Are Mexican migrants to the US adversely selected on ability?[1]

Eric R. Jensen[2]

Sarah M. Gale[3]

Paul E. Charpentier[4]

Version date: August 19, 2010

[1] We thank Frank Bean, David Card, George Grayson, Peter McHenry, David Jaeger, and Yana Rodgers for helpful comments and Dheeraj Jagadev and Shannon White for capable research assistance. Contact author is Eric Jensen (eric_jensen@wm.edu)

[2] Thomas Jefferson Program in Public Policy and Economics Department, College of William and Mary

[3] MBA candidate, Columbia University School of Business

[4] Allstate Insurance Company

Are Mexican migrants to the US adversely selected on ability?

Abstract

Recent migrants to the United States may be displaying lower earnings levels and a slower rate of earnings convergence with natives than previous immigrants. Borjas has argued that this reflects adverse selection of immigrants; others, including Card, Chiquiar and Hanson, and Smith, question this contention. Some of the ambiguity may be due to measurement problems, with schooling measured at varying levels of aggregation used in place of unobserved migrant quality. Using Mexican Migration Project data, we construct measures that isolate some sources of variation. Our findings regarding adverse selection of migrants are mixed. Echoing Smith's (2006) cautions on the use of earnings as a measure of migrant quality, we suggest that supply constraints in its provision render educational attainment alone a poor indicator of ability.

Are Mexican migrants to the US adversely selected on ability?

In an influential body of work, Borjas (1990, 1991, 1994, 1999) has extended the Roy (1951) model of endogenous selection to analyses of international migration to the United States. Roy's well known result is that for individuals choosing between two occupations, those in the left tail of the occupation with a relatively diffuse earnings distribution have an incentive to shift to the other occupation, since, with relatively smaller variance in earnings, they are closer to mean earnings in the latter occupation. Borjas applies this logic to international migration. He asserts that placement within a national income distribution is reflective variously of ability, skills or "ethnic capital," and defines ethnic capital conceptually as the quality of the ethnic environment in which a person is raised and operationally as the average educational attainment in a sending country. Borjas completes the model with the implicit assumption that the variance differential in earnings between the US and Mexico is sufficiently large to dominate the US mean earnings advantage, and asserts that a significant share of Mexican migration to the US is the result of adverse selection by migrants (*e.g.,* Borjas 1994).

In later work of Borjas and other contributors to this literature, discussion centers on the placement of migrants within the distribution of skills, with special regard for the transferability of skills acquired in the sending country, and on measurement issues. Empirical support for adverse selection of migrants to the US has been mixed. Duleep and Regets (1999), for example, predict that those with more to gain in investing in human capital, including migrating, will do so. This is partly driven by low opportunity cost and partly by high expected benefit of migrating, and so Duleep and Regets predict that earnings will grow faster for immigrants than for the native born. Schoeni (1997) finds less reassuring evidence on this count, but Smith (2006) urges caution in using wage data to make inferences regarding migrant quality. Smith suggests that increasingly young migrants, high wage premia for skilled occupations, and low earnings of illegals account for a substantial share of the observed wage differential between migrants and native-born. Chiquiar and Hanson (2005) compare counterfactual predictions of skill premia for US immigrants, had they remained in the Mexican labor market, with the actual distribution of US earnings. Rather than adversely selecting, they find that those who left Mexico for the US would have fallen in the middle and upper portion of the Mexican wage distribution. Mckenzie and Rapoport (2007) also cast doubt on the adverse selection hypothesis, confirming an inverse U-shaped relationship between the probability of migration and wealth due in part to the high cost of migration for low-wealth individuals.

Roy's self-selection model is very much a story of efficient outcomes and functioning markets. His hunters and fishers knew that they were good at one or the other activity, and the model is based fundamentally on their freedom to sort themselves by occupation (Roy 1951). However, in many developing countries, educational resources are sufficiently concentrated in urban areas that the notion that self-selection accounts for much of the observed national-level variation in educational attainment becomes difficult to credit. Behavior of migrants from rural areas is important, because it is the driving force in Mexican migration to the US. For example, Mexican census data show that although the Mexican population is roughly two-thirds urban, only 28.2% of migrants to the US during the period 1995-2000 were from areas with populations greater than 100,000 (Wong, Resano-Perez, and Martinon, 2006). Because disproportionate social spending in urban areas means that educational supply constraints are more likely to bind in rural areas, levels of educational attainment will be lower in rural than in urban areas, all else constant. Absent reason to believe that the distributions of other elements of human capital, particularly ability, are similarly skewed along rural-urban lines, the usefulness of years of schooling as a measure of quality seems on the face of it to be limited. There is additional state-to-state variation in spending between and within rural areas, so subtler geographic heterogeneity in the supply of education than would be captured by a simple rural/urban categorization also is likely to blur the connection between ability and educational attainment.

Where supply constraints bind, rather than an efficient Roy sorting mechanism, the noisy relationship between migrant quality, broadly, and presumptive measures of skill such as educational attainment allows other sources of variation to exert influence on migration. Interpreting low educational attainment of rural migrants to the US as being reflective of adverse selection may in turn lead to misleading policy recommendations. In particular, if migrant educational attainment is low because of supply constraints in the sending village or, after Smith, migrant earnings are low because migrants are young, the likely path of earnings of this generation and subsequent generations is much more favorable to migrants than one where adverse selection truly operates.

We use data from the Mexican Migration Project to replicate an apparent adverse selection result. We then construct a model that accounts for potentially supply-constrained access to education within Mexico. We find that it is individuals who are well-educated by local standards but from impoverished, education-constrained regions of Mexico, rather than a broad cross-section of the less educated, who are most likely to migrate to the United States. In the process, we offer evidence reinforcing the Chiquiar and Hanson result for earnings, with our finding that migrants are most likely to be those from the middle of sending area educational attainment distributions. Within sending communities, we confirm the Mckenzie and Rapoport result that Mexicans at the lowest levels of education are much less likely to migrate to the US than are those near the middle of the local distribution.

Negative or positive selection of immigrants?

Much of the work on this topic has been done with highly aggregated data. For example, Borjas has regressed a dummy variable representing the degree of income inequality in the sending country relative to the United States or, closely related, the ratio of earnings of the top households to the earnings of lowest-income households in the sending country, on the level and rate of change of national averages of immigrants' earnings (Borjas 1990). He finds that either of these country-level regressors is inversely associated with earnings and interprets these findings as reflective of adverse selection of migrants on ability or skill differentials. As Borjas acknowledges, either approach is an indirect means of testing of the Roy's model of individual behavior. However, in a more direct examination employing individual-level Mexican census data, Ibarraran and Lubotsky (2007) also found that migrants to the United States tended to be less educated than nonmigrants.

Figure 1, based on the Mexican Migration Project data that we employ in the present paper and describe more fully below, illustrates the relationship between educational attainment and proportions ever migrating to the US, and replicates the Ibarraran and Lubotsky result with a different data set. We limit the sample to those between 16 and 64 years of age. Excluding a small proportion of outliers reporting more than 20 years of education yields a sample of just over 90,000 observations. The lines are lowess-smoothed sample proportions ever migrating to the US as a function of educational attainment in years, and the histograms are the underlying distributions of schooling. Results are presented separately for those residing in rural and urban areas. For rural areas, roughly 18% of those with no education had ever migrated to the US.[5] The corresponding figure from urban areas was 16%. Peak migration rates in both rural and urban areas were at levels of education well below sample medians, and the inverted-U shapes of the distributions yield migration probabilities that do not fall to the levels displayed at zero educational attainment until at least ten years of schooling are attained.

