Do migrant networks affect education in source countries? Evidence from urban Mexico

Alfonso Miranda*
Centre for Economic Research, Keele University and IZA
(A.Miranda@econ.keele.ac.uk)

This version: March 2007

Abstract. This paper examines whether family and community migration experience affect the probability of high school graduation in urban Mexico. Bivariate random effects dynamic Probit models are estimated to control for the potential correlation of unobservables across migration and education decisions as well as within groups of individuals such as the family. Significant migrant network effects are detected. Having a migrant father (mother) increases the likelihood of US migration by 5 percentual points (9 p.p.). Similarly, a migrant mother (elder migrant sibling) increase the likelihood of high school graduation by 12 p.p. (6 p.p). Negative migrant selection is detected.

JEL classification: I21, F22, J61, C35.

Keywords: Education, Migration, Dynamic Bivariate Probit, Maximum Simulated Likelihood, Mexico.

*I am grateful to Gauthier Lanot, Wiji Arulampalam, Mark Stewart, Massimiliano Bratti, Leslie Rosenthal, Amanda Gosling, João Santos Silva, and seminar attendants at Keele and Kent for useful comments. Correspondence Address: Centre for Economic Research, Institute for Public Policy and Management, University of Keele, Keele, Staffordshire ST5 5BG, UK. E-mail: A.Miranda@keele.ac.uk

1 Introduction

In the last few decades international migration has become a topic of primary interest. The main destinations, Northern America and Europe, received nearly 13.1 million of new immigrants between 2000 and 2005. In contrast, Asia, Latin America and the Caribbean, the main origin areas, sent 10.1 million emigrants during the same five-year period (UNPD 2006). This intensive international flow of labour creates a number of economic, political, and social challenges that are attracting more and more the attention from policy makers and international organisations.

Traditionally, academic research in the field has focused in understanding the effects that immigration has on the labour market of the host country. In recent years, however, there is an increasing interest in learning whether international migration has impacts on poverty, accumulation of human and physical capital, economic growth, and development in source countries (see, for instance, World Bank 2006). The Mexico-US is a leading case of interest because in the last two decades the flow of labour from Mexico to the US has reached unprecedented numbers and the amount of remittances sent by migrants to their families in Mexico has increased steadily.[1]In fact, Banco de México estimates that in 2005 remittances from the USA represented nearly 2.6% of the GDP of Mexico.

The present paper intends to contribute a study on these issues. In particular, attention is focused on learning whether Mexico-US migration networks

[1]Mexico is by far the main origin country in Latin America. In fact, during the period 2000–2005 alone, Mexico sent nearly 2 million of emigrants to the United States (UNPD 2006).

affect the likelihood of high school graduation in Mexico.

Econometric work is essentially complicated by the fact that individual unobserved heterogeneity affecting migration choice is potentially correlated with unobserved traits affecting educational decisions. Unobserved skills are a good example. On the one hand, the labour economics literature stresses the fact that skilled individuals are more likely to succeed at school and to find qualified jobs (see, for instance, Miranda and Bratti 2006, Blundell et al. 2000). On the other hand, the migration literature points out that returns to education are higher in Mexico than in the US and that unqualified jobs are better paid in the American side. As a consequence, Mexican unskilled workers have strong incentives to emigrate to the US (see, for instance, Borjas 1994). A negative correlation among unobservables is therefore expected because skilled individuals are likely to study more and emigrate less. Clearly, failing to account explicitly for such a correlation may be a cause of serious bias.

Besides correlation across education and migration choices at the individual level, unobservables can be correlated within certain groups of individuals. The family is an obvious unit for this type of clustering because siblings within a family share a set of unobservable traits (say, for instance, genetic make-up or common adverse shocks) that affect their performance at school and change their likelihood of migration. Failing to account for this "intra-family clustering" can lead, once again, to serious bias.

Controlling for intra-family clustering is also important because an individual's education and migration decisions can be a function of the choices taken by elder siblings. For instance, individuals who are successful at school can create peer pressure and learning resources for his/her younger siblings and influence their school performance. Similarly, a migrant individual can help his/her younger siblings to migrate (i.e., to access migrant network resources) and/or to provide them with a successful role model of migration (i.e., to access information and 'reputation' spillovers). Finally, important dynamic cross-effects may be present because if an individual migrates younger siblings left behind can benefit from the money she/he sends home and from the contacts he/she builds up at the destination country. However, unless common sources of unobserved variation are set apart, the researcher will find impossible to distinguish between real and spurious dynamic sibling dependence — the latter being dynamic sibling dependence induced by unaccounted unobserved heterogeneity (see Arulampalam and Bhalotra 2006, Heckman 1981a).

The present paper addresses all these econometric challenges by estimating bivariate random effects dynamic probit models.

To date the literature has not fully recognised the complexity of the relationship between migration and education. There are two main strands of study. One strand is related to the analysis of social networks and its influence on migration decisions (see, for instance, Delechat 2001, Winters et al. 2001). These studies commonly use univariate dynamic probit models to disentangle the effects of migrant networks and previous migration experience on current migration events. Unfortunately, correlation of unobservables across migration and educational outcomes is not allowed. The second strand is concerned directly with the effects of migration networks on educational attaintment in origin countries and has produced relatively fewer pieces of work. In this category are McKenzie and Rapoport (2006), and Hanson and Woodruff (2003). These authors use instrumental variable techniques to control for the correlation of unobservables across migration and schooling variables. However, none of them allow for intra-family clustering. To the knowledge of the author, no previous study has addressed both potential problems simultaneously.

