CEDI

Working Paper No. 06-13

The Influence of Stocks and Flows on Migrants' Location Choices

Thomas Bauer

Gil S. Epstein

Ira N. Gang

June 2006

Centre for Economic
Development & Institutions
Brunel University
West London
http://cedi.org.uk

CEDI DISCUSSION PAPER SERIES

The Influence of Stocks and Flows on Migrants' Location Choices *

Thomas Bauer
(RWI Essen, Ruhr-Universität Bochum, IZA, Bonn and CEPR, London)
thomas.bauer@ruhr-uni-bochum.de

Gil S. Epstein
(Bar Ilan University, Ramat Gan, IZA, Bonn and CReAM, London)
epsteig@mail.biu.ac.il

Ira N. Gang
(Department of Economics, Rutgers University)
gang@economics.rutgers.edu

This version: June 9, 2006

ABSTRACT

We examine the determinants of a current migrant's location choice emphasizing the relative importance and interaction of migrant stocks and flows. We show that both stocks and flow have significant impacts on the migrant's decision of where to locate. The significance and size of the effects vary according to legal status and whether the migrant is a "new" or a "repeat" migrant.

Keywords: flow, stock, migrants, immigration, location choice
JEL classification: F22, J61

Correspondence: Gil S. Epstein, Department of Economics, Bar-Ilan University, Ramat-Gan 52900 Israel.

*We thank Julie Phillips for discussions and helping us with the data. We thank David Jaeger, John Landon-Lane, Jackline Wahba, Yoshi Yonezawa and Myeong-Su Yun for their comments. This paper was partly written while Gil Epstein and Ira Gang were visiting IZA, Bonn. Ira Gang thanks the Rutgers Research Council for partial support. Earlier versions of this paper were presented at the NEUDC Conference, Ithaca, NY; IZA Transatlantic Labor Economics Workshop, Ammersee, Germany; ESRC Workshop, Trade, Migration and Labour Markets, University of East Anglia, Norwich; Centre for Economic Development and Institutions (CEDI) Inaugural Conference, Brunel University; as well as at several university workshops.

1. INTRODUCTION

A characteristic of international migration is the clustering of immigrants in ethnic communities. Prominent examples are the concentration of Turks in Germany, Tamils in Switzerland, Moroccans in the Netherlands and Belgium, Italians in Argentina, Greeks in Australia, and Ukrainians in Canada. Clustering may be very narrow, such as when immigrants from a town or region are concentrated in a specific foreign town or region. For example, Macedonians from Skopje have come to make up a notable part of the population of Gothenburg, Sweden. In the United States, noticeable clusters of Mexican immigrants exist in California, Texas, Florida and Chicago. Fifty-eight percent of migrants from Guanajuato, the Mexican state with the highest emigration rate to the US, go to California and another 23 percent to Texas.

The prevailing explanation for immigrant clusters is the existence of beneficial network externalities when previous immigrants provide shelter and work, assistance in obtaining credit, and/or generally reduce the stress of relocating to a foreign culture (see Gottlieb, 1987; Grossman, 1989; Marks, 1989; Church and King, 1993; Carrington, Detragiache, and Vishwanath, 1996; Chiswick and Miller, 1996; Zahniser, 1999; Munshi, 2003). Ethnic networks, however, might also be associated with negative externalities. Disadvantageous network externalities may arise if immigration is subject to adverse selection, or if increases in immigrant concentration increases competition for jobs and lower immigrants' wages. Under certain conditions the tendency to cluster may lower incentives to learn the language of the host country, which in turn may "trap" migrants in poverty (Bauer, Epstein and Gang, 2005). These negative network externalities limit the benefits immigrants can obtain from clustering.

Several studies investigate the determinants of location choice by immigrants in the United States. Bartel (1989) finds that post-1964 migrants to the US tend to locate in cities with a high concentration of immigrants of similar ethnicity. She further shows that highly skilled migrants are less geographically concentrated and rely less on the location of fellow compatriots. Similarly, Jaeger (2000), who differentiates between immigrants of different admission statuses, finds that immigrants tend to locate where former immigrants of the same ethnicity are concentrated.

Migrants consider several factors in making their decisions about where to move, including the clustering of compatriots and similar folk in various localities. Ties of kinship, friendship, and village, link migrants, former migrants, and non-migrants in the home and host country. This paper contributes to the literature by investigating the differing effects of “stock” and “flow” factors on migrants' location decisions. Stock factors measure the degree to which migrants may view a US location as (ethnically) hospitable and the availability of information about specific locations. We characterize two types of stock factors, an ethnic goods component and a village migration history component. Our flow factor measures the tendency of migrants to follow the paths of very recent migrants from their own villages.

