# Self-Selection and Liquidity Constraints in Different Migration Cost Regimes

Scott C. Borger[*]

Revised: September 2010

## Abstract

This paper presents a theoretical model of the decision to migrate. A liquidity constraint arises from the inability of potential migrants to borrow against future earnings in the United States to pay the cost for 'professionally-assisted' clandestine entry. However, family members in the United States often assist unauthorized migrants with the smuggling cost and thereby alleviate the liquidity constraint. This theoretical model, therefore, incorporates both the liquidity constraint and the ability of social networks to alleviate the liquidity constraint in the decision to migrate. As the smuggling cost along the U.S.-Mexico border increased, the characteristics of who is able to migrate has evolved. Unauthorized migration is separated into two periods: a low-cost migration period before 1993 and a high-cost migration period after 1993. To test the theoretical model empirically, a method is proposed to control the returned-migrant selection bias based on the limited long-term data available on unauthorized migrations from Mexico to the United States. This paper finds evidence of a shift in the migrant-sending regions of Mexico from negative self-selection in the low-cost period to intermediate self-selection in the high-cost period. Additionally, social networks increase the propensity to migrate for resource-constrained potential migrants, but the effect is not different between the two periods.

# 1 Introduction

A liquidity constraint arises in the decision to migrate when the potential migrant does not have the savings and cannot borrow against his or her future earnings in the United States to pay the cost of migration. Yet, family members in the United States often assist unauthorized migrants with the smuggling cost and thereby alleviate the liquidity constraint. This paper explicitly models the liquidity constraint faced by potential unauthorized migrants and allows for social networks in the United States to alleviate this constraint. The standard migration model, beginning with the work of Sjaastad (1962), implicitly assumes the cost of migration can be paid out of future earnings when evaluating the decision to migrate. If the potential migrant is able to leverage future funding or cover the cost through borrowing or savings, then the standard calculation represents the 'true' decision of the migrant. However, if the the cost of migration cannot be paid in this manner, then the migrant's decision includes a liquidity constraint.

This paper proposes a constrained optimization model to incorporate the inability of potential migrants to borrow against future earnings using formal banking loans.[1] The constraint applies to all potential migrants.[2] By requiring the cost of migration to be paid in the period before wages in the destination country can be earned, this paper builds on the work of Orrenius and Zavodny (2005). In essence, the "cash-in-advance" constraint prevents the least-skilled potential migrants from migrating (Orrenius and Zavodny, 2005; Stark and Taylor, 1991).

However, this constraint is not binding if family members in the United States are able to provide a loan to pay the up-front cost of migration that will be paid back when the migrant is earning wages in the United States. It is important to note that in my model, social networks are not assumed to reduce the total cost faced by potential migrants, as assumed previously in the literature. Social networks, in that case, reduce the intangible cost of migration by providing the migrant 'social capital,' such as knowledge about crossing methods or employment opportunities in the United States (Singer and Massey, 1998; Munshi, 2003; Mckenzie and Rapoport, 2007a, 2007b). Rather, in this model, social networks are assumed to reduce the tangible out-of-pocket expenses

---

[1]Hatton and Williamson (2005) use the term poverty constraints in the context of the Great Migration in the late 1800s and early 1900s to define the inability of potential migrants with limited resources to migrate. Orrenius and Zavodny (2005) and Stark and Taylor (1991) use the term 'cash-in-advance' constraint.

[2]An extension of the model suggests that potential migrants vary in their exposure to formal credit markets. Fernandez-Huertas (forthcoming) notes this possibility in the context of differences in the characteristics of migrants from urban and rural Mexico.

faced by potential migrants before earning wages in the United States.

The costs of unauthorized migration across the United States-Mexico border fall into two categories: out-of-pocket expenses and implicit costs. The out-of-pocket expenses consist of the fee required by the *coyotaje* (people-smuggler) who assists the migrant in crossing the international border, the costs of transportation to the border and then to the migrant's destination in the United States. The implicit costs of migration include the psychological burden to the migrant of being away from home and searching for employment after arriving. While the implicit costs have been reduced by the formation of social networks in the United States (Espinosa and Massey (1997), Orrenius (1999), Munshi (2003) and Waldinger and Lichter (2003)), out-of-pocket expenses significantly increased in inflation-adjusted terms as a response to the intensification of border enforcement(Gathmann (2008), Roberts et al. (2010)). This paper will explore the impact of an increase in the out-of-pocket expenses on the decision to migrate.

Average smuggling fees have increased more than threefold in real terms since the mid-1990s, as reported by migrants in the surveys conducted by the Mexican Migration Project (MMP, 2008) and the Mexican Migration Field Research and Training Program (MMFRP) (Cornelius, Fitzgerald and Borger (2009)). This increase in the smuggling cost significantly affects the ability of potential migrants to pay for these services and in many cases forces a decision against migration. Surveys show that most migrants unable to secure loans against their future earning, and pay the cost out of savings or use their social networks in the United States to pay the smuggling cost (Cornelius, 2001; Massey, Goldring and Durand, 1994).

The decision to migrate in my theoretical model requires two standard conditions to be satisfied. The first is that the sum of the migrant's future wages in the destination country (less the cost of migration) exceeds the sum of the foregone future wages of the migrant in his country-of-origin.[3] Second, to be consistent with the evidence from migrant surveys, the cost of migration is paid from migrant's savings or through the migrant's social networks.[4]

In addition to the decision based on whether the migrant's future wages in the United States will offset the cost of migration and whether the migrant has the savings or social networks available to

---

[3]This model considers the impact of cost on unauthorized migrations, so the differentiation of wages between legal and illegal migrants in Ethier (1986) is not necessary.

[4]The dynamics of migration costs are considered in Carrington, Detragiache and Vishwanath (1996) where the costs decrease as social networks in the destination increase. However, by limiting my model to a two-period framework, migration patterns can be characterized without knowing how the social networks evolved over time.

cover the cost, the theoretical model includes a function that transfers foreign skill levels into U.S. labor markets.[5] The model assumes that migrants' skills will not fully transfer into the destination country's labor markets and that the disparity in transferring skills is higher for higher-skilled workers. This non-linear functional form is influenced by factors such as language barriers or certification requirements that prevent foreign skill levels from being transferred one-for-one. This is similar to the concept in the seminal work of Borjas (1987), building on Roy (1951). Using the intercountry differences in skill to model the self-selection of migrants, Borjas (1987) concludes that migrants with lower skill levels receive higher returns on their labor compared to those with higher skill levels. Since income disparity is greater in Mexico than in the United States, Mexican migrants in the United States would tend toward negative self-selection. However, the finding of negative self-selection in Borjas (1987) would be consistent with my model's predictions only if the smuggling cost is very low or potential migrants always has social networks and/or savings to pay the smuggling costs in the period before the migrant earns wages in the United States. But if the liquidity constraint is faced by more and more migrants as smuggling costs increase, then migration is restricted to those with mid-level skills, which is described by the term 'intermediate self-selection' (representing the middle of their country-of-origin's skill distribution).

This model considers a range of skill levels among potential unauthorized migrants, imposes the restriction of non-negative assets, and transfers foreign skills level into the U.S. skill level equivalent. With this basic setup, the model is able to isolate the impact of social networks and out-of-pocket expenses on the decision to migrate during a period (similar to the current period) when both social networks and out-of-pocket expenses are increasing.

Three primary predictions are derived from the comparative static analysis of my model. First, an increase in smuggling fees intensifies intermediate self-selection by narrowing the range of individuals who are able to migrate.[6] Second, an increase in the U.S. wage, holding the Mexican wage constant, increases migration among high-skilled individuals. Third, an increase in the number of settled migrants in the United States who help newly arriving migrants raises the migration rate

---

[5]Skill levels in this paper refer to some intrinsic talents of the migrant which are often correlated with the migrant's educational attainment or the migrant's earnings potential.

[6]This finding is consistent with Chiquiar and Hanson (2005), Orrenius and Zavodny (2005), McKenzie and Rapoport (2007a), and Kaestner and Malamud (2010), but differs with the findings in Caponi (2006) that the highest and lowest educated persons migrate, as well as the findings in Borjas (1987), Ibarraran and Lubotsky (2007) and Fernandez-Huertas (2008) of negative self-selection. The latter two papers find intermediate self-selection from rural Mexican communities.

among low-skilled individuals.

The uniqueness of the theoretical model is its ability to differentiate decisions to migrate during periods when both out-of-pocket expenses and the Mexican immigrant population increases. However, the theoretical model is not able to determine which of these two factors dominate. The migrant's decision is then tested empirically using survey-based data using the significant increase in smuggling costs to estimate whether migrant self-selection has changed over time. Two different migration-cost regimes (before and after 1993) are estimated using a structural break method on a time series of enforcement hours. A liquidity constraint is exhibited in high migration-cost regime. Moreover, the empirical estimation of who chooses to migrate suggests negative self-selection from migrant-sending communities in Mexico during a period of relatively low smuggling fees, and intermediate self-selection during a period of relatively high smuggling fees.

It should be noted that this paper considers the evolution of the characteristics of unauthorized migrants over time from mostly rural migrant-sending communities. There is recent evidence in Ferandez-Huertas (2009) that suggests a difference in the selection characteristics of Mexican migrants from rural and urban communities. However, some of this difference may be due an inability of the Mexican National Employment survey data to distinguish between unauthorized and authorized migrants to the United States. This may contribute to some of the differences among the regions in Mexico. This paper focuses on the theoretical and empirical implications of increasing smuggling costs on the self-selection characteristics of unauthorized migrants.

The structure of this paper is as follows. Section 2 describes the potential migrant's decision to migrate, establishes the lower- and upper-bound thresholds for migration, and examines the impact of a change in the smuggling cost, the wage ratio, and the social network on these thresholds. Section 3 identifies two different migration cost regimes with a structural break test of the time series data on border enforcement intensity, proposes a method for controlling non-return migrant bias in the MMP data, and tests the implications of the theoretical model by estimating the propensity to migrate in different migration cost regimes. Section 4 summarizes the findings.