The study uses data from the Mexican Migration Project (MMP). The MMP is a rich individual-level data set that contains detailed information on migrant networks and collects information about the head of the household and all her/his sons and daughters independently of the current location of the latter individuals — therefore, long-term emigrants are well covered. Further, legal and illegal border crossings are carefully recorded.

Results suggest that family and community migration networks have a significant effect on the likelihood of emigration. Similarly, ceteris paribus, a migrant mother and an elder migrant sibling increase the likelihood of high school graduation by 12 percentual points (p.p. hereafter) and 6 p.p. respectively. Negative migrant selection is detected.

2 Do migrant networks affect education? Why?

When migrants leave their home country family is commonly left behind. Once established at the destination, migrants keep close contact with their communities back home and, in many cases, send money (remittances) and help members of their kin to migrate themselves.

The money migrants send home is used in a number of ways, including helping credit-constrained individuals to achieve their desired level of education. This option is particularly attractive to those who have no plans to emigrate themselves and education offers them an opportunity to improve their standard of life at the home country. As a consequence, through its links with remittances, migrant networks are expected to increase education at the source county (see, for instance, World Bank 2006).[2]

The story, however, does not end there. A group of recipient individuals plan to leave the home country. Those individuals will use remittances to finance their education at the source if observable qualifications are broadly portable across host and source countries (for more on this argument, see Vidal 1998).[3]In contrast, if observable qualifications are non-portable, rational prospective migrants will behave in a forward looking fashion and drop out school early to avoid wastage of valuable resources (a similar argument is put forward by McKenzie and Rapoport 2006). Finally, if qualifications are 'noisily' portable, then a zero effect of migrant networks on education at the source country may be observed.

Even if migrants do not send money home they can still effect education decisions at the source country. Namely, through their networks, current successful migrants can help prospective migrants to reduce labour market uncertainties at the destination country and to increase the returns to education acquired at the source before departure. The reduction in such uncertainties

[2]Obviously, a zero effect can be observed either because households are not credit-constrained in the first place, or because the contribution of remittances do not change significantly the overall financial position of recipient households.

[3]Under such an assumption acquiring education at the origin country is an efficient way to improve the odds of a highly paid job at the destination country. This route will be attractive specially if prospective migrants have no access to education at the destination country.

will, in turn, make education at the source more attractive to individuals who plan to emigrate and have access to migrant network resources.

Clearly, the effect of migrant networks on education in the source country is a function of its effects in the two aforementioned subpopulations — i.e., a function of its effects in the eventually-migrant and the eventually-stayer subgroups. In this context, even if one is willing to assume that qualifications are non-portable across host and source countries, it is not possible to sign the direction of the effect based purely on theoretical grounds. Empirical investigation is therefore needed.

3 Data

Data from the Mexican Migration Project (MMP) are used. The MMP is a pooled cross section of migrant communities located throughout Mexico, which is collected by a joint group of researchers at Princeton University and University of Guadalajara.[4] Every year, from 1982 to 2005, members of the MMP team survey a random sample of 200 households in two to five communities in Mexico to gather a new cross-section. Such cross-section is then added to the pool. Current files, the MMP107 database, contain information at individual and community level in 107 localities.

The communities surveyed by the MMP are not selected at random. As a consequence, the data may not be argued to be National or State representative. Instead, the MMP107 is representative of the population in the 107 communities that are included in the study. Very importantly, selected

[4] Data files are freely available at http://mmp.opr.princeton.edu/

communities are chosen on the basis that they have some migrant tradition. Across the years, the MMP team has managed to survey communities in many regions of the country and with different sizes, from small rural towns to large cities.[5] Moreover, there has been some effort to select communities so that there is enough variation in terms of economic activity — from small places that specialise in mining, fishing, and farming, to large urban areas that are highly diversified.

National representative surveys commonly contain too few observations of migrant individuals to allow meaningful statistical analysis (CONAPO 2000). As a consequence, there is always a need to over-sample areas with strong migrant tradition if useful numbers of migrants are to be obtained. Moreover, it is well-documented that migrants do not come at random from all the geographical areas of Mexico. Instead, they cluster intensively in the States and areas covered by the MMP107 (CONAPO 2000). Hence, if a trend is not present in the MMP107 data, it will hardly appear in a national representative survey. From this point of view, using the MMP107 to perform exploratory analyses of Mexico-US migration issues is well justified and a number of influential papers in the field have used the survey (see, for instance, Delechat 2001, Durand et al. 1996).

The MMP107 has characteristics that made it an important source of information for the study of migration. First, and substantively, it is the only Mexico-US migration survey that covers long-term migrants. In particular, information about the head of the household and all her/his sons and

[5] Twenty States are covered: Aguascalientes, Baja California Norte, Chihuahua, Colima, Durango, Guanajuato, Guerrero, Hidalgo, Jalisco, Michoacan, Nayarit, Nuevo Leon, Oaxaca, Puebla, San Luis Potosi, Sinaloa, Tlaxcala, Veracruz, and Zacatecas.

daughters is gathered, independently of the current location or household membership status of the latter individuals. This implies that data for all sons and daughters is available even if some of them formed their own households and emigrated to the US — and haven't come back — many years before the survey. Further, an individual's emigration event is recorded regardless of her/his legal status in the United States. Date and destination of every legal or illegal border crossing in an individual's life history is carefully documented. The other two major surveys about Mexico-US migration, the ENADID and the MXLFS, do not cover long-term emigrants.[6]

Another significant advantage of the MMP107 over other sources is its special focus on migrant networks. Detailed information about migration status of the family and extended family of the household head and her/his spouse are available. Migration characteristics of friends of the head are also known. Finally, the MMP107 contains a number of community level data including the proportion of migrants in the locality.

