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

David McKenzie\textsuperscript{a} and Hillel Rapoport\textsuperscript{b}

\textsuperscript{a} World Bank, Development Economics Research Group  
\textsuperscript{b} Department of Economics, Bar-Ilan University,  
and CADRE, University of Lille II

First draft – January 2006  
Preliminary – do not quote

Abstract:  
This paper examines the role of migration networks on self-selection patterns in Mexico-U.S. migration. We first present a simple theoretical framework showing how such networks impact on migration incentives for 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 levels 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, migrant networks, education attainments, self-selection, Mexico  
JEL codes:

\textsuperscript{1} Corresponding author…. We thank….
1. Introduction
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). Borjas (1987) uses a Roy model of migration selection to show that migrants from countries with more inequality will be negatively selected. Intuitively, higher inequality means that the top of the wage and skill distribution in Mexico will be relatively better rewarded in Mexico than in the United States, lessening their incentive to migrate. Chiquiar and Hanson (2005) indeed show that returns to education are substantially higher in Mexico than they are for Mexican immigrants in the United States, a finding we expect to be particularly the case for the predominately illegal migrants in our sample. If migration were costless, one would therefore expect to find the probability of migration from Mexico to the U.S. to be decreasing in skill, as measured by education. However, Chiquiar and Hanson (2005) find migration rates to be increasing in education up to relatively high education levels, which they interpret as evidence that migration costs are important.

In practice international migration is costly, involving upfront monetary costs, information and search costs, opportunity costs in terms of income foregone while traveling and searching for work, and psychic costs (Massey, 1988). Illegal migrants often rely on smugglers (coyotes) to help them cross the border. Orrenius (1999) reports the median cost of a coyote was $619 in 1994, having fallen over time. However, the INS estimates that the cost increased substantially since then, especially following increased border enforcement after September 11, 2001, with prices of $1,500-2,000 in 2002.\(^2\) Many of these costs are constant across education levels, whereas others, such as the costs of acquiring information and of dealing with complicated paperwork for legal migration, are likely to be decreasing with education. Coupled with the presence of credit constraints which bind more heavily for the low-skilled, these costs may result in individuals with less education being unable to emigrate, despite U.S.-Mexico wage differentials being highest for them. If these costs are large enough, one may then find migrants being positively selected from the Mexican education distribution.\(^3\)

Migration networks can act to lower the costs of migrating (Massey, 1988; Orrenius, 1999) and to relax credit constraints (Genicot and Sénèschy, 2004). 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, 2004, Bauer et al., 2005). Based on these expected effects of networks on the pattern of migration, one may suspect that everything else equal, education will increase the probability of migration and result in positive self-selection of migrants in communities with small networks, and conversely the probability of migration will be decreasing with education in communities with large networks.

---


\(^3\) Or at least coming from the upper-middle of the education distribution. Those with sufficiently high education and wealth may not find it worthwhile to migrate, so we would expect to see the probability of migration decreasing with education at high levels of education.
In the following we take education decisions as exogenous and investigate how education affects the pattern of migration (the type of self-selection obtained). We first show this theoretically in a simple model of self-selection adapted from Chiquiar-Hanson (2005) and augmented to allow for network effects. 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. This is consistent with the findings of Chiquiar and Hanson (2005) who found evidence of positive and intermediate selection of Mexican immigrants in the US labor market. 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). Also, this demonstrates the instrumental role played by migration networks in determining the pattern of migration, and goes part of the way towards reconciling the conflicting results from previous studies on self-selection types.

2. The model

Using their notations, we extend the simple model of Chiquiar and Hanson (2005), itself adapted from Borjas (1987), to allow for network effects.

The wage equation in Mexico (subscript 0) is given by:
\[ \ln w_0 = \mu_0 + \delta_0 s \]  
(1)
where \( w \) is the wage, \( \mu > 0 \) is the base wage, \( \delta > 0 \) 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:
\[ \ln w_1 = \mu_1 + \delta_1 s \]  
(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 \( \delta_0 > \delta_1 \).