# 2 Theoretical Model

## 2.1 Basic Setup

The decision to migrate can be described by a discrete, two time-period model. This model includes two countries, the United States and Mexico, divided by an international border. The cost of migration is assumed to be exogenously determined by the U.S. government and its choice of border enforcement intensity. The potential migrant possesses a skill level and decides the level of consumption, the level of savings in the first period and whether or not they migrate to the United States in the second period. The potential migrant is restricted from borrowing the cost of migration with a formal bank loan against future earnings. The decision is therefore subject to the liquidity constraint. However, the borrowing constraint is partially mitigated if the potential migrant has social networks in the United States from whom the cost of migration can be borrowed. All individuals in the model are assumed to be able to find employment in either location and the second period's wages are known with certainty.

## 2.2 The Migrant's Decision

The potential migrant maximizes his consumption, savings, and decides whether or not to migrate to the United States. The potential migrant faces the following two-period well-behaved utility function with non-satiation and diminishing returns to consumption:

$$U(C_t) + \beta U(C_{t+1}) \tag{1}$$

where $C_t$ is consumption in period t and $\beta$ is the discount rate. The budget constraint for the potential migrant in period 1:

$$C_t + a_{t+1} + \Phi(1 - \eta)\Psi_t = \alpha W_t^{mex} \tag{2}$$

where $a_t + 1$ is the assets saved for period t+1, $\Phi$ is an indicator variable denoting the person's choice to migrate with the variable equal to 1 if the person migrates from Mexico to the United States and equal to 0 if the person does not migrate. The parameter $\eta$ is the fraction of the fee that can be borrowed from family and friends in the United States and is assumed to be related to the strength of the potential migrant's social network in the United States, $\Psi_t$ is the cost of

crossing the international border at the end of period t. $W_t^{mex}$ is the wage earned in Mexico and $\alpha$ is the skill measure of the potential migrant. During the second period, the budget constraint for the potential migrant is the following:

$$C_{t+1} + \Phi\eta\Psi_t = (1+r)a_{t+1} + (1-\Phi)\alpha W_{t+1}^{mex} + \Phi\Gamma^u(\alpha)W_{t+1}^{us} \tag{3}$$

where r is the return on assets, $W_{t+1}^{us}$ is the wage earned in the United States, and $\Gamma^u(.)$ is the foreign skill transferability function, described hereafter.

Figure 1: PRODUCTIVITY AS A FUNCTION OF SKILL
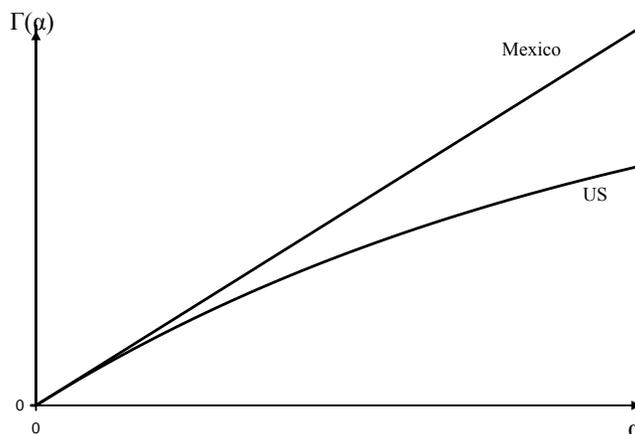


**Note:** This graph characterizes the transferability of skill level into productivity in the workforce in the United States and Mexico. Skill levels for potential migrants in Mexico translate one-for-one to a productivity measure in Mexico whereas skill translates less than one-for-one into the United States' labor market.

The function $\Gamma^u(.)$ translates foreign skill levels into skill levels in the U.S. labor markets. The function is assumed to be continuous, everywhere differentiable, and an increasing function of skill-type with less than a one-for-one transfer from Mexico to the United States and with diminishing returns:

$$\Gamma^{u'}(.) \leq 1 \qquad \Gamma^{u''}(.) < 0 \qquad \Gamma^u(0) = 0 \qquad \Gamma^{u'}(0) = 1 \tag{4}$$

The implication of the functional form assumed in equation 4 is that higher-skilled potential migrants pay a higher relative cost in skills when they migrate. For example, migrants in the agri-

cultural or construction sectors with relatively low skill levels convert their skills more readily in the United States than migrants with relatively high skill levels in the medical or legal sectors. Occupations in these sectors would require more education or re-accreditation to use these skills in the U.S. labor market.[7]

The potential migrant maximizes utility in equation 1 with respect to the level of consumption in each period, the level of savings in the first period, the decision to migrate (subject to the budget constraints in equation 2 and equation 3), and the liquidity constraint $a_t \geq 0$.

The constrained optimization problem can be solved with a Lagrangian:

$$
\begin{aligned}
L = \ & U(C_t) + \beta U(C_{t+1}) \\
& + \lambda_t \Big( \alpha W_t^{mex} - C_t - a_{t+1} - \Phi(1-\eta)\Psi_t \Big) \\
& + \lambda_{t+1} \Big( (1+r)a_{t+1} + (1-\Phi)\alpha W_{t+1}^{mex} + \Phi\Gamma^u(\alpha)W_{t+1}^{us} \\
& - C_{t+1} - \Phi\eta\Psi_t \Big) + \mu_{t+1}a_{t+1}
\end{aligned}
\tag{5}
$$

The Kuhn-Tucker conditions necessary for an optimum are the following:

$$
U'(C_t) - \lambda_t = 0
\tag{6}
$$

$$
\beta U'(C_{t+1}) - \lambda_{t+1} = 0
\tag{7}
$$

$$
\lambda_t - \lambda_{t+1}(1+r) = \mu_{t+1}
\tag{8}
$$

The complementary slackness conditions are:

$$
\lambda_t \geq 0, \quad \lambda_{t+1} \geq 0,
$$

$$
\mu_{t+1} \geq 0, \quad a_{t+1}\mu_{t+1} = 0
$$

---

[7]The non-linear restriction in this model is not necessary to predict intermediate self-selection since a linear transformation of less than one-for-one would discourage a range of high-income earners. However, the choice of a non-linear transfer of skills is more clear in the graphical depictions of the theoretical predictions. Nevertheless, other skills transformation functions could be considered with different return curves for low and high-skilled individuals.

Assets are completely consumed in the second period, so $a_{t+2} = 0$. If a potential migrant decides to migrate, his consumption in the second period must be greater than his consumption in the first period. Therefore, $\mu_{t+1}$ will be positive and $a_{t+1}$ equals zero to satisfy the complementary slackness condition. The equation $\lambda_t \geq \lambda_{t+1}(1+r)$ must hold, such that when $\mu_{t+1}$ is positive, the equation is strictly greater, and when $\mu_{t+1}$ is zero, the equation is equal.

An additional condition must be satisfied if the potential migrant chooses to migrate:

$$\lambda_{t+1}\left[\Gamma^u(\alpha)W_{t+1}^{us} - \alpha W_{t+1}^{mex} - \eta\Psi_t\right] \geq \lambda_t(1-\eta)\Psi_t \tag{9}$$

Combining equations 6, 7 and 9, a potential migrant's decision to migrate is a function of their net earnings in the United States and the cost of migration.

$$\frac{U'(C_{t+1})\beta}{U'(C_t)}\left[\Gamma^u(\alpha)W_{t+1}^{us} - \alpha W_{t+1}^{mex} - \eta\Psi_t\right] \geq (1-\eta)\Psi_t \tag{10}$$

The net US earnings is defined as the wage earned in the United States less the foregone wages in Mexico and any part of the fee that the migrant borrowed from his social networks in the first period to be able to migrate.

## 2.3 Implications of the Model

The model generates five implications about migration.[8]

### 2.3.1 Incentive to Migrate

First, there is an incentive to migrate to the United States given that for some level of skill, $\alpha^*$, in equation 10, the discounted wages in the United States in the next period–less the discounted foregone wages in Mexico and the remainder of the cost to be paid in the second period–is equal to the cost of crossing the border that is required to be paid in the first period. For any $\alpha_i \geq \alpha^*$, the benefit of migrating exceeds the cost. This is the condition required for migration in the model and corresponds to the basic decision made by economic migrants.

---

[8]The analysis examines the decision of the potential migrant in a partial equilibrium context. The impact that immigrants have on wages in the United States is not considered.

### 2.3.2 Thresholds for Migration

Second, there are lower-bound and upper-bound thresholds for migration. A potential migrant in Mexico with a low skill level has the incentive to migrate to earn higher wages in the United States and would satisfy the condition previously described in equation 10. The maximum amount able to be saved by the potential migrant in the first period is equal to $\alpha W_t^{mex} - C_t$. Initial savings are assumed to be equal to zero. Then, since the second condition of migration must be satisfied, $a_t \geq \Psi_t$, such that the level of savings are greater than or equal to the migration cost required to be paid in the first period, there is a lower-bound threshold for skill levels less than or equal to $\underline{\alpha}$. Any skill level below $\underline{\alpha}$ is a skill level in Mexico at which the amount able to be saved is less than the cost of crossing the border.

$$\alpha W_t^{mex} - C_t = (1 - \eta)\Psi_t \tag{11}$$

Figure 2: UPPER- AND LOWER-BOUND THRESHOLDS
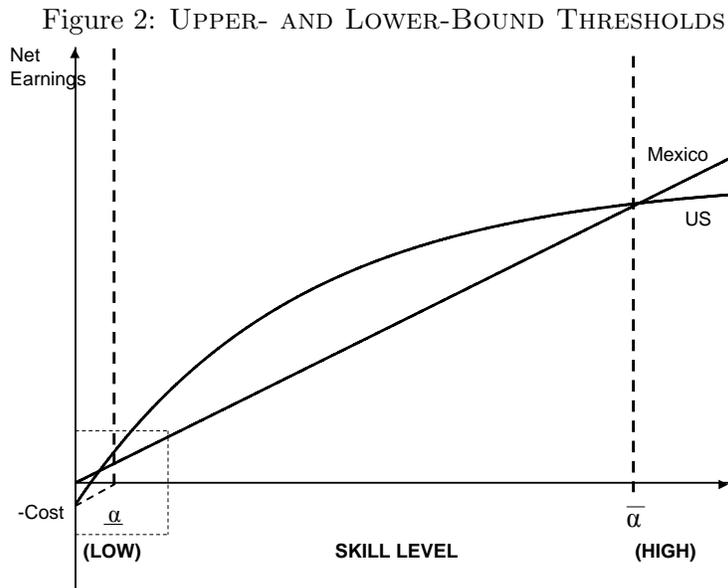


**Note:** This graph depicts the lower-bound and the upper-bound thresholds for migration. The upper-bound threshold is where the earnings in Mexico equal the net earnings in the United States. The dashed box in the lower left of the graph is expanded in Figure 3.