The present study is based on information for 5,354 siblings collected in 1,206 households in 16 communities of more than 15,000 inhabitants throughout Mexico between 1997 and 2004.[7]Siblings are clustered in families. Hence, the estimation sample can be seen as an unbalanced panel given that different families have a different number of children. Within a household the MMP107 gives information about the age and birth order of each sibling.

[6]In both cases information is collected for persons who lived in the household up to five years before the date of the survey. Anyone who left the household before that is not considered a member and no information is recorded. This is unfortunate because, most likely, many migrants do not comply with such requirements.

[7]In previous years the MMP survey did not collected data for some of the relevant variables for the analysis. For these reasons, the present study uses data gathered from 1997 onwards.

Since the focus of the paper is high school graduation, only individuals aged 18 or over at the time of the survey are included in the sample. The MMP107 contains information on whether individuals have ever migrated to the US (usmigra=1) and on whether they graduated from high school (prepa=1). These are the two dichotomous dependent variables. Eighteen per cent of the individuals have migrated to the US at least once. Similarly, twenty eight per cent of the sample are high school graduates. Migrants are clearly less educated. In fact, 17% of the migrants are graduates compared to the 30% of non-migrants. Table 1 contains summary statistics.

[Table 1 around here]

Family migration networks effects are controlled for by a number of variables. Dummy variables indicating whether the household head and her spouse have ever migrated to the US are included (husmigra=1 and spmigra=1 respectively). Similarly, the number of siblings of the head and her spouse with migration experience are also controlled for. Finally, the number of migrants in the head's (spouse's) extended family are present in the list of explanatory variables as well. The relevance of social networks is accounted for by the inclusion of controls for the number of friends of the household head with migration experience and the percentage of males in the community who have ever migrated to the United States in 1990.

Other explanatory variables include sex, age, education of the family head (and spouse), total number of children the head ever had, number of rooms of the parental household (which is taken as a proxy for wealth), percentage of community's labour force which are self-employed, unemployment rate and size of the labour force at the community's main US city/urban area destination. Finally, dummies for birthplace, region, and survey year are also included.

4 Econometric Issues

Dynamic bivariate random effects Probit models are used for the analysis. Denote by M_{ji} the variable that takes on one if, by the time of the survey, the i-th sibling in the j-th family has emigrated to the US at least once and zero otherwise. Similarly, E_{ji} indicates whether the i-th sibling in the j-th family graduated from high school ($E_{ji} = 1$) or not ($E_{ji} = 0$) by the time of the survey. Siblings within families are ordered by age so that the jk-th individual is older than the jl-th whenever $l > k$.

4.1 Dynamic equations

A latent variable framework is the natural approach. Let M^*_{ji} and E^*_{ji} be two latent continuous variables. The econometrician does not observe M^*_{ji} and E^*_{ji}. Instead two dichotomised variables, M_{ji} and E_{ji}, are available. It is supposed that the high school dummy is generated according to the following data generating process,

$$E^*_{ji} = \mathbf{x^e_{ji}}\boldsymbol{\beta}^{\mathbf{e}} + \delta_{11}E_{j,i-1} + \delta_{12}M_{j,i-1} + f^e_j + u^e_{ji}, \tag{1}$$

with $E_{ji} = 1$ if $E^*_{ji} > 0$ and zero otherwise. Notice that $\mathbf{x^e_{ji}}$ represents a vector of observed characteristics that can vary at the individual, family, and community levels. Elements of $\mathbf{x_{ji}^{e}}$ are assumed to be strictly exogenous and $\boldsymbol{\beta^{e}}$ denotes a conformable coefficient vector — including the constant term. Similarly, $\delta_1 = \{\delta_{11}, \delta_{12}\} \in \mathbb{R}^2$ represent coefficients on the migration and education outcomes of the immediately elder sibling in the j family. Finally, variables f_j^e and u_{ji}^e are random heterogeneity terms. One term, f_j^e, varies at the family level while the other term, u_{ji}^e, varies at the individual level. The equation for the migration dummy is,

$$M_{ji}^{*} = \mathbf{x_{ji}^{m}}\boldsymbol{\beta^{m}} + \delta_{21}E_{j,i-1} + \delta_{22}M_{j,i-1} + f_j^m + u_{ji}^m, \tag{2}$$

with $M_{ji} = 1$ if $M_{ji}^{*} > 0$ and zero otherwise. Following Alessie et al. (2004), f_j^m and f_j^e are specified to be jointly Normally distributed with mean vector zero and covariance matrix Σ_f,

$$\Sigma_f = \begin{bmatrix} \sigma_m^2 & \rho\,\sigma_m\sigma_e \\ \rho\,\sigma_m\sigma_e & \sigma_e^2 \end{bmatrix}.$$

In a similar fashion, u_{ji}^m and u_{ji}^e are jointly Normal with mean vector zero and covariance matrix Σ_u,

$$\Sigma_u = \begin{bmatrix} 1 & \rho_u \\ \rho_u & 1 \end{bmatrix}.$$

To close the model it is assumed that f_j^h and u_{ji}^h are independent, for $h = (m, e)$. Further, errors f_j^h and u_{jk}^h are serially uncorrelated for every j and k.