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, \( n \):
\[ C = C(n), C' < 0 \]  
(3)

Expressed in time-equivalent units, the migration cost may be written as:
\[ \pi = \pi(n, s) = \frac{C(n)}{w_0} \]  
(4)

Then, a resident of Mexico will find it beneficial to migrate to the U.S. if:
\[ \ln(w_1) - \ln(w_0 + C) \equiv \ln(w_1) - \ln(w_0) - \pi > 0 \]  
(5)

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 \( \pi = e^{n_{1}+k_{0}+k_{1}} \), with \( s \) and \( k \) independently distributed. Note also that \( k \) could be seen as positively depending on \( n \) as having landsleit 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 quality 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.

The approximation is valid if \( \pi \) is small, which we can easily accept if \( w \) is defined as the present value of a flow of future incomes.
We assume, as in Chiquiar and Hanson (2005), that time-equivalent migration costs decrease with schooling (many references that confirm this; justify also using a credit-constraint argument). In addition, as explained, migration costs and therefore time-equivalent migration costs decrease with the size of the community migration network, \( n \):

\[
\ln(\pi) = \mu - \gamma_1 s - \gamma_2 n \tag{6},
\]

so that \( \pi = e^{\mu - \gamma_1 s - \gamma_2 n} \) 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 \( A = \mu_1 + \delta s - e^{\mu_1 - \gamma_1 s} \) (see the solid line in Figure 1). In order not to rule out the possibility of positive self-selection, we also assume with Chiquiar and Hanson (2005) that \( \mu_1 - \mu_0 < e^{\mu_1} \). In words, the inter-country 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. 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 \( [s_s, \bar{s}] \).

Now, if \( s_s < s_L < \bar{s} < s_U \), then positive self-selection obtains. Conversely, if \( s_L < \bar{s} < s_U < s_s \), then negative self-selection obtains.\(^7\)

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. The new wage profile at destination is now given by \( B = \mu_1 + \delta s - e^{\mu_1 - \gamma_1 s - \gamma_2 n} \) (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:

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

\textit{Proof:} The induced change in migration incentives, which we denote by \( \Delta \), is given by the difference between \( A \) and \( B \): \( \Delta = e^{\mu_1 - \gamma_1 s} - e^{\mu_1 - \gamma_1 s - \gamma_2 n} = e^{\mu_s - \gamma_1 s} \left[ 1 - \frac{1}{e^{\gamma_2 n}} \right] > 0 \), with \( \partial \Delta / \partial n > 0 \) and \( \partial \Delta / \partial s < 0 \).

A change in migration incentives following the expansion of networks defines two new threshold values of \( s \), \( s_L \) and \( s_U \), with \( s'_s < s_L \) and \( s'_U > s_U \). As migration

\^6 To follow up on footnote 4 above, this would also require \( \gamma_3 < \frac{\mu_s}{\ln(\mu_1 - \mu_0)} \).

\^7 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. For example, positive self-selection obtains if \( \bar{s} < s_L \) and \( \bar{s} > s_U \) but the distribution of \( s \) is such that the density is highest between \( \bar{s} \) and \( s_L \).
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 \( \bar{s} < s_L < \bar{\bar{s}} < s_U \) (positive self-selection) or \( s_L < \bar{s} < s_U < \bar{\bar{s}} \) (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 now 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 \( s_L > 0, s_U < \bar{\bar{s}} \), 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.

With these understandings, we may now state:

**Proposition 2:** 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:** We must first note that \( s_L \) and \( s_U \) are solutions of the following equation:

\[
\mu_0 + \delta_0 s(n) = \mu_i + \delta_i s(n) - e^{\mu_{Li} - \gamma s(n) - \gamma i n}
\]  

(7)

Deriving (7) with respect to \( n \), we have:

\[
\delta_0 \frac{\partial s}{\partial n} = \delta_i \frac{\partial s}{\partial n} + (\gamma_i \frac{\partial s}{\partial n} + \gamma_i e^{\mu_{Li} - \gamma s(n) - \gamma i n})
\]

and therefore

\[
\frac{\partial s}{\partial n} = \frac{\gamma_i e^{\mu_{Li} - \gamma s(n) - \gamma i n}}{\delta_0 - \delta_i - \gamma_i e^{\mu_{Li} - \gamma s(n) - \gamma i n}}
\]  

(8).

