the individual: schooling, age, community migration prevalence (from the ENADID) and its interaction with education,[34] family characteristics, distance to the border and dummies for time and Mexican states (X_{it}). The dummies for the Mexican states control for time-invariant multilateral resistance to migration, that is, opportunities to migrate to alternative destinations (such as internal migration opportunities) that do not change over time. However, the inclusion of state-quarter fixed effects, which would also control for time-variant multilateral resistance to migration, does not alter the results in this case.[35]

The most interesting regressor, though, is the measure of household wealth taken from applying McKenzie (2005) index to the ENET ($asset_{it}$) in both the rural and the urban setting.

A traditional linear regression analysis of the effect of the asset index on the probability of emigration could be inappropriate if, as expected, there are non-linearities in this relationship. For this reason, a semi-parametric approach following the local linear regression method of Fan (1992) is preferred. The regression to be estimated is the following:

$$mig_{it} = G(asset_{it}) + \Gamma X_{it} + \epsilon_{it}$$

Fan (1992) shows how to apply the local linear regression method for one independent

[33]Internal migrants and international migrants to other destinations are taken out of the sample although their inclusion does not alter the results.

[34]This follows McKenzie and Rapoport (2010).

[35]See Bertoli and Fernández-Huertas Moraga (2013) for the exact definition and an empirical application of the concept of multilateral resistance to migration.

variable. Thus, the effect of all the controls must be discounted in a first step by estimating Γ. To this end, the high order differencing method of Yatchew (1998) can be used. First, the data are ordered in ascending order according to $asset_{it}$. With d_j ($j = 0...5$) denoting the optimal Yatchew (1998) differencing weights of fifth order,[36] the following ordinary least squares regression can be estimated:

$$\sum_{j=0}^{5} d_j mig_{i-j,t} = \left(\sum_{j=0}^{5} d_j X_{i-j,t}\right)' \Gamma + \varepsilon_{it}$$

The idea is that the difference between contiguous observations of the asset variable is small enough to disregard it so that $\hat{\Gamma}$ is estimated efficiently and Fan (1992) local linear regression can be run on:

$$mig_{it} - X_{it}\hat{\Gamma} = G(asset_{it}) + \eta_{it}$$

The results from estimating Fan (1992) local linear regression[37] for urban and rural Mexico can be observed in figure 5.

(Figure 5)

The two panels of figure 5 separately show the estimated functions for rural and urban Mexico with their corresponding 90 percent bootstrapped confidence intervals[38] on the two different versions of the asset index. Urban Mexico seems to fit an inverted u-shape relationship between the emigration probability and wealth while rural Mexico shows a clear,

[36] The weights are 0.9064, -0.2600, -0.2167, -0.1774, -0.1420 and -0.1103. Yatchew (1998) shows that differencing with these weights attains 91 percent efficiency relative to the asymptotic efficiency bound.

[37] The complete results from the first auxiliary regression estimating Γ are available from the author upon request. It is run on 321,537 observations for urban Mexico and 37,781 for rural Mexico, restricted to the second quarter of the sample of males aged 16 to 65 years old, and the R^2 from the regressions are 0.01 for urban Mexico and 0.03 for rural Mexico. The variables included are discussed and shown in appendix B. Quadratic terms in age and schooling years and an interaction of the network prevalence variable with schooling years are added. The signs of the coefficients coincide for the more relevant variables, notably that of the education and network interaction (negative), with those reported in McKenzie and Rapoport (2010). Following Deaton (1997), the Epanechnikov kernel is used. A bandwidth of 0.2 times the asset index range is chosen.

[38] The interval comes from the 5th and 95th percentile of the distribution originated by repeating the procedure 1000 times by randomly sampling with replacement half of the observations.

and larger, increasing relationship.[39] In the case of urban Mexico, only the poorest part of the population could be subject to some sort of wealth constraint according to this graph but it must be mentioned that the extremes are precisely the values of the asset index that Filmer and Pritchett (2001) advise to take more carefully, since it tends to over-discriminate at low wealth levels and to under-discriminate at high wealth levels. For the rest of the urban Mexico sample (more than 60 percent, by looking at the depicted wealth quintiles), if anything, there is a negative relationship between wealth and the probability of emigrating to the United States in the following quarter so that there does not seem to be any scope for the existence of wealth constraints affecting the emigrating decision.

On the contrary, rural Mexico's result is consistent with the existence of wealth constraints in the emigration decision. The probability of emigration is clearly increasing in the household wealth level (after controlling for all other observed factors in the ENET, including the wage, and the network variable from the ENADID) for 98.7 percent of the total rural population in the sample.

It would be interesting to investigate further why there are these disparate relationships between wealth and the emigration probability at the rural and urban level. A first hypothesis that could be put forth is household size. If emigrating individuals do not leave until they have accumulated enough wealth on which their family can live before they start sending remittances or in case there is a failure in getting across the border, then higher household size should be related to a greater incidence of wealth constraints. In fact, if the above estimation procedure is further divided by household size, the results point in this direction. The emigration probability of individuals belonging to households with a size above the median in the lowest wealth quintile is increasing in wealth whereas there is no relationship for individuals belonging to lower size households. The problem is that dividing the dataset too much leads to a lack of power in the estimation and the standard errors become too big to draw meaningful conclusions.[40]

A second hypothesis that could explain the rural/urban divide is the thickness of the credit market in rural and urban areas. There is some evidence that the credit market could be more developed in urban areas of Mexico than in rural ones. According to Focke (2004),

[39]When the extremes are considered, as in figure B5 in the appendix, rural Mexico also presents a u-shape relationship, consistent with the findings of McKenzie and Rapoport (2007).

[40]A more traditional analysis based on a probit model of the emigration decision, available from the author upon request, did not confirm this hypothesis.

only 6 percent of the population in rural areas had access to a bank account, in contrast to 15 percent in urban areas. In this sense, the World Bank unfolded a project (2004-2009) to develop the rural financial sector in Mexico since 74 percent of the municipalities (hosting 22 percent of the population) did not even have a bank branch in their territory.[41] Although the World Bank does not provide this information, these percentages are likely to refer to rural Mexico (22 percent of the population in 2000-2004; see table 1). These numbers refer to the formal financial sector so that it would be possible for a developed informal financial sector to fill in these gaps. However, Paxton (2006) studies semi-formal financial institutions in rural Mexico and finds that they are highly inefficient. She calculates that the Mexican rural financial sector is 50 percent less efficient than the urban one and attributes this difference to institutional factors rather than different client profiles.

Finally, a third hypothesis can be found by going back to the original meaning of the constructed asset index. The index reflects household infrastructure whose value could be notably higher in urban than in rural areas. In this sense, it would not be surprising that homes with the same amenities are more valuable and thus correspond to wealthier households in urban Mexico relative to rural Mexico. This would explain why individuals with the same asset index value could be constrained in rural areas but not in urban areas.

In addition to being significant, the relationship between wealth and the emigration probability is of a considerable magnitude in rural Mexico. Taking into account that the average emigration rate out of rural Mexico is 1.3 percent, figure 5 shows that the effect of wealth on the emigration probability could be substantive. However, the fact that wealth is associated with the emigration probability in rural Mexico and not in urban Mexico does not say anything as to the ability of wealth constraints to explain the different selection patterns in both areas. Coming back to DiNardo, Fortin, and Lemieux (1996) reweighing procedure, the McKenzie (2005) asset index can be included in the computation of counterfactual wage densities already undertaken in subsections 4.1 and 4.2 (adding networks on the second case). This is an useful accounting exercise to see to what extent wealth constraints could be relevant in shaping the degree of selection. Figure 6 shows the results.[42]

(Figure 6)

[41] Information available at the World Bank web page: www.worldbank.org. Web accessed on 10-20-2006.

[42] See subsection 4.1 for an explanation of the computation of the wage counterfactual. The weights are calculated as in footnote 23 and adding the network variable, the McKenzie (2005) asset index, their interaction and the interactions of these two variables with the schooling categories. If assets are added

Figures 6 does not seem very different from figures 3 and 4. For the case of urban Mexico, the counterfactual wage densities are basically identical to the ones obtained when adding only the network variable. The percentage of the actual average degree of selection that observables (adding assets and networks) can explain is 42 percent (s.e.=5 pp) for US-bound emigrants out of urban Mexico. Thus, approximately half of the observed degree of selection is still attributable to negative selection on unobservable characteristics. Again, this situation is not surprising taking into account the above result that the asset index had no effect on the probability of emigration to the US out of urban Mexico.

In the case of rural Mexico, despite the previous finding that the emigration probability to the US increases on the wealth index for most of the rural population (suggestive of wealth constraints), adding the wealth measure to the counterfactual wage estimation seems to have a visually negligible effect. However, looking at the averages, the implied degree of selection is now 0.28 (s.e.=0.02), much more than the actual 0.18 observed, thus implying a negative degree of selection in unobservables of -0.10 (s.e.=0.05) not much smaller than the degree of selection in unobservables calculated for urban Mexico emigrants to the US: -0.17 (s.e.=0.03) although more imprecisely estimated.

For internal migrants out of rural Mexico, adding the asset index to the counterfactual wage estimation in figure 6 makes the degree of selection in observables even smaller: 0.02 (s.e.=0.02) than in subsection 4.2 (0.06). This results on non-significant negative selection in unobservables (-0.08; s.e.=0.06). Sorting of migrants to the US with respect to internal migrants out of rural Mexico ends up being statistically zero on unobservables (-0.02) but clearly positive on observables (0.26) as the theory in subsection 2.3 predicts.

The conclusion is that, once wealth constraints and network effects are taken into account, the positive selection result for US-bound emigrants out of rural Mexico can be completely accounted for. In addition, although positive selection prevails in rural Mexico in terms of observable characteristics, there is significant negative selection on unobservables. Once networks and wealth constraints are considered, there is no need to keep trying to explain different selection patterns in emigration to the US out of urban and rural Mexico. What is left for future research to understand, though, is why the selection in the remaining unobservables is still negative.

alone, without networks, the results coincide with those of subsection 4.2. Results from the auxiliary logit regression are available from the author upon request.

In any case, the value of these results depends critically on the quality of the chosen asset measure. All the results are similar when different household infrastructure indices are taken from the ENET, which provides some confidence about the robustness of the estimation. However, it would have been desirable to find an appropriate instrumental variable for the effect of wealth in the probability of migration so that causation in addition to correlation could have been tested and to avoid omitted variable bias, especially in the case of idiosyncratic network effects or the possibility of being a remittance recipient. Nevertheless, the inclusion of controls and fixed effects for time and location suggests that the estimation procedure can be solid enough and informative for the proposed question. A supplementary dataset other than the ENET would be needed to deepen the analysis presented here.[43]

5 Conclusion

Immigration affects welfare in receiving and sending countries both through the size and the composition of migration flows, which is determined by how emigrants self-select. This paper explores three factors affecting selection, and thus the composition, of migration flows from Mexico to the United States in the period 2000-2004: wealth constraints, network effects and skill prices. There are two motivations for this. The first one is the need to explain why there are two very different patterns of selection inside Mexico: negative selection in urban Mexico (emigrants earn a lower wage and have less years of schooling than non-migrants) versus positive selection in rural Mexico (emigrants earn a higher wage and have more schooling years than non-migrants). The second motivation is that the effect of policy on the composition of migration flows will depend on the mechanisms that generate emigrant selection. For example, a more restrictive migration policy consisting of toughening border controls will generate a more negatively selected migration flow when selection is negative to begin with (if costs increase, only those with the highest return to migration will continue migrating) but its effect is theoretically ambiguous when selection is positive to begin with.

Out of the three theories that could explain the differing selection patterns in rural and urban Mexico, all of them matter but they matter differently in different areas. First, higher skill prices in urban Mexico than in the United States account for half of the observable

[43]Kaestner and Malamud (2012) also analyze wealth constraints with the richer Mexican Family Life Survey. Their main problem is that the survey is not large enough to be representative of rural Mexico.

degree of negative selection in urban Mexico, where there seems to be no role either for network effects or for wealth constraints. Nevertheless, higher skill prices in rural Mexico with respect to the United States, although relevant, are not enough to generate negative selection in rural Mexico as well. However, positive selection does not survive the removal of skill prices and network effects. In addition, when wealth constraints are added, the positive selection result becomes a negative selection result in unobservables.

To sum up, the combination of the three factors is enough to account for the mechanism of selection observed in rural Mexico. What remains to be shown is why there is negative selection both in urban and in rural Mexico once the effect of all these variables is discounted.

By addressing the effect of wealth constraints on the migration decision, this paper also contributes to understanding the structure of migration costs. Semi-parametric techniques are used to estimate a non-linear function of the probability of emigration on wealth. The result is that there is no evidence of any effect of wealth constraints in urban Mexico but there is evidence that wealth constraints could be playing a role in the migration decision of individuals living in rural areas. This would lead to the conclusion that migration numbers would increase and the degree of selection would be more negative if the wealth constraints suffered by the rural population were reduced by better banking institutions or just by an improvement in their economic condition, ceteris paribus.

The paper has also addressed the selection of internal migrants, concentrating on internal migrants out of rural Mexico who could be considering the US as an alternative destination. Although the results obtained on internal migration, particularly on sorting, were in line with the theory, future research should refine the analysis offered here with a more appropriate dataset that is able to identify the correct concept of rural-urban migration in Mexico.

As for the dataset, the ENET offers the great advantage of its sample size to address the questions investigated here on the rural-urban divide. However, although its selection results have been confirmed by other datasets, such as the Mexican Family Life Survey, there is clear scope for improvement on the quality of the data, specifically in terms of a lower attrition rate together with the ability to follow whole households emigrating at the same time and leaving no one behind.

Bibliography

Ambrosini, J. W., and G. Peri (2012): "The Determinants and the Selection of MexicoUS Migrants," *The World Economy*, 35(2), 111–151.

Bauer, T. K., G. S. Epstein, and I. N. Gang (2007): "The Influence of Stocks and Flows on Migrants Location Choices," *Research in Labor Economics*, 26, 199–229.

——— (2009): "Measuring Ethnic Linkages among Migrants," *International Journal of Manpower*, 30(1-2), 56–69.

Belot, M. V., and T. J. Hatton (2012): "Immigrant Selection in the OECD," *Scandinavian Journal of Economics*, 114(4), 1105–1128.

Bertoli, S., and J. Fernández-Huertas Moraga (2013): "Multilateral Resistance to Migration," *Journal of Development Economics*, forthcoming.

Borger, S. C. (2010): "Self-Selection and Liquidity Constraints in Different Migration Cost Regimes," *mimeo*.

Borjas, G. J. (1987): "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 77(4), 531–553.

——— (1999): "The Economic Analysis of Immigration," *In Handbook of Labor Economics, edited by Orley Ashenfelter and David Card*, 3A.

Caponi, V. (2010): "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the United States?," *Annales d'Économie et de Statistique*, 97/98, 207–234.

Carrington, W. J., E. Detragiache, and T. Vishwanath (1996): "Migration with Endogenous Moving Costs," *American Economic Review*, 86(4), 909–930.

Chiquiar, D., and G. H. Hanson (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.

Chiswick, B. R. (1978): "The Effect of Americanization on the Earnings of Foreign-Born Men," *Journal of Political Economy*, 86, 897–921.

——— (1999): "Immigration Policy and Immigrant Quality. Are Immigrants Favorably Self-Selected?," *American Economic Review*, 89(2), 181–185.