The model implies the following relationships. M^*_{ji} and M^*_{jk}, $k \neq i$, are correlated within the j-th family through the random term f^m_j. However, no such correlation exist among individuals who belong to different families. Intra-family clustering is also induced between E^*_{ji} and E^*_{jk} by the random term f^e_j. Also, at the family level, correlation between E^*_{ji} and M^*_{jk} for all i and k that belong to the j-th family is induced by correlation between f^e_j and f^m_j. Finally, at the individual level, correlation between M^*_{ji} and E^*_{ji} is created by correlation between u^m_{ji} and u^e_{ji}. True dynamic sibling dependence is present if at least one element of vector $\delta = (\delta_{11}, \delta_{12}, \delta_{21}, \delta_{22})$ is different from zero. In particular, we say that true "self" dynamic sibling dependance is present if δ_{11} and/or δ_{22} are different from zero. Similarly, true "cross" dynamic sibling dependance is present if δ_{12} and/or δ_{21} are different from zero.

4.2 Initial conditions

Given that migration and educational outcomes of different siblings within the j-th family are correlated, treating M_{j0} and E_{j0} as exogenous in system (1)-(2) will produce inconsistent estimators. This is known in the econometrics literature as the initial conditions problem. To address the problem we follow the strategy suggested by Heckman (1981b). Namely, a model for the reduced-form marginal probability of M_{j0} and E_{j0} given f^e_j and f^m_j is specified. Hence two further equations are needed,

$$E^*_{j0} = \mathbf{z^e_{j0}}\boldsymbol{\gamma}^\mathbf{e} + \lambda_{11} f^e_j + \lambda_{12} f^m_j + v^e_{j0} \tag{3}$$

$$M^*_{j0} = \mathbf{z^m_{j0}}\boldsymbol{\gamma^m} + \lambda_{21} f^e_j + \lambda_{22} f^m_j + v^m_{j0} \tag{4}$$

with $E_{j0} = 1$ if $E^*_{j0} > 0$ and $M_{j0} = 1$ if $M^*_{j0} > 0$ and zero otherwise. As usual, $\mathbf{z^e_{j0}}$ and $\mathbf{z^m_{j0}}$ represent vectors of explanatory variables that can vary at the individual, family, and community level. Notice that $\lambda = (\lambda_{11}, \lambda_{12}, \lambda_{21}, \lambda_{22}) \in \mathbb{R}^4$ are free parameters (*factors loadings*) that allow any type of correlation among E^*_{j0}, M^*_{j0}, E^*_{ji}, and M^*_{ji}. We suppose that v^h_{j0} is uncorrelated with v^h_{jk} for every j and k. As usual, v^e_{j0} and v^m_{j0} are jointly normal with mean vector zero and covariance matrix Σ_v,

$$\Sigma_v = \begin{bmatrix} 1 & \rho_v \\ \rho_v & 1 \end{bmatrix}.$$

4.3 Identification

Technically the model is identified through functional form (see Heckman 1978). However, in the absence of exclusion restrictions identification may be 'tenuous' (in the context of the multinomial probit model see Keane 1992). Hence, specifying exclusion restrictions to help identification is a good practise.

Using information from the MMP survey one can identify the main US city/urban area destination of each community in the sample between 1990 and 2000. Similarly, local area unemployment rates and labour force statistics in the US are available from the Bureau of Labor Statistics (BLS). Hence, it is possible to obtain an average unemployment rate (laur) and size of the labour force (lforce) between 1990 and 2000 for each local area reported by the BLS and match such information with the MMP data. Both laur and lforce are indicators of the labour market characteristics of the main US city/urban area destination of the MMP communities included in the sample.

Variables laur and lforce enter the migration equations but are excluded from the schooling equations. Clearly, unemployment rate at the community's main US destination is a good indicator of how difficult is for new immigrants to find a job at arrival. The higher laur is the less attractive migration will be for prospective migrants. Similarly, large cities have complex economies and are more capable of absorbing people with different skills and backgrounds than small urban areas. As a consequence, one can expect migration to be more attractive as lforce becomes larger. Both laur and lforce are unlikely to affect education decisions in Mexico and, if they do, it is exclusively through their impact on migration. These two variables are, therefore, good candidates for imposing exclusion restrictions to help identification.

Conditional on the migration status of the head and the spouse, it is likely that the education of head and/or spouse will affect children's probability of high school graduation but have no bearing on children's probability of migration. In such a context, the education of the head and/or the spouse can be included in the education equations but be excluded from the migration equations. Over-identification tests are performed to check the validity of this hypothesis.

4.4 Estimation strategy

The model is estimated by Maximum Simulated Likelihood (see, for instance, Train 2003). The contribution of the j-th family to the likelihood is,

$$
\begin{aligned}
L &= \int\int \Phi_2\left(q_1 w_{11}, q_2 w_{12}, q_1 q_2 \rho_v\right) \\
&\quad \times \prod_{j=1}^{J} \Phi_2\left(q_1 w_{21}, q_2 w_{22}, q_1 q_2 \rho_u\right) g\left(f^e, f^m, \Sigma_f\right) df^e df^m \qquad (5)
\end{aligned}
$$

where $g(.)$ represents the bivariate normal density of the family random effects, $q_1 = 2E_{ji} - 1$, and $q_2 = 2M_{ji} - 1$. Finally, w_{11} and w_{12} are the right-hand side of equations (3) and (4) excluding u^e_{ji} and u^m_{ji} respectively. Variables w_{21} and w_{22} are defined in the same fashion using equations (1) and (2).

