Centre for Research and Analysis of Migration

CReAM

Discussion Paper Series

CDP No 01/07

Self-selection patterns in Mexico-U.S. migration: The role of migration networks

David McKenzie and Hillel Rapoport

Centre for Research and Analysis of Migration
Department of Economics, University College London
Drayton House, 30 Gordon Street, London WC1H 0AX

CReAM Discussion Paper No 01/07

Self-selection patterns in Mexico-U.S. migration: The role of migration networks

David McKenzie* and Hillel Rapoport†

*World Bank, Development Economics Research Group.

† Department of Economics, Bar-Ilan University, EQUIPPE-CADRE, University of Lille II and CReAM, University College London

Non-Technical Abstract

This paper examines the role of migration networks in determining self-selection patterns of Mexico-U.S. migration. We first present a simple theoretical framework showing how such networks impact on migration incentives at different education levels and, consequently, how they are likely to affect the expected skill composition of migration. Using survey data from Mexico, we then show that the probability of migration is increasing with education in communities with low migrant networks, but decreasing with education in communities with high migrant networks. This is consistent with positive self-selection of migrants being driven by high migration costs, as advocated by Chiquiar and Hanson (2005), and with negative self-selection of migrants being driven by lower returns to education in the U.S. than in Mexico, as advocated by Borjas (1987).

Centre for Research and Analysis of Migration
Department of Economics, Drayton House, 30 Gordon Street, London WC1H 0AX
Telephone Number: +44 (0)20 7679 5888
Facsimile Number: +44 (0)20 7916 2775

Self-selection patterns in Mexico-U.S. migration: The role of migration networks*

David McKenzie[a] and Hillel Rapoport[b]

[a] World Bank, Development Economics Research Group
[b] Department of Economics, Bar-Ilan University, EQUIPPE-CADRE, University of Lille II and CReAM, University College London

Abstract:

This paper examines the role of migration networks in determining self-selection patterns of Mexico-U.S. migration. We first present a simple theoretical framework showing how such networks impact on migration incentives at different education levels and, consequently, how they are likely to affect the expected skill composition of migration. Using survey data from Mexico, we then show that the probability of migration is increasing with education in communities with low migrant networks, but decreasing with education in communities with high migrant networks. This is consistent with positive self-selection of migrants being driven by high migration costs, as advocated by Chiquiar and Hanson (2005), and with negative self-selection of migrants being driven by lower returns to education in the U.S. than in Mexico, as advocated by Borjas (1987).

Keywords: Migration, migration networks, educational attainments, self-selection, Mexico
JEL codes: O15, J61, D31.

*Corresponding author: David McKenzie, MSN MC3-300, The World Bank, 1818 H Street N.W., Washington D.C., USA 20433. Email: dmckenzie@worldbank.org. We thank Gordon Hanson and participants at the 2006 meeting of the European Society for Population Economics, Verona, for helpful comments.

1. Introduction

The skill level of Mexicans migrating to the United States is an issue of important policy relevance on both sides of the border. In the U.S. an important element of opposition to immigration centers on the extent to which low-skilled Mexicans depress the wages of low-skilled natives. The effects of emigration on Mexico's development will vary according to whether those who leave are less skilled than those who remain, helping to reduce poverty and inequality, or more skilled, heightening already high inequality levels.

However, a series of recent papers have produced conflicting results as to whether Mexican migrants are positively or negatively selected in terms of educational skills. Chiquiar and Hanson (2005) find migration rates to be increasing in education up to relatively high education levels, that is, positive selection. Cuecuecha (2005b) and Mishra (2007) also find positive selection. These findings have been challenged by Ibarraran and Lubotsky (2006) and Fernández-Huertas (2006), who conclude that there is negative selection, with migrants tending to be less educated than non-migrants. Orrenius and Zavodny (2005) find intermediate selection, with migrants more likely to be in the middle of the skill distribution than in the low or high end, compared to non-migrants. In contrast, Caponi (2006) finds a U-shaped relationship, with the highest and lowest educated tending to migrate more than the middle educated.

The contribution of this paper is to show that migration networks can in part reconcile some of these conflicting findings, with positive self-selection occurring in communities with low migration networks and negative self-selection occurring in communities with high migration networks. Different data sets of communities will thus yield different answers as to the average direction of selection.

Self-selection is driven primarily by wage differentials net of migration costs (Sjaastad, 1962). Thus, in theory various self-selection patterns with respect to education and skills may be observed depending on whether the wage-skill profile is steeper at origin or destination and on whether migration costs increase or decrease with skills. Borjas (1987) concentrates on the wage side, and famously argues that individuals migrating from countries with high earnings inequality to countries with low earnings inequality will tend to be negatively self-selected. Income inequality is substantially higher in Mexico than in the United States. The Gini index of income in 2000 was 0.41 in the U.S. and 0.55 in Mexico, while in the same year the income share of the highest 10 percent was 43 percent in Mexico, compared to 30 percent in the U.S. (World Bank, 2004). All else equal, one would therefore predict negative self-selection among Mexican emigrants.

This prediction assumes that all migration costs are proportional to wages at home and therefore do not determine self-selection patterns. However, in practice international migration is costly, involving upfront monetary costs, search and information costs, and psychological costs. Such costs are unlikely to be constant across education levels, but instead be decreasing in skills (Chiquiar and Hanson, 2005; Cuecuecha, 2005a). For example, fixed costs of migration represent fewer hours of work and can be met with no or lower borrowing costs by more educated individuals, and education can help in seeking information. As Chiswick (1999, p. 182) puts it, the more able are also more efficient in migration. If migration costs are large enough, and credit constraints sufficiently binding, one should therefore expect to see positive selection in terms of education as individuals with low education find it too costly to move.

Given this, the pattern of self-selection should depend on how costly migration is from a given community. Migration networks act to lower the costs of migrating in a number of ways: they provide information on job opportunities and labor market conditions at destination, information on border crossing, including on ways to find and deal with smugglers, assistance in job search and housing, and help relaxing credit constraints (Massey, 1988; Orrenius, 1999, Orrenious and Zavodny, 2005, Dolfin and Genicot, 2006). These effects are likely to benefit low-skill migrants the most. In part this is due to them being more likely to be credit-constrained, but may be due also to the fact that ethnic enclaves provide services mainly to migrants with low skills in general and low levels of host-language fluency in particular (Borjas, 1999, Chiswick and Miller, 2005, Bauer et al., 2005).[1] As a result we should expect to see a greater degree of negative self-selection in communities with larger migration networks.

This paper therefore examines the role of migration networks in shaping the self-selection pattern of Mexico-U.S. migration. We begin by augmenting the simple theoretical model in Chiquiar-Hanson (2005) to allow for network effects, and use this to determine the impact of increasing network size on selectivity. Using survey data from Mexico we then show that in communities with small migration networks, education increases the probability of migration, giving rise to positive self-selection among migrants. However, for communities with large migration networks, education is shown instead to decrease migration propensities, giving rise to negative self-selection, as conjectured by Borjas (1987). We use historic migration networks to instrument for current networks in this analysis. The results are found to be robust to

[1] Indeed, Bauer et al. (2005) show that enclaves selectively attract people with limited language skills while all else equal migrants with better host-country language proficiency choose destinations with smaller home-country networks.

attempts to account for the undercounting of migrants who move with their entire households to the United States, and to allowing education to be itself affected by the prospect of migration.

These findings further demonstrate the pivotal role played by migration networks in determining the pattern of migration, and go part of the way towards reconciling the conflicting evidence on migrant selectivity arising from the recent literature.

The remainder of the paper is structured as follows. Section 2 lays out a model of self-selection which includes migration networks. Section 3 describes our identification strategy and empirical methodology, while Section 4 discusses the data used. Section 5 provides the main results and Section 6 robustness to undercounting and endogenous education formation. Section 7 concludes.

2. The model

2.1 Wage Equations

Starting with Sjaastad (1962), migration has been modeled as an investment decision where prospective migrants make their decision based on the net discounted value of income streams across locations. Given that migration incentives and costs vary according to age, gender, education, and other individual characteristics, immigrants self-select out of the general population non-randomly. Following Borjas (1987), a series of recent papers have adapted Roy's (1951) model of self-selection to the issue of Mexican immigration to the U.S. (e.g., Chiquiar and Hanson, 2005, Orrenious and Zavodny, 2005, Ibarraran and Lubotsky, 2005). We extend these models to allow for network effects. Using the notation of Chiquiar and Hanson (2005), the wage equation in Mexico (subscript 0) may be written as:

$$\ln w_0 = \mu_0 + \delta_0 s \tag{1}$$

where w is the wage, $\mu > 0$ is the minimal wage level paid in the absence of schooling, δ>0 is the return to schooling and s is the level of schooling.