The marginal effect of an increase in network size on the two critical schooling thresholds is therefore given by:
\[
\frac{\partial s_L}{\partial n} = \frac{\gamma_2 e^{\mu_1 - \gamma_1 \mu_1 - \gamma_1 n}}{\delta_0 - \delta_1 - \gamma_1 e^{\mu_1 - \gamma_1 \mu_1 - \gamma_1 n}} \quad (9)
\]

\[
\frac{\partial s_U}{\partial n} = \frac{\gamma_2 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_2 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}} \quad (10)
\]

Note that having \( \frac{\partial s_L}{\partial n} < 0 \) and at the same time \( \frac{\partial s_U}{\partial n} > 0 \) requires
\[
e^{\mu_1 - \gamma_1 \mu_1 - \gamma_1 n} < \frac{\delta_0 - \delta_1}{\gamma_1} < e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}.
\]
Combining (9) and (10), it is straightforward to see that
\[
\frac{1}{\frac{\partial s_L}{\partial n}} \left| \frac{\partial s_L}{\partial n} \right| = \frac{\gamma_2 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_2 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}} \quad \frac{\delta_0 - \delta_1 - \gamma_1 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}}{\gamma_2 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}} = \frac{e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}}{\delta_0 - \delta_1 - \gamma_1 e^{\mu_2 - \gamma_2 \mu_2 - \gamma_2 n}} \quad > 1
\]
which, together with the non-increasing density assumption, concludes our proof.

2. Data

This paper uses data from 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.\(^8\) 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. We restrict our analysis to rural communities, defined broadly here to be municipalities which are outside of cities of population 100,000 or more. In order to examine the impact of community migration networks, we require at least 50 households to be surveyed from each municipality. This gives a sample of 214 rural municipalities across all Mexican states. Within these communities we have a sample of 26,197 households, of which 9,758 households contain at least one child aged 12 to 18 years.

The ENADID asks whether household members have ever been to the United States in search of work. This question is asked of all household members who normally live in the household, even if they are temporarily studying or working elsewhere. Additional questions ask whether any household members have gone to live in another country in the past five years, capturing migration for study or other non-work purposes in addition to work related migration. We define a household as having a migrant if they have a member aged 19 and over who has ever been to the U.S. to work, or who has moved to the U.S. in the last five years for any other reason. We then define the community migration prevalence ratio as the proportion of all individuals aged 19 and over in the community who have ever migrated. We use a cutoff of age 19, rather than the age 15 threshold proposed by Massey, Goldring, and Durand (1994) in order to exclude migration by children aged 15-18. However as there are relatively few migrants in this age group, our results are robust to the use of this lower threshold.

Table 1 provides summary statistics for the key variables used in this study. The ENADID questions on migration within the last five years are identically worded.

\(^8\) Survey methodology, summary tables, and questionnaires are contained in INEGI (1999).
to those used in the 2000 Mexican Census, which does not capture migration by household members outside of a five-year window.

The ENADID asks migrants who have ever been to the U.S. for work a set of additional questions about their migrant experience, including the number of trips they have ever made, and whether they had legal documentation to work. Approximately 50 percent of all migrants have made more than one trip, with a mean of 2.8 trips per migrant. The vast majority of migrants in our sample had no legal documentation to work, especially on their first trip. Over 91 percent of first-time migrants who went to work in the U.S. had no legal documentation to do so.