Two uncorrelated Halton sequences of dimension R are first obtained. Then random draws from density $g(.)$ are simulated using the Halton sequences, a Cholesky decomposition, and the inverse cumulative normal distribution. Next, for each draw (which is a two dimension vector), the conditional likelihood of the j-th family is evaluated. Finally, an average of the R simulated conditional likelihoods is taken. This average is the contribution of the j-th family to the overall simulated likelihood — an approximation of the double integral in (5). Halton sequences have been shown to achieve higher precision with fewer draws than uniform pseudorandom sequences because it have a better coverage of the $[0, 1]$ interval (for more on this topic see Train 2003).

Maximum simulated likelihood is asymptotically equivalent to ML as long as R grows faster than $\sqrt{N}$ (Gourieroux and Monfort 1993). Following Alessie et al. (2004) maximisation is performed on the basis of the BHHH algorithm. At convergence, numerical second derivatives are obtained to calculate the robust covariance matrix.

5 Empirical Results

Table 2 presents the results. For comparison reasons, table 2 contains results from univariate dynamic probit models for usmigra and prepa along with the estimates from the bivariate dynamic model. Regressions were initially estimated using 200 Halton draws. Then, 50 draws were successively added until no significant differences in coefficients and log-likelihood were detected. In all cases 400 Halton draws were enough to achieve high precision. Marginal effects (MEs) are calculated at the means of the independent variables and standard errors are obtained using the delta method.

[Table 2 around here]

Let us start the discussion with the results from the univariate models. Exlusion Wald tests at the bottom left of Table 2 confirm that the spouse's education dummies can be excluded from the migration equation but not from the education equation. In contrast, the education dummies of the head cannot be excluded from any of the two univariate models. Clearly, Table 2 show that the more educated a head and his/her spouse are the more chances that their children will graduate from high school and less chances that their children will choose to emigrate. Estimates for σ_e and σ_m are significant at 1%. Hence, itra-family clustering is present in both migration and education equations.

Variables laur and lforce affect significantly the probability of migration. In fact, the unemployment rate on the community's main US destination is detected to have, as expected, a negative marginal effect on the probability of migration of about 8.5 percentual points (p.p. hereafter). This marginal effect is significantly different from zero at a 1% level. A similar story can be told for the size of the labour force at the community's main US destination. A positive marginal effect of lforce on usmigra of about 4 p.p is detected, and such marginal effect is significantly different from zero at a 1% significance level.

Migrant network variables have a significant impact on the likelihood of high school graduation. In fact, a Wald test for the exclusion of all migrant network variables in the univariate model for usmigra rejects the null hypothesis at a 5% significance level (p-val = 0.03). This is a test for the joint exclusion of: hmigra, hsbus, hexfus, frevus, spmigra, sbilevus, spexfus, mratio90. A similar conclusion is obtained from the univariate model for usmigra. Notice however that, unlike the schooling equation, the marginal effect on the lagged dependent variable in the usmigra equation is found to be insignificantly different from zero. In other words, migrant network effects are found in both education and migration equations but true self dynamic sibling dependence is present only in the schooling variable.

Among other results, the univariate model for prepa suggests that a migrant mother — the spouse of the head is in most cases the mother — increases the likelihood of high school graduation by 11.4 p.p. Interestingly, the migration status of the family head does not affect significantly children's likelihood of high school graduation. This result is consistent with findings in the intra-household resource allocation literature showing that income in the hands of a mother has much higher effect on children's health and education outcomes than income managed by the father (see, for instance, Thomas 1990). The univariate model for usmigra suggests that both mother and father migration status increase children's likelihood of migration — by 9 p.p. and 4 p.p. respectively. Yet again, it is mother's migration status the factor that affects the most usmigra.

Let us now move to discuss empirical results from the bivariate random effects dynamic probit (right panel of Table 2). A likelihood ratio test for the null of $\rho_u = \rho_v = \rho = \delta_{12} = \delta_{21} = 0$ is provided at the bottom right of Table 2. This is a test for the relevance of the bivariate model over the information already provided by the univariate models. The null hypothesis is easily rejected with a $\chi^2(5) = 18$ and a p-val $= 0.006$.

Like in the two univariate models, over-identification test in the bivariate model show that the spouse's education dummies can be excluded from the usmigra equation ($\chi^2(5) = 8.42$, p-val $= 0.59$) but not from the prepa equation ($\chi^2(5) = 20.26$, p-val $= 0.03$). In contrast, the education dummies of the family head cannot be excluded from any of the two equations. As before, laur and lforce are highly significant in the migration equations. In fact, marginal effects on laur and lforce on the usmigra equation are significant at 5% and have their expected signs. Finally, estimates for σ_e and σ_m are significantly different from zero at all standard significance levels. Hence, there is strong evidence that intra-family clustering is present in both migration and schooling equations.