There is also an upper-bound threshold, in which the skill level of the potential migrant in Mexico is high and corresponds to a relatively low productivity level in the United States. This would result in the earnings of the potential migrant in Mexico exceeding earnings in the United

States less the cost of migration. Figure 3 expands the area in the dashed box in the lower left corner of Figure 2. The shaded region in Figure 3 represents skill levels for which the potential migrant's earnings in the United States less the cost of migration would exceed the potential migrant's earnings in Mexico. Thus the first condition of migration would be satisfied. However, the potential migrant is liquidity-constrained since his earnings in Mexico in the first period are not enough to pay the migration cost.

By rearranging equation 10 and normalizing the equation by the wage in Mexico, the migration decision requires the following equation to hold:

Figure 3: Liquidity Constraint and Lower-Bound Thresholds
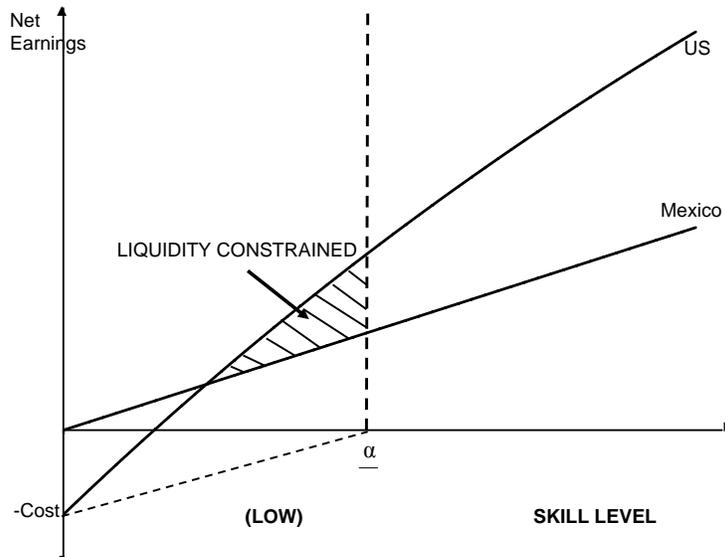


**Note:** This graph is an expanded version of the dashed box in the lower left-hand corner of Figure 2. The shaded region of the graph depicts the liquidity constrained potential migrants.

$$\left(\frac{W_{t+1}^{us}}{W_{t+1}^{mex}}\right)\Gamma^u(\alpha) - \alpha - \frac{\eta\Psi_t}{W_{t+1}^{mex}} \geq \frac{(1-\eta)\Psi_t}{W_{t+1}^{mex}}\frac{U'(C_t)}{U'(C_{t+1})\beta} \tag{12}$$

Equation 13 then transforms equation 12 by setting the wage ratio between the United States and Mexico at some constant B and combining the costs of migration variables. For the purposes of depicting the thresholds, social networks are assumed to be zero and the individual is required to pay the entire cost in the first period. Figure 2 depicts the upper-bound threshold using equation 13. The upper-bound threshold is $\alpha = \overline{\alpha}$ such that the following equation is satisfied:

11

$$B * \Gamma^u(\overline{\alpha}) - \overline{\alpha} = \frac{\Psi_t}{W_{t+1}^{mex}} \frac{U'(C_t)}{U'(C_{t+1})\beta} \qquad (13)$$

For any $\alpha$ such that $\alpha \geq \overline{\alpha}$, equation 13 would not be satisfied and the person would choose to stay in Mexico.

### 2.3.3 Wage Ratio and Migration

Figure 4: IMPACT OF INCREASED US WAGE ON MIGRATION
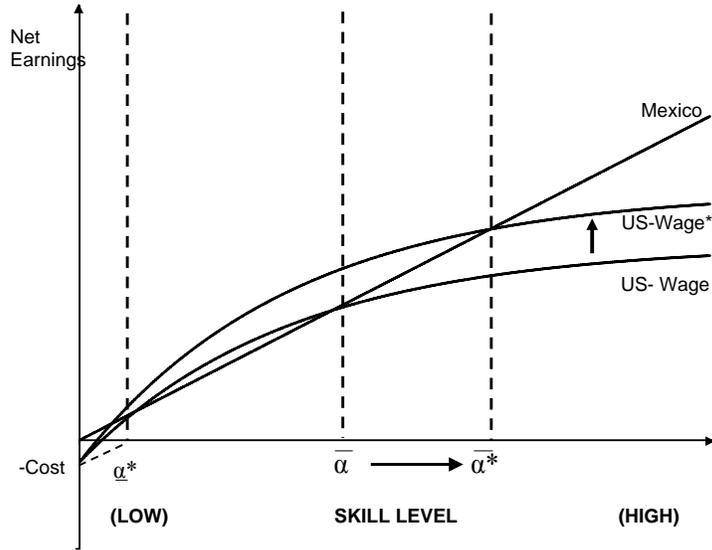


**Note:** This graph illustrates the effect on the upper-bound threshold of an increase in wage ratio and shifts the U.S. Wage curve up from $Wage$ to $Wage^*$, thus increasing $\overline{\alpha}$ to $\overline{\alpha}^*$

Third, an increase in the wage ratio between the United States and Mexico increases the number of migrants, but only those individuals constrained by the upper bound threshold. Figure 4 demonstrates the effect of increasing the wage ratio from $US - Wage$ to $US - Wage^*$ through an increase in the wage in the United States holding the wage in Mexico constant. The incentive to migrate increases and therefore the upper-bound constraint increases, $\overline{\alpha}^* > \overline{\alpha}$, such that higher skill types are more likely to migrate. However, the constraint on the lower bound threshold still binds since the cost of migration is held constant. Therefore, an increase in the wage ratio has no impact on the lower-bound constraint, $\underline{\alpha}$. The implication of the model that an increase in the wage ratio corresponds to an increase in migration is an important aspect of migration from Mexico to

the United States. Empirically, changes in the U.S.-Mexico wage ration has affected migration. (Hanson and Spilimbergo (1999), Borger (2009)).

### 2.3.4 Cost and Migration

The fourth implication of the model is that an increase in the cost of migration decreases the number of potential migrants who choose to migrate at the lower- and upper-bound thresholds. For a given wage ratio and skill level in equation 13, an increase in the cost of migration decreases migration for skill types for both the lower- and upper-bound thresholds.

Figure 5: IMPACT OF INCREASED COST ON MIGRATION
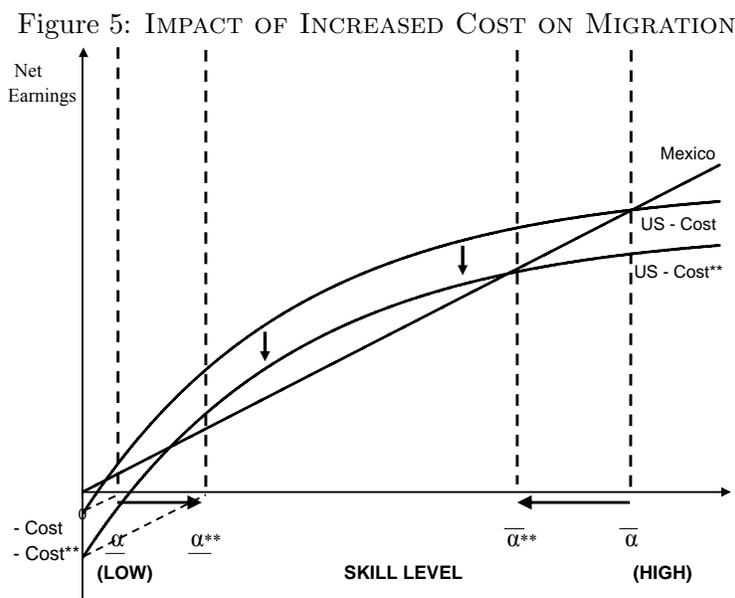


**Note:** The increase in the cost of migration constrains more low wage earners and reduces the incentive to migrate for more high wage earners. The increase in the migration cost concentrates migration among middle wage earners.

The increase in the migration cost raises the lower-bound threshold in equation 11, $\underline{\alpha}' > \underline{\alpha}$, meaning that lower-skilled individuals are less likely to migrate. Migration in a higher-cost environment requires the earnings potential of higher-skilled individuals to pay the cost in the first period. Figure 5 exhibits the increase in the lower-bound threshold from an increase in the cost, which shifts the U.S. net earnings curve down. Note that lower-bound is not where the U.S. net earnings cross the Mexican earnings. Because the earnings to pay the cost of migration must be made in the first period, the cost is paid by Mexican earnings and therefore a parallel line from the intercept term where the individual does not earn a wage is drawn. The interest rate is set at zero

for the purposes of depicting the lower-bound.

The increase in cost also decreases the upper-bound threshold in equation 13. For a given earnings level of a migrant, the increase in the cost of migration reduces the migrant's net earnings, creating an $\overline{\alpha}^{**} < \overline{\alpha}$, and the condition in equation 13 is satisfied for fewer potential migrants. Figure 5 illustrates the intensification of intermediate self-selection from an increase in the migration cost.