Cragg, M. I., and M. Epelbaum (1996): "Why has wage dispersion grown in Mexico? Is it the incidence of reforms or the growing demand for skills?," *Journal of Development Economics*, 51, 99–116.

Cuecuecha, A. (2005): "The Characteristics of the Immigrants from Mexico to the US," *Working Paper, ITAM-CIE.*

Deaton, A. (1997): *The Analysis of Household Surveys: A Microeconometric Approach to Development Policy.* The Johns Hopkins University, Baltimore.

DiNardo, J., N. M. Fortin, and T. Lemieux (1996): "Labor Market Institutions and the Distribution of Wages. 1973-1992: A Semiparametric Approach," *Econometrica*, 64, 1001–1044.

Fan, J. (1992): "Design-adaptive Non-parametric Regression," *Journal of the American Statistical Association*, 87(420), 998–1004.

Fernández-Huertas Moraga, J. (2011): "New Evidence on Emigrant Selection," *The Review of Economics and Statistics*, 93(1), 72–96.

Filmer, D., and L. H. Pritchett (2001): "Estimating Wealth Effects without Expenditure Data–Or Tears: An Application to Educational Enrollments in States of India," *Demography*, 38(1), 115–132.

Focke, K. (2004): "Mexico. Rural Financial System Consolidation Program (ME-0243). Loan Proposal," Discussion paper, Inter-American Development Bank.

Grogger, J., and G. H. Hanson (2011): "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 95(1), 42–57.

Hanson, G. H. (2006): "Illegal Migration from Mexico to the United States," *Journal of Economic Literature*, 44(4), 869–924.

Heckman, J. J., L. L. Lochner, and P. E. Todd (2003): "Fifty years of Mincer earnings regressions," *NBER Working Paper Series*, 9732.

Ibarrarán, P., and D. Lubotsky (2007): "Immigration and Self-Selection: New Evidence from the 2000 Mexican Census," *In Mexican Immigration to the United States, edited by George J. Borjas.*

INEGI (2005): *Encuesta Nacional de Empleo Trimestral (2000-2004).* Instituto Nacional de Estadística, Geografía e Informática. Aguascalientes, México.

Kaestner, R., and O. Malamud (2012): "Self-Selection and International Migration: New Evidence from Mexico," *The Review of Economics and Statistics*, forthcoming.

Lacuesta, A. (2010): "A Revision of the Self-selection of Migrants Using Returning Migrants Earnings," *Annales d'Économie et de Statistique*, 97/98, 235–260.

Leibbrandt, M., J. Levinsohn, and J. McCrary (2005): "Incomes in South Africa since the fall of Apartheid," *NBER Working Paper Series*, 11384.

McKenzie, D., and H. Rapoport (2011): "Can migration reduce educational attainment? Evidence from Mexico," *Journal of Population Economics*, 24(4), 1331–1358.

McKenzie, D. J. (2005): "Measuring Inequality with Asset Indicators," *Journal of Population Economics*, 18, 229–260.

McKenzie, D. J., and H. Rapoport (2007): "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico," *Journal of Development Economics*, 84(1), 1–24.

——— (2010): "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks," *The Review of Economics and Statistics*, 92(4), 811–821.

Mincer, J. (1958): "Investment in Human Capital and Personal Income Distribution," *Journal of Political Economy*, 66(4), 281–302.

Mishra, P. (2007): "Emigration and wages in source countries: Evidence from Mexico," *Journal of Development Economics*, 82(1), 180–199.

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the US Labor Market," *Quarterly Journal of Economics*, pp. 549–599.

Neumark, D., and D. Kawaguchi (2004): "Attrition bias in labor economics research using matched CPS files," *Journal of Economic and Social Measurement*, 29, 445–472.

Orrenius, P. M., and M. Zavodny (2005): "Self-selection among undocumented immigrants from Mexico," *Journal of Development Economics*, 78, 215–240.

Paxton, J. (2006): "Technical Efficiency in the Rural Financial Sector: Evidence from Mexico," *The Journal of Developing Areas*, 39(2), 101–119.

Robertson, R. (2000): "Wage Shocks and North American Labor Market Integration," *American Economic Review*, 90(4), 742–764.

Rosenzweig, M. (2007): "Education and Migration: A Global Perspective," *Mimeo, Yale University*.

Roy, A. D. (1951): "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 3, 135–146.

Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. K. Hall, M. King, and C. Ronnander (2004): *Integrated Public Use Microdata Series: Version 3.0 [Machine-readable database]*. Minneapolis, MN: Minnesota Population Center [producer and distributor]. Web accessed on 7-15-2006: http://www.ipums.org.

Silverman, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

Sjaastad, L. A. (1962): "The Costs and Returns of Human Migration," *Journal of Political Economy*, 70(1), 80–93.

Yatchew, A. (1998): "Non-Parametric Regression Techniques in Economics," *Journal of Economic Literature*, 36, 669–721.

A Data Appendix

Attrition rates in the ENET[44] are displayed in table A1. The average attrition rate is 11 percent after one quarter. This is mostly due to the increase in the attrition rate happening

[44] Percentages are computed using survey weights. The results are basically identical without using them.

in 2003, when the average is as high as 17 percent. As it is noted in the text, the main characteristics of the missing observations do not differ from the non-missing ones in the first quarter in which they are all recorded.

(Table A1)

The ENET enumerates both individuals who will emigrate in the following quarter to the United States and those who claim to have returned from the United States to a household included in the sample. The total numbers by quarter can be observed in figure A1 with their corresponding 95 percent confidence intervals.

(Figure A1)

The implied average emigration rate is 0.25 percent of the population per quarter whereas the implied return migration is one third of this number. Both figures are surely an underestimation since, as discussed in the text, the ENET does not gather information about neither emigrants who leave nobody behind (migrating with their entire family) nor on return migrants who do not come back to an established household. If a US migrant returns and creates a new household, then it is not recorded as a return migrant. An additional problem in the dataset is the absence of migrants (their observations were deleted in the available file) from the first quarter of 2004.

Table A2 presents the same summary statistics as table 1 but it divides urban Mexico between three subcategories: localities with more than 100,000 inhabitants, between 15,000 and 100,000 and between 2,500 and 15,000. Comparing table A2 and table 1, rural Mexico and localities with more than 100,000 inhabitants, where more than half of the Mexican population lives (52 percent) represent polar cases in term of selection, with the intermediate locality sizes presenting intermediate, though always negative, selection patterns. Still, once controls for observables are added as in figures 3, 4 and 6, the three urban Mexico subcategories show equivalent results, while such an equivalence only appears for rural Mexico in figure 6.[45]

(Table A2)

The construction of the index takes advantage from six questions about housing characteristics in the ENET. In total, there are thirty-six characteristics from which dummy variables are constructed taking a value of 1 if it is present and 0 otherwise, except for the

[45]Additional calculations available from the author upon request.

number of rooms and bedrooms. Filmer and Pritchett (2001) show that such a set of dummy variables can be used to construct an index through principal components analysis which does a good job in approximating household wealth by comparing the distribution arising from the index from that arising from traditional measures of wealth in expenditure and income household surveys in Indian states. Filmer and Pritchett (2001) methodology has been extensively used. In particular, for the case of Mexico, McKenzie (2005) shows that such a household wealth index performs well in approximating measures of wealth taken from the ENIGH (Encuesta Nacional de Ingresos y Gastos de los Hogares), the official Mexican income and expenditure household survey, in 1998. Table A3 shows that both the McKenzie (2005) and the All ENET index produce a reasonable ordering that can be a good approximation of wealth. This does not matter for the results since all the six versions of the asset index are very closely related, as it can be seen in table A4.

(Table A3, Table A4)

Figure A2 displays the histograms of the chosen asset indices, the rural and urban version of the index in McKenzie (2005). The asset index has a well shaped distribution for rural Mexico but still a skewed one for urban Mexico, particularly in the top two wealth quintiles.

(Figure A2)

B Additional Material

Figure B1 rewrites the result offered in figure 1 as the difference between the density of migrant wages and the density of non-migrant wages, both for urban and rural Mexico. The concentration of positive mass to the left of the median wage level (vertical solid line in figure 1) reflects negative selection in the urban Mexico graph both for internal and US migrants whereas the opposite is true for emigrants to the US out of rural Mexico.

(Figure B1)

Table B1 shows summary statistics for the ACS from 2000 to 2004 (used in table 2 and figure 2) and its comparison with the values from the ACS itself for 2000 (which has only 244 observations) and the more reliable 2000 US Census. The picture that the ACS offers for recent Mexican immigrants in the US (arrived a year earlier) does not statistically differ from that offered by the 2000 US Census.

(Table B1)

If the wage substitutes the log wage in the regression represented in the first panel in figure 2, figure B2 is obtained. Figure B2 shows the average experience-adjusted wage obtained by Mexicans of different schooling levels in urban and rural Mexico and it compares it with that of recent Mexican immigrants in the United States. The figure suggests that, ignoring migration costs (that could be varying by skill levels), college graduates have much more incentives to emigrate from Mexico than any other group in the population, which would lead to positive selection both out of rural and out of urban Mexico. Given that selection is negative out of urban Mexico, the linear utility model does not seem able to explain the differing selection patterns out of urban and rural Mexico to the United States.

(Figure B2)

To see more clearly the differences between the actual and counterfactual wage distributions, figure B3 is constructed in the same way as figure B1 but the difference in the densities comes from figure 3.

(Figure B3)

Figure B4 comes from figure 4 and is computed in the same way as figures B1 and B3.

(Figure B4)

The summary statistics for the data that are used in the estimation procedure of figures

5 and B5 are presented in table B2. The estimation is restricted to men aged 16 to 65 years old in order to be consistent with the rest of the paper.

(Table B2)

There are two main differences between the summary statistics in table 1 and those in table B2. First, table B2 refers only to observations recorded in the second quarter of every year from 2000 to 2004 whereas table 1 refers to all available quarters. Second, table B2 only provides summary statistics for the observations that are actually used in the regression analysis, that is, those not including missing values on any of the variables. The main difference here is the exclusion of those individuals not receiving a wage. All the regressions in section 4.3 have also been run dropping the wage variable and, if anything, the results are strengthened.[46]

Table B2 confirms that working-age males are much more likely to emigrate from rural Mexico (12.9 versus 4.8 emigrants per thousand in urban areas). Male individuals aged 16 to 65 in rural Mexico earn notably lower wages, which is consistent with an average education level that shows four less years of schooling than what is prevailing in urban Mexico. From the other categories, the most relevant information is that rural households tend to have more members than urban households (5.5 versus 4.8). Finally, network prevalence is much higher on average in rural (10.5 percent) than in urban Mexico (7.8 percent).

Figure B5 combines the rural and urban local linear regression estimation for a clearer comparison than in the main text, employing the overall version of the asset index computed in column 1 of table A3. The quintiles are those of the national wealth distribution, which represent 98.7 percent of the total rural population in the sample. The big drop in the rural function is only due to the effect of a very tiny fraction of the rural population (1.3 percent), which would be consistent with them being landowners[47] although the number of observations is too small and, as mentioned in the main text, the wealth index may present problems at the extremes. McKenzie and Rapoport (2007) find similarly a u-shape relationship between wealth and the probability of emigrating in the Mexican Migration Project and attribute this decreasing zone to the presence of landowners, who can obtain rents from their lands in Mexico that they cannot get in the United States so they have lower incentives to emigrate than those reflected in the typical selection model in section 2 of the paper.

[46] Results available from the author upon request.

[47] McKenzie (2005) suggests that the household wealth index is also well correlated with land ownership.

(Figure B5)

Figure B6 comes from figure 6 and is computed in the same way as figures B1, B3 and B4.

(Figure B4)

C Tables and Figures

Table 1:

Summary Statistics

Population aged 16 to 65		*Mexico*			*Urban Mexico*			*Rural Mexico*		
		Non-Migrants	*Internal Migrants*	*US Emigrants*	*Non-Migrants*	*Internal Migrants*	*US Emigrants*	*Non-Migrants*	*Internal Migrants*	*US Emigrants*
Percent Male		47%	64%	81%	47%	62%	78%	47%	67%	86%
		(0.0004)	(0.0061)	(0.0056)	(0.0005)	(0.0076)	(0.0081)	(0.0010)	(0.0103)	(0.0074)
Males aged 16 to 65										
Age	Average	35.0	28.3	29.4	34.8	29.2	29.6	35.9	26.8	29.2
		(0.0175)	(0.1756)	(0.1726)	(0.0194)	(0.2175)	(0.2422)	(0.0397)	(0.2931)	(0.2441)
	Median	33	25	27	33	26	27	34	23	27
Schooling Years	Average	8.5	8.3	7.2	9.4	9.4	8.1	5.4	6.5	6.1
		(0.0060)	(0.0693)	(0.0551)	(0.0066)	(0.0883)	(0.0788)	(0.0106)	(0.0941)	(0.0689)
	Median	9	9	6	9	9	9	6	6	6
Labor force participation		87%	87%	89%	85%	84%	87%	92%	92%	91%
		(0.0004)	(0.0052)	(0.0051)	(0.0005)	(0.0070)	(0.0076)	(0.0007)	(0.0073)	(0.0065)
Usable wage observations		68%	62%	60%	68%	62%	61%	70%	62%	58%
		(0.0006)	(0.0077)	(0.0078)	(0.0007)	(0.0095)	(0.0109)	(0.0013)	(0.0132)	(0.0112)
Unemployment rate		1.8%	3.2%	3.9%	2.1%	4.4%	5.3%	0.8%	1.3%	2.2%
		(0.0002)	(0.0027)	(0.0036)	(0.0002)	(0.0039)	(0.0058)	(0.0002)	(0.0029)	(0.0033)
Hourly wage in 2006 dollars										
	Average	2.05	1.60	1.40	2.34	2.05	1.64	1.03	0.85	1.08
		(0.0039)	(0.0358)	(0.0332)	(0.0048)	(0.0519)	(0.0538)	(0.0041)	(0.0233)	(0.0282)
	Median	1.43	1.10	1.14	1.62	1.40	1.28	0.80	0.77	0.95
Live in Rural Area		22%	37%	45%						
		(0.0005)	(0.0079)	(0.0079)						
Observations		4,252,646	19,471	12,649	3,764,680	16,134	9,150	487,966	3,337	3,499
Females aged 16 to 65										
Age	Average	35.3	26.0	28.2	35.2	27.1	29.0	35.3	23.6	26.5
		(0.0161)	(0.2248)	(0.4083)	(0.0180)	(0.2832)	(0.5381)	(0.0361)	(0.3408)	(0.5575)
	Median	34	22	24	34	23	25	33	20	23
Schooling Years	Average	7.9	8.7	8.5	8.7	9.3	9.1	5.0	7.3	7.1
		(0.0057)	(0.0863)	(0.1412)	(0.0063)	(0.1067)	(0.1890)	(0.0099)	(0.1357)	(0.1787)
	Median	9	9	9	9	9	9	6	8	6
Labor force participation		42%	45%	39%	45%	49%	41%	32%	38%	33%
		(0.0006)	(0.0105)	(0.0160)	(0.0007)	(0.0127)	(0.0199)	(0.0012)	(0.0184)	(0.0270)
Usable wage observations		28%	32%	24%	31%	36%	26%	17%	22%	20%
		(0.0005)	(0.0100)	(0.0137)	(0.0006)	(0.0124)	(0.0169)	(0.0010)	(0.0157)	(0.0237)
Unemployment rate		1.0%	2.6%	2.0%	1.2%	3.5%	2.4%	0.3%	0.9%	1.1%
		(0.0001)	(0.0028)	(0.0037)	(0.0002)	(0.0039)	(0.0048)	(0.0001)	(0.0026)	(0.0057)
Hourly wage in 2006 dollars										
	Average	1.92	1.39	1.49	2.06	1.56	1.74	1.06	0.79	0.86
		(0.0047)	(0.0491)	(0.1027)	(0.0053)	(0.0604)	(0.1359)	(0.0069)	(0.0359)	(0.0624)
	Median	1.32	1.02	1.05	1.42	1.17	1.15	0.79	0.72	0.71
Live in Rural Area		22%	32%	33%						
		(0.0005)	(0.0100)	(0.0154)						

Source: ENET (2005). Standard errors in smaller font and in parentheses computed using the svy linearized option in Stata with the survey weights. 2000 only includes the last three quarters and 2004 only the first three quarters. The construction of wages follows the lines of Chiquiar and Hanson (2005). The ENET asks Mexicans for their wage in the previous week to that in which the survey is performed or, if the individual did not work that particular week, for the usual wage. The figure is then brought to the monthly level. In order to prevent wages to refer to different time periods, the observations for individuals who reported usual rather than actual wage income are dropped. I follow Chiquiar and Hanson (2005) in dropping observations of individuals who worked more than 84 hours or less than 20 hours per week. Finally, the observations for people who worked in the United States (mostly border workers) are also dropped. Real wages are constructed with inflation data from the INPC series, Mexican CPI, in Banxico (www.banxico.org.mx), the Mexican central bank. These are quarterly averages based on June 2002 and brought to December 2005 with an index of 116.301. The exchange rate, from the International Financial Statistics of the IMF, corresponds to the 1 January 2006 and it is 10.7777 pesos per dollar. Following Chiquiar and Hanson (2005), hourly wages are computed by dividing the monthly wage income reported in the ENET by 4.5 times the number of hours worked in the previous week. Individuals are considered to live in a rural area when their locality has less than 2,500 inhabitants according to the 2000 Mexican Census. Internal migrants are defined as those who move to a different state within Mexico.