(Figure 1 here)

This sort of result has been contentious, in part because adverse selection is often seen as an unexpected finding in human capital modeling of migration. Chiswick (1999) notes that the Roy model is a special case of the standard human capital model of migration (*e.g.*, Sjaastad 1962). According to Chiswick, adverse selection is a tempering influence on what is likely, because returns to human capital investments are greater for those with higher ability, to be positive selection overall. Regarding the specific mechanism through which positive selection would operate, Chiquiar and Hanson (2005) point out that border crossing entails lower costs for the more educated, because they have less difficulty obtaining the needed cash. In their model, high ability individuals, able to bear the financial cost of migration yet still benefiting from higher wages in the United States, are drawn there.

In Roy's (1951) model, fishing is relatively difficult, and so a reasonable distribution of skill differentials over the population implies that some individuals will have large potential earnings from fishing, but others will catch few fish. Conversely, the less demanding occupation of hunting will have a narrower dispersion in potential earnings.[6] Roy's key insight was that

---

[5] Note that the MMS responses indirectly include some individuals in the US at the time of the survey.

[6] If potential earnings in the two occupations are positively correlated, as is required for adverse selection to occur, fishers will also earn more than hunters, on average. This is because the dispersion itself causes selection on ability. In contrast, Mexican earnings have slightly higher variance but a lower mean in comparison with the US.

someone equally bad at hunting and fishing would be more likely to choose hunting, because of the lower responsiveness of hunters' earnings to skill differentials. The underlying occupational decision is based upon the conditional expected value of earnings, however, and there are other sources of earnings variation than the dispersion of the underlying distributions. For example, as Roy discusses, if the price of meat increases, we would expect a reallocation of the labor force into hunting, as the lowest-earning fishermen leave that field. Differences in variances are in this sense a second-order source explaining occupational self-selection.

Turning back to Mexicans who self-select to migrate to the US, it seems unlikely on the face of it that relatively small differences in the variances of earnings would dominate the very large differences existing in earnings levels between the US and Mexico. The World Bank (World Bank, 2010) reports that average purchasing power parity in the US was more than three times that in Mexico in 2009, while the OECD reports mid-2000s Gini coefficients (after taxes and transfers) of 0.38 for the US and 0.46 for Mexico (OECD 2010). It perhaps is not surprising in this context that when Chiquiar and Hanson project as counterfactuals the Mexican earnings of US immigrants, they find them to be solidly in the middle of the Mexican earnings distribution. The adverse selection evidence of Borjas, Ibarraran and Lubotsky, and others, including the relationship we depict in Figure 1, therefore constitutes a puzzle.

In the remainder of this section, we lay out graphically the logic of a model, more fully developed in subsequent sections, that relies on educational constraints to reconcile what seem to be very different findings regarding migrant quality. We begin in Figure 2 by roughly replicating Figure 1, except that community means for educational attainment are used in place of individual values. The pattern apparent in Figure 1 is even more striking in Figure 2. Migration probabilities are high for individuals from communities where mean education levels are below average, peak at levels well below the mean level of educational attainment, and are strikingly low from those communities with high levels of education. [7] This graph represents individual migration probabilities solely as a function of community-level averages, and so, while its symmetry with Figure 1 is appealing, the behavioral analogy to adverse selection is flawed. "Adverse selection" as the term is commonly used is not really occurring here, because this is not, strictly speaking, an individual's decision based on his or her own characteristics. Rather, the relationship depicted in Figure 2 relies completely on between-community variation in mean educational attainment to predict individual migration probabilities.

(Figure 2 here)

Also modeled after Figure 1, Figure 3 replaces individual years of schooling with the difference in individual educational attainment, measured in years, from the community mean. By construction, this measure of relative schooling has a mean of zero. The results depicted in this Figure are very different from those in Figure 1. In both rural and urban areas, individuals around the local average level of schooling are those most likely to migrate to the US. To the extent that educational attainment and placement in the local earnings distribution are related, this mirrors the Chiquiar and Hanson result from Mexican census data.

Figure 4 is identical to Figure 3, except that we measure relative schooling as the proportional difference from community mean educational attainment. A value of 1 for this variable represents an individual with twice the local average level of education, someone with half the local mean has a value of -0.5, and by construction, those with no education have a value of -1.0. In comparison to the absolute deviation measure, the proportional deviation measure

[7] This is a familiar pattern in migration to the US. For example, in the period 1901-1914, over 70% of transoceanic Italian emigrants were workers from the agricultural South, though this region accounted for just 38% of the Italian population at the turn of the century (Del Boca and Venturini, 2003).

weights a given individual increment in years of schooling more heavily where average educational attainment is low, and less heavily where average years of schooling is high. The pattern of migration is similar, but the proportional formulation highlights the increasing fraction of highly educated individuals who migrate to the US with increases in relative education.

(Figures 3 and 4 here)

In Figures 3 and 4, probabilities of migration are strikingly low for those who, locally, have little schooling. Perhaps because increasing income makes the direct costs of migration more affordable, migration probabilities continue to rise until the local average level of schooling is attained. Rather than adverse selection, within local areas, this parameterization shows clear evidence of positive selection on educational attainment at levels of schooling up to and including the local mean. At the highest levels of schooling, additional education may increase (Figure 4) or decrease (Figure 3) migration probabilities. In both Figures, community-level mean educational attainment has been differenced out, and so both graphs rely solely on within-community variation to predict migration probabilities. In this sense, one interpretation of the adverse selection evident in Figure 1 is that it is an artifact of a high level of aggregation, resulting from the geographic variation in local average educational attainment swamping a fundamentally neutral (Figure 3) or perhaps even a positive (Figure 4) individual selection mechanism.

While on net, either adverse selection or the process detailed in the last three Figures can lead in the aggregate to migrants who have lower than average levels of education, the policy implications differ markedly. Education, training, and experience supplement ability in the formation of human capital. If selection truly is adverse, the human capital package of migrants is uniformly unappealing. If, on the other hand, selection merely appears to be adverse as an artifact of aggregation, the issue is more complex. Migrants come disproportionally from disadvantaged geographic regions and so have low average levels of schooling compared to the entire population. However, given local educational funding and other constraints, migrants have roughly the same educational attainment as nonmigrants. At least since Becker (1975) and arguably since Roy (1951), there has been the expectation that ability and schooling choices are positively correlated. It is at the local level that one sees these choices being made. Absent a reason to believe otherwise for Mexico, the last three figures therefore provide little support for the more sweeping claim of adverse selection on ability by Mexican migrants to the United States.

In the remainder of the paper, we construct and estimate a model to allow us to examine the relative importance of variation between and within geographic areas in more detail.

A model of Mexican-US migration

There are three characteristics that drive our model, none controversial:

1. Migrants choose to migrate where the wage differential, net of migration costs, is positive.
2. Human capital accretions increase wages in both sending and receiving labor markets, though at varying rates.
3. Human capital levels depend on an initial endowment and on subsequent accumulation, and such accumulation is costly.

We assume that the wage received by an individual in Mexico, w_0, is a function of the base Mexican wage μ_0, human capital accumulation h, and the return to human capital in

Mexico, $\delta_0(h)$, where $\delta_0'>0$.[8] Human capital accumulation is a function of ability a and skill accumulation s, with skills contributing to human capital accumulation at constant rate π_0 in Mexico. Skills are less costly for those with high ability or for those living in areas with lower cost of education (τ) to obtain, implying $\frac{\partial s}{\partial a} > 0$ and $\frac{\partial s}{\partial \tau} < 0$. Thus:

(1) $w_0 = \mu_0 + \delta_0(h)$

(2) $h = a + \pi_0 s$

(3) $s = s(\tau, a)$

The within-Mexico skill premium per unit of skill accumulated (in Mexico) is $\delta'_0\pi_0$. The wage impact of a unit change in education cost, τ, for a Mexican worker in Mexico is

(4) $\frac{\partial w_0}{\partial \tau} = \delta_0^{'}\pi_0 \frac{\partial s}{\partial \tau}$

and comparably, for a unit change in ability, the Mexican wage impact is

(5) $\frac{\partial w_0}{\partial a} = \delta_0^{'}(1 + \pi_0 \frac{\partial s}{\partial a})$.