Similarly, the wage equation in the U.S. (subscript 1) may be written as:

$$\ln w_1 = \mu_1 + \delta_1 s \tag{2}$$

As minimum wages are higher in the U.S. and relative returns to schooling are higher in Mexico, we assume $\mu_1 > \mu_0$ and $\delta_0 > \delta_1$.[2]

2.2 Migration Costs

Let C be the migration cost. In line with the migration networks literature (Massey et. al., 1987, Carrington et al., 1996, Bauer et al., 2002, Munshi, 2003, Kanbur and Rapoport, 2005), we assume that it is decreasing with the size of the community migration network, n:

$$C = C(n), C' < 0 \tag{3}$$

Expressed in time-equivalent units, the migration cost[3] may be written as:

$$\pi = \pi(n,s) = \frac{C(n)}{w_0} \tag{4}$$

Then, a resident of Mexico will find it beneficial to migrate to the U.S. if[4]

[2] Chiquiar and Hanson (2005), Ibarraran and Lubotsky (2005) and Cuecuecha (2005) all provide evidence supporting these assumptions.

[3] To explain why people with similar networks and schooling levels may make different migration decisions, we could add a source of unobserved heterogeneity such as preferences for consuming in one's origin country. This may be captured for example by adding a psychological component to the migration cost in the form of an individual taste parameter k_i distributed on [0,1] with density *g(k)* such *that* $C=C(n,k)$, $C'<0$ and $\pi = e^{\mu_\pi - \gamma_1 s - \gamma_2 n - \gamma_3 k}$, with s and k independently distributed. Note also that k could be seen as positively depending on n as having *landsleit* (people from the same town or village) certainly makes it not only cheaper to move, but also easier to adapt to the new destination, thus reinforcing the role of networks. However, this would not change the qualitative predictions of our results (which may be seen as valid "for a given k" and easily generalize to all k) and we therefore keep this additional component implicit.

$$\ln(w_1) - \ln(w_0 + C) \cong \ln(w_1) - \ln(w_0) - \pi > 0 \quad (5)$$

We assume that time-equivalent migration costs decrease with schooling. This occurs as a result of higher wages requiring less hours of work to pay a given fixed fee, and is also consistent with recent evidence provided by Cuecuecha (2005a), who describes a number of other channels leading to this decreasing relationship, including the better ability of more educated individuals to bargain with smugglers. In addition, migration costs and therefore time-equivalent migration costs decrease with the size of the community migration network, *n*:

$$\ln(\pi) = \mu_\pi - \gamma_1 s - \gamma_2 n \quad (6)$$

so that $\pi = e^{\mu_\pi - \gamma_1 s - \gamma_2 n}$ with $\gamma_1, \gamma_2 > 0$.

Assume first an initial migration network of a given size, which we normalize to zero without loss of generality. Prospective migrants face a wage profile by schooling level at destination which is given by $A = \mu_1 + \delta_1 s - e^{\mu_\pi - \gamma_1 s}$ (see the solid line in Figure 1). In order not to rule out the possibility of positive self-selection, we also assume, following Chiquiar and Hanson (2005), that $\mu_1 - \mu_0 < e^{\mu_\pi}$. That is, the inter-country minimum wage differential is not high enough to warrant migration for people with very high migration costs (i.e., people with no schooling and no migration network to rely on). For a given size of the migration network, one can then distinguish two schooling thresholds between which people will want to migrate: s_L, below which migration costs are so high that they make migration not profitable, and s_U, above which returns to schooling in Mexico are high enough to discourage migration.

[4] The approximation is valid if π is small, which appears the case if w is defined as the present value of a flow of future incomes.

Chiquiar and Hanson then discuss the pattern of self-selection in U.S.-Mexico migration by reference to these two thresholds. Assume that the support of s is $[\underline{s}, \overline{s}]$. Now, if $\underline{s} < s_L < \overline{s} < s_U$, then positive self-selection obtains. Conversely, if $s_L < \underline{s} < s_U < \overline{s}$, then negative self-selection obtains. For any other ranking we need to know the distribution of schooling before we can make a judgment on the type of self-selection obtained. [5]

2.3 The effect of a larger migrant network

The effect of expanding (or introducing) migration networks is to decrease migration costs at all schooling levels. Diagrammatically, this means an upward shift of the wage-schooling profile at destination following introduction or expansion of migration networks. In addition, schooling and networks are substitutes in lowering the cost of migration. The new wage profile at destination is now given by $B \equiv \mu_1 + \delta_1 s - e^{\mu_\pi - \gamma_1 s - \gamma_2 n}$ (see the dashed line in Figure 1), with the two profiles A and B converging at high levels of schooling as the reduction in migration costs is strongest at low schooling levels. This can be stated formally as:

Proposition 1: Larger migrant networks increase migration incentives (i) at all schooling levels, and (ii) more so at low schooling levels.

[5] For example, positive self-selection obtains if $\underline{s} < s_L$ and $\overline{s} > s_U$ but the distribution of s is such that the density is highest between $\underline{s}$ and s_L.

Proof: The induced change in migration incentives, which we denote by Δ, is given by the difference between *A* and *B*: $\Delta = e^{\mu_\pi - \gamma_1 s} - e^{\mu_\pi - \gamma_1 s - \gamma_2 n} = e^{\mu_\pi - \gamma_1 s}\left[1 - \frac{1}{e^{\gamma_2 n}}\right] > 0$, with $\partial\Delta / \partial n > 0$ and $\partial\Delta / \partial s < 0$.

Following the expansion of networks, a change in migration incentives (i.e., in wages at destination net of migration costs) defines two new threshold values of *s*, $s_L^{'}$ and $s_U^{'}$, with $s_L^{'} < s_L$ and $s_U^{'} > s_U$. As migration networks expand, more people are willing to migrate at both ends of the migrants' schooling distribution. How will this translate in terms of self-selection patterns? In all likelihood, larger networks will reinforce, or increase the chances of obtaining, negative self-selection.

To show why this is the case, consider first the two configurations for which it is obvious that there is either positive or negative self-selection, independently of the exact schooling distribution; that is, in the case where $\underline{s} < s_L < \overline{s} < s_U$ (positive self-selection) or $s_L < \underline{s} < s_U < \overline{s}$ (negative self-selection). In the first event, all additional migrants have schooling levels below s_L and it is therefore clear that the average level of schooling among migrants decreases. Positive self-selection still obtains as by construction non-migrants are at the lower end of the schooling distribution, but in a less pronounced way. In the second event, the effect of networks is to increase average schooling levels both among migrants and non-migrants, but more so among the later so that negative self-selection still prevails but in a more pronounced way.

Consider then the more general case where there is intermediate self-selection of migrants; assuming that there will always be non-migrants at the two ends of the schooling distribution (i.e., the support of s is $[0, \bar{s}]$ and $s_L^{'} > 0$, $s_U^{'} < \bar{s}$), are we sure that migration networks reinforce, or increase the chances of obtaining, negative self-selection? In this configuration, we know that networks will act to increase the number of migrants and that the additional migrants will come from the two intervals $(s_L - s_L^{'})$ and $(s_U - s_U^{'})$. The impact in terms of migrants' skills relatively to non-migrants will depend on which of these two intervals is longer and on the density of the schooling distribution on the two intervals. In the following, we focus on the length of the two segments and rule out the possibility that the density of the schooling distribution is higher on $(s_U - s_U^{'})$ than on $(s_L - s_L^{'})$, which is quite realistic (and increasingly so for larger and larger networks). Hence, our results hold true for any distribution for which the density is not increasing in schooling (including, obviously, the uniform distribution) and for other distributions as well providing that the above restriction holds.

In particular, under this configuration, we have:

Proposition 2: With intermediate self-selection, where the support of s is $[0, \bar{s}]$ *and* $s_L^{'} > 0$, $s_U^{'} < \bar{s}$,

(a) An increase in the migration network increases the range of lower schooling levels that wants to migrate more than it increases the range of higher schooling levels that wants to migrate. i.e. $|s_L - s_L'| > |s_U - s_U'|$.