Table 2:

Mincer Regressions: Mexican males aged 16 to 65 (2000-2004)

Dependent variable: Log of the hourly wage	*Urban Mexico*			*Rural Mexico*			*US*
	All	*Migrants*		*All*	*Migrants*		*Migrants*
		Internal	*US*		*Internal*	*US*	
Schooling Years	**0.0932**	**0.0920**	**0.0630**	**0.0855**	**0.0566**	**0.0374**	**0.0320**
	(0.0003)	**(0.0044)**	**(0.0079)**	**(0.0009)**	**(0.0129)**	**(0.0122)**	**(0.0065)**
Experience	**0.0325**	**0.0319**	**0.0199**	**0.0030**	-0.0001	**-0.0173**	**0.0155**
	(0.0003)	**(0.0047)**	**(0.0059)**	**(0.0007)**	(0.0093)	**(0.0078)**	**(0.0057)**
Experience^2	**-0.0005**	**-0.0004**	-0.0003	**-0.0002**	-0.0002	0.0002	-0.0000
	(0.0000)	**(0.0001)**	(0.0002)	**(0.0000)**	(0.0002)	(0.0002)	(0.0001)
Constant	**-0.6923**	**-0.7495**	**-0.4491**	**-0.8094**	**-0.7763**	**-0.2736**	**1.6075**
	(0.0035)	**(0.0591)**	**(0.0873)**	**(0.0112)**	**(0.1385)**	**(0.1252)**	**(0.0782)**
Observations	1,734,619	6,232	4,297	233,098	1,375	1,651	1,264
R^2	0.27	0.26	0.08	0.13	0.08	0.05	0.05

Source: ENET (2005) for urban and rural Mexico and ACS for Mexican immigrants in the United States arrived a year earlier. Standard errors in smaller font and in parentheses. Coefficients in bold if significant at a 95 per cent confidence level. Experience is computed as age – 16 – (schooling years – 9).

Table 3:

Summary Statistics (2000-2004)

Urban Mexico

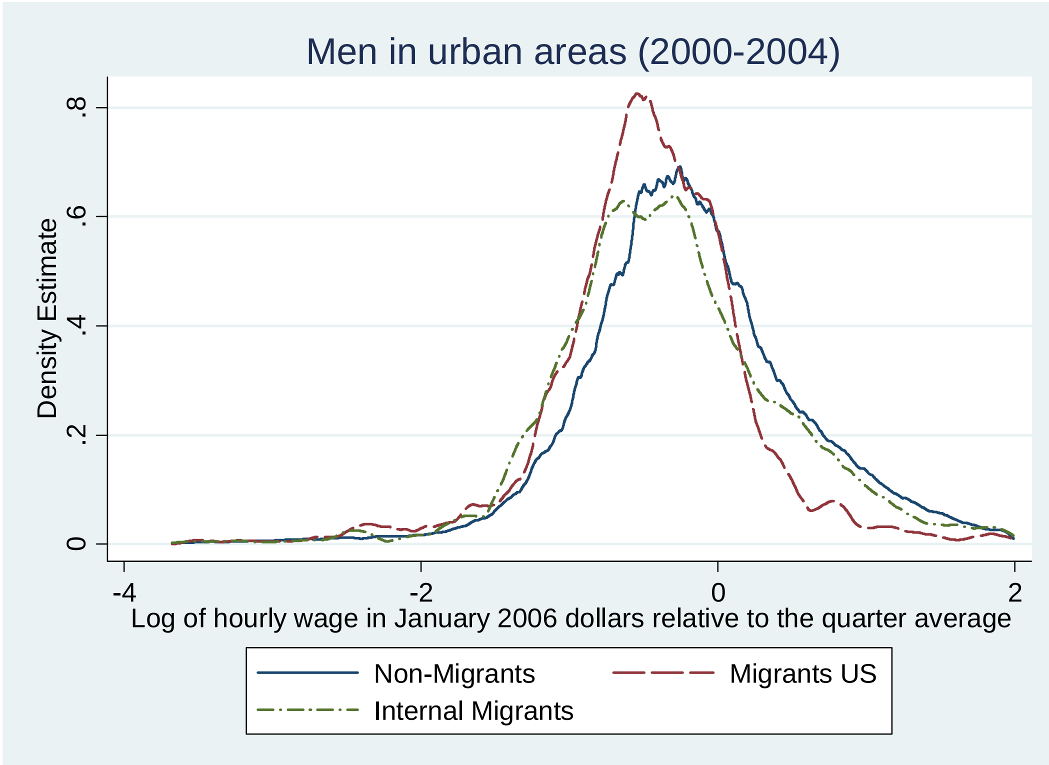
Individuals aged 16 to 65		*High Network Prevalence*			*Low Network Prevalence*		
		Non-Migrants	*Internal Migrants*	*US Migrants*	*Non-Migrants*	*Internal Migrants*	*US Migrants*
Percent Male		47%	60%	77%	47%	63%	79%
		(0.0006)	(0.0098)	(0.0094)	(0.0008)	(0.0113)	(0.0166)
Males							
Age	*Average*	34.7	29.3	30.0	34.8	29.1	28.7
		(0.0248)	(0.2963)	(0.2904)	(0.0296)	(0.3128)	(0.4569)
	Median	33	25	27	33	26	27
Schooling years	*Average*	9.1	9.1	7.7	9.7	9.7	9.1
		(0.0086)	(0.1153)	(0.0881)	(0.0099)	(0.1295)	(0.1620)
	Median	9	9	8	9	9	9
Usable Wage Observations		68%	61%	59%	67%	63%	65%
		(0.0009)	(0.0120)	(0.0127)	(0.0011)	(0.0140)	(0.0224)
Hourly wage in 2006 dollars							
	Average	2.48	2.16	1.64	2.26	2.00	1.71
		(0.0065)	(0.0719)	(0.0423)	(0.0072)	(0.0750)	(0.1621)
	Median	1.75	1.51	1.34	1.52	1.32	1.19
Network prevalence		14%	14%	21%	3%	2%	3%
		(0.0002)	(0.0025)	(0.0031)	(0.0000)	(0.0004)	(0.0007)
Migration rate			0.5%	0.7%		0.6%	0.2%
			(0.0001)	(0.0002)		(0.0002)	(0.0001)
Observations		2,166,536	8,715	7,278	1,544,663	7,098	1,730

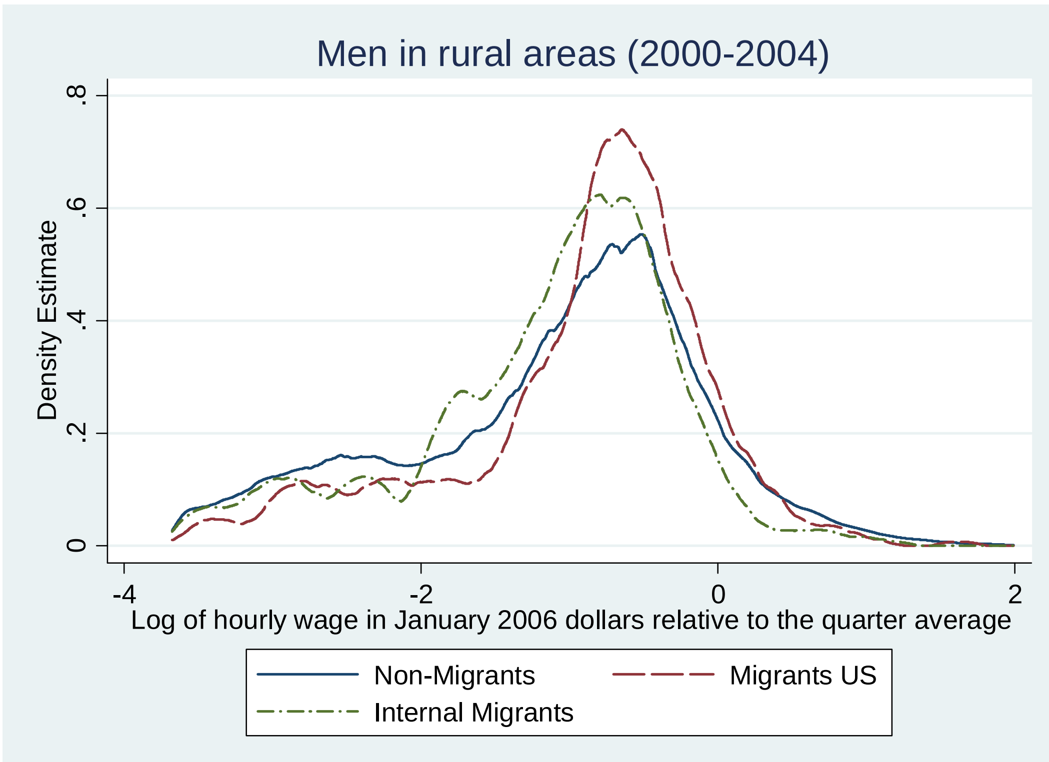
Rural Mexico

Individuals aged 16 to 65		*High Network Prevalence*			*Low Network Prevalence*		
		Non-Migrants	*Internal Migrants*	*US Migrants*	*Non-Migrants*	*Internal Migrants*	*US Migrants*
Percent Male		45%	63%	86%	48%	70%	82%
		(0.0013)	(0.0159)	(0.0083)	(0.0016)	(0.0145)	(0.0246)
Males							
Age	*Average*	36.5	28.0	29.5	35.3	26.2	28.4
		(0.0565)	(0.4739)	(0.2768)	(0.0641)	(0.4034)	(0.7388)
	Median	35	24	27	33	22	27
Schooling years	*Average*	5.5	6.6	6.0	5.3	6.3	6.5
		(0.0151)	(0.1453)	(0.0757)	(0.0171)	(0.1306)	(0.2403)
	Median	6	6	6	6	6	6
Usable Wage Observations		69%	60%	56%	71%	63%	72%
		(0.0018)	(0.0202)	(0.0125)	(0.0021)	(0.0185)	(0.0323)
Hourly wage in 2006 dollars							
	Average	1.26	1.02	1.13	0.85	0.72	0.88
		(0.0065)	(0.0413)	(0.0290)	(0.0059)	(0.0278)	(0.1037)
	Median	1.03	0.90	0.99	0.62	0.62	0.68
Network prevalence		21%	19%	29%	1%	1%	2%
		(0.0005)	(0.0049)	(0.0037)	(0.0001)	(0.0005)	(0.0009)
Migration rate			0.8%	2.4%		1.5%	0.4%
			(0.0003)	(0.0006)		(0.0006)	(0.0003)
Observations		237,765	1,464	2,796	194,062	1,552	378

Source: ENET (2005). See table 1 for details. The network prevalence variable is built as the percentage of individuals older than 15 in a municipality that have migrated to the US according to the ENADID 1997, following the definition in McKenzie and Rapoport (2010). High network prevalence refers to a value above the median Mexican prevalence (5.4 per cent).