Let w_1 and similarly subscripted variables represent equivalent quantities in the United States. After Roy (1951), Sjaastad (1962), Borjas (1999) and Chiquiar and Hanson (2005), the decision to migrate is straightforward, based on a comparison of migration costs C with expected wage benefits from migration. The individual chooses to migrate if

[8] For the sake of simplicity, we suppress an individual-specific subscript throughout, and because we focus on those in the left tail of the education distribution, ignore the likely negative second derivative of $\delta_0(h)$.

$$(6) \quad w_1 - w_0 > C \quad .$$

The distinguishing feature of our model is the explicit consideration of access to education, τ. Consider the impact of a change in Mexican schooling costs or ability, and subsequent within-Mexico generation of human capital, on subsequent US earnings for migrants:

$$(7) \quad \frac{\partial w_1}{\partial \tau} = \delta_1^{'} \pi_0 \frac{\partial s}{\partial \tau} \qquad \text{and}$$

$$(8) \quad \frac{\partial w_1}{\partial \tau} = \delta_1^{'} (1 + \pi_0 \frac{\partial s}{\partial \tau}) \quad .$$

Here, the US wages of migrants follow the same general pattern as they would have in Mexico, increasing in ability and decreasing in education costs, but wage levels differ according to differences in returns to human capital ($\delta_j(h)$) in each setting. Defining the net benefit to migration to be $B = w_1 - w_0 - C$ and assuming constant migration costs C, we have the following expression for variation in net benefits to migration:

$$(9) \quad \partial B = (\delta_1^{'} - \delta_0^{'}) \left[(1 + \pi_0) \frac{\partial s}{\partial a} \partial a + \pi_0 \frac{\partial s}{\partial \tau} \partial \tau \right].$$

B is monotonically related to the probability of migration, and the term ($\delta_1' - \delta_0'$) is the difference in marginal returns to human capital in the US and Mexico. The first term in brackets shows the impact of ability (∂a) and aptitude ($\frac{\partial s}{\partial a}$) differences, and the second term the impact of variation in schooling costs ($\partial \tau$) and a term representing the transformation of costs into schooling outcomes, ($\frac{\partial s}{\partial \tau}$). Where the difference $(\delta_1^{'} - \delta_0^{'})$ is positive, (9) shows that the probability that an individual migrates from Mexico to the US increases with ability but, because $\frac{\partial s}{\partial \tau}$ is negative, the probability of migration decreases with increases in the cost of education.

The stylized adverse selection result is that low-skill Mexicans are more likely to come to the United States than are high-skill Mexicans. This can occur in the Roy model where the migration wage premium $(\delta_1^{'} - \delta_0^{'})$ is positive for those in the tail of the more diffuse distribution, but becomes negative as skills increase. The engine in the Roy model, as previously discussed, is the wage differential, and the wage differential is likely to have more direct drivers than the variance in earnings. In particular, average wage differentials between the US and Mexico obviously have not been competed away for large proportions of the potential population of migrants. It is unlikely that $(\delta_1^{'} - \delta_0^{'})$ could be negative for much of the Mexican population, and especially for those below mean levels of income. Flows of low-ability migrants may in some sense be "high", but the human capital model unambiguously predicts that these flows will be even higher for those with more human capital, assuming increasing rewards for human capital at the destination.

It is still possible for (9) to yield negative values of ∂B where $(\delta_1^{'} - \delta_0^{'})$ is positive. This is when $(1+\pi_0)\frac{\partial s}{\partial a}\partial a < \left|\pi_0 \frac{\partial s}{\partial \tau}\partial \tau\right|$, that is, when the adverse effects of increasing costs of skill accumulation τ or inefficient skill production $\frac{\partial S}{\partial \tau}$ exceed the migration-increasing effects of ability ∂a or apititude $\frac{\partial S}{\partial a}$. Within a geographic area, binding supply constraints in education render τ fixed. In this case, ability and aptitude variation alone drives the skill-accumulation function $s(a,\tau)$, and the model predicts positive selection of migrants on ability.[9] However, between geographic areas, where τ varies, variations in schooling costs may be sufficient to reverse the sign of (9) even though migrants select positively on ability. While superficially consistent with adverse selection, in that a flow of poorly schooled immigrants would result, the migrants would be of high ability. A graph like Figure 1 would be the expected result. In small geographic areas, constant costs are a credible assumption, and graphs like Figures 3 and 4 are interpreted, within the model, as the impact of ability variations on migration.

Our model also allows for intermediate findings like those of Chiquiar and Hanson (2005), where the combined ability and accumulation effects yield a mix of migrants—those with some human capital, either through native ability or training, are more likely to come to the US than are those with very little or those with a great deal of human capital. A possible causal mechanism at the low end of the scale is the impediment created by migration expense. The "cash in advance" nature of migration cost (Orrenius and Zavodny 2005) is likely to have the largest deterrent effect on migrants from poor areas, and the steep upward slope of the migration probability function in Figures 3 and 4, especially for relatively poor rural areas, reinforces the case for these arguments. Costly migration has the potential to increase the force of selection on ability from poor areas, as only high-ability migrants may be able to generate positive net

[9] In this respect, the model is somewhat like that of Card's (2001), where individual heterogeneity in the marginal cost of education plays a key role in determining educational attainment.

migration returns. This also suggests the importance of factors influencing the cost of migration, such as access to migration networks, as control variables in estimating the model.

An empirical model using Mexican data

To identify our model, we need to differentiate between baseline ability (a) and skill accumulation (s). Fortunately for our purposes, access to education, and so the cost of skill accumulation, varies widely within Mexico. Education through at least high school is easily available in urban areas, but educational access is severely constrained in rural areas. We treat these variations between geographic areas as exogenously determined with regard to international migration, and so interpret geographic differentials in educational attainment as the outcomes of natural experiments, presumably based on geographic variation in social spending generally and in education funding in particular. We assume that public resources allocated to education are constant within sample clusters, which constitute very small geographic areas, so that each individual within a cluster is subject to the same education funding constraint. We then are able to interpret differences in educational attainment within these small areas as reflective of ability, motivation and other innate differences in human capital, holding cost of educational attainment constant.

a. Data source and variable definitions

The Mexican Migration Project is an ongoing survey of households in a large number of communities in Mexico[10]. The survey inquired of one or more household members about other, potentially migrant, members and so partially alleviated the problem of interviewing only those

[10] See Massey *et al*. 1994 for a more complete description of these data.

who reside in Mexico at the time of the survey [11]. For all household members, including adult children not currently resident, general demographic information and brief migration measures were collected. Data included age, sex, relationship to head of household, marital status, schooling, current economic indicators, and characteristics of the first and last trips made to the US or to other Mexican locations. Each year, the survey was administered during the winter months surrounding the Christmas holidays, in an effort to capture those migrants who live most of the year in the U.S. and returned to Mexico for the holidays. A different set of sample clusters was surveyed each year. Data collection began in 1982, and we use information collected between 1982 and 2009. The unit of analysis is the individual, and each surveyed household can contribute multiple observations. Those individuals who reportedly had their first US migration experience prior to age 16 and any children in the household at the time of the survey under age 16 were dropped as unique observations, as they were not likely to have engaged in an independent choice to migrate. These children were, however, included in calculations of household size. Also excluded were those older than 65 years and, because they are so different from most of the sample, a small number of individuals reporting more than 20 years of education. Finally, because we were trying to obtain estimates of impact of schooling based on childhood place of residence but observe only residence at the survey date, we excluded those individuals who reportedly resided in a state other than the one in which they were born. [12] The

[11] Some endogenous selection of the sample may still occur, to the extent that entire households may no longer have resided in Mexico at the time of the survey.