Interestingly, correlation between the random terms u^e and u^m, ρ_u, is insignificant. Therefore, at the individual level, unobservables in the migration and the education equations are independent. A similar observation is valid for ρ_v (see right panel of Table 2). Correlation between the family random effects f^e and f^m, ρ, is negative and marginally significant at 10% (p-value = 0.055). As a consequence, one can conclude that family unobservable traits that increase the likelihood of migration are associated as well with reductions in the likelihood of high school graduation.

A negative ρ implies that individuals who study more migrate less. This is consistent with previous work on the Mexico-US literature suggesting that migrants to the US are drawn from the bottom tail of the skills distribution — a phenomenon commonly known as "negative migrant selection" (see, for example, Borjas 1994). However, empirical evidence suggests that what induces negative correlation among education and migration decisions are unobserved traits that affect all siblings in the family rather than individual specific unobserved heterogeneity. So, the idea that individuals are selected into migration on the basis of skills should be taken with care as other factors may be at work.[8]

For example, if a family is badly hit by an adverse event such as illness or unemployment of the family head (which may be a specially common event during recessions) all siblings in the family may be obliged to leave school and to migrate. Such adverse events are common shocks to all siblings in a given family and will generate a negative correlation between the family random heterogeneity terms f^e and f^m. Clearly, this is relevant new evidence that univariate models cannot deliver.

[8]Significant negative ρ_u or ρ_v would be strong evidence of selection on the basis of skills and, in particular, negative migrant selection. Data, however, do not seem to support this view.

Another advantage of the bivariate model over univariate ones is its ability to test for the presence of true cross dynamic sibling dependance, which occurs whenever δ_{12} and/or δ_{21} are different from zero in equations (1) and (2). Table 2 shows that, at a 1% significance level, usmigra$_{j,i-1}$ has a positive marginal effect on prepa of 0.06 (which implies a higly significant δ_{12}). Hence, empirical evidence shows that having a migrant elder sibling increases the likelihood of high school graduation by 6 p.p. No evidence was found to suggest that the education of an elder sibling affects the odds of a migration event. Finally, in line with findings from univariate models, self dynamic sibling dependance is significant only in the prepa equation.

Wald tests for the exclusion of the migration variables in the schooling equations clearly reject the null hypothesis at a 1% of significance level. Therefore, as in the univariate case, here there is strong evidence of migrant network effects affecting education decisions.

Looking at the marginal effects on the marginal probability of prepa=1, the reader can conclude that a migrant mother increases significantly the odds of high school graduation by 12 p.p. A migrant head has no significant marginal effect on the marginal probability of prepa=1. Similarly, a community's male migration ratio in 1990 is found to have a significantly negative marginal effect on the probability of prepa=1. This is consistent with findings reported by McKenzie and Rapoport (2006) using ENADID data and IV methods to control for the endogeneity of migration in schooling equations. However, the size of the marginal effect on the marginal probability of prepa=1 found here is rather negligible (less than 1 p.p.). Hence, it seems this is a second order effect. Given the evidence, and at least for now, policy makers should focus on designing strategies to ensure source countries benefit the most from the positive effects that go from migration to education rather than trying to minimise migration's negative side effects on education.

In the case of the migration equation, significant migrant networks are also detected. In particular, marginal effects on the marginal probability of usmigra=1 indicate that a migrant father (mother) increase the chances of migration by 5 p.p. (9 p.p.).

Marginal effects on marginal probabilities from the bivariate model are similar to those calculated from the the univariate models — a fact that is expected given that the correlation between unobservables across equations is rather low. Estimating the bivariate model, however, delivers important new pieces of information. Fist, it is found that negative migrant selection is driven by correlation of unobservables at the family level and that factors other than skills may be at work. Hence, further research on the field is needed to explore and identify what other factors may generate negative migrant selection. This conclusion is policy relevant because up to now the literature on Mexico-US migration has always stressed the idea that negative migrant selection is based on unobservable skills. Second, the bivariate model finds that, along with significant positive migrant network effects on education and education choices, there are significant positive dynamic feedbacks from migration to education.

6 Conclusions

The present paper enquires about the potential links between family and community migration and the probability of high school graduation in urban Mexico. Bivariate dynamic Probit models for panel data are estimated to account for the fact that unobservables can be correlated across migration and education decisions as well as within groups of individuals such as the family. Maximum simulated likelihood techniques are used for the analysis.

The study shows that a migrant mother increases by 12 percentual points (p.p. hereafter) the likelihood that her children will be high school graduates. Similarly, a migrant elder sibling increases the likelihood of high school graduation by 6 p.p. These are good news showing that there are previously unaccounted significant positive feedbacks going from migration to education that, eventually, may help source countries to increase their accumulation of human capital. Interestingly, the migration status of the family head is found to have no bearing on the odds of high school graduation. These results are consistent with findings in the intra-household resource allocation literature showing that income in the hands of a mother affects more children's health and education than income in the hands of a father.

In line with previous studies on the Mexico-US migration literature, community and family migration experience are found to increase the likelihood of migration — i.e., there are important migrant network effects. In particular, it is found that a migrant family father (mother) increases the likelihood that his/her children will migrate by 4 p.p. (9 p.p.). Hence, like in the case of education, evidence suggests that mothers' outcomes are the factors that affect the most children's migration.