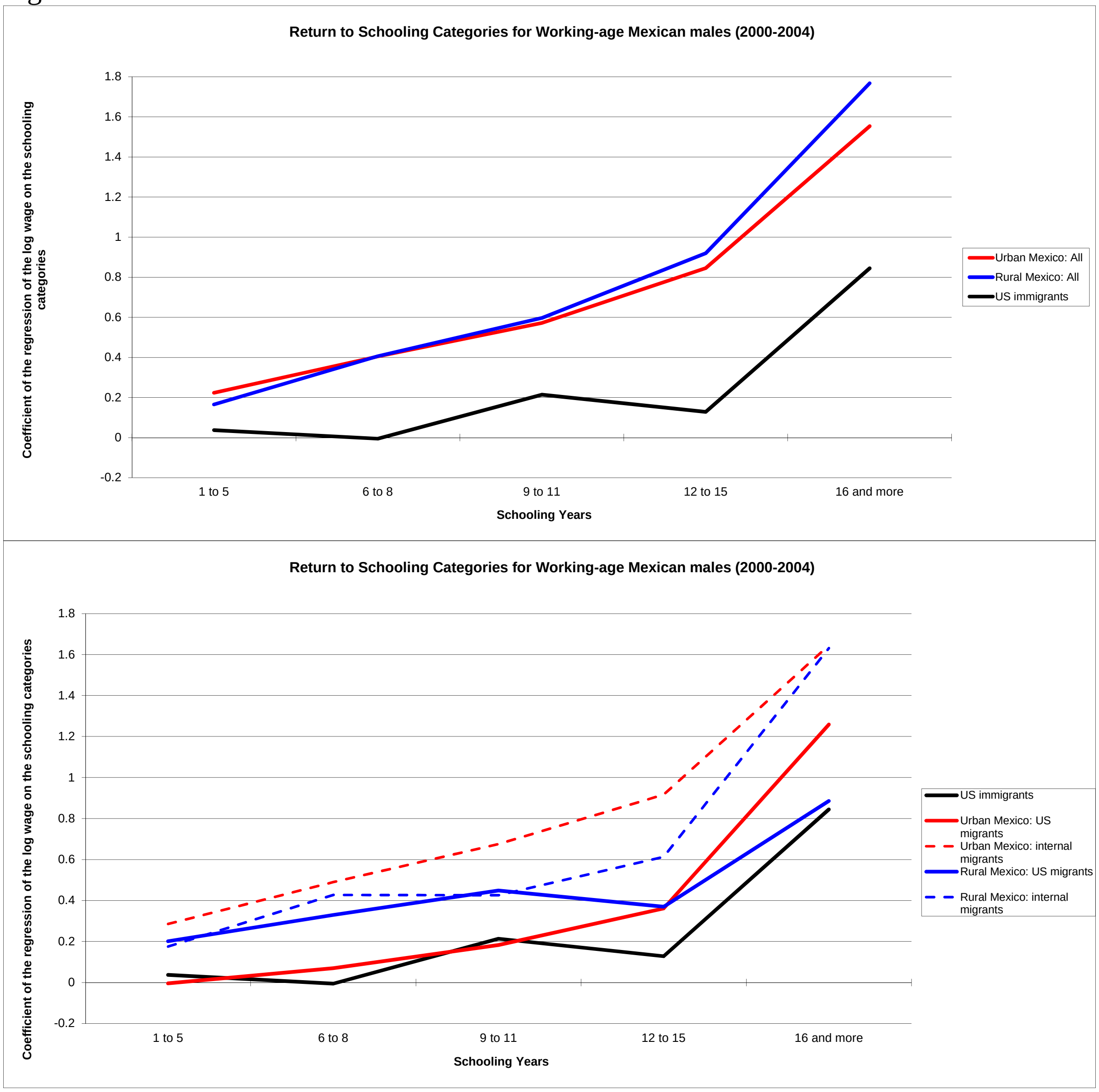
Figure 1:





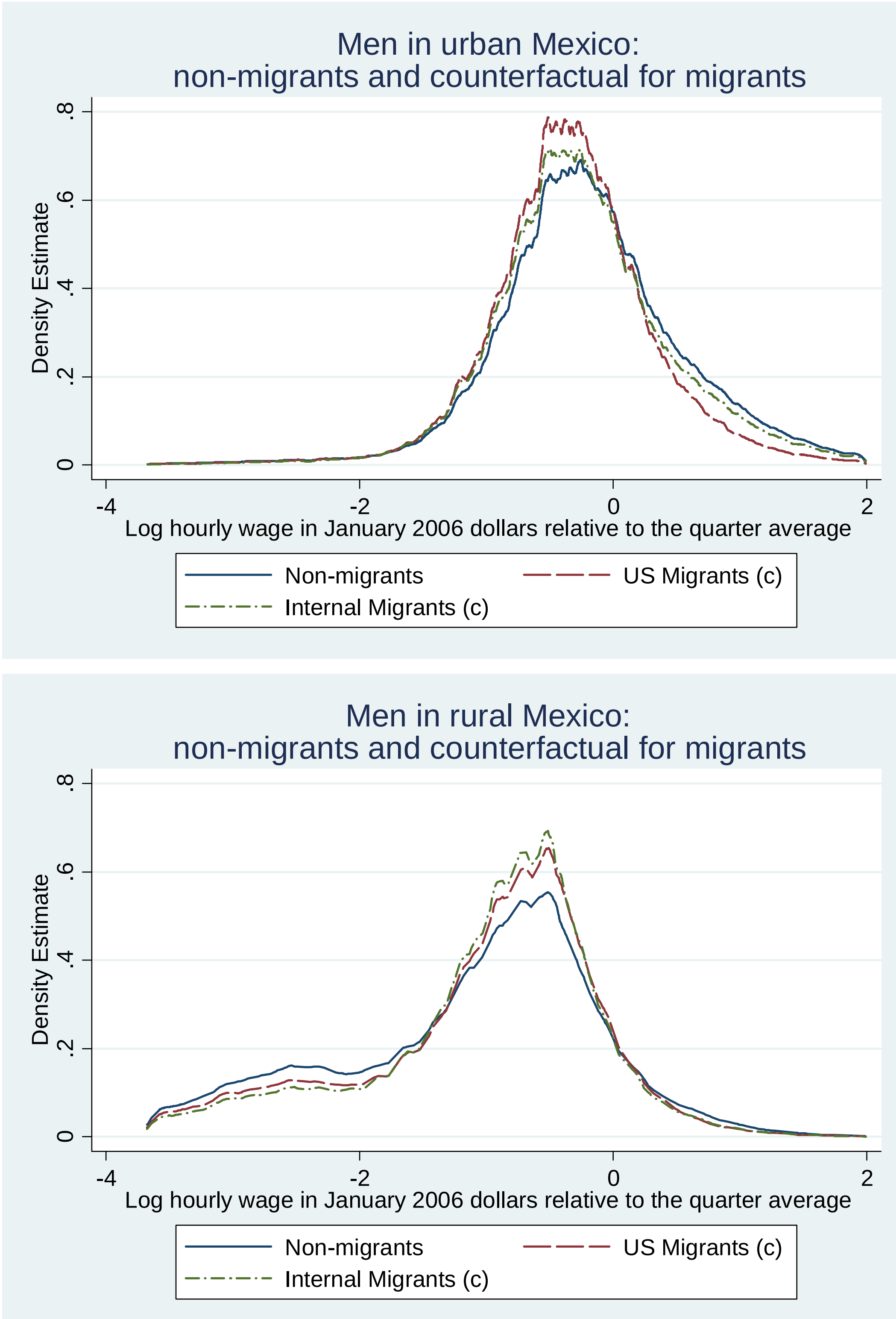
Source: ENET (2005). Log of the hourly wage relative to the quarter average. See table 1 for the construction of wages. For the estimation of the kernel densities, I use an Epanechnikov kernel (Silverman (1986)). To prevent over-smoothing, I follow Leibbrandt, Levinsohn, and McCrary (2005) in using a bandwidth which is .75 times the optimal. I follow Chiquiar and Hanson (2005) in dropping the highest and lowest 0.5 percent of observations to eliminate outliers.

Figure 2:



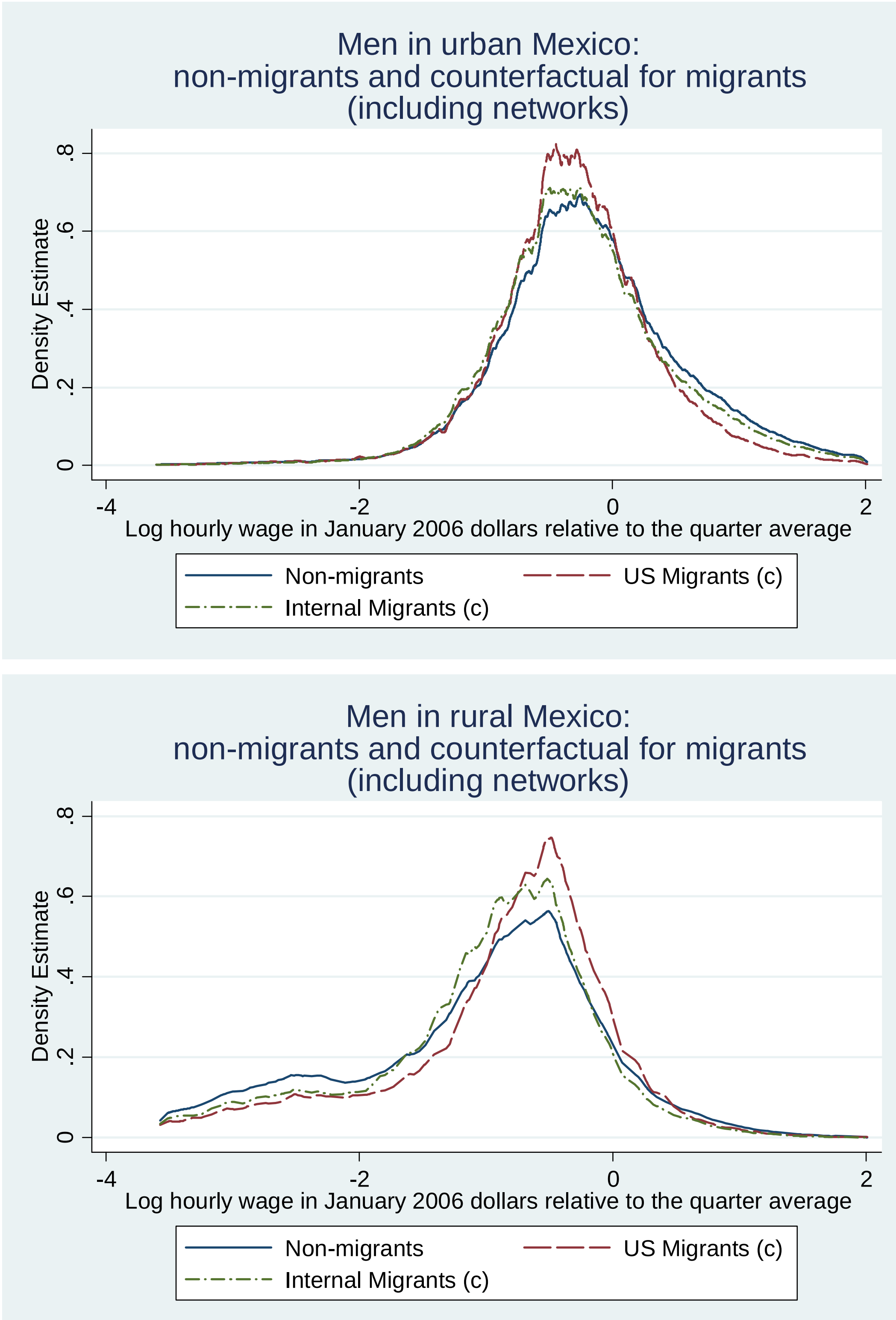
Source: ENET (2005) for urban and rural Mexico and ACS for Mexican immigrants to the US. The figure represents the coefficients from the same regressions as in table 2 but this time substituting the schooling years variable for schooling category dummies, where 0 years of schooling is the excluded category.

Figure 3:



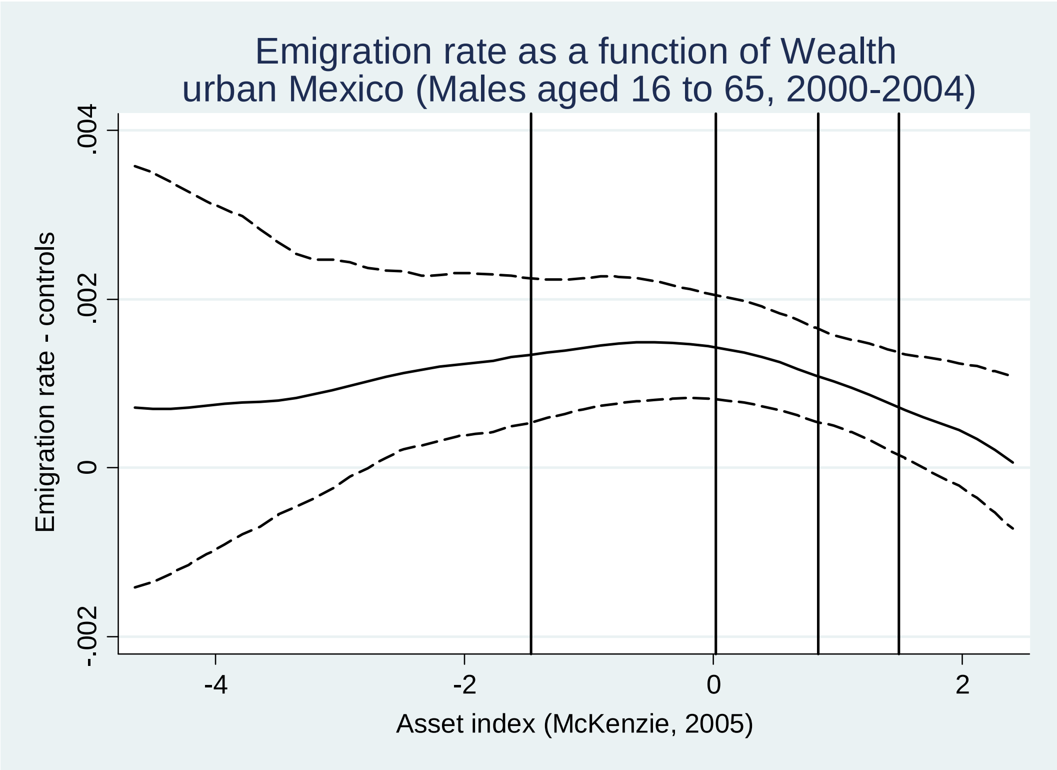
Source: ENET (2005). The counterfactual (emigrant wages based only on their observable characteristics) is estimated following Chiquiar and Hanson (2005).

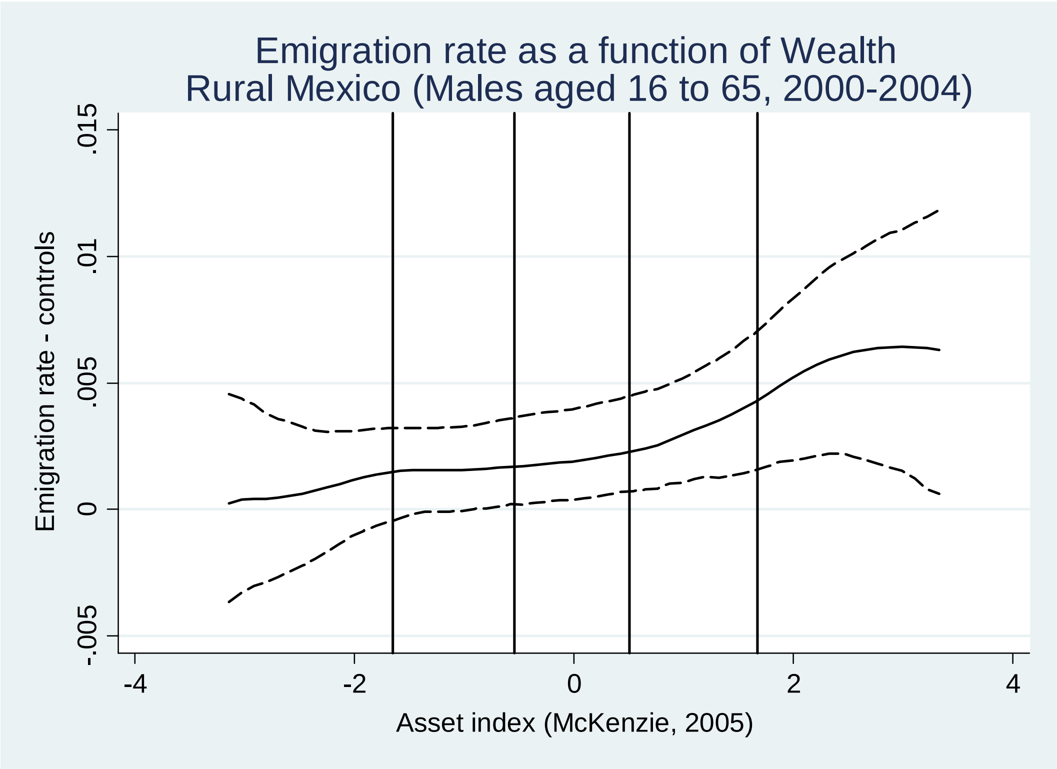
Figure 4:



Source: ENET (2005). See figure 3 for an explanation. The calculation of the counterfactual includes the network variable and its interaction with schooling groups.