[12] Eliminating 3420 such cases surveyed in a rural area reduces average schooling by 0.6% in rural areas, and eliminating 6568 currently urban cases reduces average schooling by 0.1 % in urban areas, so there is little evidence that previous migration across state lines is systematically related to schooling.

resulting dataset has information from 128 communities for 90,839 individuals, with about 60% of the sample from rural areas.

Summary statistics are presented in Table 1. The variables largely are self-explanatory, with the possible exceptions of the relative education and wealth variables. The wealth variable is a factor score based on a set of thirteen selected indicators of wealth or financial wellbeing[13]. We take the factor score approach because the thirteen indicators of wealth we employ are likely to be highly correlated with one another. Using factor scores allows us to preserve as much information from these variables as we can, and so, we hope, credibly to control for wealth variation in the regressions we run.

(Table 1 here)

The key variables in our analysis are relative educational attainment, measured either as the ratio of individual years to the community mean of years of schooling attained or the simple difference in years of these two values, and raw years of schooling. In poor communities, the mean number of years of school attended was low—the value for the 25th percentile of educational attainment over the full sample was 4 years—so that even those with fairly low levels of schooling could be construed to be relatively well educated. Cross-sectional variation in educational attainment over the entire sample represents the impact of a combination of supply side constraints, shared by all in a locality, and individual variations in aptitude, motivation, and so forth. Our goal in employing either relative education measure is to remove variation between localities in access to schooling, and therefore to focus on other components of skill in the remaining within-locality variation.

---

[13] The variables are indicators for ownership of land; a stove, refrigerator, washing machine, sewing machine, radio, television, stereo, telephone, or motor vehicle, and whether the respondent's dwelling has running water, electricity, and flush toilets. The factor score represents a weighted averaged of these indicator variables, with mean zero. See Montgomery et al. (2000) for more detail.

We have two concerns regarding the use of relative education as a proxy for ability. First, relative education is calculated based on the community of residence for the respondent at the time of the survey, but its theoretical justification rests on constraints operating during childhood and adolescence. Domestic migration is fairly common in Mexico, and this may lead to some measurement error in relative education. We have eliminated those surveyed in a state other than their state of birth, but there still are likely to be within-state migrants in the sample. Jensen and Ahlburg (2004) show that migration between areas of comparable urbanization accounts for almost half of internal moves in the Philippines. If this sort of migration also occurs in Mexico, and, because the moves are within the same state, these observations remain in our sample, we would expect the standard errors-in-variables bias toward zero in our estimated effects. Where this error occurs more systematically because of past rural-urban migration, the relative education we calculate is based on urban means. Because this results in an understatement of relative education for this group of internal migrants, we expect that this also will bias our results in the direction of conservatism.

The second concern is more speculative. Our analysis rests on the claim that relative education reflects underlying ability or motivation more accurately than does overall years of education attained, because it measures educational attainment in the context of local supply constraints. The validity of this claim depends on the degree to which relative educational attainment reflects merit over local privilege or some other allocative mechanism. If relative educational attainment is merely a noisy proxy for ability, we again appeal to the the standard measurement error result to claim that we once again are likely to underestimate the true effect of ability through the use of relative education. The problem may be more severe than simple noise. For example, a more systematic effect may occur if the locally privileged have both easier access to education and reduced likelihood of leaving a comfortable situation. In this case, we are probably again underestimating the true impact of relative education, biasing our result by including those who have high levels of relative education but lower expected migration gains due to a better original condition, and so low probabilities of migrating. While in either case, the estimated effect of relative education is likely to be conservatively estimated, we recognize that relative education remains an imperfect measure of ability or motivation.

Schooling has served as the dominant measure of quality in the literature, and so we include total years of education attained as a covariate. We include several demographic measures, including age, marital status, and sex. We use family size and our factor score for wealth as indicators of per-capita household resource availability. Regarding migration costs, three proxies for the availability of migration networks and opportunity costs appear as dummy variables: whether the individual had migrated domestically, and whether they had immediate family or extended family in the United States.

b. Models

We estimate several models regressed on a common set of covariates. In every case, the dependent variable is a binary variable indicating ever-migrating to the US. Table 2 shows marginal effects evaluated at sample means based on estimated probit coefficients. We refer to the difference between individual and local mean schooling, expressed as a proportion of local mean schooling, as "relative education." We include both relative education and schooling attainment in years in the model. Squares of both variables are incorporated to capture the parabolic shapes evident in the preceding Figures. The first column of results is based on the entire estimation sample. The results from estimating the model separately for those living in urban and rural areas and for males and females constitute the remainder of the table.

In every sample, the relatively better educated are more likely to have migrated to the US at least once, but the absolutely better educated are less likely to have done so. For males the effect is particularly striking. Based on our estimated marginal effects, being one standard deviation (0.57) above mean relative education increases the probability of migrating by 0.14 points. In contrast, our estimated marginal effect for years of schooling implies that being one standard deviation (4.23 years) above mean absolute education decreases migration probability by 0.09. Both are sizable effects compared to the mean probability of ever migrating for males of 0.30. Age is not an important determinant of US migration, but those from wealthier households and with access to migration networks display higher probability of ever migrating. Males are much more likely to migrate than females, and, perhaps because a nontrivial share of female migrants are trailing spouses (Chattopadhyay 1997), the marginal effects for females are uniformly smaller in magnitude. We therefore focus our discussion subsequently on male migrants.

(Table 2 here)

It is instructive to compare the magnitudes of the education effects in two stylized settings. In the first, comparable to rural southern Mexico, assume that adults average three years (the actual 25th percentile for rural areas) of education completed. In the second, as might be the case in a large urban area, adults average eleven years (the actual 75th percentile value for urban areas) of completed education. Consider someone with the rural median of six completed years of education. If they lived in the stylized rural area, they would have relative education of 2.0; if in the urban area, relative education of 0.55. A simple calculation of the net effects for males of our two measures of education shows that an otherwise average rural male with six years of education has a migration probability 0.19 greater than an otherwise comparable rural male with the rural median level of education. An urban male with the same six years of education would have a predicted probability of migrating to the US that is 0.02 higher than the median urban male. The evidence from Table 2 regarding adverse selection therefore is slightly mixed. Rural residents appear to select positively by schooling into migration, but the urban result is consistent with a small degree of adverse selection. Table 2 shows that the estimated marginal effect of relative education based only on urban residents was almost identical to that for rural residents, so it seems likely that what appears to be adverse selection from urban areas actually reflects education supply differentials between urban areas. In any case, since the estimated effect in urban areas is very small and nearly three-fourths of Mexican migrants come from rural areas, selection of male migrants to the US on ability appears to be fairly strongly positive, in this stylized case.