### 2.3.5  Social Networks and Migration

The final implication of the model is that an increase in social networks decreases the lower-bound threshold. Although there are other benefits of social networks for the migrant, the fact that this model captures the ability of social networks to partially relax the liquidity constraint faced by potential migrants is an important dynamic in understanding migration to the United States. In the model, stronger social networks imply a lesser likelihood that equation 11 binds for those with lower skill-levels. Figure 6 illustrates this effect on the lower-bound threshold.

Figure 6: IMPACT OF INCREASED SOCIAL NETWORKS ON MIGRATION
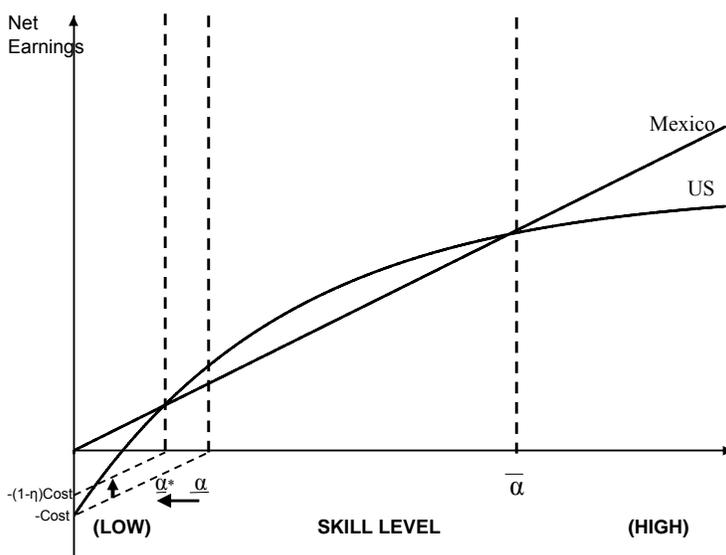


**Note:** This graph illustrates the effect on the lower-bound of an increase in social networks that allows the migrant to pay for part of the cost in the second period, decreasing $\underline{\alpha}$ to $\underline{\alpha}^{*}$. There is a slight effect (not shown) on the upper-bound due to the benefit of paying part of the cost in the second period.

There is a benefit to the migrant for paying part of the fee in the second period. Figure 6 abstracts from this benefit without the loss of generality. A characterization of the effect of paying

part of the fee in the second period can be found in the appendix. Moreover, if the interest rate charged by the social network is equal to the discount rate of the potential migrant, the net benefit would not impact the lower-bound threshold.

## 3  Empirical Analysis

### 3.1  Different Migration Cost Regimes

The previous section examined the theoretical impact of the increased cost and social networks on the thresholds for migration. The following describes how unauthorized migrations from Mexico to the United States can be separated into two migration cost periods due to a change in border enforcement intensity and the associated increase in the cost of clandestine entry.

The intensity of border enforcement along the Southwest border in the United States varies over time, and there has been a significant increase in the total number of hours U.S. Border Patrol agents spend patrolling the border ('linewatch hours') in recent years. The linewatch hours data along the U.S. Southwest border were recorded by U.S. Customs and Border Protection. The structural break estimation uses monthly linewatch hours from October 1976 to September 2004.

Figure 7: LINEWATCH HOURS (1977-2004)
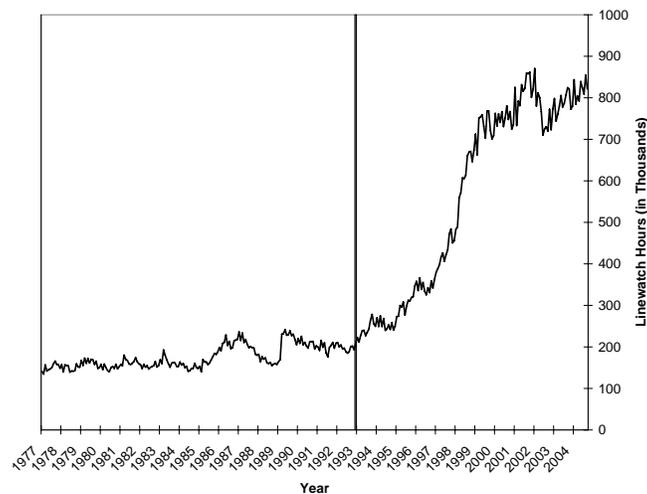


**Note:** Linewatch hours from 1976:10 to 2004:9 are the total number of hours agents patrolled the Southwest border. The vertical lines represent structural breaks in the data as estimated herein.

The timing of the shift in border enforcement intensity is determined by estimating the structural breaks in the linewatch hours data series using the multiple break model of Bai and Perron

15

Figure 8: Coyote Fees (1970-2006)



**Note:** The smuggling fees are calculated from the Mexican Migration Project dataset (MMP124), which estimates the average amount paid (in 2007 dollars) by migrants entering clandestinely into the United States in a given year. Adjustments of the fees in real terms use CPI inflation rates in the United States. Smuggling fees are almost always paid in U.S. dollars. The structural break in the border enforcement intensity for 1992:12 is highlighted with the solid vertical line.

(1998) and the sequential estimation approach of Bai (1997). [See appendix for details.] Although two breakdates are considered likely, the first (May 1985) is a minor change compared to the full sample period. The second likely breakdate (December 1992) corresponds to the period when new policies were enacted to increase enforcement across the Southwest border. Moreover, the increase in enforcement during the period after 1993 corresponds to an increase in the smuggling fee. Gathmann (2008) concludes that higher smuggling fees during this period are the result of both the enforcement effect, which increases the probability that a smuggler might be apprehended, and the diversion effect, in which border enforcement efforts force migrants to use more remote crossing locations. The latter effect increases fees to compensate for the longer time required to cross the border and the greater physical risk to the smuggler.

Figure 8 calculates the median real smuggling fee reported by first-time undocumented male migrants over time. The smuggling fees are estimated using the Mexican Migration Project (MMP) dataset, a long-term research project now based at Princeton University that has surveyed a large, geographically diverse, but not nationally representative set of migrant-sending communities in

16

Mexico. The surveys are primarily conducted in Mexico, which limits the number of recent observations since many non-returned migrants could not participate in the survey administered in Mexico. Note that the smuggling fees are reported in dollars since almost all of the transactions take place in dollars and the financing often comes from relatives receiving US wages. The change in smuggling fees and subsequently the change in the cost of migration is dramatic over the the entire sample from 1970-2006. The first breakdate in 1985–characterized by a small increase in linewatch hours and a subsequent decrease–had little impact on smuggling fees. However, the period after the breakdate in 1992 witnessed both an increase in linewatch hours and an increase in the smuggling cost.[9] This paper will contrast these two migration-cost regimes (pre- and post-1993) to test the impact of border enforcement intensification and the increased cost on migration.

## 3.2   Estimating the Propensity to Migrate

The intensity of border enforcement and the subsequent increase in the out-of-pocket migration cost has an impact on who migrates, how long they stay, and who chooses to remain in Mexico. The theoretical model provides some insight into the impact the increased cost has on who migrates. Acknowledging the limited data on unauthorized migrants, the following questions should nonetheless determine whether the two tenets of the theoretical model are supported empirically: (1) Is there a quantifiable liquidity constraint that prevents potential migrants who would otherwise benefit from migrating to the United States from paying the cost of migration? And (2) Do social networks assist in alleviating that liquidity constraint for potential migrants with limited resources?

### 3.2.1   Data

To test the implications of the model, data on the propensity to migrate would need to include the periods before and after 1993. However, such data over time is limited. Data that are nationally representative (such as that found in Ecuesta Nacional de la Dinmica Demogrfica (ENADID) or the Mexican Census) are confined to discrete periods and do not allow the econometrician to control for variations in border enforcement intensity. Data that identifies migrations from Mexico to the United States over time (such as that found in the Mexican Migration Project (MMP)), are not nationally representative and the data are subject to a selection bias from non-returned migrants.

---

[9]Gathmann (2008) and Roberts et al. (2010) estimate the elasticity of the smuggling cost and linewatch hours. Gathmann (2008) finds an one percent increase linewatch hours increases the smuggling cost by 0.3%. Roberts et al. (2010) find an on percent increase in linewatch hours increases the smuggling cost by 0.35%.

Therefore, the MMP data set requires a method to control for the returned migrant selection bias and this paper will propose such a method with an inverse mills ratio estimated from the ENADID data set. However, the estimation results provided by the MMP survey should be considered representative of only the migrant-sending regions in Mexico.

The MMP survey has been administered since 1982. Its sample is based primarily of interviews conducted in Mexico and supplemented by interviews of migrants in the United States. The survey provides migrant histories for the heads of households or for another member of the household if the head of the household has never migrated. The survey also provides demographic and socioeconomic data for individuals and households.

The ENADID survey, a nationally representative Mexican-based survey, was administered by the government of Mexico in 1992, 1997 and 2006. The survey records different socioeconomic characteristics and migration histories at the household level. The data also provide information on members of the household who had migrated in the previous five years, but who resided in the United States at the time of the survey.

In the theoretical model, potential migrants differ in their skill levels. These skills directly account for the different earnings of the individual in the first period, where the earnings are equal to the product of the skill level and the average wage. The skill-differentiated earnings in Mexico are directly related to the resources available to the potential migrant to pay the cost of migration from savings. In the empirical estimation approach, then, this paper uses household resources to characterize how migrants self-select.