Our main measure of education is based on years of schooling attained by adults. Elementary education (grades 1 to 6) is compulsory in Mexico and is normally provided to children aged 6 to 14. Lower secondary education (grades 7 to 9) became compulsory in 1993 and is generally given to children aged 12 to 16 years who have completed elementary education. This is followed by three years of upper secondary schooling (grades 10 to 12) and higher studies. Despite education being compulsory, there is still far from complete compliance and a lack of infrastructure in some of the more rural areas (SEP, 1999).

3. Identification Strategy

We follow Woodruff and Zenteno (2001) and a number of subsequent studies in using historic state-level migration rates as an instrument for current migration stocks. In particular, we use the U.S. migration rate from 1924 for the state in which the household is located, taken from Foerster (1925). These historic rates can be argued to be the result of the pattern of 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 become self-perpetuating. Hildebrandt and McKenzie (2004) show that the historic migration rate is a strong predictor of current migration rates, with a first-stage F-statistic of over 30.

4. Results

To test our predictions on the role of networks in determining the pattern of self-selection into migration, we estimate the following equation separately for males and females to predict whether individual $i$ in community $c$ migrates for the first time in the period 1996-97, conditional on never having previously migrated:

$$ M_{i,c} = \beta_0 + \beta_1 educ_{i,c} + \beta_2 educ_{i,c}^2 + \beta_3 network_c + \beta_4 educ_{i,c} \times network_c + \phi' X_{i,c} + \lambda' Z_c + \varepsilon_{i,c} $$

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

---

9 Hanson and Woodruff (2003); McKenzie and Rapoport (2004); López-Córdoba (2004); and Hildebrandt and McKenzie (2004) all employ historic migration rates as instruments for current migration.

10 Thanks to Chris Woodruff for supplying these historic rates.
characteristics respectively. Although $M_{i,c}$ is a binary variable, we use ordinary least squares for estimation in order to provide for easier interpretation of the magnitudes of the coefficients. We supplement this with two-stage least square estimates in which the 1924 historic migration rate and its interaction with years of schooling are used to instrument community migration prevalence and its interaction with education. We obtained similar results using probit and IV-probit estimation. The interaction term between education and community migration prevalence allows for the impact of education on the likelihood of migration to differ 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 $\beta_4$ should be negative.

Table 2 presents the results of estimating equation (1). For both males and females we find a significant positive coefficient on the linear term in education and a significant negative coefficient on the quadratic term, so that the probability of migration is hump-shaped with respect to education. Migration is more likely in communities with larger networks, with this effect much larger in magnitude for males, and insignificant for females after instrumenting. The interaction between education and network size is significant and negative for males, whereas we find no evidence of an interaction effect for females. Most of the community control variables are insignificant, although we do find slightly higher migration rates from communities in states that had more unequal distributions of schooling in 1960.

To better explore the predicted relationship between education, network size, and the probability of migration, Figure 2 plots the predicted probabilities from 2SLS estimation at different percentiles of the community migration prevalence distribution. The results for males in Figure 2a are very striking and provide strong support for the theoretical predictions discussed above. In communities with small migration networks, where costs of migrating are likely to be high, one finds the probability of migration to be increasing in completed years of schooling up to nine or ten years of schooling, which occurs at the 75th percentile of the education distribution. These results concur with the finding of Chiquiar and Hanson (2005) that migrants tend to be selected from the upper-middle of the education distribution, resulting in positive selection. However, as the migrant network grows, lowering the costs of migration and reducing credit constraints, we find a reversal in this pattern, with the highest probabilities of migration occurring at education levels of two to five years, resulting in negative selection. That is, when migration costs become less important and less binding due to large migration networks, the relationship between migration propensities and education is in line with 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. Our results expand on the finding of Winters, de Janvry and Sadoulet (2001) that household variables representing skill levels are important determinants of the decision to migrate in communities with small networks, but not for those with large networks. They also complement the findings of McKenzie and

---

11 Winters, de Janvry and Sadoulet (2001) split their sample into small and large network communities. They include some municipal level and ejido-level controls, but conditional on this treat network size as exogenous in their calculations.
Rapoport (2004), who find similar patterns in the migration and wealth relationship, whereby migrants tend to come from the upper-middle of the wealth distribution when networks are small, and from the lower end of the wealth distribution when networks are large in size.