Table 3 presents as a baseline the last column of Table 2, and goes on to present three alternative parameterizations. We present only results for males in this table, because, for reasons outlined above, we expect these decisions to be more the result of individual choice than for females. Each main column of results is paired. The first shows marginal effects for each element of the underlying regression, and the second (in bold) shows only the relevant sample-average marginal effect for variables that include squared terms. In the second main column of results, rather than as a proportion, relative education is expressed as the simple difference in years between an individual observation's schooling and the community mean. The advantage of this formulation is that the marginal effects of relative schooling and schooling can be netted out simply by adding them together. The theoretical disadvantage of this linearity is that an additional year of schooling where average schooling levels are low is presumed to have the same effect on the margin as an additional year where average schooling levels are high. The empirical disadvantage is the probit index function is linear as well, and so the only unique information the relative education term imparts is the community mean. On net, the marginal effect of an additional year of education almost exactly offsets the marginal effect of an additional year of relative education. This formulation of the model therefore shows no evidence of adverse selection on schooling.

(Table 3 here)

Excluding community-level differences in average schooling, the third column of results seems to show a small degree of adverse selection on schooling. Migration rates initially increase with schooling, but after only about three years of schooling, migration rates begin to decrease at an increasing rate. This is the expected result of ignoring community-level effects, and echoes Figure 1. As we have previously argued, this reflects geographic supply variations that are not captured in a model that includes only years of schooling. Finally, the last column of Table 3 presents marginal effects from a conditional fixed effects logit, evaluated at zero for each of the community-level fixed effects. The coefficient of relative education is identified in a fixed effects model only when this variable is in its proportional form, which is an arbitrary restriction, and the coefficient of urban is not identified in the fixed effects model. Therefore, these variables are excluded. The estimated marginal effect of education again initially is positive. The inflection point is at a somewhat higher level of education compared to the results of the previous column, but still below the mean level of education. The averaged marginal effect of schooling is slightly negative overall. This suggests that the simple linear differencing of the fixed effects model helps, but is not sufficient to fully capture nonlinear impacts (per Figure 3 or Figure 4) of local educational choices on migration.

Discussion

Our empirical findings are in some ways consistent with previous work indicating negative selection on education. As other have, we find that immigrants have low absolute levels of education. Borjas finds that migrants are more likely to come from countries with higher degrees of income inequality, and we indirectly extend this to the subnational level, with the finding that those from Mexican communities with lower average levels of education are much more likely to migrate to the US than are individuals from areas where education levels are higher. We find that individuals with low levels of education are likely to migrate from Mexico to the US, and that, except at very low levels of educational attainment, the probability that they migrate to the US decreases as levels of education increase. Our differences with models reporting negative self-selection of migrants to the US emerge as we look more closely at measures of immigrant quality.

Most notably, we find that Mexicans with levels of education at local mean levels are most likely to migrate to the US, and much more likely than those with little education (by local standards) to do so. We attribute the apparent adverse selection by schooling to constraints in the availability of education. Because most Mexicans coming to the US are from rural areas, where average levels of schooling are low, the average schooling levels of Mexican immigrants is also low. We argue that this is misleading evidence regarding overall migrant quality, and that, because ability and schooling choices are closely tied at the local level, examining educational choices in comparison to peers offered the same choice set is a better measure of migrant quality. We estimate a model that allows for both within- and between-community variation. Figure 5 presents predictions from this model, using the estimates presented in the first column of Table 1, and contrasts these predictions to the comparable (Figure 1) plot of educational attainment in years and migration probabilities. As Figure 5 illustrates, our model generates predictions that are consistent with the data. The underlying model is one of positive selection at the local level on schooling, so the apparent adverse selection pattern evident in our Figure 1 seems to be due to uncontrolled geographic variation in access to education.

(Figure 5 here)

To the extent that relative education captures traits like ability or motivation, a case can be made that the more motivated or innately able are migrating to the US, and so that Mexican migrants are selecting positively on ability, motivation, or other hard-to-measure attributes. To the further extent that this ability is heritable, the findings of Schoeni (1997), Card (2005), and Smith (2006) on the labor market advantages enjoyed by migrants' children are consistent with the contention that the migrants themselves are of high innate ability.

It is likely that Mexico represents the rule rather than an exception amongst developing countries in this regard. If so, educational supply constraints in sending countries may have the effect of masking relatively high underlying ability of immigrants. Viewed in this light, the Borjas (1990) result that those from sending countries where income is more unequally distributed are more likely to come to the US may not really be evidence of adverse selection. The more unequally income is distributed in countries of origin, the more that some immigrants of high ability, faced with constraints in obtaining education at home, have to gain by immigrating to the US. Models of adverse immigrant selection often are used to justify reducing the number of immigrants of particular origin. Our results support a more nuanced policy, recognizing the historically familiar immigration stream of relatively able but poorly trained migrants to the US.

References

Becker, Gary S. 1975. **Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education, 2nd ed**. Cambridge, MA: National Bureau of Economic Research.

Borjas, George J. 1990. *Friends or Strangers: The Impact of Immigrants on the US Economy.* New York: Basic Books.

Borjas, George J. 1991. *Ethnic Capital and Intergenerational Mobility.* **Quarterly Journal of Economics** 107 (1): 123-150.

Borjas, George J. 1994. *The Economics of Immigration.* **Journal of Economic Literature** 32 (4): 1667-1717.

Borjas, George J. 1999. **Heaven's Door: Immigration Policy and the American Economy**. Princeton: Princeton University Press.

Card, David. 2001. *Estimating the Returns to Schooling: Progress on some Persistent Economic Problems*. **Econometrica** 69 (5): 1127-1160.

Card, David. 2005. *Is the New Immigration Really So Bad?* **The Economic Journal** 115 (507): F300-F323

Chattopadhyay, Arpita. 1997. *Family migration and the economic status of women in Malaysia.* **International Migration Review** 31(2): 338-352.

Chiquiar, David and Gordon Hanson. 2005. *International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States.* **Journal of Political Economy** 113: 239–281.

Chiswick, Barry R. 1999. *Are Immigrants Favorably Self-Selected? An Economic Analysis.* **American Economic Review** 89 (2):181-185.

Del Boca, Daniella and Alessandra Venturini. 2003. *Italian Migration.* Institute for the Study of Labor Discussion Paper No. 938. Bonn: IZA.

Duleep, Harriet Orcutt and Mark C. Regets. 1999. *Immigrants and Human Capital Investment*. **American Economic Review** 89 (2): 186-191.

Grogger, Jeffrey and Gordon H. Hanson. 2008. *Income Maximization and the Selection and Sorting of International Migrants*. Processed, University of California-San Diego.

Ibarraran, Pablo and Darren Lubotsky. 2007. *Mexican Immigrants and Self Selection: New Evidence from the 2000 Mexican Census*. In Jorge Borjas, ed., **Mexican Immigration to the United States**. National Bureau of Economic Research Conference Report. Chicago: University of Chicago Press.

Jensen, Eric R. and Dennis A. Ahlburg. 2004. *Why Does Migration Decrease Fertility? Evidence from the Philippines*. **Population Studies** 58 (2): 219-231.

Massey, Douglas S., Joaquin Arango, Ali Koucouci, Adela Pelligrino, and J. Edward Taylor. 1994. *An Evaluation of International Migration Theory: The North American Case.* **Population and Development Review** 20: 699-752.

Mckenzie, David and Hillel Rapoport. 2007. *Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico.* **Journal of Development Economics** 84: 1-24

Montgomery, Mark R., Michele Gragnolati, Kathleen A. Burke and Edmundo Paredes. 2000. *Measuring Living Standards with Proxy Variables.* **Demography** 37 (2): 155-174

OECD. 2010. OECD Stat Extracts. Online retrieval from http://stats.oecd.org.

Orrenius, Pia M. and Madeline Zavodny. 2005. Self-selection among undocumented immigrants from Mexico. **Journal of Development Economics** 78: 215-240.

Roy, Andrew D. 1951. “Some Thoughts on the Distribution of Earnings,” **Oxford Economic Papers** 3: 135-146.