A *Resource Index* variable provides the basis of comparison between individuals who choose to migrate and those who choose to remain in Mexico. The variable approximates the level of resources available to the potential migrant's household by constructing a de-trended index of durable good consumption. The index is calculated using a primary factor analysis of the household's reported ownership of durable goods and accounts for the increased likelihood over the sample period that households would report in the affirmative. McKenzie (2005) uses Mexican-based household surveys to demonstrate that this methodology provides a good approximation of household resource distributions. For the purposes of this paper, the absolute level of the resource index is not as important as the relative distribution of the resource index over the given group of communities. However, the number of durable items owned by a household increases over time. The resource in-

18

dex is accordingly de-trended to account for any variation of reported ownership during the sample period.[10]

### 3.2.2 Methodology

The characteristics of undocumented migrants are estimated from the propensity to migrate given different household resources, network capacities, demographic characteristics and the level of border enforcement. The analysis is restricted to the head of the household and excludes those with the legal documentation to enter the United States, since the analysis estimates the impact of smuggling cost on the propensity to migrate. Migrations must have occurred during the year of the survey or the previous year, increasing the likelihood that the resources available to the migrant at the time of migration are reflected by the estimated resource index. The empirical model allows for a non-linear relationship between migration and the potential migrant's resources (e.g., Affluent migrants are able to pay the smuggling fee, but the most affluent potential migrants remain in Mexico since their earning potential in Mexico exceeds their earning potential in the United States, less the cost of migration.) From the prediction in the theoretical model, I expect the relationship to exhibit an inverse U-shaped migration rate over the resource index distribution. An additional prediction of the theoretical model is that social networks in the United States alleviate the liquidity constraint for potential migrants with limited resources. Accordingly, social networks would have a greater impact on low-resource potential migrants than on high-resource potential migrants.

The baseline estimate of the propensity to migrate uses a linear probability model represented by the following equation:[11]

$$M_i = \alpha + \beta_1 I_i + \beta_2 I_i^2 + \beta_3 \delta_i + \beta_4 (I * \delta) + \gamma \chi_{\mathbf{i}} \tag{14}$$

where $M_i$ is whether the individual migrated during the current or previous year, $I$ is the level of the potential migrant's household resource index, $\delta$ is the social network available to the migrant in the United States, $I * \delta$ is an interaction effect variable that considers whether there is a different effect of social networks on different levels of resources, and *chi* are the demographic, border enforcement, inverse mills ratio and yearly control variables. The control for border enforcement uses the log of

---

[10]McKenzie and Rapoport (2007b) use this variable. However, the variation in household goods ownership over time is not accounted for in their constructed variable.

[11]Probit and logit models were also estimated and did not differ significantly from the predictions in the linear probability model.

sector level linewatch hours, or the number of hours the U.S. Border Patrol report monitoring the border in the sector where the migrant crossed the border. The linewatch hours were recorded as the linewatch hours for the San Diego sector for non-migrants in the sample.

The benefits provided by social networks to the migrant are both observed and indirect. Historical rates of migration have been used to approximate future migrations, since communities with social networks in the United States have access to information, resources, and support when the potential migrant arrives in the United States. This provides an indirect measure of the likelihood that a migrant has family members or friends in the United States able to assist with the cost of clandestine entry. Traditionally, this likelihood is assumed to be related to the historical rates of migration between the United States and different states in Mexico. Woodruff and Zenteno (2007) construct the instrument for social networks using the migration rates for different Mexican states during 1924 and the period 1954-1959. McKenzie and Rapoport (2007b) use these migration rates to demonstrate the impact of social networks on the propensity to migrate. However, this measure better characterizes the intangible 'social capital' of social networks than it does the tangible out-of-pocket expenses provided by family members in the United States. Therefore, in addition to historical migration rates, I also consider whether a person reports having a sibling or parent currently living in the United States. This increases the likelihood that a potential migrant has someone in the United States able to assist in the cost of clandestine entry or reduce the cost of migration by providing shelter, employment opportunities, or transportation once the migrant is in the United States.

The propensity to migrate during the previous year is estimated with and without time-fixed effects. There are many factors that contribute to the probability that a potential migrant will migrate in a given year, including business cycle conditions in the migrant's country-of-origin and in the destination country. The time-fixed effect controls for any variation due to changing economic and political trends unrelated to border enforcement intensity. The propensity to migrate also controls for the age of the potential migrant, and clustered standard errors are estimated for each community.

Due to the different impacts that cost has on who is able to migrate, equation 14 is modified to allow for a possible structural break in the parameters between the low-cost and high-cost period. This break is estimated using a dummy variable for migrations that occurred before and after 1993.

The following equation is estimated:

$$M_{i,t} = \alpha + \beta_{1,i,t}I + \beta_{2,i,t}I^2 + \beta_3\delta + \beta_{4,i,t}(I * \delta) + \gamma\chi \tag{15}$$

where t = 1,2 based on whether the migration would occur in the period between 1982-1992 or in the period between 1993-2004.

### 3.2.3 Returned Migrant Selection Bias

The MMP survey provides information on migration behavior over a long period of time, which enables the estimation process to control for the intensity of border enforcement. However, since most interviews in the MMP survey are conducted in Mexico, migration information is mostly derived from returned migrants. This introduces a bias into the data since it over-samples returned migrants and under-samples migrants who reside in the United States. The returned migrant selection bias is illustrated using the percent of non-returned migrants in the MMP survey and the ENADID survey for migrations that occurred in the previous or current year. The ENADID survey provides information on non-returnees by recording information from household members in Mexico who recently migrated to the United States but have yet to return.

Table 1: Non-Returned Migrants in Surveys

|  | ENADID | | MMP | |
|  | 1992 | 2006 | Pre-1993 | 1993-2006 |
| --- | --- | --- | --- | --- |
| Non-Returned Migrants in Sample | 61.9% | 82.4% | 12.3% | 21.3% |

**Note:** Migrants in this sample are considered non-returned if the migrant has not returned to Mexico within following year after the migration occurred. The returned migrant selection bias in the MMP survey results from only a fraction of the total number who migrated in the previous year being in the survey sample.

Table 1 depicts a much lower rate of non-returned migrants in the MMP sample during both periods than in the ENADID sample. Migrants are not returning frequently enough in the MMP sample to represent the full sample of migrants. Moreover, if there is a systematic difference in the resources or social networks of migrants who remain in the United States, the estimates from the MMP data will be biased. However, since the MMP survey data provide more information on direct social networks and migration patterns over a longer time horizon than the ENADID survey,

21

I correct for the returned migrant selection bias in the MMP sample by using information gathered from the non-returned migrants in the ENADID survey.

To control for the returned migrant selection bias, I use the procedure developed in Heckman (1979) to estimate a probit model to predict whether a migrant returns to Mexico by the following year in the ENADID surveys. This estimate is then used to construct an Inverse Mills Ratio (IMR) for different quintiles of the resource index and network distributions in the MMP data. In addition to the resource index variables and the network variables, the IMR is also estimated with a variable of whether the potential migrant is married. A married potential migrant is not more likely to migrate to the United States than a non-married potential migrant. However, the probability that a married person who has migrated to the United States will return to Mexico is positive and statistically significant. The marriage variable, therefore, increases the predictive content of the IMR variable in controlling for the bias.

## 3.3   Results

The parameter estimates for MMP data are reported in Table 2.[12] An inverse U-shaped relationship is evidenced between the household resource index and the propensity to migrate, with a positive and statistically significant coefficient on $Index$ and a negative, statistically significant coefficient on the $Index^2$ term over the full sample period. This is consistent with the findings in Mckenzie and Rapoport (2007b) of intermediate self-selection of migrants from Mexico to the United States. However, when I control for the non-migrant return bias in column (1b) of Table 2, the network variable used by Mckenzie and Rapoport (2007b) does not have a statistically significant impact on the propensity to migrate. A similar result, not shown, is found when using the historical migration rate in 1924 as the proxy for social networks.

The impact of parents and siblings in the United States is shown in column (2a) of Table 2. This measure of social networks has a statistically significant impact on the propensity to migrate. Each additional family member in the United States increases the propensity to migrate by 3.9 percentage points at the average resource index level, and a family member increases the migration rate more than 9 percentage points at the lowest resource levels. The first column does not adjust for the returned migrant bias, while the second column includes the Inverse Mills Ratio (IMR)

---

[12]The parameters are also estimated with the ENADID data set despite the survey only representing discrete time periods. The estimates are reported in the Appendix.

variable, as previously described, to control for the non-migrant return bias. The coefficients on the proxy for network density are statistically insignificant when the non-migrant bias is controlled.

The direct measure of social networks is estimated by the model in column (2a) by including whether the potential migrant has parents or siblings currently living in the United States. This measure of social networks has a statistically significant impact on the propensity to migrate. Each additional family member in the United States increases the propensity to migrate by 3.9 percentage points at the average resource index level, and a family member increases the migration rate more than 9 percentage points at the lowest resource levels.

Figure 9: PREDICTED MIGRATION RATE WITH DIFFERENT NETWORK DENSITIES
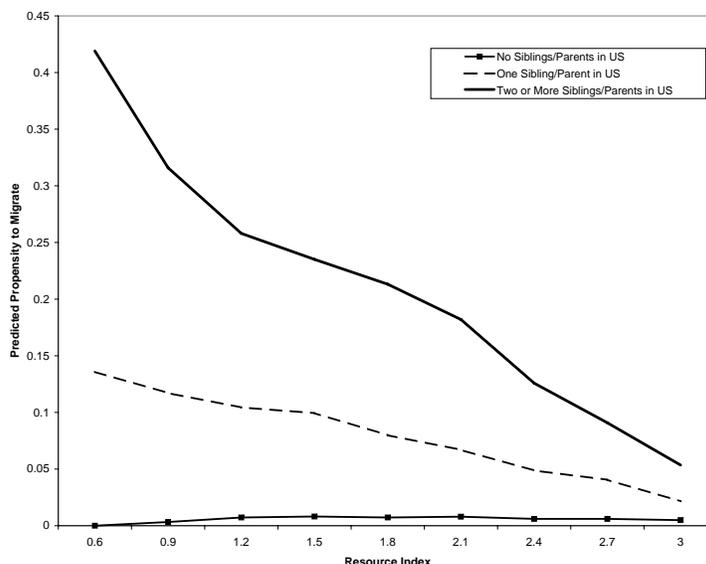


**Note:** This graph depicts the predicted propensity to migrate at different levels of the direct measure of social networks. The solid line is the propensity to migrate for potential migrants with two or more siblings or parents in the United States. The dashed line is the propensity to migrate for potential migrants with one sibling or parent in the United States. The squared line is the propensity to migrate for potential migrants without any siblings or parents in the United States.