In contrast, while we find a small increase in the propensity to migrate for females with larger networks, there is no evidence that migration networks change the relationship between education and the likelihood of migrating. The highest migration propensity occurs for females with between eight and nine years of education, which lies between the 61st and 83rd percentiles of the female education distribution. Therefore female migrants tend to be positively selected on average. Female migration is relatively rare compared to male migration: 82 percent of all migrants in our sample are male. As a result, we have a much smaller sample of female migrants with which to detect effects. Nevertheless, our finding that the marginal effect of the network is larger for males than females is consistent with Davis and Winters (2001), who find this result using data from Mexican ejidos. They summarize a number of potential reasons why female migration patterns may differ from male migration patterns. Their survey found significant differences between male and females in terms of occupational choice, with female migrants much less likely to work in agriculture and relatively more likely to work in skilled employment and other types of non-agricultural work. It is likely that education matters more, and earns a higher return, in these occupations than in agricultural work. These patterns coupled with less support from the network in terms of lowering the cost barriers to migration may account for the continued positive selection of female migrants, even in communities with large networks.

5. 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.

5. References


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


Figure 1: Migration networks and self-selection patterns

\[ A = \mu_1 + \delta S - e^{\mu_s \gamma_s} \]

\[ B = \mu_1 + \delta S - e^{\mu_s \gamma_s} \]

\[ \ln w \]

\[ s \]
FIGURE 2: PROBABILITY OF FIRST-TIME MIGRATION IN 1996-97 ACCORDING TO EDUCATION AND NETWORK SIZE

Figure 2a: Males aged 15–49

Figure 2b: Females aged 15–49

Predicted probabilities of migration obtained from instrumental variables regressions in columns 3 and 6 of Table 3 are plotted at the 10th, 30th, 50th, and 70th percentiles of the community migration prevalence distribution.
<table>
<thead>
<tr>
<th>TABLE 1: SUMMARY STATISTICS OF KEY VARIABLES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Observations</td>
</tr>
<tr>
<td>-------------------------</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Household Variables (for households with a child aged 12 to 18)</strong></td>
</tr>
<tr>
<td>Proportion of Households with a migrant</td>
</tr>
<tr>
<td>Proportion of Households with a migrant by census definition</td>
</tr>
<tr>
<td>Proportion receiving remittances</td>
</tr>
<tr>
<td>Percentage share of income from remittances</td>
</tr>
<tr>
<td><strong>Individual Variables</strong></td>
</tr>
<tr>
<td>Years of Schooling of Mother for children aged 12 to 18</td>
</tr>
<tr>
<td>Years of Schooling of Males 12 to 15</td>
</tr>
<tr>
<td>Years of Schooling of Males 16 to 18</td>
</tr>
<tr>
<td>Years of Schooling of Females 12 to 15</td>
</tr>
<tr>
<td>Years of Schooling of Females 16 to 18</td>
</tr>
<tr>
<td><strong>Community level variables</strong></td>
</tr>
<tr>
<td>Number of communities</td>
</tr>
<tr>
<td>Community migration prevalence</td>
</tr>
<tr>
<td>Gini of Years of Education for Males 12 to 15</td>
</tr>
<tr>
<td>Gini of Years of Education for Males 16 to 18</td>
</tr>
<tr>
<td>Gini of Years of Education for Females 12 to 15</td>
</tr>
<tr>
<td>Gini of Years of Education for Females 16 to 18</td>
</tr>
<tr>
<td>State migration rate in 1924</td>
</tr>
<tr>
<td>Proportion of rural households owning land in 1910</td>
</tr>
<tr>
<td>Male School Attendance in 1930 (6 to 10 year olds)</td>
</tr>
<tr>
<td>Female School Attendance in 1930 (6 to 10 year olds)</td>
</tr>
<tr>
<td>Gini of Household Income in 1960</td>
</tr>
<tr>
<td>Number of Schools per 1000 population in 1930</td>
</tr>
<tr>
<td>Gini of Years of Schooling for Males 15-20 in 1960</td>
</tr>
<tr>
<td>Gini of Years of Schooling for Females 15-20 in 1960</td>
</tr>
<tr>
<td>Average Male Years of Schooling in 1960</td>
</tr>
<tr>
<td>Average Female Years of Schooling in 1960</td>
</tr>
</tbody>
</table>