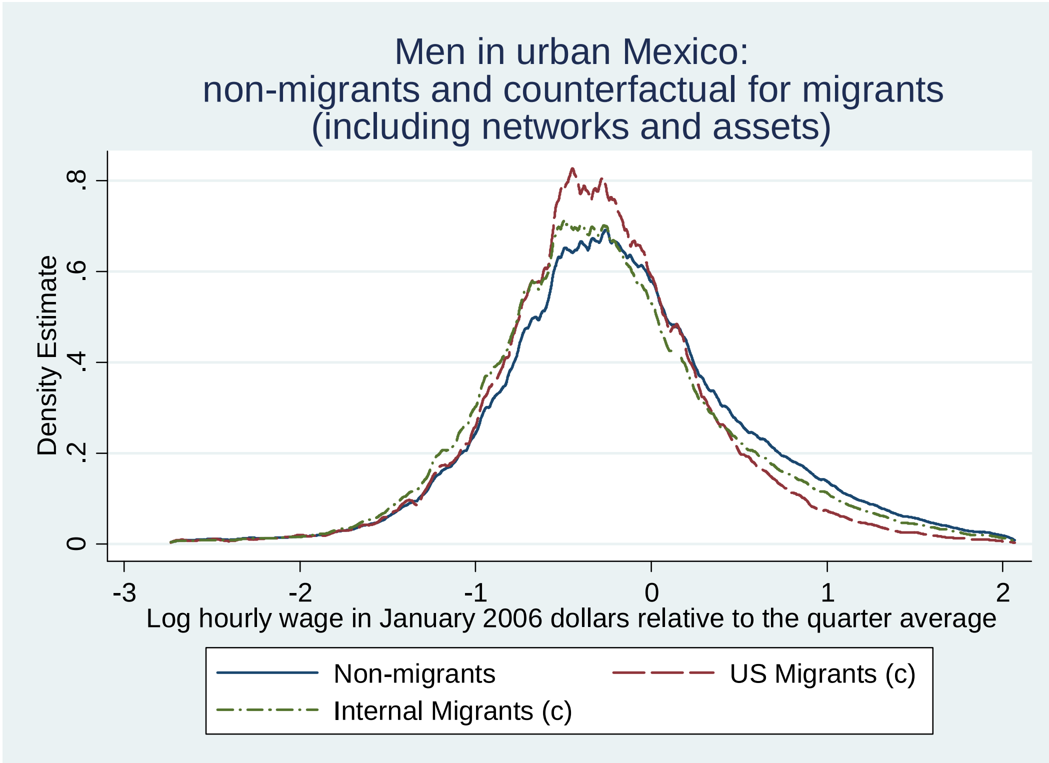
Figure 5:

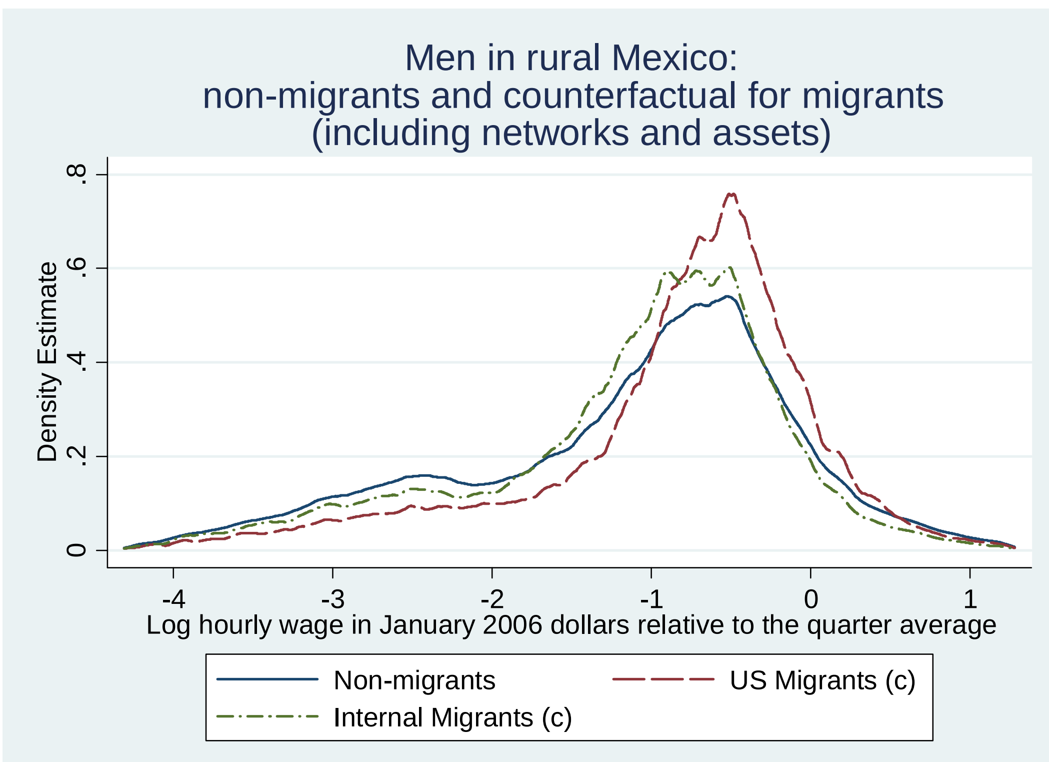




Source: Fan (1992) local linear regression of the emigration rate net of other controls (see table B2) on the McKenzie (2005) versions of the asset index for urban and rural Mexico (see table A3). The dotted lines represent the 90 per cent confidence interval obtained from bootstrapping the procedure 1000 times by randomly sampling with replacement half of the observations. The vertical solid lines represent the position of wealth quintiles in urban and rural Mexico. Following Deaton (1997), the Epanechnikov kernel is used. A bandwidth of 0.2 times the asset index range is chosen. Although they are used for the calculations, the representation drops 2.5 per cent of the highest and lowest asset values.

Figure 6:





Source: ENET (2005). See figure 3 for an explanation. The calculation of the counterfactual includes the network variable, the McKenzie (2005) asset index in the rural and urban version, respectively, their interaction and their interactions with schooling groups.

D Appendix A Tables and Figures

Table A1:

Year	Quarter	After one quarter	After two quarters	After three quarters	After four quarters
		Attrition (non matched individuals from quarter to quarter)			
2000	2	12%	18%	26%	27%
	3	10%	18%	20%	27%
	4	11%	14%	16%	19%
2001	1	7%	12%	16%	19%
	2	7%	12%	16%	19%
	3	8%	12%	18%	20%
	4	8%	13%	23%	29%
2002	1	9%	20%	25%	21%
	2	15%	19%	18%	32%
	3	7%	10%	24%	32%
	4	6%	21%	31%	33%
2003	1	19%	29%	32%	39%
	2	23%	26%	30%	29%
	3	14%	19%	17%	30%
	4	12%	13%	27%	17%
2004	1	7%	24%	15%	
	2	21%	13%		
	3	9%			
Average		11%	17%	22%	26%

Source: ENET (2005)

Table A2:

Summary Statistics for Urban Mexico

Population aged 16 to 65		*More than 100,000 inhabitants*			*15,000-100,000 inhabitants*			*2,500-15,000 inhabitants*		
		Non-Migrants	*Internal Migrants*	*US Emigrants*	*Non-Migrants*	*Internal Migrants*	*US Emigrants*	*Non-Migrants*	*Internal Migrants*	*US Emigrants*
Percent Male		47%	60%	75%	46%	62%	78%	46%	66%	83%
		(0.0006)	(0.0095)	(0.0104)	(0.0014)	(0.0199)	(0.0205)	(0.0012)	(0.0153)	(0.0119)
Males aged 16 to 65										
Age	Average	34.7	30.2	30.1	34.8	28.6	29.0	35.3	27.5	29.3
		(0.0225)	(0.2807)	(0.3135)	(0.0547)	(0.5349)	(0.5931)	(0.0487)	(0.4212)	(0.3791)
	Median	33	26	28	33	26	26	34	23	27
Schooling Years	Average	10.1	10.4	9.2	8.6	8.7	7.5	7.3	7.9	7.1
		(0.0074)	(0.1099)	(0.1123)	(0.0182)	(0.2314)	(0.1762)	(0.0155)	(0.1575)	(0.1139)
	Median	9	9	9	9	9	7	6	9	6
Labor force participation		84%	83%	84%	86%	86%	87%	88%	86%	90%
		(0.0006)	(0.0088)	(0.0108)	(0.0014)	(0.0174)	(0.0193)	(0.0011)	(0.0145)	(0.0100)
Usable wage observations		67%	60%	60%	69%	67%	62%	70%	62%	63%
		(0.0008)	(0.0119)	(0.0145)	(0.0019)	(0.0240)	(0.0267)	(0.0016)	(0.0189)	(0.0165)
Unemployment rate		2.3%	4.9%	6.0%	1.9%	4.6%	7.3%	1.4%	3.1%	2.8%
		(0.0003)	(0.0049)	(0.0071)	(0.0006)	(0.0117)	(0.0169)	(0.0004)	(0.0060)	(0.0051)
Hourly wage in 2006 dollars										
	Average	2.64	2.46	2.03	2.01	1.79	1.54	1.56	1.37	1.29
		(0.0065)	(0.0764)	(0.1172)	(0.0094)	(0.1097)	(0.0896)	(0.0071)	(0.0756)	(0.0409)
	Median	1.80	1.64	1.48	1.45	1.22	1.22	1.19	1.06	1.11
Population share			52%			13%			13%	
			(0.0005)			(0.0004)			(0.0004)	
Observations		2,912,746	11,550	5,931	469,005	2,526	1,466	382,929	2,058	1,753
Females aged 16 to 65										
Age	Average	35.3	27.9	29.9	35.0	26.1	29.2	35.1	25.6	27.0
		(0.0212)	(0.3575)	(0.5255)	(0.0496)	(0.6943)	(1.5268)	(0.0439)	(0.5973)	(0.7385)
	Median	34	24	26	34	22	24	33	22	23
Schooling Years	Average	9.5	9.6	10.3	8.0	9.3	7.9	6.7	8.5	8.4
		(0.0071)	(0.1368)	(0.1740)	(0.0169)	(0.2457)	(0.4874)	(0.0145)	(0.2307)	(0.2907)
	Median	9	9	10	9	9	9	6	9	9
Labor force participation		46%	52%	46%	45%	48%	36%	41%	42%	39%
		(0.0008)	(0.0155)	(0.0235)	(0.0019)	(0.0330)	(0.0466)	(0.0016)	(0.0268)	(0.0363)
Usable wage observations		32%	39%	32%	30%	36%	17%	26%	29%	24%
		(0.0007)	(0.0155)	(0.0231)	(0.0017)	(0.0320)	(0.0312)	(0.0014)	(0.0242)	(0.0320)
Unemployment rate		1.4%	3.3%	2.9%	0.9%	2.8%	1.7%	0.7%	4.5%	2.4%
		(0.0002)	(0.0047)	(0.0062)	(0.0003)	(0.0085)	(0.0076)	(0.0003)	(0.0107)	(0.0130)
Hourly wage in 2006 dollars										
	Average	2.28	1.77	2.17	1.73	1.16	1.07	1.44	1.22	1.18
		(0.0067)	(0.0857)	(0.2098)	(0.0115)	(0.0892)	(0.0867)	(0.0112)	(0.0929)	(0.1895)
	Median	1.58	1.31	1.44	1.17	0.95	1.03	0.98	0.95	0.91
Population share			51%			14%			13%	
			(0.0005)			(0.0004)			(0.0003)	

Source: ENET (2005). See table 1 for details.

Table A3:

Asset Index Construction (ENET 2000-2004)

	Mexico				*Urban Mexico*				*Rural Mexico*			
	Scoring factors				Scoring factors				Scoring factors			
Characteristics	McKenzie (2005)	All ENET	Mean	Standard Deviation	McKenzie (2005)	All ENET	Mean	Standard Deviation	McKenzie (2005)	All ENET	Mean	Standard Deviation
Home owner	-0.05	-0.02	0.70	0.46	0.04	0.08	0.64	0.48	-0.01	0.02	0.86	0.34
Number of rooms	0.30	0.29	3.87	1.71	0.34	0.33	4.09	1.74	0.32	0.33	3.19	1.40
Bathroom	0.23	0.24	0.88	0.33	0.23	0.26	0.92	0.27	0.25	0.25	0.74	0.44
Adobe walls	-0.21	-0.15	0.08	0.28	-0.20	-0.10	0.04	0.20	-0.17	-0.08	0.22	0.41
Brick walls	0.33	0.27	0.83	0.38	0.31	0.21	0.91	0.28	0.37	0.30	0.56	0.50
Cardboard or asbestos roof	-0.32		0.24	0.42	-0.36		0.16	0.37	-0.31		0.47	0.50
Brick roof	0.36	0.30	0.69	0.46	0.39	0.28	0.79	0.41	0.39	0.28	0.39	0.49
Dirt floor	-0.29	-0.23	0.10	0.29	-0.25	-0.16	0.04	0.18	-0.36	-0.34	0.28	0.45
Wood floor	0.30	0.27	0.38	0.48	0.32	0.28	0.47	0.50	0.25	0.14	0.09	0.29
Electricity	0.17	0.13	0.98	0.15	0.10	0.07	0.99	0.08	0.22	0.18	0.93	0.26
Water	0.24	0.19	0.91	0.29	0.18	0.12	0.97	0.18	0.25	0.20	0.72	0.45
Sewerage	0.30	0.25	0.83	0.38	0.25	0.18	0.93	0.25	0.27	0.22	0.50	0.50
Phone	0.30	0.26	0.38	0.49	0.31	0.25	0.48	0.50	0.23	0.17	0.08	0.26
Other utilities	0.19	0.16	0.15	0.36	0.21	0.17	0.19	0.39	0.05	0.04	0.04	0.19
Loft		0.00	0.00	0.02		-0.01	0.00	0.02		-0.01	0.00	0.01
Communal appartment		-0.08	0.06	0.24		-0.19	0.07	0.25		-0.05	0.03	0.17
Appartment Building		0.11	0.11	0.32		0.05	0.15	0.36		-0.01	0.01	0.09
House		-0.04	0.83	0.38		0.07	0.78	0.41		0.05	0.96	0.19
Lent house		-0.05	0.09	0.29		-0.11	0.09	0.29		-0.03	0.08	0.28
Rented house		0.03	0.12	0.33		-0.06	0.16	0.36		0.02	0.02	0.13
Not full ownership		0.06	0.09	0.28		0.05	0.10	0.31		0.00	0.04	0.19
Kitchen		0.17	0.84	0.37		0.24	0.84	0.37		0.19	0.83	0.37
Number of bedrooms		0.26	1.99	1.15		0.30	2.10	1.15		0.31	1.68	1.08
No bathroom		-0.22	0.08	0.27		-0.16	0.03	0.16		-0.24	0.23	0.42
Collective bathroom		-0.09	0.04	0.21		-0.20	0.05	0.22		-0.05	0.02	0.16
Cardboard walls		-0.05	0.00	0.06		-0.07	0.00	0.06		-0.04	0.00	0.05
Metal or asbestos walls		-0.05	0.01	0.07		-0.07	0.00	0.07		-0.04	0.01	0.08
Wooden walls		-0.20	0.07	0.26		-0.15	0.03	0.18		-0.26	0.19	0.39
Cardboard roof		-0.14	0.04	0.19		-0.15	0.03	0.17		-0.16	0.07	0.26
Asbestos roof		-0.21	0.20	0.40		-0.21	0.13	0.34		-0.11	0.40	0.49
Wooden roof		-0.11	0.07	0.25		-0.08	0.05	0.22		-0.11	0.12	0.33
Cement floor		-0.13	0.53	0.50		-0.22	0.50	0.50		0.23	0.63	0.48
House older than 20 years		0.05	0.32	0.47		0.06	0.32	0.47		0.05	0.30	0.46
House 10 to 20 years		0.04	0.33	0.47		0.04	0.34	0.47		0.02	0.31	0.46
House 5 to 10 years		-0.04	0.20	0.40		-0.04	0.19	0.39		-0.03	0.22	0.41
House 1 to 5 years		-0.07	0.11	0.31		-0.08	0.10	0.30		-0.06	0.14	0.35
House less than 1 year		-0.02	0.01	0.09		-0.02	0.01	0.09		-0.03	0.01	0.10
McKenzie (2005) Asset Index			0	2.09			0	1.92			0	1.80
All ENET Asset Index			0	2.37			0	2.30			0	2.09
Observations:	2,760,359	2,760,359			2,399,821	2,399,821			360,538	360,538		
Eigenvalue for first component	4.3873	5.6388			3.6881	5.2986			3.2232	4.3722		
Share of variance	0.3134	0.1566			0.2634	0.1471			0.2302	0.1215		

Source: ENET (2005). Observations from the second quarter for the period 2000-2004. Principal components analysis and construction of a household wealth index.