Figure 9 illustrates the propensity to migrate for different levels of social networks in the United States. Although it is logical to assume that the presence of family members in the United States would increase the propensity to migrate, the actual effect family members have on migration varies markedly according to the family's resource levels. Suppose that instead of an economic motivation for migration, which is assumed in the theoretical model, the primary motive for migration is family reunification. Then the potential migrants with family members in the United States would likely

Table 2: Propensity to Migrate, Networks and Border Enforcement Intensity

| Propensity of head of household to migrate in previous year or the year of the survey | | | | |
|---|---|---|---|---|
| | (1a) | (1b) | (2a) | (2b) |
| $Index$ | $0.018^{***}$ | $0.022^{***}$ | $0.017^{***}$ | $-0.024^{**}$ |
| | (0.005) | (0.005) | (0.005) | (0.011) |
| $Index^2$ | $-0.004^{***}$ | $-0.008^{***}$ | $-0.006^{***}$ | $0.007^{**}$ |
| | (0.001) | (0.002) | (0.002) | (0.003) |
| $Network\ Density$ | $0.800^{***}$ | 0.293 | | 0.286 |
| | (0.301) | (0.328) | | (0.284) |
| $Index * Network$ | $-0.202^{**}$ | 0.069 | | -0.002 |
| | (0.096) | (0.136) | | (0.112) |
| $ln(linewatch)$ | $-0.079^{***}$ | $-0.076^{***}$ | $-0.053^{***}$ | -0.007 |
| | (0.021) | (0.021) | (0.020) | (0.005) |
| $Age$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ |
| | (0.0001) | | (0.0001) | (0.0001) |
| $Family\ in\ US$ | | | $0.091^{***}$ | $0.092^{***}$ |
| | | | (0.020) | (0.020) |
| $Index * Family$ | | | $-0.027^{***}$ | $-0.027^{***}$ |
| | | | (0.008) | (0.008) |
| $Index * I_{pre-93}$ * | | | | $0.074^{***}$ |
| | | | | (0.018) |
| $Index^2 * I_{pre-93}$ | | | | $-0.031^{***}$ |
| | | | | (0.008) |
| $Index * Net * I_{pre_93}$ | | | | $0.427^{***}$ |
| | | | | (0.153) |
| $Index * Family * I_{pre_93}$ | | | | -0.001 |
| | | | | (0.004) |
| | | | | |
| Social Network Variable | A | A | B | A/B |
| Year Dummy Variable | Yes | Yes | Yes | No |
| IMR Variable | Yes | Yes | Yes | Yes |
| Observations | 20489 | 20489 | 20489 | 20489 |

**Note:** Clustered standard errors are in parentheses at the community level. $^{***}$ significant at the 1% level; $^{**}$ significant at the 5% level; $^*$ significant at the 10% level. The data are from migrations that occurred between 1982, 1987 to2004 with the latter date limited by the linewatch data series. The total number of communities in the MMP data is 102 through 2004. Social network variable A is the migration rate between the years 1955-1959 for the state in which the migrant resides. Social network variable B is the number of migrant's siblings or parents currently living in the United States and is represented as $Family\ in\ US$ in the table. The Inverse Mills Ratio (IMR) is included in the regressions to control for returned migrant selection bias. $Index$ is the resource index constructed from reported household durable consumption. $Index^2$ allows for a non-linear relationship between resources and migration. $Network\ Density$ refers to the historical migration rate variable. $Index * Network$ is the interaction effect variable that allows the effect of historical migration rates on the propensity to migrate to vary by the resource level of the potential migrant. $ln(Linewatch)$ is the log-level of border enforcement linewatch hours. $Index * Family$ is the interaction effect variable that allows the effect of family members on the propensity to migrate to vary by the resource level of the potential migrant. $I_{pre93}$ is an indicator variable of whether the migration occurred before or after 1993.

increase migration rates, even with different resource levels. However, if social networks alleviate

barriers to migration, such as liquidity constraints, employment opportunities, or housing costs,

Figure 10: PREDICTED MIGRATION RATE WITH DIFFERENT MIGRATION COST REGIMES
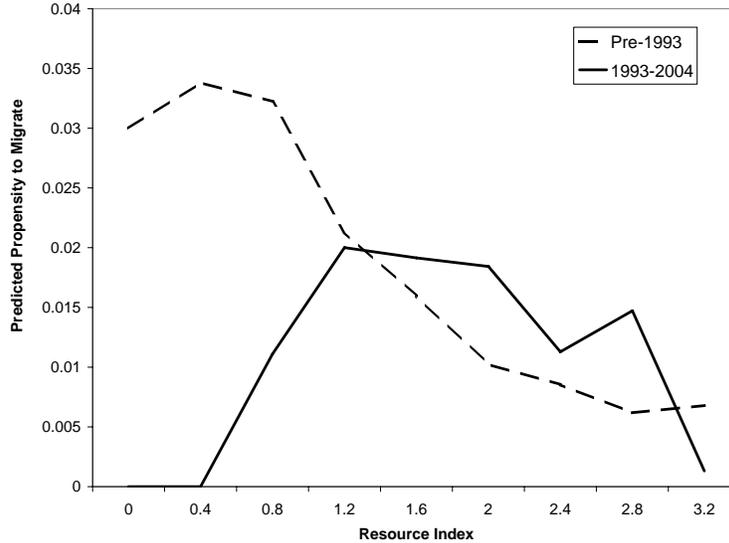


**Note:** The graph depicts the predicted probability at the median social network density using the 1955-1959 historical migration rates for different migration cost regimes. The dashed line is the period before 1993 when the smuggling cost was low relative to the current period. The solid line is the period after 1993 when the smuggling cost was high relative to the previous period. The increased cost intensified the intermediate self-selection of migrants, steepening the inverse U-shaped relationship between migration rates and the resource index. Moreover, the lower-bound migration threshold is binding in the later period, but not binding in the earlier period. This corresponds to the prediction in the theoretical model of who migrates with an increase in the cost.

then an increase in the propensity to migrate for limited-resource potential migrants would be observed. The propensity to migrate is highest among these limited-resource potential migrants with more than one family member in the United States. This is consistent with the assumptions in the theoretical model that potential migrants at low resource levels might not be able to afford the cost of migration even though their incomes would be higher in the United States.

Column (2b) of Table 2 tests whether there is a change in self-selection over time by allowing for a change in the U-shaped relationship between resources and the propensity to migrate during the low-cost and high-cost periods of migration. The effect that a family member in the United States had on the propensity to migrate did not differ between these two periods. However, the propensity to migrate among those surveyed did differ in their time-trend adjusted household resource level.

Figure 10 depicts the propensity to migrate in the earlier, low-cost period relative to the later, high-cost period of migration. The migration rate in the earlier period is characterized by negative self-selection and in the later period by intermediate self-selection. The significant shift from

limited-resource potential migrants to median-resource potential migrants suggests a change in who is able to migrate to the United States over time. This paper proposes that one possible reason for this shift is the liquidity constraint imposed on an increasing number of migrants by the increase in the smuggling cost.

# 4   Conclusion

This paper models the impact of liquidity constraints faced by potential unauthorized migrants when they evaluate their decision to migrate. The model's predictions are consistent with intermediate self-selection if the increased migration cost constrains lower-skilled individuals, and consistent with negative self-selection if social networks enable lower-skilled individuals to migrate.

Unauthorized migration from Mexico to the United States is divided into two migration cost periods. Before 1993, unauthorized migration was characterized by low border enforcement intensity and relatively low smuggling costs. In the period since 1993, unauthorized migration has been characterized by high border enforcement intensity and relatively high smuggling costs.

This paper proposes a method to control the returned migrant selection bias given the limited long-term data on unauthorized migration patterns from Mexico to the United States. It uses this data to test the theoretical model's predictions by estimating the propensity to migrate against the resources available to the household de-trended during the sample period. The estimated parameters suggest evidence of a shift in migrant selection characteristics from negative self-selection in the low-cost period before 1993 to intermediate self-selection in the high-cost period after 1993. Further, liquidity constraints appear to bind the least-skilled potential migrants in the current period.

This paper demonstrates that social networks play an important role in reducing the cost of migration. However, in addition to financial assistance for the cost of migration, social networks also reduce the cost of migration by providing assistance in finding employment.[13] Although the primary mechanism through which this paper considers the cost of migration is the constraint imposed by an increase in the smuggling fee, I leave to future research a possible secondary mechanism that explicitly models the probability of finding U.S. employment based primarily on the migrant's network density.

---

[13]See Espinosa and Massey (1997), Munshi (2003) and Waldinger and Lichter (2003).

# References

[1] Bai J (1997) "Estimating Multiple Breaks One at a Time." *Econometric Theory,* 13:315-352.

[2] Bai J and Perron P (1998) "Estimating and Testing Linear Models with Multiple Structural Changes." *Econometrica,* 66(1): 47-78.

[3] Borger SC (2009) "Estimates of the Cyclical Inflow of Undocumented Migrants to the United States." CCIS Working Paper Series. Working Paper 181.

[4] Borjas GJ (1987) "Self-Selection and the Earnings of Immigrants." *The American Economic Review,* 77(4): 531-553.

[5] Caponi V (2006) "Heterogeneous Human Capital and Migration: Who Migrates from Mexico to the U.S.?" IZA Discussion Paper No. 2446.

[6] Carrington WJ, Detragiache E, Vishwanath T (1996) "Migration with Endogenous Moving Costs." *The American Economic Review,* 84(4): 909-930.

[7] Chiquiar D and Hanson G (2005) "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy,* 113(2): 239-281.

[8] Clemens MA, Montenegro CE, Pritchett L (2008) "The Place Premium: Wage Differences for Identical Workers across the US Border," Policy Research Working Paper Series 4671, The World Bank, Washington, D.C.

[9] Cornelius WA (2001) "Death at the Border: Efficacy and Unintended Consequences of US Immigration Control Policy." *Population and Development Review,* 27(4): 661-685.

[10] Cornelius W, Fitzgerald D, Borger S, eds. (2009) *Four Generations of Nortenos: New Research from the Cradle of Mexican Migration.* La Jolla, CA and Boulder, CO: Center for Comparative Immigration Studies, UCSD, and Lynne Rienner Publications.