Schoeni, Robert F. 1997. *New Evidence on the Economic Progress of Foreign-Born Men in the 1970s and 1980s*. **The Journal of Human Resources** 32 (4): 683-740.

Sjaastad, Larry A. 1962. *The costs and returns of human migration.* **Journal of Political Economy** 70 (Supplement): 80-93.

Smith James P. 2006. *Immigrants and the Labor Market.* **Journal of Labor Economics** 24: 203-233

Stark, Oded and J. Edward Taylor. *Migration Incentives, Migration Types: The Role of Relative Deprivation*. **Economic Journal** 101 (408): 1163-1178.

Wong, R., Resano-Perez, E. and Martinon, L. 2006. *Una constante cambiante: la migracion de la poblacion Mexicana hacia Estados Unidos de Norteamerica*. Universidad Nacional del Estado de Mexico. Mexico.

World Bank. 2010. World Development Indicators database. (revised 9 July 2010).

Table 1: Descriptive Statistics

Variable	Description	Full Sample	Rural	Urban
Migrant	Ever migrated to the United States	0.20 (0.40)	0.22 (0.42)	0.17 (0.37)
Relative Education	Individual years of schooling as a proportion of community mean schooling	0.00 (0.57)	0.00 (0.60)	0.00 (0.53)
Educational Attainment	Years of schooling	7.10 (4.23)	6.44 (4.03)	7.88 (4.33)
Urban	Indicator; unity when resident in census metropolitan category 1 or 2	0.46 (0.50)		
Age	Age in years	35.44 (15.06)	35.88 (15.41)	34.92 (14.61)
Wealth	Factor score for asset ownership	0.00 (0.85)	-0.19 (0.83)	0.23 (0.80)
Married	Indicator; unity for currently married at survey	0.69 (0.46)	0.69 (0.46)	0.68 (0.47)
Male	Indicator; unity for males	0.48 (0.50)	0.48 (0.50)	0.48 (0.50)
Household Size	Number of individuals residing in household	5.08 (2.70)	5.13 (2.76)	5.02 (2.62)
Domestic Migrant	Indicator; unity when individual has made one or more domestic migrations	0.18 (0.39)	0.20 (0.40)	0.16 (0.37)
Family Network	Indicator; unity when one or more immediate family members has migrated to the US	0.73 (0.44)	0.75 (0.43)	0.71 (0.45)
Extended Network	Indicator; unity when one or more extended family members has migrated to the US	0.89 (0.31)	0.91 (0.29)	0.87 (0.34)
Observations		90,839	49,149	41,690

Source: Authors' calculations from MMP data. Standard deviations in parentheses.

Table 2: Determinants of Migration as Probit Marginal Effects

Variable	Full sample	Rural	Urban	Male	Female
Relative Education (as proportion of local mean)	0.291	0.299	0.287	0.440	0.159
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Relative Education Squared	-0.079	-0.079	-0.091	-0.129	-0.037
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Educational Attainment	-0.054	-0.058	-0.052	-0.088	-0.026
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Educational Attainment Squared	0.001	0.001	0.001	0.001	0.000
	[0.03]	[0.12]	[0.08]	[0.00]	[0.41]
Urban	-0.009			-0.026	0.004
	[0.56]			[0.23]	[0.71]
Age	0.000	0.000	0.001	0.001	0.000
	[0.23]	[0.58]	[0.00]	[0.01]	[0.11]
Wealth	0.055	0.068	0.038	0.068	0.043
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Married	0.078	0.086	0.068	0.105	0.048
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Male	0.204	0.241	0.159		
	[0.00]	[0.00]	[0.00]		
Household Size	-0.004	-0.004	-0.003	0.000	-0.007
	[0.00]	[0.02]	[0.05]	[0.84]	[0.00]
Domestic Migrant	-0.002	-0.012	0.010	-0.001	-0.012
	[0.78]	[0.23]	[0.34]	[0.96]	[0.08]
Family Network	0.130	0.146	0.109	0.174	0.087
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
Extended Network	0.057	0.045	0.068	0.093	0.020
	[0.00]	[0.01]	[0.00]	[0.00]	[0.09]
Observations	90,839	49,149	41,690	43,942	46,897

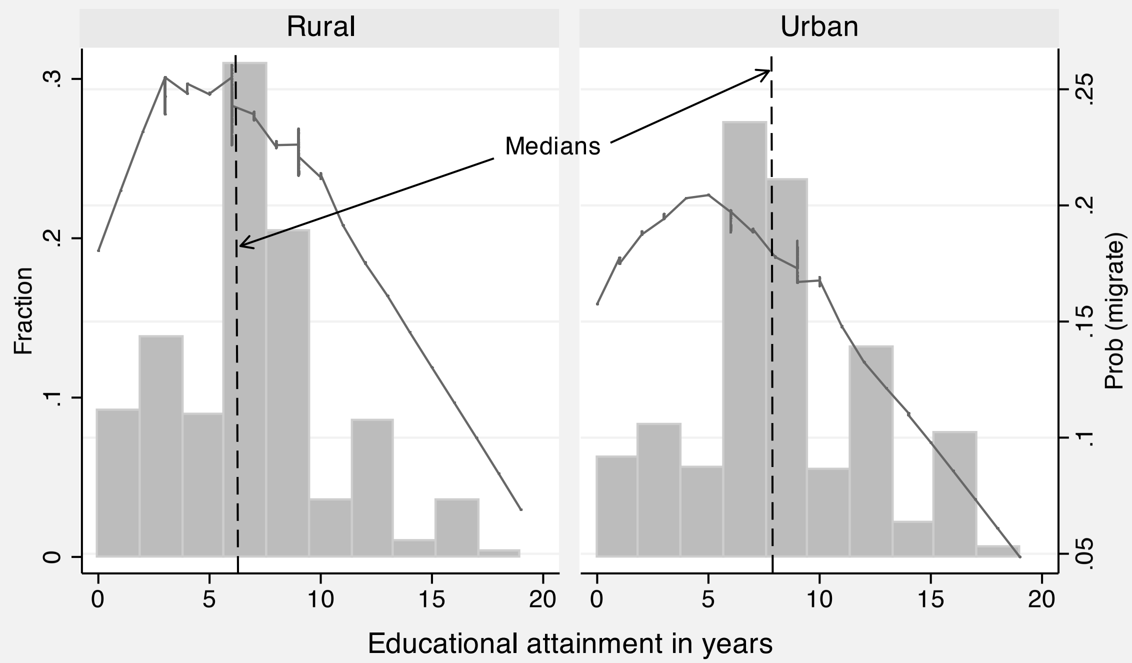
Source: Authors' calculations based on MMS data. Robust p values in brackets, based on community level clustering. Entries of 0.00 or 0.000 represent values that round to less than 0.01 or 0.001.