Table A4:

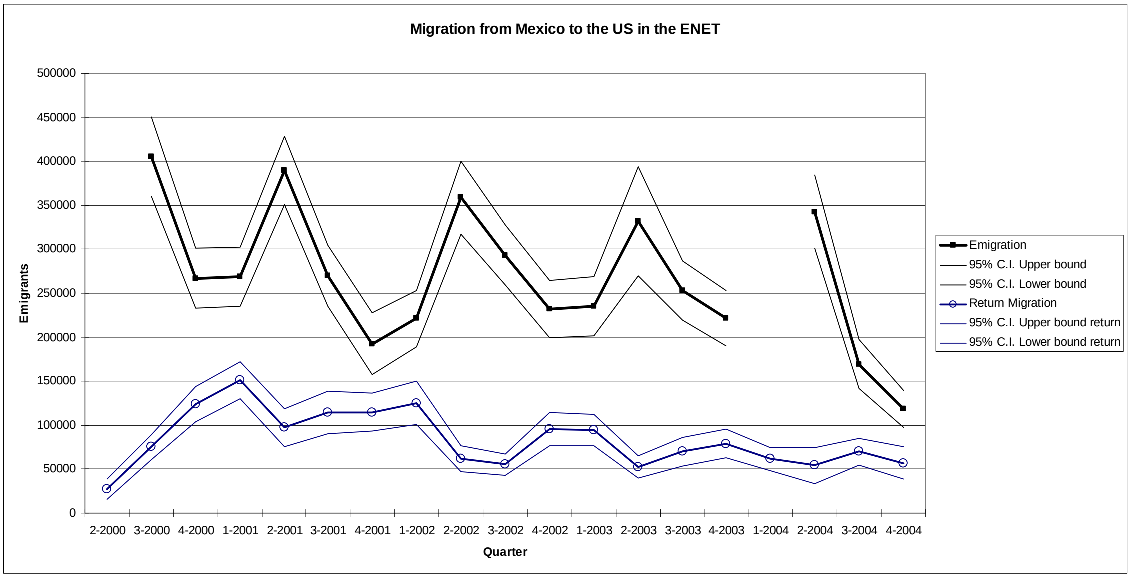
Correlation between different versions of the asset index

Urban Mexico		Overall Mexico		Urban Mexico	
		All ENET	McKenzie (2005)	All ENET	McKenzie (2005)
Overall	All ENET	1.00			
Mexico	McKenzie (2005)	0.97	1.00		
Urban	All ENET	0.98	0.93	1.00	
Mexico	McKenzie (2005)	0.96	1.00	0.93	1.00

Rural Mexico		Overall Mexico		Rural Mexico	
		All ENET	McKenzie (2005)	All ENET	McKenzie (2005)
Overall	All ENET	1.00			
Mexico	McKenzie (2005)	0.96	1.00		
Rural	All ENET	0.97	0.92	1.00	
Mexico	McKenzie (2005)	0.96	0.99	0.93	1.00

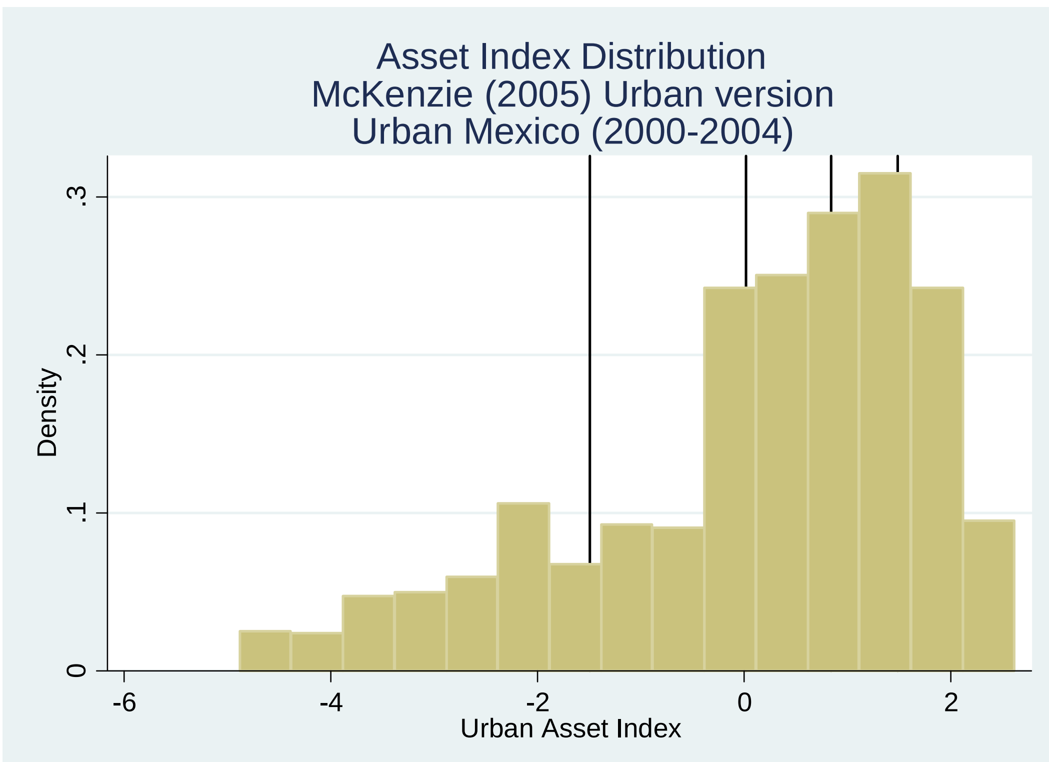
Source: ENET (2005). See table A.3 for details on the construction of the different versions of the index. The two versions that are used in the text are underlined.

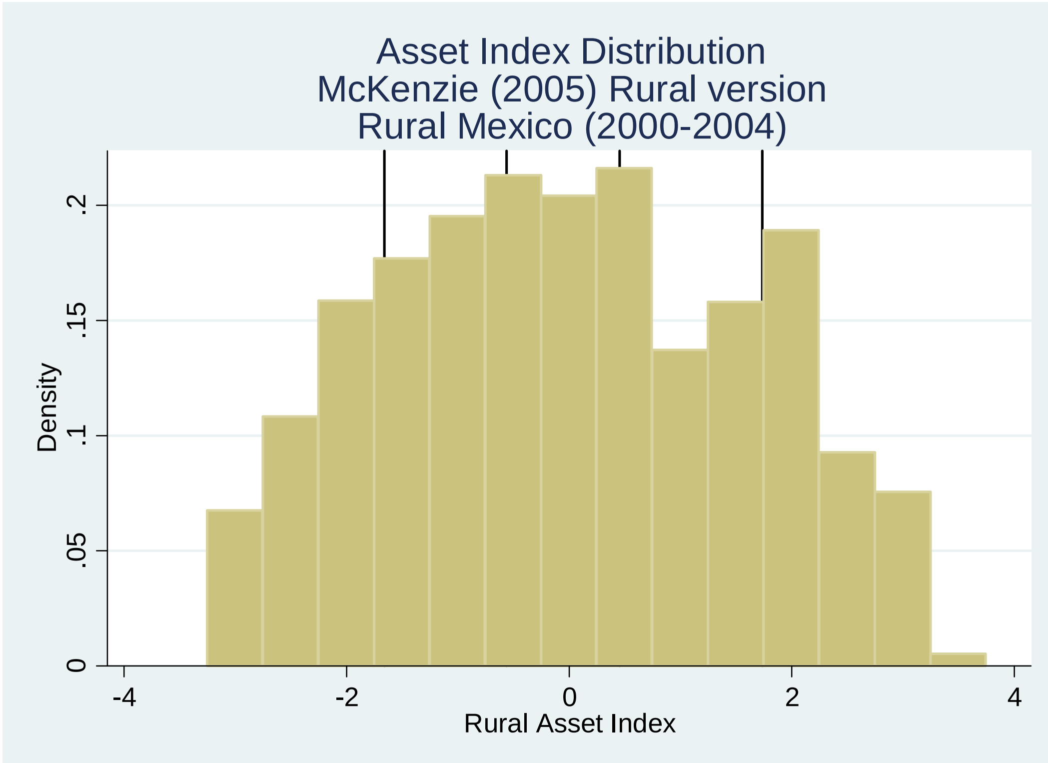
Figure A1:



Source: ENET (2005)

Figure A2:





Source: ENET (2005). See table A3 for the construction of the two versions of the asset index. The vertical solid lines represent the position of wealth quintiles in urban and rural Mexico.

E Appendix B Tables and Figures

Table B1:

Summary Statistics (US sources on recent Mexican immigrants)

Individuals aged 16 to 65		US Census 2000 *Arrived a year earlier*	ACS 2000 *Arrived a year earlier*	ACS 2000-2004 *Arrived a year earlier*
Percent Male		62%	66%	64%
		(0.00)	(0.03)	(0.01)
Men				
Age				
	Average	26.6	27.2	28.0
		(0.09)	(0.89)	(0.30)
	Median	24	24	25
Schooling years				
	Average	8.6	8.9	8.9
		(0.04)	(0.33)	(0.12)
	Median	9	9	9
Hourly wage in 2006 dollars				
	Average	10.16	8.10	9.75
		(0.12)	(0.44)	(0.28)
	Median	7.94	7.14	7.75
Observations		21,930	244	2,658

Source: Ruggles et al. (2004)

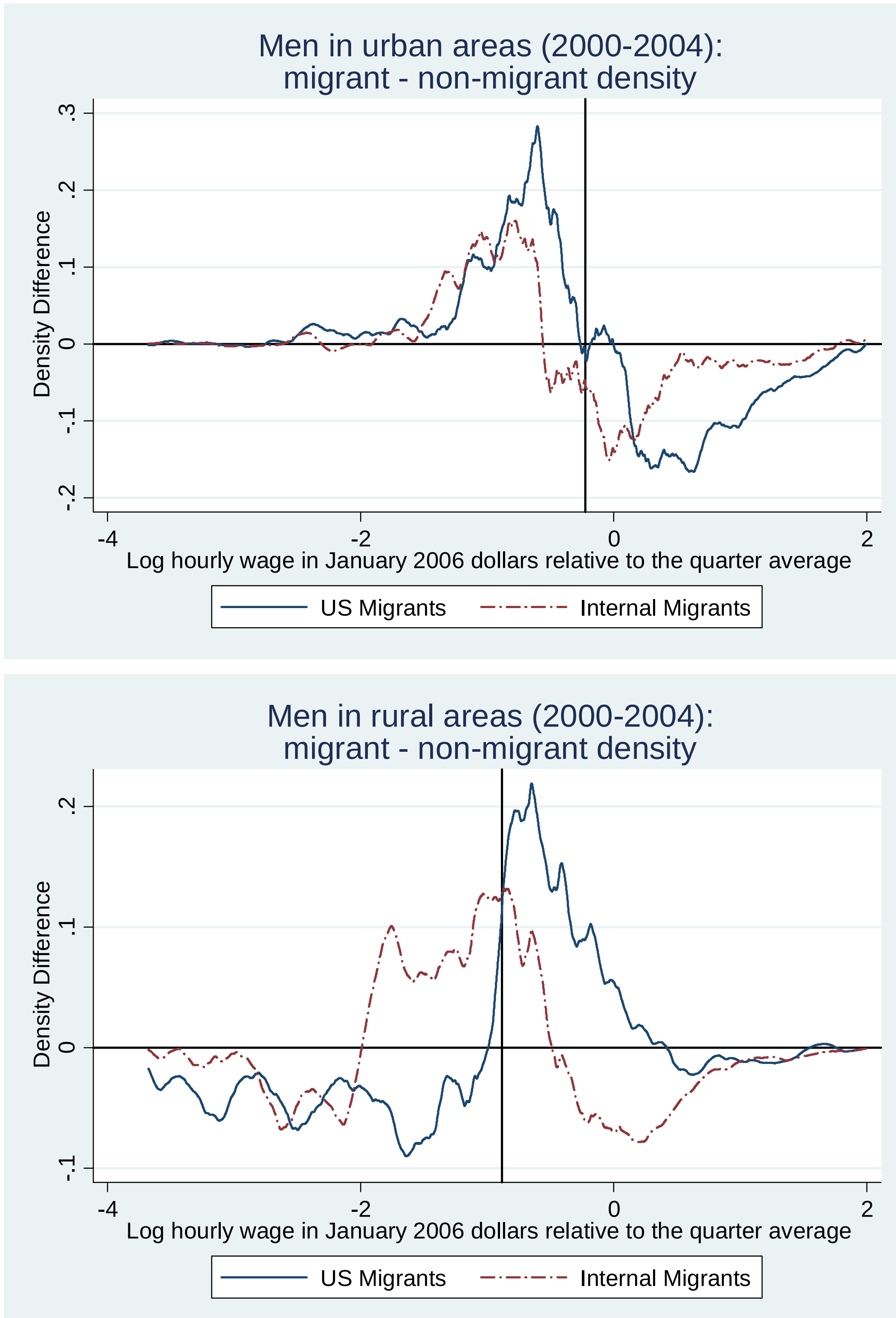
Table B2:

Summary Statistics

Mexican Men aged 16 to 65 years old	Urban Mexico		Rural Mexico	
	Mean or Proportion	Standard Deviation	Mean or Proportion	Standard Deviation
Emigrant to the US the following quarter	0.0048	0.0002	0.0129	0.0008
Emigrant to another state the following quarter	0.0054	0.0003	0.0117	0.0008
Household Asset Index (McKenzie, 2005)	0.0023	0.0066	0.0136	0.0123
Log hourly wage	0.5434	0.0027	-0.4292	0.0077
Schooling years	9.1734	0.0155	5.0405	0.0261
Age	35.8566	0.0412	37.9716	0.0931
Network prevalence (ENADID 1997)	0.0782	0.0003	0.1052	0.0009
Metropolitan Area	0.4554	0.0017	0.0167	0.0007
Rural Area	0.0000	0.0000	1.0000	0.0000
Distance to the border (km.)	640.3462	0.8385	747.8026	1.6307
Married	0.7142	0.0015	0.7632	0.0029
Household Size	4.8465	0.0071	5.4530	0.0173
Household head	0.6742	0.0016	0.7152	0.0031
Spouse	0.0117	0.0004	0.0108	0.0007
Offspring	0.2433	0.0015	0.2267	0.0029
Other household members	0.0708	0.0009	0.0473	0.0015
Quarters:				
2-2000	0.2234	0.0014	0.2867	0.0031
2-2001	0.1987	0.0013	0.1941	0.0028
2-2002	0.2209	0.0014	0.2323	0.0030
2-2003	0.1846	0.0014	0.1793	0.0028
2-2004	0.1723	0.0014	0.1077	0.0020
States:				
Aguascalientes	0.0095	0.0001	0.0080	0.0002
Baja California	0.0382	0.0004	0.0141	0.0005
Baja California Sur	0.0057	0.0001	0.0054	0.0002
Campeche	0.0080	0.0001	0.0102	0.0003
Coahuila	0.0311	0.0004	0.0097	0.0004
Colima	0.0065	0.0001	0.0041	0.0001
Chiapas	0.0239	0.0005	0.0810	0.0023
Chihuahua	0.0373	0.0006	0.0193	0.0007
Distrito Federal	0.1178	0.0013	0.0010	0.0001
Durango	0.0134	0.0002	0.0207	0.0006
Guanajuato	0.0432	0.0005	0.0618	0.0019
Guerrero	0.0223	0.0004	0.0533	0.0013
Hidalgo	0.0145	0.0005	0.0441	0.0011
Jalisco	0.0605	0.0008	0.0308	0.0014
México	0.1625	0.0016	0.0789	0.0022
Michoacán	0.0294	0.0007	0.0534	0.0017
Morelos	0.0158	0.0002	0.0098	0.0004
Nayarit	0.0072	0.0001	0.0136	0.0004
Nuevo León	0.0589	0.0006	0.0106	0.0004
Oaxaca	0.0190	0.0004	0.0701	0.0019
Puebla	0.0461	0.0006	0.0536	0.0018
Querétaro	0.0126	0.0002	0.0207	0.0006
Quintana Roo	0.0130	0.0002	0.0089	0.0003
San Luis Potosí	0.0182	0.0003	0.0349	0.0010
Sinaloa	0.0245	0.0004	0.0316	0.0010
Sonora	0.0246	0.0004	0.0161	0.0006
Tabasco	0.0165	0.0003	0.0520	0.0011
Tamaulipas	0.0297	0.0004	0.0163	0.0007
Tlaxcala	0.0100	0.0002	0.0107	0.0003
Veracruz	0.0517	0.0010	0.1259	0.0030
Yucatán	0.0207	0.0003	0.0160	0.0005
Zacatecas	0.0079	0.0002	0.0135	0.0005
Observations	321,541		37,786	

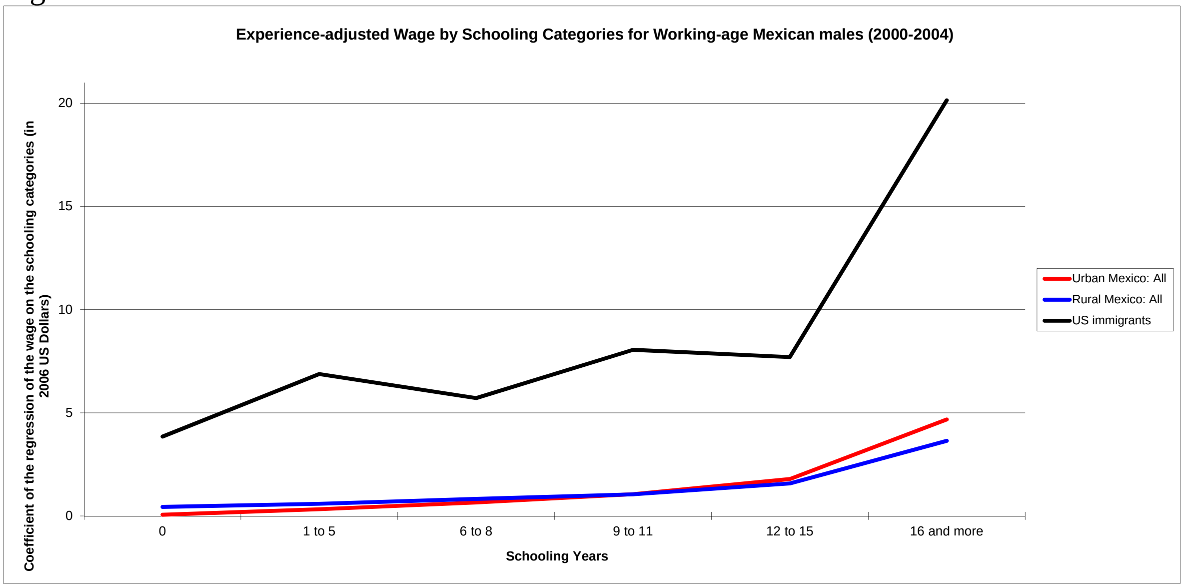
Source: ENET (2005) and ENADID 1997 for the network variable. Distance to the border calculated with data from the Center for International Earth Science Information Network (CIESIN), Columbia University, 2000. US-Mexico DDViewer, 3.1. Palisades, NY: CIESIN, Columbia University. Available at: http://plue.sedac.ciesin.org/plue/ddviewer/ddv30-USMEX/. The household asset index corresponds to the urban and rural versions, respectively, calculated in columns 5 and 9 of table A3. The difference in the summary statistics is due to the fact that only observations with valid values for all the variables (wages in particular) are included in this table.

Figure B1:



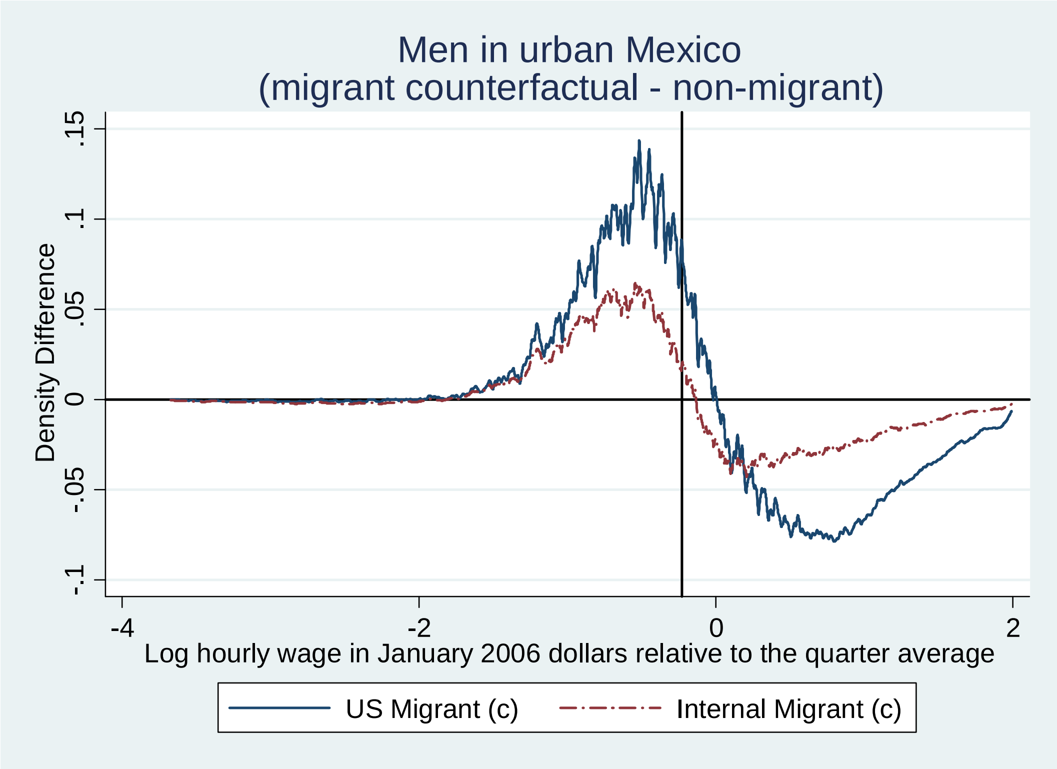
Source: ENET (2005). Migrant minus non-migrant wage densities computed in figure 1. See figure 1 for an explanation. The solid black vertical line represents the median of the log of the relative wage distribution.

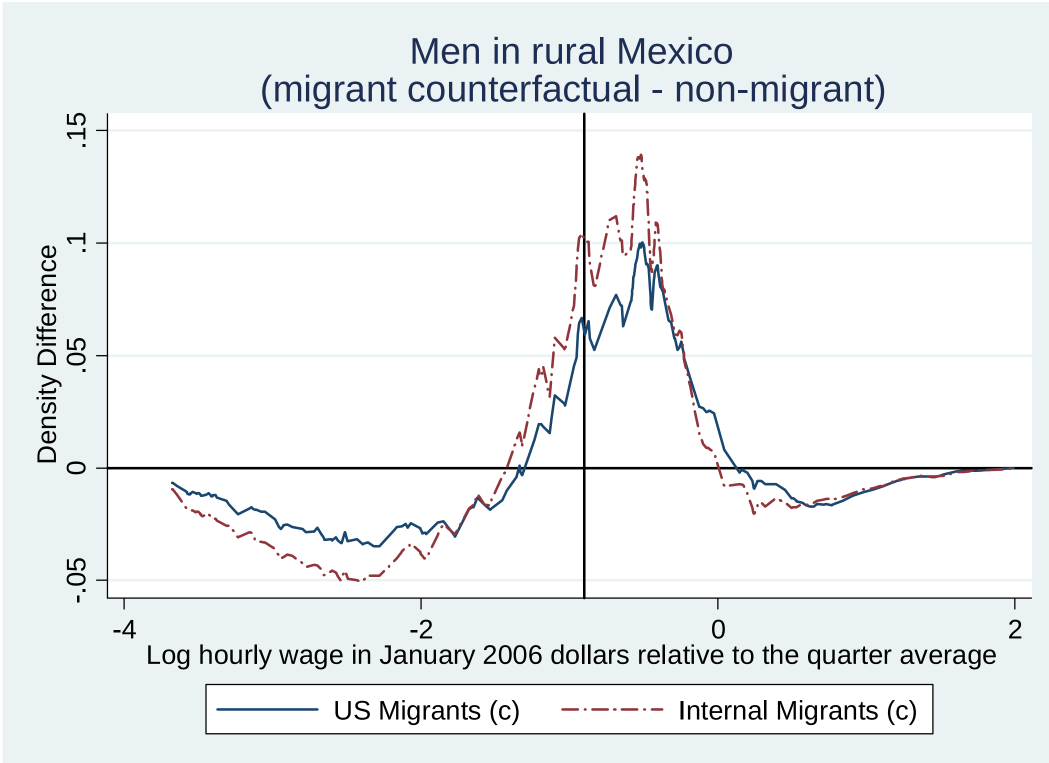
Figure B2:



Source: ENET (2005) for urban and rural Mexico and ACS for Mexican immigrants to the US. The figure represents the coefficients (adding the constant) from the same regressions as in figure 2 but this time substituting the log wage for the absolute wage as the dependent variable.

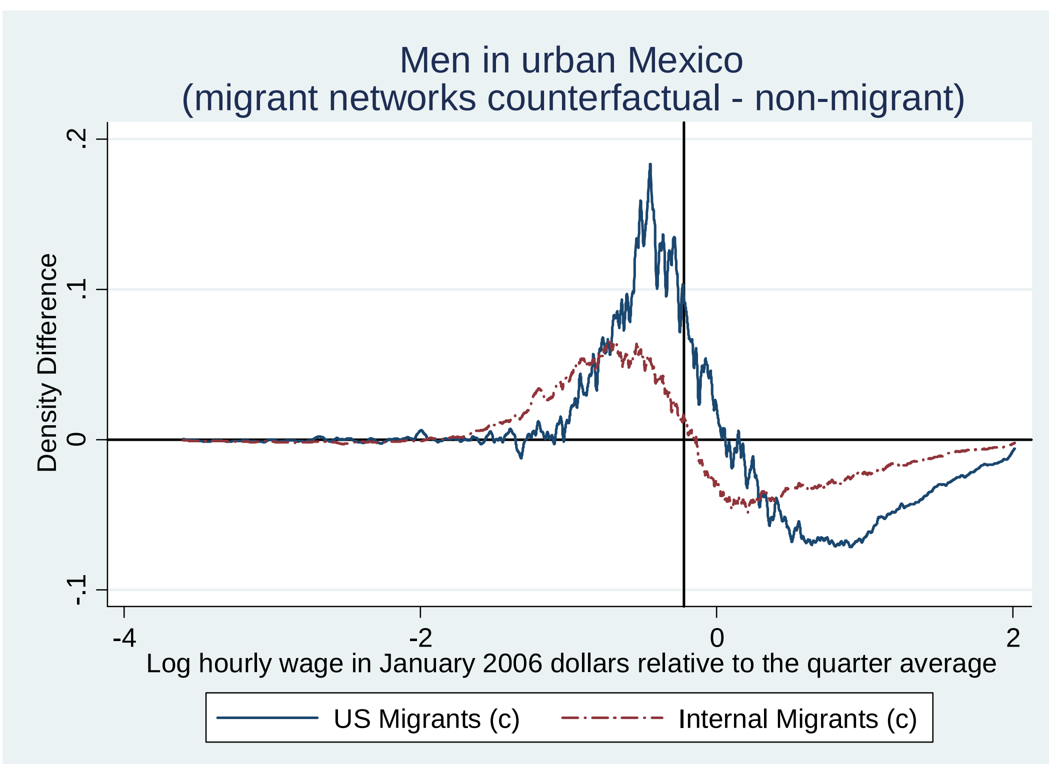
Figure B3:

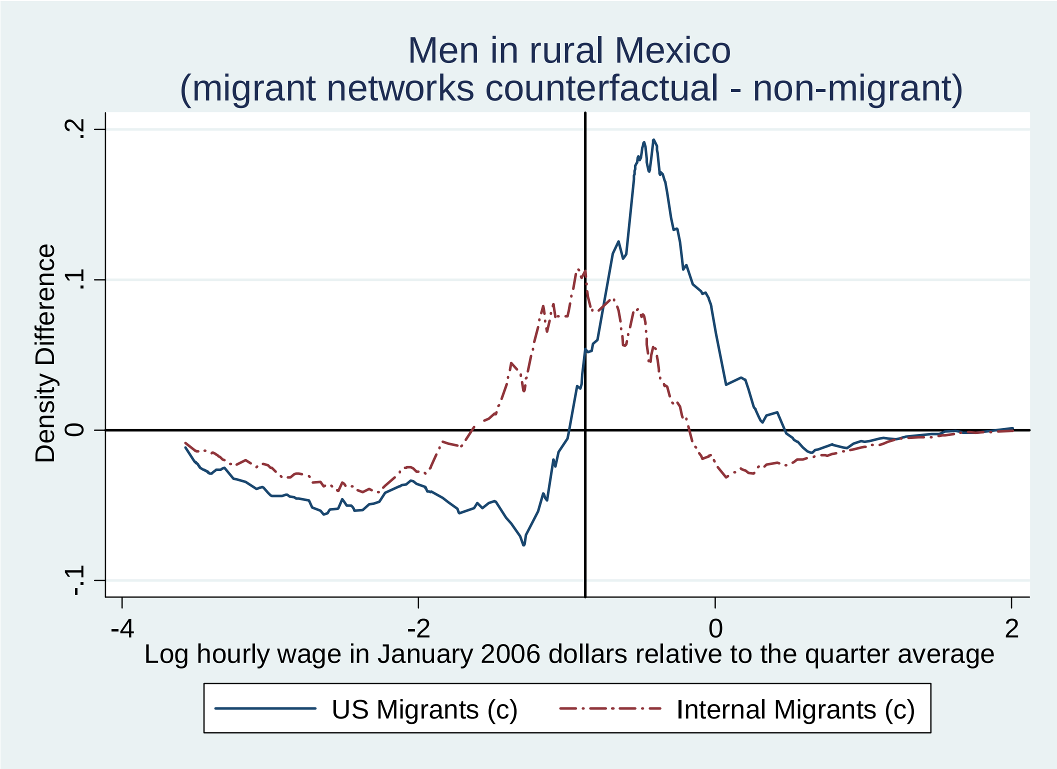




Source: ENET (2005). Counterfactual migrant wage density minus actual non-migrant wage density computed in figure 3. See figure 3 for an explanation. The solid black vertical line represents the median of the log of the relative wage distribution.