[11] Espinosa K, Massey D (1997) "Undocumented Migration and the Quantity and Quality of Social Capital." *Soziale Welt,* 12: 141-162.

[12] Ethier WJ (1986) "Illegal Immigration." *The American Economic Review,* 76(2): 258-262.

[13] Fernandez-Huertas J (forthcoming) "New Evidence on Emigration Selection." *Review of Economics and Statistics.*

[14] Fernandez-Huertas J (2009) "Wealth Constraints, Skill Prices, or Networks?: What Determines Emigrant Selection." Mimeo, Instituto de Analisis Economico.

[15] Gathmann G (2008) "Effects of Enforcement on Illegal Markets: Evidence from Migrant Smuggling along the Southwestern Border." *Journal of Public Economics,* 92: 1926-1941.

[16] Hansen BE (2001) "The New Econometrics of Structural Change: Dating Breaks in U.S. Labor Productivity." *Journal of Economic Perspectives,* 15(4): 117-128.

[17] Hanson GH, Spilimbergo A. (1999): "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border." *The American Economic Review,* 89(5): 1337-1357.

[18] Hatton TJ, Williamsonm JG (2005): *Global Migration and the World Economy.* Cambridge, MA: The MIT Press.

[19] Heckman JJ (1979): "Sample Selection Bias as a Specification Error." *Econometrica,* 47(1): 153-161.

[20] Ibararran P, Lubotsky, D (2007): "The Socioeconomic Status of Mexican Migrant Families: New Evidence from the 2000 Mexican Census." In George J. Borjas, ed. *Mexican Immigration to the United States.* Chicago: The University of Chicago Press.

[21] Kaestner R, Malamud O (2010): "Self-Selection and International Migration: New Evidence from Mexico." NBER Working Paper 15765.

[22] Massey DS, Goldring L, Durand J (1994) "Continuities in Transnational Migration: An Analysis of Nineteen Mexican Communities." *American Journal of Sociology,* 99(6): 1492-1533.

[23] Mckenzie D (2005) "Measuring Inequality with Asset Indicators." *Journal of Population Economics,* 18(2): 229-260.

[24] Mckenzie D, Rapoport H (2007a) "Self-Selection Patterns in Mexico-U.S. Migration: The Role of Migration Networks." World Bank Research Working Paper 4118.

[25] Mckenzie D, Rapoport H (2007b) "Network Effects and the Dynamics of Migration and Inequality: Theory and Evidence from Mexico." *Journal of Development Economics,* 84:1-24.

[26] Mexican Migration Project (2008) "Border Crossing Costs. 1980-2008." available at http://mmp.opr.princeton.edu/results/001costs-en.aspx.

[27] Munshi K (2003) "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market." *The Quarterly Journal of Economics,* 118(2): 549-599.

[28] Orrenius PM (1999) "The Role of Family Networks, Coyote Prices and the Rural Economy in Migration from Western Mexico: 1965-1994." Federal Reserve Bank of Dallas Research Department Working Paper 9910.

[29] Orrenius PM, Zavodny M (2005) "Self-Selection among Undocumented Immigrants from Mexico." *Journal of Development Economics,* 78(1): 215-240.

[30] Sjaastad LA (1962) "The Costs and Returns of Human Migration." *The Journal of Political Economy,* 70(5): 80-93.

[31] Stark O, Taylor JE (1991) "Migration Incentives, Migration Types: The Role of Relative Deprivation." *Economic Journal,* 101: 1163-1178.

[32] Roberts B, Hanson G, Cornwell D, Borger S (2010) "An Analysis of Smuggling Costs along the Southwest Border." Office of Immigration Statistics Working Paper Series.

[33] Roy AD (1951) "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers,* 3(2): 135-146.

[34] Waldinger R, Lichter M (2003) *How the Other Half Works: Immigration and the Social Organization of Labor.* Berkeley, CA: University of California Press.

[35] Woodruff C, Zenteno R (2007): "Migration Networks and Microenterprises in Mexico." *Journal of Development Economics,* 82: 509-528.

# A  Appendix

## A.1  Data

Data for linewatch hours were compiled originally by the U.S. Immigration and Naturalization Service and now are made available through U.S. Customs and Border Protection. The data from 1963:7 to 2004:9 are available at http://irps.ucsd.edu/faculty/ faculty-directory/gordon-hanson.htm on Gordon Hanson's webpage.

The *Resource Index* variable uses the data in the MMP survey on household ownership of certain durable items, such as televisions, radios, refrigerators, stoves, stereos, phones, etc. A principal factor component is then derived from these data and puts more weight on household goods that vary more in the sample. The index is then constructed by multiplying the factor weight and the indicator variable according to whether the household owns a particular durable item. The index is then de-trended to eliminate variation over time.

## A.2  Effect of Second Period Payment

There is some benefit to the migrant of paying the fee in the second period when their social network provides the financing for the migration cost in the first period. Since the effect of the migration cost does not include this benefit, the following formulation characterizes the difference between the net earnings for potential migrants with different levels of social networks. This benefit is equal to:

$$\frac{\Psi_t}{W_t^{mex}}\left[1 - \frac{U'(C_t)}{U'(C_{t+1})\beta}\right](\eta - \eta') \tag{A1}$$

As the discount rate of future earnings (the second term in the brackets) approaches unity, the impact of paying the second period is zero.

## A.3  Structural Break Test

The timing of the shift in border enforcement intensity is estimated in a simple dynamic model of border enforcement that follows an AR(1) process:

$$l_t = \alpha + \rho l_{t-1} + e_t \tag{A2}$$

where $l_t$ is the border enforcement log-level at time period $t$ and is related to the border enforcement level in the previous period by the parameter $\rho$, and $e_t$ is a time series of serially uncorrelated shocks. A structural break occurs if the parameters $\alpha$ and $\rho$ change at some date during the sample period. With the possibility that border enforcement activities can lag behind a policy change, the timing of the structural change in the parameters is not known with certainty *a priori*. Moreover, multiple changes in policy over the sample period could exhibit multiple structural breaks in the parameters. Therefore, the structural breaks are estimated using the multiple break model of Bai and Perron (1998) and the sequential estimation approach of Bai (1997).

Bai and Perron (1998) propose a method to test the null hypothesis that some number $l$ characterizes the number of breaks in the data against the alternative hypothesis that the data series is better described by $l+1$ breaks. For the current analysis, I determine whether the data are consistent with two structural breaks ($l = 2$) over the sample period. Using the F-test and critical values proposed in Bai and Perron (1998) at the 95% critical value, two structural breaks are determined to be more likely than one. Three structural breaks are not likely, based on a 95% critical value. Therefore, the timing of the two most likely structural breaks are estimated and uses the notation developed in Hansen (2001) to describe that estimation procedure.

The estimation of the breakdate uses the sum of the squared residuals to construct the residual variance and plots the residual variance at each possible breakdate in the sample time period. Both global and local minima of the residual variance are of interest to the estimation procedure, and the most likely structural breaks occur when the residual variance over a subperiod forms a strong v-shape. The sum of the squared residuals are estimated by the following:

$$S_T(k) = \sum_{t=1}^{k} (Y_t - \bar{Y}_k)^2 + \sum_{t=k+1}^{T} (Y_t - \bar{Y}_k^*)^2 \tag{A3}$$

where $\bar{Y}_k$ is the mean of the first $k$ observations, $\bar{Y}_k^*$ is the mean of the last $T - k$ observations, and the value of $\hat{k}$ is the estimated break point where $\hat{k} = argmin_k S_t(k)$. The residual variance is then calculated as $\hat{\tau} = \hat{k}/T$. The residual variance is estimated at each possible breakdate $k$ with the data divided into sample periods before and after date $k$. The global and local minima of the residual variances are treated as possible breakdates. The top-left panel in figure A1 depicts the residual variance for the whole sample period from 1977:1 to 2004:9, with possible break points at global minima in May 1985 and local minima in December 1992 and December 1994. The sample

Figure A1: Residual Variance Breakdate Estimation
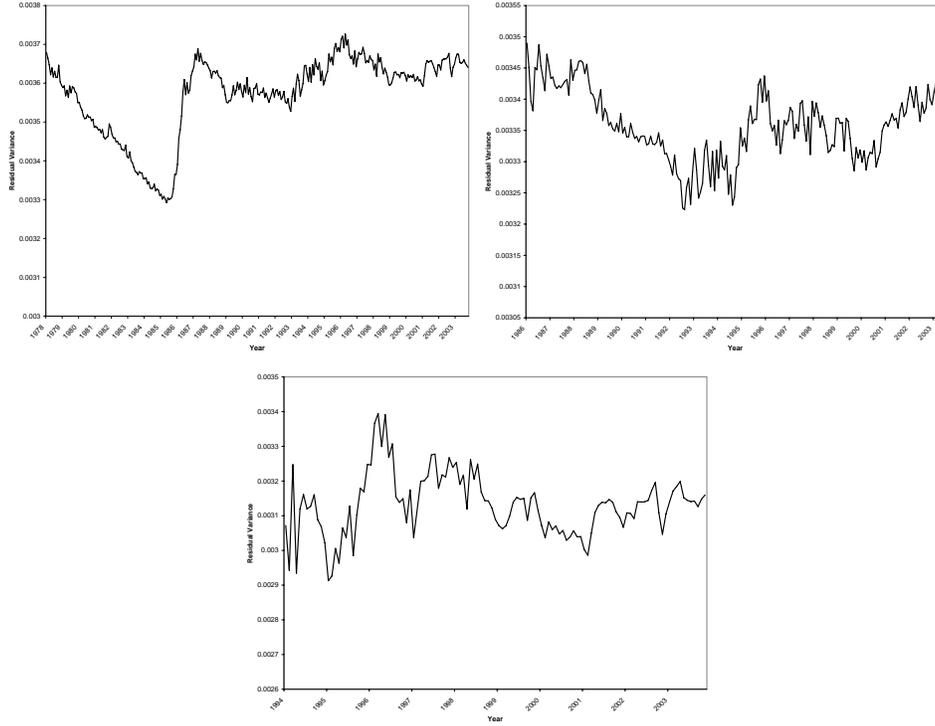