(b) Providing that the density of the schooling distribution is not increasing in schooling, larger migration networks reduce average levels of schooling among

migrants (and increase average levels of schooling among non-migrants), therefore increasing the likelihood and/or degree of migrants' negative self-selection.

Proof: see appendix.

3. Data

The main source of data is the 1997 Encuesta Nacional de la Dinámica Demográfica (ENADID) (National Survey of Demographic Dynamics) conducted by Mexico's national statistical agency (INEGI) in the last quarter of 1997.[6] The ENADID is a large nationally representative demographic survey, with approximately 2000 households surveyed in each state, resulting in a total sample of 73,412 households.

3.1 Which migrants to look at?

The survey asks whether household members have ever been to the U.S. in search of work. This information is collected for all individuals who normally live in the household, even if they are temporarily studying or working elsewhere, and includes information on the number of times an individual has been abroad, and the date of the last visit. Secondly, additional questions are asked about migration during the last five years, with detailed questions collected on the last trip. Finally, the survey also asks whether there are any individuals who were living in the household five years ago who have moved abroad, regardless of whether or not they are currently considered part of the household.

With this information we can examine whether a 15 to 49 year old individual migrated for the first time during the period 1996-97. This restriction is made for several reasons. Firstly, the role of networks is believed to be less important for

[6] Survey methodology, summary tables, and questionnaires are contained in INEGI (1999).

individuals making a repeat migration trip (Massey, Goldring and Durand, 1994), and therefore we wish to model the initial migration decision. However, since we only know the number of trips an individual has made, and the date of last migration, we can only model the initial migration decision for individuals making their first trip in the recent past.[7] Secondly, by comparing migration from different communities in a single two-year period, we are much less concerned about the interaction between community networks and macroeconomic shocks. Thirdly, this short period makes recall bias which varies with education much less of a concern than looking at migration over a lifetime.

As with other Mexico-based surveys of migration[8], the ENADID will only capture data on migrants who have either returned to Mexico, or who have at least one household member remaining in Mexico. As a result it will tend to underrepresent permanent migrants (Hanson, 2006) who are likely to take their whole household. In Table 1, we use the 5% public use sample of the 2000 U.S. Census (Ruggles et al., 2004) to examine the marital status of 18-45 year old Mexican migrants who arrived in the U.S. within two years of the Census. We see that 14.4 percent of male migrants are married, with their spouse present. These individuals are likely not to be reported on in Mexico-based surveys. However, the majority of migrants are either single, or married with a spouse remaining in Mexico, and so should be reported on from

[7] In particular, an individual is defined to be a first-time migrant if they migrated during 1996 or 1997, and have only made one trip to the United States.

[8] The two other main Mexican surveys used to look at migrant selection are the Mexican Migration Project (MMP) and the Mexican Census. Neither of these is suitable for our study. The MMP does contain information on the age of first migration, but is not nationally representative, and in particular, samples only a small number of communities with small migration networks (see Figure 2 in McKenzie and Rapoport, 2006a). Since most communities have reasonably large networks, this makes it difficult to see how migration varies with network size starting from a low level. The Mexican Census only collects data on whether an individual has migrated within the last five years, preventing its use for looking at the determinants of first-time migration.

Mexico. The problem is much worse with females, with 48 percent of all recent migrants in the U.S. being married with spouse present.

There are two main concerns with this undercount for our analysis. The first is that the U.S. census provides no information on the community of origin, and so no network variables are known for the unmeasured individuals. The second concern is that the education levels of migrants who are not reported on in the ENADID may differ from the education levels of migrants who are reported on. There is strong evidence to suggest that this is the case. For example, 16.0 percent of the male migrants in Table 1 who have a spouse present in the U.S. have post-high school education, compared to only 8.3 percent of those in the other marital categories. Ibarraran and Lubotsky (2006) compare the U.S. and Mexican censuses and likewise conclude that migrants that are excluded in the Mexican census are likely to be more educated.

Given these concerns, we do not look at self-selection among female migrants – with only half of all female migrants likely to be reported on in the ENADID, we consider the likely bias from doing so to be too severe. Additionally, since female migration is so closely tied to the migration of the spouse, the theory above is less directly applicable for females. For males the ENADID is likely to measure 86 percent of migrants, and we will carry out robustness exercises to see how sensitive our results are to those who are undercounted.

3.2 Measuring the Community Network

We follow Massey, Goldring and Durand (1994) in measuring the community migration network by the proportion of all individuals aged 15 and over in a given community who have ever migrated. We restrict our analysis to municipalities in which at least 50 households were interviewed in the survey in order to measure the community migration network. This results in data on 55,848 males aged 15-49 in 288 communities for our analysis. Sample statistics are shown in Table 2. Since the role of networks is likely to be greater outside of large cities, we will carry out the majority of our analysis for individuals living in locations of less than 100,000 population. This reduces the sample to 25,531 individuals in 254 communities.

Table 2 also summarizes the main variables of interest separately for the group of migrants. On average male migrants aged 15-49 have less schooling than the average male of this age, and are more likely to be married and the head of a household. We now discuss our identification strategy and econometric methodology before examining how schooling interacts with migration networks in shaping the migration decision.

4. Methodology and Identification Strategy

4.1 Identification

Our empirical work will examine how the degree of educational selectivity in migration varies with the size of a community's migrant network. However, this raises the concern that there are unobserved community characteristics which drove migration in the past, continue to drive migration today, and which could also be correlated with education levels in the community. For example, a community with poor schooling infrastructure may have low levels of education and a lot of individuals migrating to seek better lives for their children. This would lead us to spuriously find that negative selection on education occurs more in high migration communities. However, if the first migrants from a community are positively selected, as would be the case if education is needed to learn about opportunities abroad or to adapt to a new land without a network around, we might expect large networks to arise in places with better education, biasing us towards finding positive selection on education occurs more in high education communities. This view is consistent with Feliciano (2001), who finds that in 1910, Mexican immigrants were more likely to be literate than the general Mexican population.

To account for this concern we follow Woodruff and Zenteno (2007) and a number of subsequent studies in using historic state-level migration rates as an instrument for current migration networks.[9] In particular, we use the U.S. migration rate from 1924 for the state in which the migrant household is located, taken from Foerster (1925). Likewise, we will use the interaction between education and historic migration networks as an instrument for the interaction between education and community migration prevalence. The main argument used to justify the use of this instrument is that these historic migration rates were the result of the pattern of the arrival of the railroad system in Mexico, coupled with changes in U.S. demand conditions for agricultural labor. As migration networks lower the cost of migration for future migrants, they then become self-perpetuating.[10]

These instruments are strong predictors of current migration prevalence and its interaction with education. Appendix table A1 shows first-stage F-statistics between 40.5 and 54.1. To justify the exclusion restriction, we need to argue that these historic rates affect current migration decisions only through current migration networks. A potential threat to this instrument is that communities which responded more to the expansion of the railroad may have been ones with historically poor schooling infrastructure and inequality, or that the development of the railroads ushered in the

[9] Hanson and Woodruff (2003); McKenzie and Rapoport (2006a, 2006b); López-Córdoba (2005); and Hildebrandt and McKenzie (2005) all employ historic migration rates as instruments for current migration.

[10] Hildebrandt and McKenzie (2005) and McKenzie and Rapoport (2006a) provide more detailed discussion of the historic processes involved .

expansion of infrastructure such as schooling facilities. This could affect current education through the intergenerational transmission of schooling, and through the inertia in schooling infrastructure. To allow for this possibility, we control for a number of historic variables that capture schooling access, achievement, and equality, and for historic measures of inequality. Even after controlling for these variables, the historic migration rates remain strong predictors of current community migration prevalence.

4.2 Estimation

To test our predictions on the role of networks in determining the pattern of self-selection into migration, we estimate the following equation for whether a 15-49 year old male i in community c migrates for the first time in the period 1996-97, conditional on never having previously migrated:

$$\begin{aligned} M_{i,c} &= \beta_0 + \beta_1 educ_{i,c} + \beta_2 educ_{i,c}^2 + \beta_3 network_c + \beta_4 educ_{i,t} \times network_c \\ &\quad + \phi' X_{i,c} + \lambda' Z_c + \varepsilon_{i,c} \end{aligned} \tag{7}$$

where $M_{i,c}$ is an indicator variable taking the value one if individual i migrates and zero if they do not migrate, $educ_{i,c}$ is the completed years of schooling of individual i, $network_c$ is the community migration prevalence in community c (our measure of the network), and $X_{i,c}$ and Z_c are control variables capturing individual and community characteristics respectively.