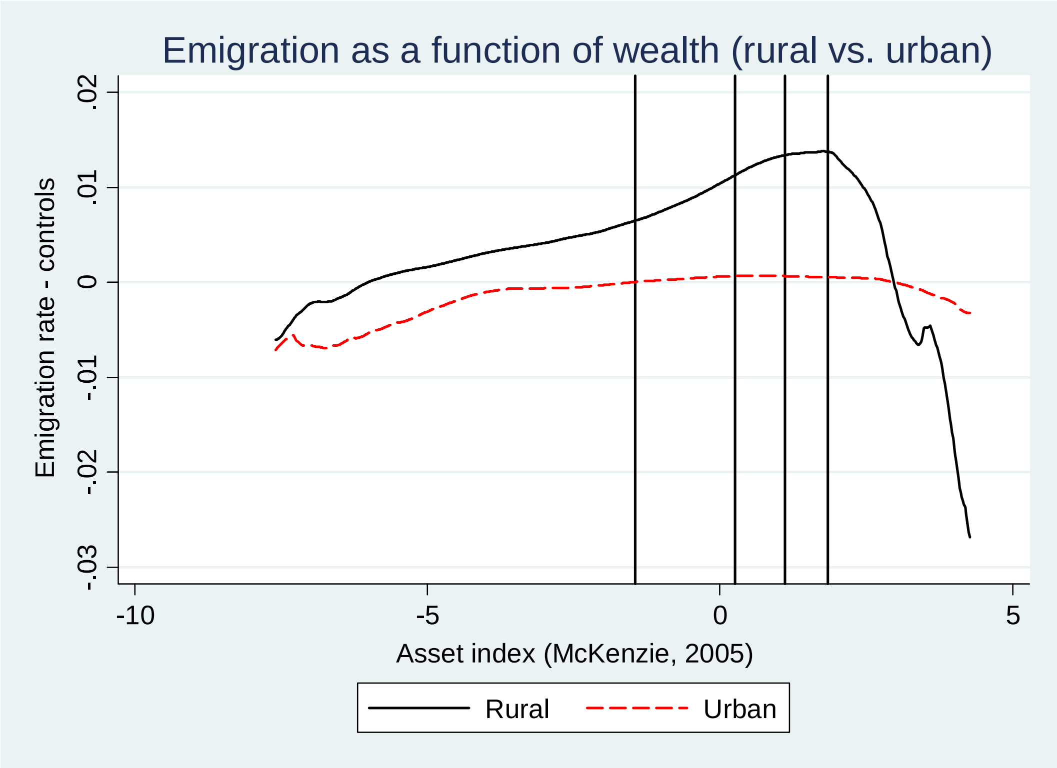
Figure B4:





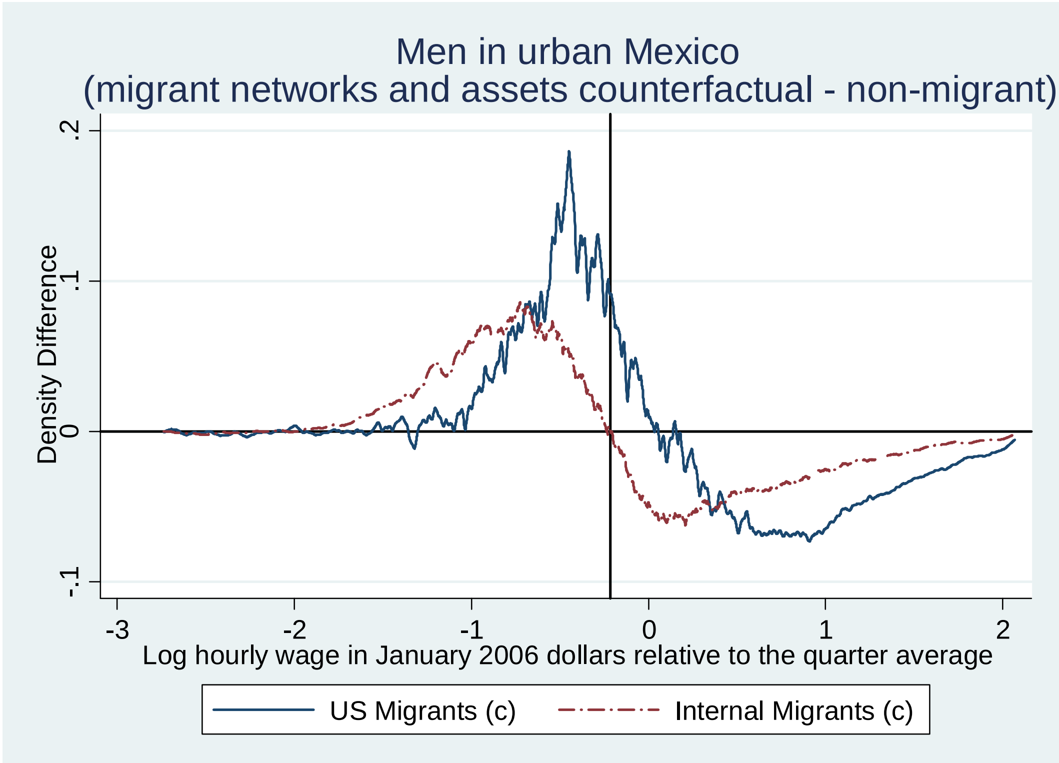
Source: ENET (2005). Counterfactual migrant wage density minus actual non-migrant wage density computed in figure 4. See figure 4 for an explanation. The solid black vertical line represents the median of the log of the relative wage distribution.

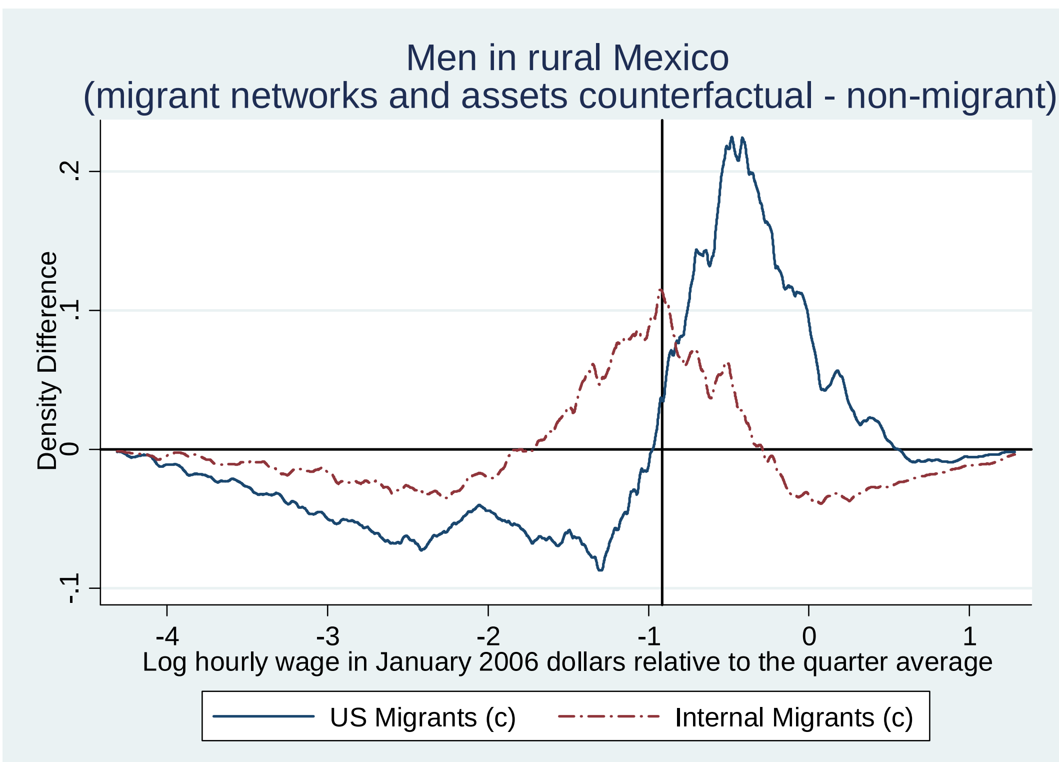
Figure B5:



Source: Fan (1992) local linear regression of the emigration rate net of other controls (see table B2) on the McKenzie (2005) overall Mexico version of the asset index (see table A3). Following Deaton (1997), the Epanechnikov kernel is used. A bandwidth of 0.2 times the asset index range is chosen. The solid vertical lines represent the situation of the wealth quintiles for Mexico as a whole.

Figure B6:





Source: ENET (2005). Counterfactual migrant wage density minus actual non-migrant wage density computed in figure 6. See figure 6 for an explanation. The solid black vertical line represents the median of the log of the relative wage distribution.

ÚLTIMOS DOCUMENTOS DE TRABAJO

2013-02: "Understanding Different Migrant Selection Patterns in Rural and Urban Mexico", **Jesús Fernández-Huertas Moraga**.
2013-01: "Publicizing the results of standardized external tests: Does it have an effect on school outcomes?, **Brindusa Anghel, Antonio Cabrales, Jorge Sainz y Ismael Sanz**.
2012-12: "Visa Policies, Networks and the Cliff at the Border", **Simone Bertoli, Jesús Fernández-Huertas Moraga**.
2012-11: "Intergenerational and Socioeconomic Gradients of Child Obesity", **Joan Costa-Fonta y Joan Gil**.
2012-10: "Subsidies for resident passengers in air transport markets", **Jorge Valido, M. Pilar Socorro, Aday Hernández y Ofelia Betancor**.
2012-09: "Dual Labour Markets and the Tenure Distribution: Reducing Severance Pay or Introducing a Single Contract?", **J. Ignacio García Pérez y Victoria Osuna**.
2012-08: "The Influence of BMI, Obesity and Overweight on Medical Costs: A Panel Data Approach", **Toni Mora, Joan Gil y Antoni Sicras-Mainar**.
2012-07: "Strategic behavior in regressions: an experimental", **Javier Perote, Juan Perote-Peña y Marc Vorsatz**.
2012-06: "Access pricing, infrastructure investment and intermodal competition", **Ginés de Rus y M. Pilar Socorro**.
2012-05: "Trade-offs between environmental regulation and market competition: airlines, emission trading systems and entry deterrence", **Cristina Barbot, Ofelia Betancor, M. Pilar Socorro y M. Fernanda Viecens.**
2012-04: "Labor Income and the Design of Default Portfolios in Mandatory Pension Systems: An Application to Chile", **A. Sánchez Martín, S. Jiménez Martín, D. Robalino y F. Todeschini**.
2012-03: "Spain 2011 Pension Reform", **J. Ignacio Conde-Ruiz y Clara I. Gonzalez**.
2012-02: "Study Time and Scholarly Achievement in PISA", **Zöe Kuehn y Pedro Landeras**.
2012-01: "Reforming an Insider-Outsider Labor Market: The Spanish Experience", **Samuel Bentolila, Juan J. Dolado y Juan F. Jimeno**.
2011-13: "Infrastructure investment and incentives with supranational funding", **Ginés de Rus y M. Pilar Socorro**.
2011-12: "The BCA of HSR. Should the Government Invest in High Speed Rail Infrastructure?", **Ginés de Rus**.
2011-11: "La rentabilidad privada y fiscal de la educación en España y sus regiones", **Angel de la Fuente y Juan Francisco Jimeno**.
2011-10: "Tradable Immigration Quotas", **Jesús Fernández-Huertas Moraga y Hillel Rapoport**.
2011-09: "The Effects of Employment Uncertainty and Wealth Shocks on the Labor Supply and Claiming Behavior of Older American Workers", **Hugo Benítez-Silva, J. Ignacio García-Pérez y Sergi Jiménez-Martín**.
2011-08: "The Effect of Public Sector Employment on Women's Labour Martket Outcomes", **Brindusa Anghel, Sara de la Rica y Juan J. Dolado**.
2011-07: "The peer group effect and the optimality properties of head and income taxes", **Francisco Martínez-Mora**.
2011-06: "Public Preferences for Climate Change Policies: Evidence from Spain", **Michael Hanemann, Xavier Labandeira y María L. Loureiro.**
2011-05: "A Matter of Weight? Hours of Work of Married Men and Women and Their Relative Physical Attractiveness", **Sonia Oreffice y Climent Quintana-Domeque**.
2011-04: "Multilateral Resistance to Migration", **Simone Bertoli y Jesús Fernández-Huertas Moraga**.
2011-03: "On the Utility Representation of Asymmetric Single-Peaked Preferences", **Francisco Martínez Mora y M. Socorro Puy**.
2011-02: "Strategic Behaviour of Exporting and Importing Countries of a Non-Renewable Natural Resource: Taxation and Capturing Rents", **Emilio Cerdá y Xiral López-Otero**.
2011-01: "Politicians' Luck of the Draw: Evidence from the Spanish Christmas Lottery", **Manuel F. Bagues y Berta Esteve-Volart**.
2010-31: "The Effect of Family Background on Student Effort", **Pedro Landeras**.
2010-29: "Random–Walk–Based Segregation Measures", **Coralio Ballester y Marc Vorsatz**.
2010-28: "Incentives, resources and the organization of the school system", **Facundo Albornoz, Samuel Berlinski y Antonio Cabrales**.
2010-27: "Retirement incentives, individual heterogeneity and labour transitions of employed and unemployed workers", **J. Ignacio García Pérez, Sergi Jimenez-Martín y Alfonso R. Sánchez-Martín**.
2010-26: "Social Security and the job search behavior of workers approaching retirement", **J. Ignacio García Pérez y Alfonso R. Sánchez Martín**.
2010-25: "A double sample selection model for unmet needs, formal care and informal caregiving hours of dependent people in Spain", **Sergi Jiménez-Martín y Cristina Vilaplana Prieto**.
2010-24: "Health, disability and pathways into retirement in Spain", **Pilar García-Gómez, Sergi Jiménez-Martín y Judit Vall Castelló**.