**Note: All panels:** The residual variance is the sum of the squared residuals for both sample periods, which are divided at each possible breakdate. **Top-Left panel:** The residual variance estimate for the whole sample period of 1977:1 to 2004:9. **Top-Right panel:** The residual variance estimate for the period 1985:6 to 2004:9, which is the sample after the possible break date of 1985:5. **Bottom panel:** The residual variance estimate for the period 1993:1 to 2004:9, which is the sample after the possible break date of 1992:12.

is trimmed to the period after the first break point from 1985:6 to 2004:9 in the top-right panel in figure A1. The likely structural break occurs during the period between December 1992 and December 1994, and the global minima occurs in December 1992. The sample is again trimmed to

Table A1: Structural Breaks in Linewatch Hours

| Estimation Period | $\rho$ | $\alpha$ | Breakdate |
|---|---|---|---|
| 1977:1 to 1985:5 | **0.461**\*\*\* | **6.445**\*\*\* | May 1985 |
| | (0.087) | (1.037) | |
| 1985:4 to 1992:12 | **0.867**\*\*\* | **1.616**\*\* | Dec 1992 |
| | (0.053) | (0.644) | |
| 1993:1 to 2004:9 | **0.978**\*\*\* | **0.305**\*\*\* | |
| | (0.010) | (0.131) | |

**Note:** Linewatch hours are estimated in log-levels. The autoregressive parameters are reported for each period, when the sample is divided by the estimated breakdates.

the period after the second possible break point from 1993:1 to 2004:9 in the bottom panel in figure A1. The residual variance estimate does not exhibit the strong v-shape, indicating an unlikely breakdate. Other samples were considered, but without evidence for additional structural breaks. Therefore, the two breakdates considered likely occurred in May 1985, during a fiscal year in which a significant number of new field agents were hired and trained. This was also followed a period of minimal variation in linewatch hours. The second likely breakdate occurred in December 1992, which corresponds to the period when new policies were put in place to increase enforcement in the El Paso sector. Enhanced enforcement continued along the Southwest border throughout the late 1990s.

### A.3.1   ENADID Estimation Results

As an alternative estimate to the MMP results, I test the theoretical model using the ENADID 1992 and ENADID 2006 survey data, which is more nationally representative, but limited to discrete periods of migrations. The 1992 survey is used to estimate the propensity to migrate during the period of low border enforcement intensity and low smuggling costs. The 2006 survey is used to estimate the propensity to migrate during the period of high border enforcement intensity and high smuggling costs. A difference between migration rates for potential migrants with varying resources and social network suport would provide evidence of liquidity constraint.

The ENADID survey is not subject to the returned migrant selection bias that was corrected for in the MMP survey data. However, these data present other difficulties. For example, migrations during the earlier period occurred either during or immediately after the 1991 recession in the United States, whereas migrations during the later period occurred at the end of an economic boom in the United States. If the impact of social networks differs at various points during the business cycle, then the estimates of the impact of social networks on resource-constrained potential migrants could be biased. In addition, these data do not directly measure members of the household who migrated more than five years ago and therefore cannot directly measure social networks.

The liquidity constraint on resource-constrained households is tested by estimating the distribution of household resources around the time of the migration with the primary factor approach as previously described. ENADID does not have a direct measure of social networks whose members migrated more than five years ago, so the estimate uses the historical migration rates.

Table A2: ENADID: Propensity to Migrate, Networks, and Border Enforcement Intensity

| Propensity to migrate in previous year or survey year | | | | |
|---|---|---|---|---|
| | (1a) | (1b) | (2a) | (2b) |
| | 1992 | 2006 | 1992 | 2006 |
| *Index* | **0.596***\*** | **0.094***\*** | **0.607***\*** | **0.098***\*** |
| | (0.074) | (0.009) | (0.074) | (0.009) |
| *Index*$^2$ | **−0.474***\*** | **−0.088***\*** | **−0.473***\*** | **−0.093***\*** |
| | (0.059) | (0.008) | (0.060) | (0.008) |
| *Network Density* | **1.072***\** | **0.817***\*** | **0.171***\** | **0.083***\*** |
| | (0.440) | (0.111) | (0.086) | (0.021) |
| *Index ∗ Network* | **1.696***\*** | **−0.533***\*** | **0.254***\** | **−0.051***\* |
| | (0.670) | (0.152) | (0.131) | (0.029) |
| *Constant* | **−0.151***\*** | **−0.010***\*** | **−0.149***\*** | **−0.006***\** |
| | (0.023) | (0.003) | (0.022) | (0.003) |
| | | | | |
| Migration Lag Variable | A | A | B | B |
| Observations | 43095 | 140296 | 43095 | 140296 |
| **Probability of Migrating** | | | | |
| Low Network ($10^{th}$ Pctl) | | | | |
|    Low-Resources | 0.0% | 0.3% | 0.6% | 0.8% |
|    Middle-Resources | 3.6% | 1.5% | 4.6% | 1.9% |
|    High-Resources | 0.7% | 0.0% | 2.0% | 0.0% |
| High Network ($90^{th}$ Pctl) | | | | |
|    Low-Resources | 5.8% | 2.3% | 4.8% | 2.0% |
|    Middle-Resources | 12.5% | 2.7% | 9.9% | 2.7% |
|    High-Resources | 11.1% | 1.0% | 8.4% | 0.4% |

**Note:** Robust standard errors in parentheses. *** significant at the 1% level; ** significant at the 5% level; * significant at the 10% level. Migration lag variable A is the average migration rate between 1955-1959 for the state in which the community resides and predicts the future rate of migration to the United States. Migration lag variable B is the migration rate in 1924 for the state in which the community resides.

The estimate for the propensity to migrate during these two migration-cost regimes is in Table A2. The rate of migration dropped significantly in the second period relative to the first period, with a similar inverse U-shaped relationship between resources and the probability of migration.

Figure A2 illustrates the propensity to migrate for the two different periods with different densities of social networks. In the earlier period, when the estimate was based on the probability of migration in 1991 and 1992, migration rates were higher for potential migrants from historically migrant-sending regions in Mexico, and social networks had a positive impact on higher-income migrants. In the later period, when the migration rate was estimated for 2005 and 2006, the probability of migrating decreased significantly as the resource index rose, and social networks had a positive impact for resource-constrained potential migrants. The different sign on the network-index interaction effect coefficient during the two different periods indicates that social networks
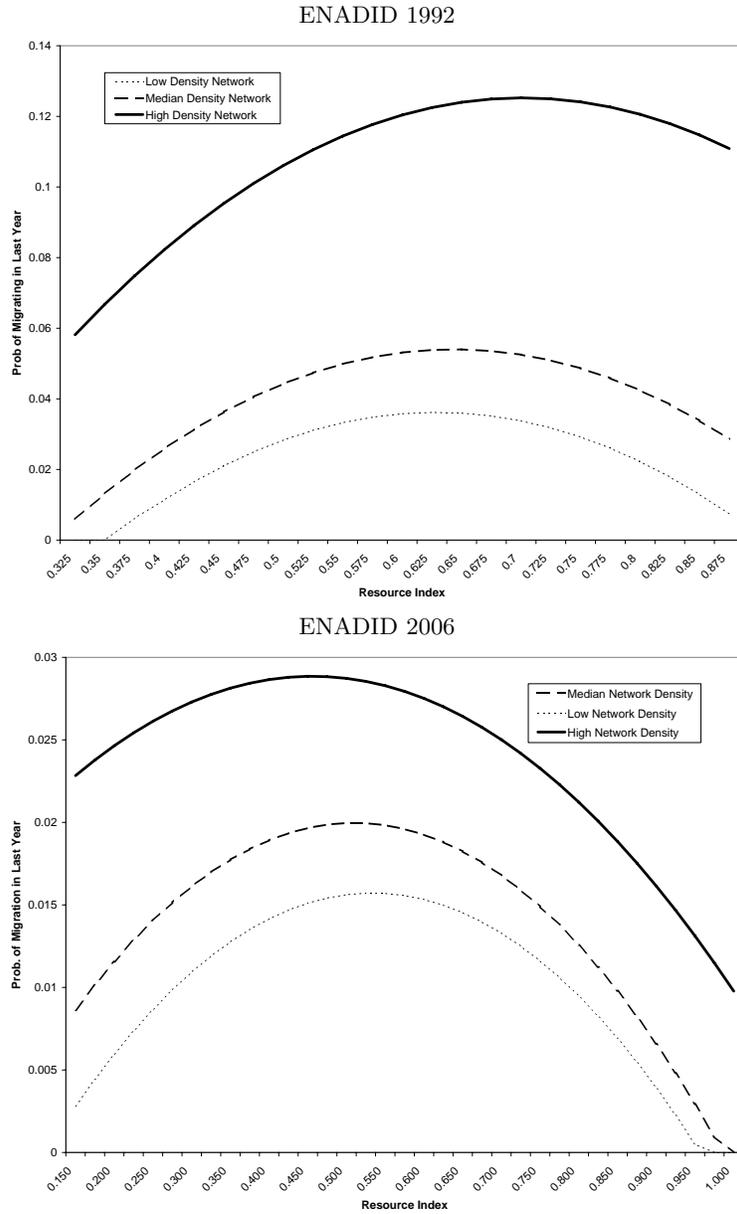
ENADID 1992

ENADID 2006

Figure A2: PROBABILITY OF MIGRATION IN DIFFERING PERIODS OF BORDER ENFORCEMENT
INTENSITY (1992, 2006)

**Note: All panels:** Estimated probability of migration during different periods of border enforcement intensity.
Reported in the figures are the network densities at the 10th, 50th and 90th percentiles, as measured by the average
rate of migration to the United States from a state in Mexico between 1955-1959. Solid line is high network density,
the dashed line is the median network density, and the dotted line is the low network density.

indeed assist liquidity-constrained potential migrants.