Although $M_{i,c}$ is a binary variable, we use ordinary least squares (OLS) and two-stage least squares (2SLS) for estimation. The main reason for this choice is that our primary coefficient of interest is β_4, which shows how the impact of education on the likelihood of migration varies according to the size of the network. Our prior is that migration networks lower the costs of migrating more for the less-educated, so that β_4

should be negative. Interpretation of this coefficient is clear in the linear model, whereas in logit and probit models, the marginal effect of the interaction term differs from, and can even be of opposite sign to, the interaction term (Ai and Norton, 2003). Justification for using 2SLS with binary outcomes can be found in Angrist (1991), who provides conditions under which linear instrumental variables estimation will consistently estimate average treatment effects, and Monte Carlo evidence to argue that these conditions may hold approximately, so that 2SLS can perform well in practice.

5. Results

Table 3 presents the results of estimating equation (7). Columns 1 through 4 present OLS estimates, and 5 through 9 present 2SLS estimates. The first column shows the OLS results for all population areas combined. The second column then interacts population size with the migration network. As we would expect, the effect of the community network is less in large cities. We therefore consider only individuals in areas with less than 100,000 population in column 3, and for robustness, show results for less than 15,000 population in column 4. Across all population sizes the OLS results show a significant positive coefficient on the linear term in education, and a significant negative coefficient on the quadratic term. Migration is more likely for individuals who are married, household heads, and in communities with larger networks. The interaction between education and network size is negative, as theory predicts, and significant at the 5% level in communities with less than 100,000 population. These results continue to hold after instrumenting for community migration networks in columns 5 through 7, with some changes in the magnitudes of the coefficient. The first-stage F-statistics are all above 30, see Appendix table A1 for the full first-stage results for column 6.

Columns 8 and 9 examine the robustness of our key parameter of interest, the interaction between education and community networks, to the inclusion of municipality fixed effects. Including municipality fixed effects will control for any differences across communities which might affect the level of migration, such as differences in infrastructure and attitudes towards education and migration. We see that the interaction coefficients remain significant and negative, with minor changes in the size of the coefficient. For example, for communities under 100,000, the interaction changes from -0.55 in column 6 to -0.43 in column 9, whereas for all communities the interaction only changes from -0.46 to -0.47.

Comparing the OLS and 2SLS results, we see that the interaction between education and the migrant network becomes more negative after instrumentation. This suggests that communities with higher migration have unobservable factors such as better schooling infrastructure or an education culture that leads, ceteris paribus, to higher education. This accords with the evidence in Feliciano (2001), who found migrants in 1910 (a time when migrant networks were first forming) to have higher literacy rates than non-migrants.

Interpreting the fitted relationship between education, network size and the probability of migration is most easily done graphically. For illustration purposes, and because networks are found to be more powerful outside of large cities, we show this for the regressions which consider individuals in areas of less than 100,000 population. Figure 2 plots the predicted probability of migration from OLS (left panel) and 2SLS (right panel) at different percentiles of the community migration prevalence distribution. Recall from Table 1 that mean years of schooling is 7.3, and the median is 7. Both the OLS and 2SLS results show clearly that migrants becomes progressively more negatively selected in terms of education as network size grows, with this result stronger after instrumenting for migration networks.

The 2SLS results in Figure 2 show that in communities with small migration networks, where the costs of migrating are thus likely to be high, the probability of migration is increasing in completed years of schooling up to nine or ten years of schooling, around the 75th percentile of the education distribution. That is, first-time male migrants from communities with small networks are likely to be positively selected. However, as the migrant network grows, lowering the costs of migration and reducing credit constraints, we find a reversal in this pattern. An individual in a community at the 70th percentile of migration networks, where 25 percent of households in the community have someone who has ever migrated to the U.S., has the highest probability of migrating if he has two to five years of schooling.

Thus we find that in communities with low migration networks, migrants tend to be selected from the upper-middle of the education distribution, which concurs with Chiquiar and Hanson's (2005) description of selectivity. However, in high network communities, where the cost of migrating is less binding, we find negative selection, which is what would be predicted by Borjas (1987) based on wage differentials. As a result, over time, as origin communities accumulate migration experience, one should expect to see a gradual worsening in the relative skill level of migrants. This finding is consistent with more aggregate evidence provided by Feliciano (2001), who finds a decline in the relative skill level of Mexican immigrants over the course of the 20th century.

6. Robustness

6.1 Robustness to Undercounting of Migrants

As discussed in Section 3, one concern with the ENADID and other Mexico-based migration surveys is that they do not capture migrants who move with their whole families, who are more educated on average than those who have family members remaining. The omission of such individuals is thus likely to bias against finding positive selection, since some highly educated migrants are not included in the regressions.

To investigate the robustness of our results to this issue, we use the public use sample of the U.S. Census to obtain the educational breakdown of the 14.4% of recent male migrants who have migrated with their spouse present, and are hence least likely to be reported on from Mexico.[11] Our sample of 55,848 males in the ENADID contains 568 recent first-time migrants. Based on an undercount of 14.4%, there should be an additional 95 migrants in the analysis. We therefore assign the education distribution observed in the U.S. Census to these 95 additional migrants, and then reweight the ENADID migrant sample so that it reflects the educational breakdown of the combined 663 migrants. For example, 17 of the ENADID sample have 16 or 17 years of education, and we estimate from the U.S. Census that 7 out of the additional 95 migrants would also have this education level. Migrants in the ENADID with 16 or 17 years education are therefore assigned a weight of 24/17 = 1.41.

[11] The U.S. Census does not provide information on whether these individuals are coming from high or low migration networks. Our weighting procedure implicitly assumes the unobserved individuals are drawn from communities in the same proportions as observed individuals. Given the relatively low level of undercount for males, and the fact that our fitted probabilities don't change by much, we believe the general pattern of our results is also robust to alternative assumptions about the network these undercounted individuals are drawn from.

Table 4 then provides the estimated 2SLS results after adjusting for undercount. Columns 1 and 2 replicate columns 5 and 6 of Table 3 for ease of comparison. Columns 3 and 4 then provide the undercount-adjusted coefficients. We see that our main coefficient of interest, the interaction between education and network size, becomes slightly more negative, and is still significant. Figure 3 plots the predicted probabilities from column 4, and compares them to the predicted probabilities for the unadjusted results in column 2. The results show that adjusting for undercounting leads to only small changes in the fitted curves. The adjusted curves are all above the unadjusted curve, reflecting that the likelihood of migration at any given education level is higher than estimated with just the ENADID data, since some migrants aren't recorded. Secondly, we see that the effect of the larger interaction term is to shift the curves rightward slightly, resulting in more positive selection or less negative selection for a given community network. However, we still find positive selection at low levels of community migration prevalence, and negative selection at high levels, showing our results are robust to undercounting of some migrants.

6.2 Robustness to Endogenous Education Formation

In keeping with the literature on migrant selection, we have thus far treated education as exogenous when modeling selectivity into migration. However, there are number of reasons to believe that migration prospects can in turn affect education levels acquired. Most first-time migrants from Mexico travel to the U.S. without documentation, including 91% of first-time migrants in the ENADID sample. Kossoudji and Cobb-Clark (2002) argue that there is no return on schooling in the first job for unauthorized workers, and using a sample of individuals becoming legalized during an amnesty program, find that finishing high school or having post-high school education only pays off in terms of wages about four years after legalization. As a result, as demonstrated by Chiquiar and Hanson (2005), the return to schooling is higher in Mexico than it is for Mexican migrants in the United States. As a result, the prospect of future education may actually lower the incentive to invest in education.