Source: own calculation from ENADID 1997 communities with population <100,000 and 50 or more households sampled. Education Ginis are only reported for communities with 20 or more children in the given age category.
<table>
<thead>
<tr>
<th></th>
<th>Males</th>
<th>Females</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) OLS</td>
<td>(2) OLS</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>0.0013</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>(4.40)**</td>
<td>(4.64)**</td>
</tr>
<tr>
<td>Years of Schooling Squared</td>
<td>-0.0001</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>(4.50)**</td>
<td>(4.82)**</td>
</tr>
<tr>
<td>Community Migration Prevalence</td>
<td>0.0860</td>
<td>0.0847</td>
</tr>
<tr>
<td></td>
<td>(7.38)**</td>
<td>(6.98)**</td>
</tr>
<tr>
<td>Years Schooling * Community Migration Prevalence</td>
<td>-0.0020</td>
<td>-0.0021</td>
</tr>
<tr>
<td></td>
<td>(-1.68)</td>
<td>(1.66)</td>
</tr>
<tr>
<td>Age</td>
<td>0.0036</td>
<td>0.0040</td>
</tr>
<tr>
<td></td>
<td>(4.31)**</td>
<td>(4.59)**</td>
</tr>
<tr>
<td>Age Squared</td>
<td>-0.0001</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>(-4.27)**</td>
<td>(-4.53)**</td>
</tr>
<tr>
<td>Proportion of rural households owning land in 1910</td>
<td>0.0002</td>
<td>0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.37)</td>
<td>(0.42)</td>
</tr>
<tr>
<td>School Attendance in 1930 (6 to 10 year olds)</td>
<td>-0.0000</td>
<td>-0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.46)</td>
<td>(0.54)</td>
</tr>
<tr>
<td>Gini of Income in 1960</td>
<td>-0.0040</td>
<td>-0.0052</td>
</tr>
<tr>
<td></td>
<td>(0.59)</td>
<td>(0.86)</td>
</tr>
<tr>
<td>Number of Schools per 1000 population in 1930</td>
<td>0.0012</td>
<td>0.0015</td>
</tr>
<tr>
<td></td>
<td>(0.53)</td>
<td>(0.74)</td>
</tr>
<tr>
<td>Gini of Years of Schooling for 15-20 year olds in 1960</td>
<td>0.0321</td>
<td>0.0341</td>
</tr>
<tr>
<td></td>
<td>(1.72)</td>
<td>(1.84)</td>
</tr>
<tr>
<td>Average Years of Schooling in 1960</td>
<td>0.0022</td>
<td>0.0022</td>
</tr>
<tr>
<td></td>
<td>(1.10)</td>
<td>(1.11)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0534</td>
<td>-0.0782</td>
</tr>
<tr>
<td></td>
<td>(4.63)**</td>
<td>(3.61)**</td>
</tr>
<tr>
<td>Observations</td>
<td>23226</td>
<td>22078</td>
</tr>
<tr>
<td>Communities</td>
<td>214</td>
<td>210</td>
</tr>
</tbody>
</table>

Notes:
1. These variables are for historic male schooling in columns 1-3, and historic female schooling in columns 4-6.
T-statistics are in parentheses with standard errors clustered at the state level.
Instruments are 1924 state-level migration rate and its interaction with years of schooling.
* significant at 5%; ** significant at 1%