Significant intra-family clustering affecting both schooling and migration decisions is detected. In line with previous work in the Mexico-US migration literature, evidence of negative migrant selection is found. The present study finds, however, that what drives negative correlation between education and migration decisions are correlated unobservable treats at the family rather than at the individual level. As a consequence, the idea that individuals are selected into migration mainly on the basis of skills should be taken with care as other factors may be at work — adverse family shocks (say, unemployment of the family head during a recession) are capable of inducing negative correlation among family unobserved treats affecting migration and education decisions and create the type of negative migrant selection detected here.

Table 1. Descriptive Statistics

Variable	Description	Obs	Mean	Std. Dev.	Min	Max
Individual characteristics						
sex	=1 if male	5354	0.51	0.50	0	1
age	age in years	5354	31.14	9.36	18	60
usmigra	=1 if ever migrated to the US	5354	0.18	0.38	0	1
prepa	=1 if completed high school	5354	0.28	0.45	0	1
Head of household						
hsex	=1 if female	5354	0.22	0.41	0	1
hmigra	=1 if ever migrated to the US	5354	0.27	0.44	0	1
hnchild	No. of children ever born	5354	7.17	3.18	1	19
hsbus	No. of migrant siblings	5354	0.55	1.18	0	11
hexfus	No. of migrants in extended family	5354	7.45	13.07	0	121
frevus	No. of migrant friends	5354	2.36	11.07	0	200
Head's spouse						
spmigra	=1 if ever migrated to the US	5354	0.28	0.45	0	1
sbilevus	No. of migrant siblings	5354	0.77	1.96	0	20
spexfus	No. of migrants in extended family	5354	0.95	1.93	0	15
Head's education						
hedug1	Less than primary	5354	0.45	0.50	0	1
hedug2	Primary	5354	0.25	0.43	0	1
hedug3	Secondary	5354	0.06	0.24	0	1
hedug4	High school or higher	5354	0.07	0.25	0	1
Head's spouse education						
sedug1	Less than primary	5354	0.33	0.47	0	1
sedug2	Primary	5354	0.22	0.41	0	1
sedug3	Secondary	5354	0.04	0.20	0	1
sedug4	High school or higher	5354	0.04	0.20	0	1
sedug88	Missing	5354	0.25	0.43	0	1
Head of household wealth						
prooms	No. of rooms parental household	5354	4.96	2.01	1	18
Community						
self90	% of self-employed in 1990	5354	19.14	7.30	9.07	38.5
mratio90	% of male migrant population in 1990	5354	24.46	16.47	0.54	67.9
laur	Unemployment rate (%) in main US destination	5354	6.06	0.92	4.1	6.9
lforce	Labour force (millions) in main US destination	5354	4.16	2.18	0.17	8.52
Birthplace						
North	North	5354	0.36	0.48	0	1
Centre	Centre	5354	0.29	0.45	0	1
CentreP	Centre Pacific	5354	0.20	0.40	0	1
South	South	5354	0.08	0.27	0	1
Survey year						
yr1998	1998	5354	0.25	0.43	0	1
yr1999	1999	5354	0.08	0.28	0	1
yr2000	2000	5354	0.16	0.37	0	1
yr2001	2001	5354	0.17	0.37	0	1
yr2003	2003	5354	0.17	0.37	0	1
yr2004	2004	5354	0.08	0.27	0	1

Table 2. Random effects dynamic Probit results — Marginal effects

	Univariate models				Bivariate Model			
	prepa		usmigra		prepa		usmigra	
Variable	ME	RSE	ME	RSE	ME$^{(a)}$	RSE	ME$^{(a)}$	RSE
Individual characteristics								
sex	0.028†	0.015	-0.131††	0.013	0.028†	0.015	-0.133††	0.013
age	0.001	0.001	0.002††	0.001	0.001	0.001	0.002††	0.001
Head of household								
hsex	0.016	0.055	-0.056††	0.023	0.028	0.058	-0.058††	0.023
husmigra	0.023	0.027	0.044††	0.018	0.018	0.028	0.046††	0.019
hnchild	-0.010††	0.004	-0.002	0.003	-0.009††	0.004	-0.001	0.003
hsbus	0.003	0.010	0.005	0.006	0.001	0.01	0.006	0.006
hexfus	0.001	0.001	-0.001	0.001	0.001	0.001	-0.001	0.001
frevus	-0.001	0.001	0.001	0.001	-0.001	0.001	0.001	0.001
Head's spouse								
spmigra	0.114††	0.051	0.089††	0.035	0.117††	0.052	0.089††	0.036
sbilevus	-0.007	0.005	0.001	0.005	-0.006	0.005	0.001	0.005
spexfus	0.008	0.007	0.031††	0.004	0.002	0.007	0.030††	0.004
Head's education								
hedug1	0.082††	0.034	0.036†	0.019	0.080††	0.035	0.036†	0.020
hedug2	0.226††	0.050	0.030	0.025	0.220††	0.052	0.026	0.026
hedug3	0.288††	0.081	0.044	0.042	0.285††	0.083	0.040	0.043
hedug4	0.494††	0.085	-0.048†	0.027	0.501††	0.087	-0.054††	0.027
Head's spouse education								
sedug1	0.050	0.042			0.037	0.046		
sedug2	0.086†	0.051			0.075	0.054		
sedug3	0.135	0.084			0.122	0.086		
sedug4	0.288††	0.101			0.244††	0.101		
spedg88	-0.036	0.067			-0.060	0.064		
Head of household wealth								
prooms	0.036††	0.006	0.002	0.004	0.034††	0.006	0.001	0.004
Community								
self90	-0.002	0.002	0.001	0.002	-0.002	0.002	0.002	0.002
mratio90	-0.003††	0.001	0.002††	0.001	-0.003††	0.001	0.002††	0.001
laur			-0.085††	0.033			-0.078††	0.034
lforce			0.036††	0.015			0.033††	0.015
Birthplace and year dummies								
Birthplace	yes		yes		yes		yes	
Year	yes		yes		yes		yes	
Lagged Dependent Variables								
$\text{prepa}_{j,i-1}$	0.151††	0.036			0.154††	0.035	0.022	0.022
$\text{usmigra}_{j,i-1}$			0.018	0.017	0.060††	0.030	0.020	0.017
σ_e	0.728††	0.098					0.744††	0.092
σ_m			0.663††	0.069			0.663††	0.072
ρ_u							-0.064	0.061
ρ_v							0.211	0.144
ρ							-0.263†	0.137
Exclusion Wald tests								
Head edu$^{(b)}$		59.61 (0.00)		18.98 (0.01)		54.54 (0.00)		18.68 (0.02)
Spouse edu$^{(b)}$		21.80 (0.00)		9.16 (0.52)		20.26 (0.03)		8.42 (0.59)
Migr. vars.$^{(c)}$		29.91 (0.03)		132.92 (0.00)		28.86 (0.04)		125.26 (0.00)
Model relevance								
$\rho_u = \rho_v = \rho = \delta_{12} = \delta_{21} = 0$					$\chi^2(5) = 18$ (pval = 0.006)			
Model information								
No. Halton draws		400		400				400
No. families		1206		1206				1206
No. observations		5354		5354				5354
Log-likelihod		-2165.5		-1971.9				-4128.4