This effect is likely to be largest for individuals with a migrant parent, who are themselves then much more likely to migrate. Moreover, although the migrant parent can provide remittances which can ease in credit constraints restricting schooling, the absence of a parent can also have a detrimental effect on schooling. McKenzie and Rapoport (2006b) estimate that the net effect of having a migrant parent is to lower the level of schooling attained by 16 to 18 year old males. Using 2SLS they estimate the size of the effect to be a 1.3 year reduction in years of schooling attained. They also carry out estimation using a censored ordered probit (COP), which accounts for the fact that some 16 to 18 year olds are still completing their education and so have censored schooling levels, and for the fact that distribution of years of schooling tends to clump at 6, 9 and 12 years, representing finishing different levels of schooling. This results in estimates of the extent to which having a migrant parent changes the probability of attaining different levels of schooling. The net effect of these is approximately a 1.8 year reduction in predicted years of schooling.

To investigate the robustness of our results to endogenous education formation, we use the point estimates described to adjust schooling levels. Unfortunately the ENADID only allows identification of migrant parents for individuals still living with their parents, preventing us knowing whether many individuals in their late twenties, thirties and forties had migrant parents. Therefore to make an extreme assumption to examine the sensitivity of our results, we will assume that all migrants had migrant parents, and no non-migrants had migrant parents. We will then adjust upwards the schooling of all migrants with less than 12 years of education by either 1.3 years (2SLS) or 1.8 years (COP) and then reestimate the instrumented version of equation (7).

Columns 5 and 6 of Table 4 present the new coefficient estimates using the 2SLS adjustment, and columns 7 and 8 using the COP adjustment. Doing this adjustment does not change the sign or significance of any of the variables in our model, but does lead to sizeable changes in magnitudes of the coefficients. In particular, the interaction between years of education and community migration prevalence becomes larger in absolute value, showing a stronger negative interaction effect. Figure 4 plots the predicted probability of migration against education years, comparing the unadjusted predictions from column 2 of Table 4 to the adjusted predictions in column 6. The shape of selection changes towards more positive selection in low and medium-network communities. Intuitively, those who migrate would have had more education in the absence of the prospect of migration or the migration of their parents. However, the pattern is still strongly one of negative selection in high network communities. The unadjusted results showed migration to peak at very low levels of education, so adjusting upwards these low levels results in a peak at 2 to 5 years, still a very low level. Thus while adjusting for endogenous education formation increases the degree of positive selectivity, our main results remain robust to this adjustment.

7. Conclusion

We find that in communities with small migration networks, the probability of migration is increasing in education up to reasonably high levels of schooling, resulting in positive selection of migrants. This is consistent with high costs of migration being the determining factor of who migrates in these communities. In contrast, in communities with large networks, where migration costs are lower, we find migration propensities to be decreasing in education, consistent with lower returns to schooling in the U.S. than in Mexico. These results are found to be robust to accounting for the undercount of some migrants in Mexican data, and for the possibility of endogenous formation of education with respect to migration.

Our results help in part to reconcile conflicting accounts of the direction of education-selection amongst migrants from Mexico found in the existing literature. Since the direction of selectivity depends on the level of migration prevalence in a community, studies which estimate the average direction of selection will give different estimates if they draw on surveys from communities with differing levels of networks, coupled with differences in the extent to which they account for undercounting of more educated migrants who move with their entire families.

The results of this paper also suggest that as migration networks continue to develop, we should expect to see more negative educational selection of migrants from Mexico, due to the low returns to education in illegal jobs in the U.S. In contrast, if proposals to allow more legal immigration of unskilled workers are enacted, the higher returns to education should result in greater degrees of positive selectivity.

Appendix

Proof of Proposition 2:

To prove (a), note first that s_L and s_U are solutions of the following equation:

$$\mu_0 + \delta_0 s(n) = \mu_1 + \delta_1 s(n) - e^{\mu_\pi - \gamma_1 s(n) - \gamma_2 n} \qquad (8)$$

Differentiating (8) with respect to *n*, we have:

$$\delta_0 \frac{\partial s}{\partial n} = \delta_1 \frac{\partial s}{\partial n} + (\gamma_1 \frac{\partial s}{\partial n} + \gamma_2) e^{\mu_\pi - \gamma_1 s - \gamma_2 n},$$

and therefore

$$\frac{\partial s}{\partial n} = \frac{\gamma_2 e^{\mu_\pi - \gamma_1 s - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s - \gamma_2 n}} \qquad (9)$$

The marginal effect of an increase in network size on the two critical schooling thresholds is therefore given by:

$$\frac{\partial s_L}{\partial n} = \frac{\gamma_2 e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}} \qquad (10)$$

$$\frac{\partial s_U}{\partial n} = \frac{\gamma_2 e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}} \qquad (11)$$

Note that having $\frac{\partial s_L}{\partial n} < 0$ and at the same time $\frac{\partial s_U}{\partial n} > 0$ requires $e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n} < \frac{\delta_0 - \delta_1}{\gamma_1} < e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}$. Combining (10) and (11), it is then straightforward to see that

$$\frac{\left|\partial s_L / \partial n\right|}{\partial s_U / \partial n} = \frac{\gamma_2 e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}}{\left|\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}\right|} \cdot \frac{\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}}{\gamma_2 e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}} = \frac{e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}}{e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}} \cdot \frac{\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_U - \gamma_2 n}}{\left|\delta_0 - \delta_1 - \gamma_1 e^{\mu_\pi - \gamma_1 s_L - \gamma_2 n}\right|} > 1$$

which proves (a). Coupling this with the non-increasing density assumption then proves (b).

References

Ai, Chunrong and Edward Norton (2003) “Interaction terms in logit and probit models”, *Economics Letters* 80: 123-129.

Angrist, Joshua (1991) “Instrumental Variables Estimates of Average Treatment Effects in Econometrics and Epidemiology”, Working Paper 115, Cambridge, MA: National Bureau of Economic Research.

Bauer, Thomas, Gil S. Epstein and Ira N. Gang (2002): Herd effects or migration networks? The location choice of Mexican immigrants in the US, IZA Discussion Paper No 551, August.

Bauer, Thomas, Gil S. Epstein and Ira N. Gang (2005): Enclaves, language and the location choice of immigrants, *Journal of Population Economics*, 18, 4: 649-62.

Borjas, George J. (1987): Self-selection and the earnings of immigrants, *American Economic Review*, 77, 4: 531-53.

Borjas, George J. (1999): *Heaven’s door: immigration policy and the American economy*, Princeton, N.J.: Princeton University Press.

Caponi, Vincenzo (2006) “Heterogeneous Human Capital and Migration: who migrates from Mexico to the U.S.?”, *IZA Discussion Paper No. 2446.*

Carrington, William J., Enrica Detragiache and Tara Vishwanath (1996): Migration with endogenous moving costs, *American Economic Review*, 86, 4: 909-30.

Chiquiar, Daniel and Gordon 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: 239-81.

Chiswick, Barry (1999): Are immigrants favorably self-selected?, *American Economic Review – Papers and Proceedings*, 89, 2: 181-85.

Chiswick, Barry and Paul W. Miller (2005): Do Enclaves Matter in Immigrant Adjustment?, *City and Community*, 4, 1: 5-35.

Cuecuecha, Alfredo (2005a): The immigration of educated Mexicans: the role of informal social insurance and migration costs, Mimeo, ITAM..

Cuecuecha, Alfredo (2005b) "The Educational Characteristics of the Immigrants from Mexico to the United States", Mimeo. ITAM.

Davis, Benjamin and Paul Winters (2001) "Gender, Networks and Mexico-US Migration", *Journal of Development Studies* 38(2): 1-26.

Dirección General de Estadística (DGE) (1941) *Anuario Estadístico de los Estados Unidos Mexicanos 1939.* Secretaría de la Economía Nacional, DGE: Mexico City.

Dolfin, Sarah and Garance Genicot (2006) "What do networks do? The role of networks on migration and 'coyote' use", Mimeo. Georgetown University.

Feliciano, Zadia M. (2001): The skill and economic performance of Mexican immigrants from 1910 to 1990, *Explorations in Economic History*, 38, 386-409.

Fernández-Huertas Moraga, Jesús (2006) "New Evidence on Emigrant Selection", Mimeo. Columbia University.

Foerster, Robert F. (1925) *The Racial Problems Involved in Immigration from Latin America and the West Indies to the United States.* Washington D.C.: The United States Department of Labor.

Hanson, Gordon H. (2006): Illegal migration from Mexico to the United States, *Journal of Economic Literature*, forthcoming.

Hanson, Gordon H. and Christopher Woodruff (2003): Emigration and educational attainment in Mexico, Mimeo., University of California at San Diego.