Table 3: Determinants of Migration as Marginal Effects, Males

<table>
<tr><th>Variable</th><th colspan="2">Relative Education as Proportion</th><th colspan="2">Relative Education in Years</th><th colspan="2">Schooling Years Only</th><th colspan="2">Fixed Effects</th></tr>
<tr><td>Relative Education</td><td>0.440
[0.00]</td><td rowspan="2">0.444
[0.00]</td><td>0.064
[0.00]</td><td rowspan="2">0.064
[0.00]</td><td colspan="2" rowspan="2"></td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Relative Education Squared</td><td>-0.129
[0.00]</td><td>-0.001
[0.02]</td></tr>
<tr><td>Educational Attainment</td><td>-0.088
[0.00]</td><td rowspan="2">-0.071
[0.00]</td><td>-0.050
[0.00]</td><td rowspan="2">-0.067
[0.00]</td><td>0.011
[0.01]</td><td rowspan="2">-0.012
[0.00]</td><td>0.022
[0.00]</td><td rowspan="2">-0.005
[0.00]</td></tr>
<tr><td>Educational Attainment Squared</td><td>0.001
[0.00]</td><td>-0.001
[0.00]</td><td>-0
[0.00]</td><td>-0.002
[0.00]</td></tr>
<tr><td>Urban</td><td colspan="2">-0.026
[0.23]</td><td colspan="2">-0.013
[0.56]</td><td colspan="2">-0.082
[0.00]</td><td colspan="2"></td></tr>
<tr><td>Age</td><td colspan="2">0.001
[0.01]</td><td colspan="2">0.001
[0.00]</td><td colspan="2">0.000
[0.73]</td><td colspan="2">0.001
[0.00]</td></tr>
<tr><td>Wealth</td><td colspan="2">0.068
[0.00]</td><td colspan="2">0.072
[0.00]</td><td colspan="2">0.054
[0.00]</td><td colspan="2">0.059
[0.00]</td></tr>
<tr><td>Married</td><td colspan="2">0.105
[0.00]</td><td colspan="2">0.099
[0.00]</td><td colspan="2">0.105
[0.00]</td><td colspan="2">0.091
[0.00]</td></tr>
<tr><td>Household Size</td><td colspan="2">0.000
[0.84]</td><td colspan="2">-0.001
[0.56]</td><td colspan="2">0.005
[0.01]</td><td colspan="2">-0.002
[0.00]</td></tr>
<tr><td>Domestic Migrant</td><td colspan="2">-0.001
[0.96]</td><td colspan="2">-0.002
[0.88]</td><td colspan="2">0.009
[0.43]</td><td colspan="2">-0.007
[0.10]</td></tr>
<tr><td>Family Network</td><td colspan="2">0.174
[0.00]</td><td colspan="2">0.171
[0.00]</td><td colspan="2">0.196
[0.00]</td><td colspan="2">0.107
[0.00]</td></tr>
<tr><td>Extended Network</td><td colspan="2">0.093
[0.00]</td><td colspan="2">0.093
[0.00]</td><td colspan="2">0.092
[0.00]</td><td colspan="2">0.046
[0.00]</td></tr>
<tr><td>Observations</td><td colspan="2">43,942</td><td colspan="2">43,942</td><td colspan="2">43,942</td><td colspan="2">43,942</td></tr>
</table>

Source: Authors' calculations based on MMS data. Robust p values in brackets. Probit marginal effects with variance estimates based on community level clustering in the first three columns. The last column reports marginal effects based on conditional logit estimates with community-level fixed effects, holding fixed effects at zero. Entries of 0.00 or 0.000 represent values that round to less than 0.01 or 0.001.

Figure 1: Probability of Migration to the US
by years of schooling



Note: From MMS data per text, for respondents aged 15 to 65 reporting 20 or fewer years of education
Lines are lowess-smoothed migration probabilities and histograms are distributions of the education measure