Note. Marginal effects are evaluated at the mean of the independent variables; robust standard errors (RSE) are reported. †† (†) Significant at 5% (10%). $^{(a)}$ Marginal effects on Marginal probabilities. $^{(b)}$ Joint test for exclusion of the education dummies in dynamic and initial conditions equations (p-values in brackets). $^{(c)}$ Joint exclusion test of migration variables in dynamic and initial conditions equations. This is a test for the exclusion of: hmigra, hsbus, hexfus, frevus, spmigra, sbilevus, spexfus, mratio90, and $\text{usmigra}_{j,i-1}$ when relevant (p-values in brackets). Results from initial conditions are available from the author upon request.

References

Alessie, R., Hochguertel, S., Van Soest, A., 2004. Ownership of stocks and mutual funds: A panel data analysis. The Review of Economics And Statistics 86, 783–796.

Arulampalam, W., Bhalotra, S., 2006. Sibling death clustering in india: Genuine scarring vs unobserved heterogeneity. The Journal of the Royal Statistical Society, Series A (forthcoming).

Banco de México, 2005. Informe anual.

Blundell, R., Dearden, L., Goodman, A., Reed, H., 2000. The returns to higher education in Britain: Evidence from a british cohort. Economic Journal 110 (461), 82–100.

Borjas, G., 1994. The economics of immigration. Journal of Economic Literature 32 (4), 1667–1717.

CONAPO, 2000. Migracion Mexico-EU presente y futuro. CONAPO.
URL http://www.conapo.gob.mx/

Delechat, C., 2001. International migration dynamics: The role of experience and social networks. Labour 15 (3), 457–486.

Durand, J., Kandel, W., Parrado, E., Massey, D., 1996. International migration and development in mexican communities. Demography 33 (2), 249–264.

Gourieroux, C., Monfort, A., 1993. Simulation-based inference: A survey with special reference to panel data models. Journal of Econometrics 59, 533.

Hanson, G., Woodruff, C., 2003. Emigration and educational attainment in Mexico, University of California at San Diego, manuscript.

Heckman, J., 1981a. Heterogeneity and state dependence. In: Rosen, S. (Ed.), Studies in Labor Markets. University of Chicago Press, pp. 91–139.

Heckman, J., 1981b. The incidental parameters problem and the problem of initial conditions in estimating a discrete time-discrete data stochastic process. In: Manski, C., McFadden, D. (Eds.), Structural Analysis of Discrete Data with Econometric Applications. MIT Press, pp. 179–195.

Heckman, J. J., 1978. Dummy endogenous variables in a simultaneous equation system. Econometrica 46 (4), 931–959.

Keane, M. P., 1992. A note on identification in the multinomial probit model. Journal of Business & Economic Statistics 10 (2), 193–200.

McKenzie, D., Rapoport, H., 2006. Can migration reduce educational attainment? evidence from Mexico, World Bank policy research working paper No. 3952.

Miranda, A., Bratti, M., 2006. Non-pecuniary returns to higher education: The effects on smoking in the UK, IZA Discussion Paper No. 2090.

Thomas, D., 1990. Intra-household resource allocation an inferential approach. Journal of Human Resources 25 (4), 635–664.

Train, K., 2003. Discrete Choice Methods with Simulation. Cambridge University Press.

UNPD, 2006. International migration wall chart 2006, United Nations Population Division.

Vidal, J.-P., 1998. The effect of emigration on human capital formation. Journal of Population Economics 11 (4), 589–600.

Winters, P., De Janvry, A., Sadoulet, E., 2001. Family and community networks in Mexico-U.S. migration. Journal of Human Resources 36 (1), 159–184.

World Bank, 2006. Economic implications of remittances and migration, Global Economic Prospects.