Hildebrandt, Nicole and David J. McKenzie (2005) "The Effects of Migration on Child health in Mexico", *Economia*, 6, 1.

Ibarraran, Pablo and Darren Lubotsky (2005): Mexican immigration and self-selection: new evidence from the 2000 Mexican Census, NBER Working Paper No 11456, June.

Instituto Nacional de Estadística, Geografía e Informática (INEGI) (1999). *ENADID: Encuesta Nacional de la Dínamica Demográfica 1997.* INEGI: Aguascalientes.

Kanbur, Ravi and Hillel Rapoport (2005): Migration selectivity and the evolution of spatial inequality, Journal of Economic Geography, 5 (1): 43-57.

Kossoudji, Sherrie and Deborah Cobb-Clark (2002) "Coming out of the Shadows: Learning about legal status and wages from the legalized population", *Journal of Labor Economics* 20(3): 598-628.

López-Córdoba, Ernesto (2004) "Globalization, Migration, and Development: The Role of Mexican Migrant Remittances", *Economia*, 6, 1.

Massey, Douglas S. (1988) "Economic Development and International Migration in Comparative Perspective", *Population and Development Review* 14(3): 383-413.

Massey, D.S, L. Goldring and J. Durand (1994): Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities, *American Journal of Sociology*, 99, 6: 1492-1533.

McKenzie, David and Hillel Rapoport (2006a) "Network effects and the dynamics of migration and inequality: theory and evidence from Mexico", forthcoming, *Journal of Development Economics.*

McKenzie, David and Hillel Rapoport (2006b) "Can Migration Reduce Educational Attainment? Evidence from Mexico", *World Bank Policy Research Working Paper No. 3952.*

Mishra, Prachi (2007) "Emigration and Wages in Source Countries: Evidence from Mexico", *Journal of Development Economics*, 82(1): 180-99.

Munshi, K. (2003): Networks in the modern economy: Mexican migrants in the US labor market, *Quarterly Journal of Economics*, 118, 2: 549-99.
Orrenius, Pia M. (1999) "The Role of Family Networks, Coyote Prices, and the Rural Economy in Migration from Western Mexico: 1965-1994", Federal Reserve Bank of Dallas Working Paper No. 9910.
Orrenius, Pia M. and Madeline Zavodny (2005) "Self-selection among undocumented immigrants from Mexico", *Journal of Development Economics*, 78, 1: 215-40.
Roy, A.D. (1951): Some thoughts on the distribution of earnings, *Oxford Economic Papers*, 3, 2: 135-46.
Ruggles, Steven, Matthew Sobek, Trent Alexander, Catherine A. Fitch, Ronald Goeken, Patricia Kelly Hall, Miriam King, and Chad Ronnander (2004) *Integrated Public Use Microdata Series: Version 3.0* [Machine-readable database]. Minneapolis, MN: Minnesota Population Center, http://www.ipums.org.
Schultz, Theodore W. (1980) "Nobel Lecture – The economics of being poor", *Journal of Political Economy*, 88, 4: 639-51.
Sjaastad, Larry A. (1962): The costs and returns of human migration, *Journal of Political Economy*, 70, 4, supplement, pp. 80-93.
Winters, Paul, Alain de Janvry and Elisabeth Sadoulet (2001) "Family and Community Networks in Mexico-US Migration", *Journal of Human Resources* 36(1): 159-84.
World Bank (2004) Global Development Finance and World Development Indicators Central Database, Version August 2004.
Woodruff, Christopher, and Rene M. Zenteno (2007): Migration networks and microenterprises in Mexico, *Journal of Development Economics*, 82(2): 509-28..

Figure 1: Migration networks and self-selection patterns

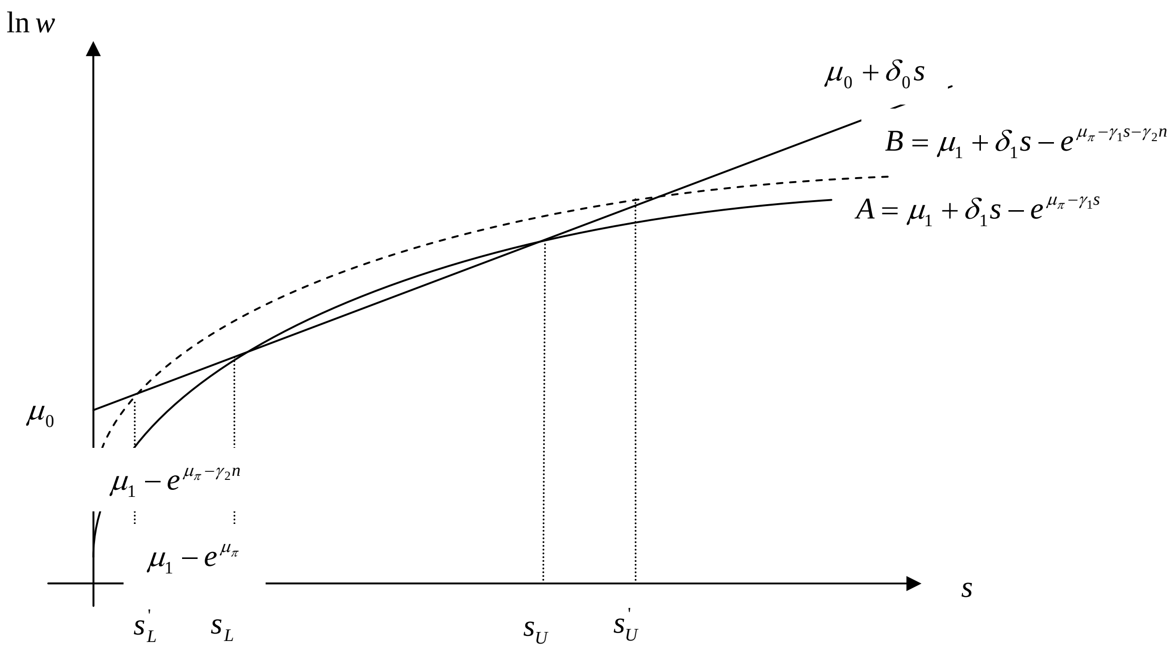
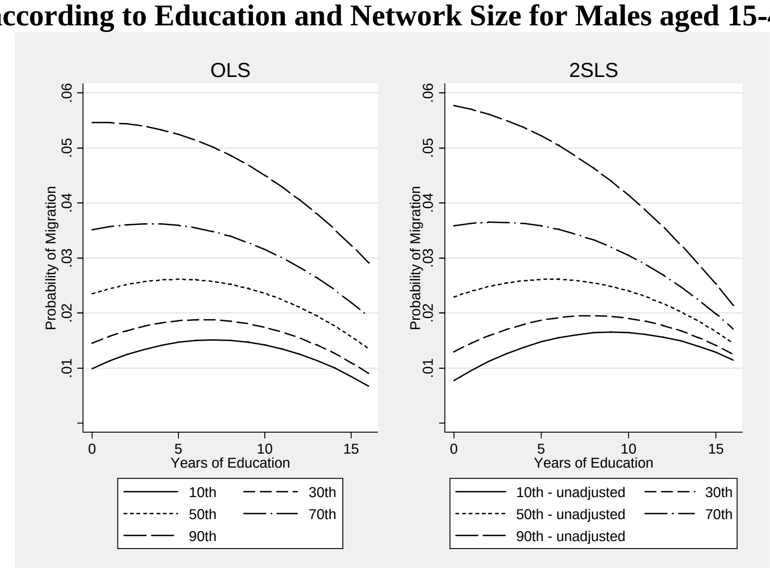
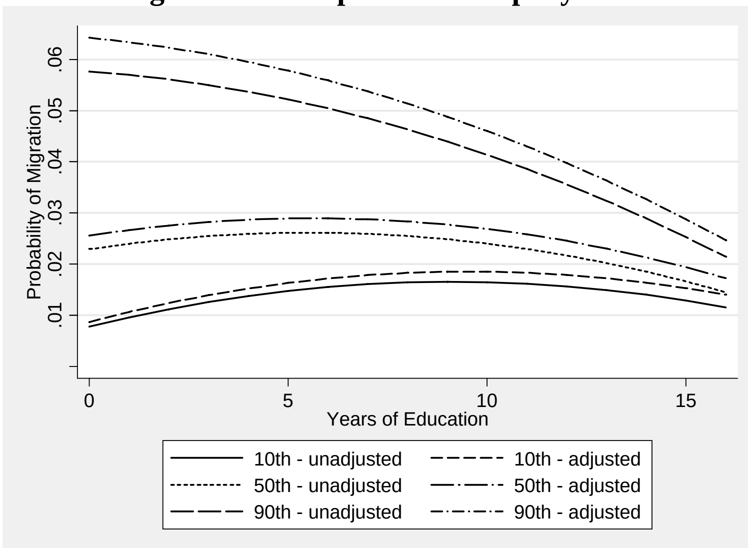