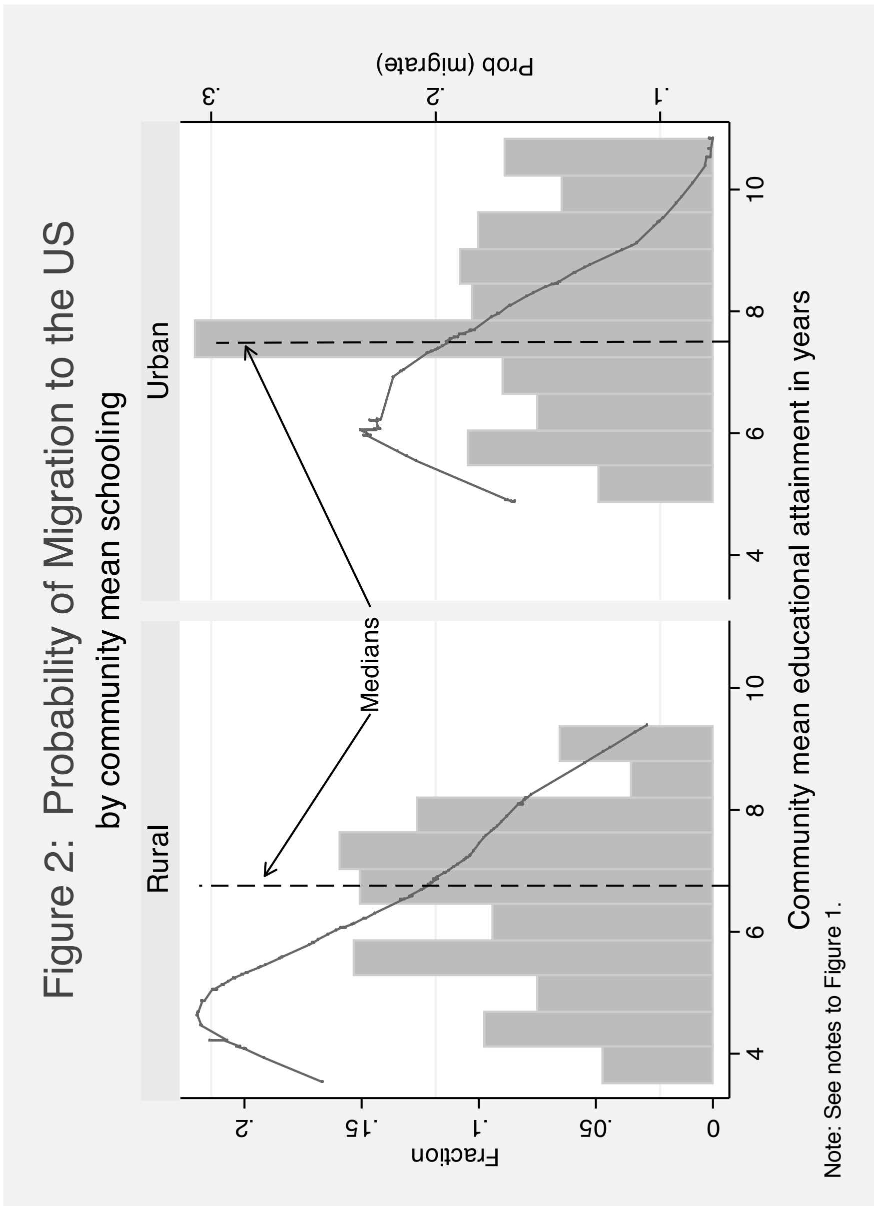


Figure 2: Probability of Migration to the US
by community mean schooling

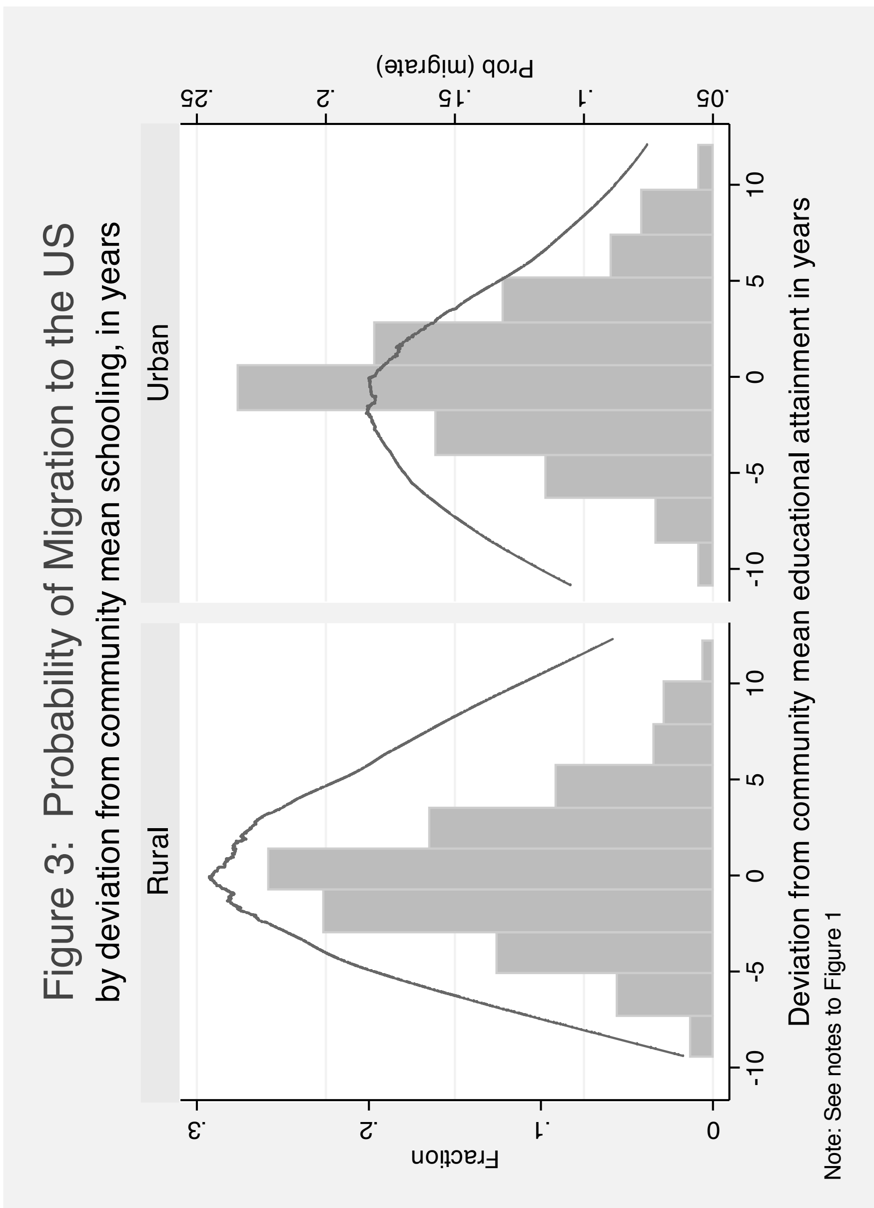
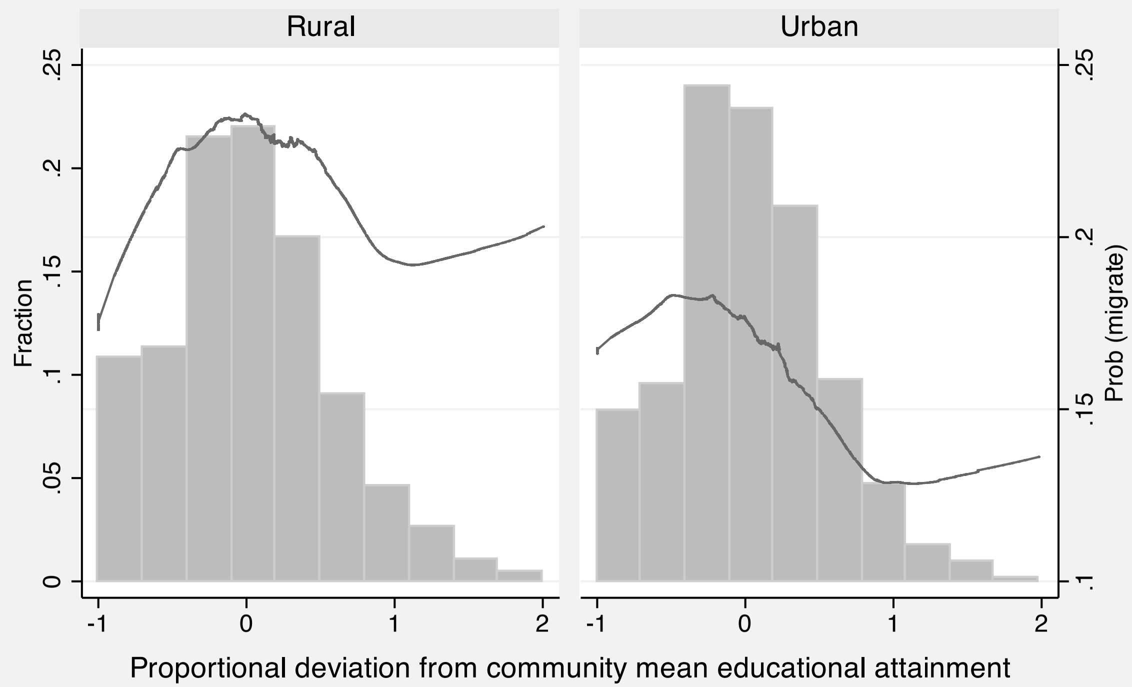


Figure 3: Probability of Migration to the US
by deviation from community mean schooling, in years

Figure 4: Probability of Migration to the US

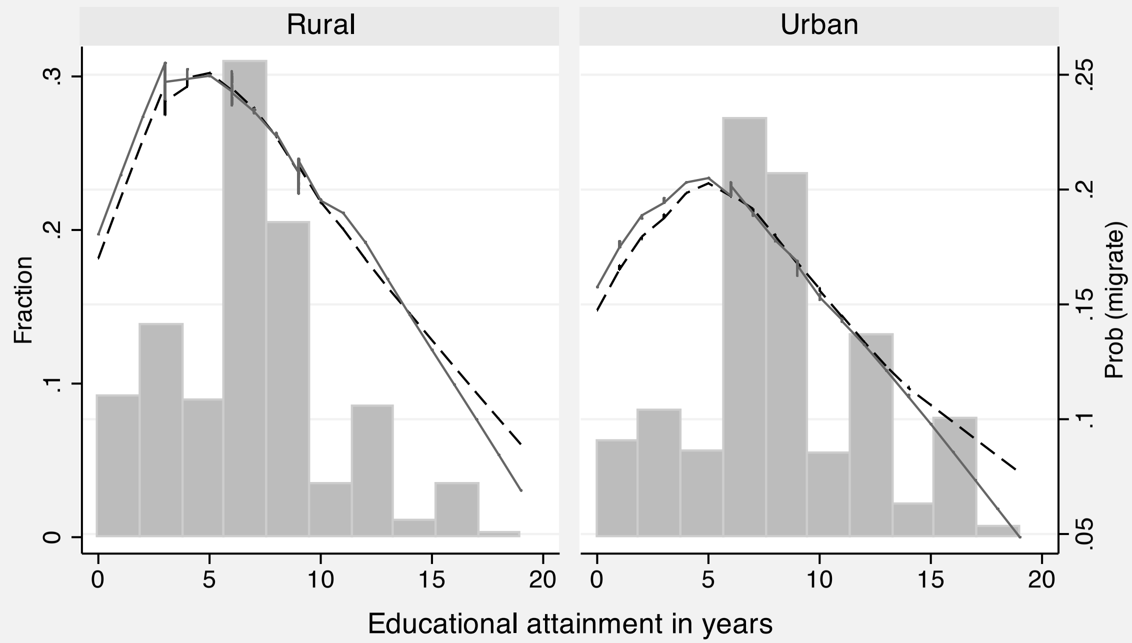
by proportional deviation from community mean schooling



Note: See notes to Figure 1.

Figure 5: Probability of Migration to the US

Predicted and actual, by years of schooling



Note: See notes to Figure 1. Solid lines are actual and dashed lines predictions based on probits