Figure 2: Probability of First-time Migration in 1996-97 according to Education and Network Size for Males aged 15-49



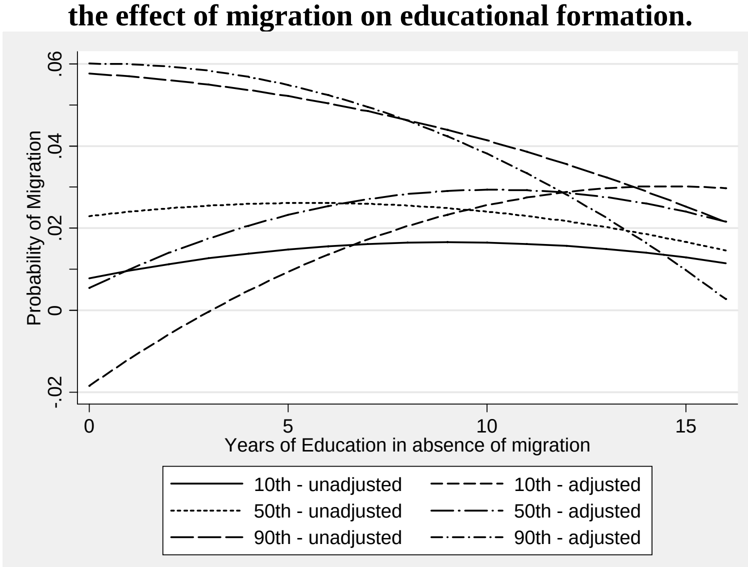
Predicted probabilities obtained from OLS and 2SLS regressions in columns 3 and 6 of Table 3, for males in communities with population less than 100,000. Probabilities are plotted at the 10^{th}, 30^{th}, 50^{th}, 70^{th}, and 90^{th} percentiles of the community migration prevalence distribution.

Figure 3: Probability of Migrating, adjusting for undercounting of married migrants whose spouses accompany them to the U.S.



Predicted probabilities obtained from columns 2 and 4 of Table 3,for males in communities with population less than 100,000. Probabilities are plotted at the 10th, 30th, 50th, 70th, and 90th percentiles of the community migration prevalence distribution.

Figure 4: Probability of Migrating, adjusting for the effect of migration on educational formation.



Predicted probabilities obtained from columns 2 and 6 of Table 3,for males in communities with population less than 100,000. Probabilities are plotted at the 10th, 30th, 50th, 70th, and 90th percentiles of the community migration prevalence distribution.

Table 1: Marital Status of Recent Mexican Immigrants in the U.S. Census

	%	
	Males	Females
Married, with spouse present	14.4	48.3
Married, with spouse absent	26.9	8.7
Separated	1.8	3.7
Divorced	1.5	2.4
Widowed	0.2	0.7
Never married/Single	55.2	36.1

Source: US Census 5% public use sample (Ruggles et al, 2004)
Immigrants born in Mexico who migrated to the U.S. in last two years

Table 2: Summary Statistics of Key Variables

	Number of	All males 15-49		Migrants	
Individual level variables	Observations	Mean	Std dev.	Mean	Std dev.
ALL COMMUNITIES					
Proportion migrating in the last two years	55848	0.0099		1	
Age	55848	28.6	9.7	28.4	7.9
Years of Education	55848	8.7	4.4	7.0	3.6
Proportion married	55848	0.55		0.75	
Proportion who are household heads	55848	0.50		0.65	
Years of Education*Community migration prevalence	55848	1.26	1.3	1.90	1.4
COMMUNITIES WITH 100,000 OR LESS POPULATION					
Proportion migrating in the last two years	25531	0.014			
Age	25531	28.2	9.8	28.0	8.0
Years of Education	25531	7.3	4.1	6.3	3.2
Proportion married	25531	0.55		0.74	
Proportion who are household heads	25531	0.49		0.67	
Years of Education*Community migration prevalence	25531	1.31	1.55	2.21	1.55

Community level variables	Number of	All communities		> median prevalence	
(communities with 100,000 or less population)	Communities	Mean	Std dev.	Mean	Std dev.
Community migration prevalence	254	0.205	0.191	0.355	0.155
State migration rate in 1924	254	0.0068	0.0083	0.010	0.008
Percent of rural households owning land in 1910	254	2.89	2.17	3.41	2.29
Number of schools per 1000 population in 1930	254	1.22	0.43	1.14	0.36
Male school attendance in 1930 (% of 6 to 10 year olds)	254	44.64	11.60	44.61	11.23
Gini of schooling years for males 15-20 in 1960	254	0.51	0.10	0.51	0.10
Average years of schooling of males in 1960	254	2.92	0.81	3.01	0.83
Gini of household income in 1960	254	0.76	0.09	0.75	0.08

Source: 15-49 year old males in ENADID 1997, non-migrants and those migrating in last 2 years for the first time only.

Table 3: Determinants of First-time Migration for Male 15-49 year olds

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	OLS	OLS	OLS	OLS	2SLS	2SLS	2SLS	2SLS	2SLS
Years of Education	0.0672**	0.0548*	0.1485***	0.1801***	0.0800*	0.1936***	0.1969***	0.0782*	0.2012***
	(0.0327)	(0.0315)	(0.0477)	(0.0557)	(0.0442)	(0.0564)	(0.0633)	(0.0441)	(0.0559)
Years of Education Squared	-0.0037**	-0.0036**	-0.0105***	-0.0132***	-0.0040**	-0.0106***	-0.0132***	-0.0042**	-0.0117***
	(0.0017)	(0.0017)	(0.0029)	(0.0036)	(0.0017)	(0.0029)	(0.0035)	(0.0017)	(0.0026)
Community Migration Network	10.1308***	10.0671***	9.8967***	9.3965***	11.4788***	11.0387***	9.5639***		
	(1.0325)	(1.0770)	(1.1400)	(1.1865)	(2.2570)	(1.9387)	(1.9690)		
Education*Network	-0.4120***	-0.3510***	-0.3082**	-0.2453*	-0.4637**	-0.5521**	-0.3441	-0.4669**	-0.4318**
	(0.1022)	(0.1105)	(0.1240)	(0.1343)	(0.2185)	(0.2248)	(0.2445)	(0.1906)	(0.2120)
Age	0.1714***	0.1716***	0.1994***	0.2282***	0.1739***	0.1944***	0.2243***	0.1771***	0.2027***
	(0.0324)	(0.0324)	(0.0561)	(0.0692)	(0.0328)	(0.0560)	(0.0696)	(0.0272)	(0.0494)
Age Squared	-0.0034***	-0.0034***	-0.0042***	-0.0047***	-0.0034***	-0.0041***	-0.0047***	-0.0035***	-0.0043***
	(0.0005)	(0.0005)	(0.0009)	(0.0011)	(0.0005)	(0.0009)	(0.0011)	(0.0004)	(0.0008)
Married dummy	0.8508***	0.8571***	0.9057***	1.0039***	0.8526***	0.9058***	1.0015***	0.8709***	0.9184***
	(0.1645)	(0.1639)	(0.2799)	(0.3279)	(0.1650)	(0.2786)	(0.3257)	(0.1418)	(0.2348)
Head of household dummy	0.3493**	0.3466**	1.0114***	1.0580***	0.3502**	1.0213***	1.0624***	0.3265**	1.0424***
	(0.1673)	(0.1668)	(0.2608)	(0.3147)	(0.1669)	(0.2612)	(0.3134)	(0.1498)	(0.2525)
Population 100,000 +	-0.2981**	0.0855			-0.2442				
	(0.1450)	(0.1355)			(0.1670)				
Population 20,000-99,999	-0.2756*	-0.3790*			-0.2713				
	(0.1637)	(0.2037)			(0.1661)				
Population 15,000-19,999	-0.1106	-0.1904			-0.1737				
	(0.3081)	(0.3045)			(0.3119)				
Proportion of rural households owning land in 1910	-0.0000	0.0000	0.0005	0.0005	-0.0001	0.0006	0.0006		
	(0.0003)	(0.0003)	(0.0006)	(0.0007)	(0.0003)	(0.0006)	(0.0007)		
Male school attendance rate in 1930	0.0000	0.0000	-0.0001	-0.0001	-0.0000	-0.0001	-0.0000		
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)		
Gini of household income 1960	-0.0007	0.0028	0.0047	0.0066	-0.0024	0.0041	0.0075		
	(0.0050)	(0.0056)	(0.0077)	(0.0099)	(0.0057)	(0.0078)	(0.0107)		
Schools per 1000 in 1930	-0.0004	-0.0008	0.0018	0.0011	0.0003	0.0014	0.0009		
	(0.0018)	(0.0018)	(0.0026)	(0.0034)	(0.0016)	(0.0024)	(0.0032)		
Gini of Male schooling in 1960	0.0125	0.0081	0.0286	0.0352	0.0138	0.0309	0.0365		
	(0.0134)	(0.0129)	(0.0204)	(0.0279)	(0.0138)	(0.0199)	(0.0275)		
Average Male years of schooling 1960	0.0013	0.0012	0.0034	0.0042	0.0014	0.0036	0.0045		
	(0.0016)	(0.0016)	(0.0024)	(0.0032)	(0.0016)	(0.0024)	(0.0032)		
Population 100,000+ * Network		-2.8767**							
		(1.3228)							
Population 20,000-99,999 * Network		0.5835							
		(1.5875)							
Population 15,000-19,999 * Network		0.2056							
		(2.0647)							
Constant	-0.0346**	-0.0345**	-0.0598***	-0.0723**	-0.0373***	-0.0625***	-0.0739**		
	(0.0135)	(0.0134)	(0.0216)	(0.0288)	(0.0136)	(0.0210)	(0.0287)		
Population range covered	All	All	<100K	<15K	All	<100K	<15K	All	<100K
First-stage F-statistics:									
Community Migration Network					34.69	54.14	59.32		
Network * Years of Education					34.93	40.51	56.79	3084.73	2183.4
Municipality Fixed Effects	no	no	no	no	no	no	no	yes	yes
Observations	55848	55848	25531	18097	55848	25531	18097	55848	25531
Number of Communities	288	288	254	228	288	254	228	288	254

Notes:

Robust standard errors in parentheses with standard errors clustered at the municipality (community) level (columns 1 to 7)

Columns 8 and 9 include municipality fixed effects, and use White-corrected standard errors.

All coefficients apart from the constant are multiplied by 100 for display purposes

Instruments used are 1924 state migration rate, and its interaction with years of education.

*, ** and *** indicate significance at the 10%, 5% and 1% levels respectively.

Table 4: Robustness to Undercounting and Endogenous Education

	Unadjusted as per Table 3		Adjusting for Undercounting		Adjusting for Endogenous Education 2SLS adjustment		Adjusting for Endogenous Education COP adjustment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
Years of Education	0.0800*	0.1936***	0.0793	0.2076***	0.5175***	0.6735***	0.6354***	0.7936***
	(0.0442)	(0.0564)	(0.0520)	(0.0652)	(0.0775)	(0.0939)	(0.0929)	(0.1096)
Years of Education Squared	-0.0040**	-0.0106***	-0.0036*	-0.0108***	-0.0164***	-0.0232***	-0.0191***	-0.0251***
	(0.0017)	(0.0029)	(0.0020)	(0.0034)	(0.0023)	(0.0034)	(0.0025)	(0.0036)
Community Migration Network	11.4788***	11.0387***	12.7339***	12.2867***	18.9076***	17.3689***	21.3500***	19.4326***
	(2.2570)	(1.9387)	(2.5931)	(2.1293)	(2.9422)	(2.3138)	(3.3127)	(2.5343)
Education*Network	-0.4637**	-0.5521**	-0.5076**	-0.6203**	-1.3764***	-1.4589***	-1.6772***	-1.7545***
	(0.2185)	(0.2248)	(0.2568)	(0.2487)	(0.2890)	(0.2945)	(0.3374)	(0.3337)
Age	0.1739***	0.1944***	0.1923***	0.2133***	0.1579***	0.1654***	0.1507***	0.1532***
	(0.0328)	(0.0560)	(0.0365)	(0.0627)	(0.0321)	(0.0551)	(0.0319)	(0.0549)
Age Squared	-0.0034***	-0.0041***	-0.0038***	-0.0046***	-0.0030***	-0.0034***	-0.0028***	-0.0032***
	(0.0005)	(0.0009)	(0.0006)	(0.0010)	(0.0005)	(0.0008)	(0.0005)	(0.0008)
Married dummy	0.8526***	0.9058***	0.9518***	0.9911***	0.8868***	0.9368***	0.9040***	0.9556***
	(0.1650)	(0.2786)	(0.1862)	(0.3060)	(0.1657)	(0.2797)	(0.1664)	(0.2809)
Head of household dummy	0.3502**	1.0213***	0.3854**	1.1707***	0.3600**	1.0698***	0.3617**	1.0839***
	(0.1669)	(0.2612)	(0.1859)	(0.2882)	(0.1678)	(0.2634)	(0.1682)	(0.2648)
Population range:	All	<100K	All	<100K	All	<100K	All	<100K
First-stage F-statistics:								
Community Migration Network	34.69	54.14	34.72	54.19	34.69	54.14	34.92	55.07
Network * Years of Education	34.93	40.51	34.98	40.59	34.93	40.51	35.07	41.06
Observations	55848	25531	55848	25531	55848	25531	55848	25531
Number of Communities	288	254	288	254	288	254	288	254

Notes:
Robust standard errors in parentheses with standard errors clustered at the municipality (community) level
All coefficients are multiplied by 100 for display purposes
All regressions also include a constant and the state-level controls included in Table 3
Instruments used are 1924 state migration rate, and its interaction with years of education.
Columns 5 though 8 adjust education using 2SLS and Censored Ordered Probit (COP) estimates of the impact of migration of parents on individual education levels, assuming all migrants have migrant parents and no migrants do (see text).
*, ** and *** indicate significance at the 10%, 5% and 1% levels respectively.

Table A1: First-stage results

For communities with 100,000 or less population (column 6, Table 3)

	Migration Network	Migration Network *Education
Instruments		
1924 state migration rate	0.1656***	0.3447***
	(0.0160)	(0.1060)
1924 state migration rate * years of education	-0.0035**	0.0834***
	(0.0015)	(0.0258)
Regression controls		
Years of Education	-0.0011	0.1304***
	(0.0016)	(0.0248)
Years of Education Squared	0.0001	-0.0007
	(0.0001)	(0.0012)
Age	-0.0036***	-0.0269***
	(0.0009)	(0.0073)
Age Squared	0.0001***	0.0004***
	(0.0000)	(0.0001)
Married dummy	-0.0037	-0.0352
	(0.0039)	(0.0317)
Head of household dummy	-0.0033	-0.0018
	(0.0047)	(0.0381)
Proportion of rural households owning land in 1910	0.0001**	0.0009***
	(0.0000)	(0.0003)
Male school attendance rate in 1930	0.0000	0.0001
	(0.0000)	(0.0001)
Gini of household income 1960	-0.0001	-0.0033
	(0.0009)	(0.0069)
Schools per 1000 in 1930	0.0007**	0.0050**
	(0.0003)	(0.0022)
Gini of Male schooling in 1960	0.0101***	0.0781***
	(0.0025)	(0.0201)
Average Male years of schooling 1960	0.0003	0.0029
	(0.0003)	(0.0022)
Constant	-0.0062***	-0.0573***
	(0.0023)	(0.0182)
Observations	25531	25531
R-squared	0.32	0.42
Shea Partial R-squared	0.27	0.24
F-statistic on instruments	54.14	40.51

Notes:

Robust standard errors in parentheses with standard errors clustered at the municipality (community) level

All coefficients apart from the constant are multiplied by 100 for display purposes

*, ** and *** indicate significance at the 10%, 5% and 1% levels respectively.