These factors offer different information to a potential migrant. The ethnic goods component sends signals to the migrant about the possibility of living in a culturally similar environment, i.e., speaking his native language, listening to his music, reading his own newspapers, and eating ethnic food. The ethnic goods factor reduces the monetary and psychic costs of migrating. The village migration history component largely captures information about the host region received in the home village. This includes, for example, information on the labor and housing market, and information on specific employers in a region. In addition, the migrant may be able to count on contacts in a specific location established by former migrants from the same village. This factor reflects the probability of receiving help from compatriots. The flow factor represents potential herd behavior by migrants, a sort of “peer emulation effect.” Following the argument by Epstein (2002), migrants may choose a location on the supposition that recent migrants had information that he does not have. We examine the relative importance of migrant stocks and flows as explanations of immigrant location choice, also accounting for several other determinants.

Until the appearance of the paper by Polachek and Horvath (1977) much of migration theory treated migration as an individual investment decision. Family members other than the household head are not always explicitly considered. However, other members are clearly influential in migration decisions.– Polachek and Horvath (1977) established the foundations for models of location choice that take into consideration all the different type of considerations. They do so by adopting a life cycle approach used in human capital theories of earnings accumulation, accounting for household considerations in both a general theoretical and empirical model. More importantly, migration was analyzed within a nonstochastic framework and remigration was endogenously explained.

We describe our data and define and characterize the variables we employ in Section 2.

Section 3 presents our empirical results, while Section 4 offers a theoretical model explaining our results. Section 5 concludes.

2. THE GEOGRAPHIC DISTRIBUTION OF MEXICAN MIGRANTS IN THE US

In absolute numbers, the US is the world's largest country of immigration; Mexico is the world's major country of emigration; migration from Mexico to the United States is the largest sustained flow of migration in the world. Empirical evidence suggests that there exist strong network effects in Mexican migration (Bustamante, 1998; Munshi, 2003; Winters, de Janvry and Sadoulet, 2001). We explore the stock and flow effects of clustering on migrants' location choices using individual and village level data on Mexican-US migration available through the Mexican Migration Project.[1] The data comprise more than 7,000 households in 52 communities in the states of Colima, Guanajuato, Guerrero, Jalisco, Michoacán, Nayarit, San Luis Potosí, and Zacatecas. The data set provides information on the socioeconomic characteristics of household heads, such as age, education and marital status, their migration histories including information on year of migration, costs of border crossing, documentation and location in the United States. In calculating our flow variable and one of our stock variables, we use an event-history file containing detailed labor and family histories of each household head, such as information on the number of trips to the United States, the duration of each trip, and related information, for each year from the birth of the household head until the year of the survey.

We calculate for each year t $(t=1,\ldots,T)$ the cumulative migration experience (in months) for each migrant I $(i=1,\ldots,N)$ from the Mexican community m $(m=1,\ldots,M)$ in each US location j $(j=1,\ldots,J)$.[2] The cumulative migration experience of community m in US location j, EXP_{mjt}, is

$$EXP_{mjT} = \sum_{t=1}^{T} \sum_{i=1}^{N} M_{mjit} \quad , \qquad (1)$$

where M_{mjit} is a dummy variable that takes the value 1 if an individual i in the Mexican community m is in US location j at year t.

[1] See the Appendix A for a description of the data and its reliability.

[2] We do not discount months over time, or for those who have returned to their village in Mexico. Although their knowledge of current labor market conditions may deteriorate, they provide key links and support for the network. The differential impact of more recent migrants is

We define the *Village Migration Experience*, VME_{mjt}, as the cumulative migration experience for each migrant *i* from the Mexican community *m* in each US location *j, relative to* the total experience of that village in the US. The measure captures the Mexican village's migration experience in a US location at the time a person makes his migration decision, and is calculated as

$$Village\ Migration\ Experience = VME_{mjt} = \frac{EXP_{mjt}}{\sum_{j=1}^{J} EXP_{mjt}} \cdot 100 . \qquad (2)$$

In addition to the migration experience of a particular Mexican village, we use the *Mexican Share of the Total Population* in a US location (see the Appendix for a description of the calculation of this variable). This second stock variable disregards specific village information, instead capturing the concentration of ethnic goods in a location relative to other locations. Adding this second stock variable helps distinguish a generalized stock effect from village-specific links.[3]

We also examine the impact of the *flow* of migrants during the year before an individual migrates, calculated as the *Change in Village Migration Experience*, $FLOW_{mjt}$, in the year before an individual migrates,

$$FLOW_{mjt} = VME_{mjt} - VME_{mj(t-1)} , \qquad (3)$$

where $VME_{mjt} \geq VME_{mj(t-1)} \geq 0.$ We visualize that the person makes his/her decision at the end of period t. This enables us to see how the relative flow of migrants between *t-1* and *t* affects the probability of migrating to a particular location at time *t*. Since we are interested in the flow to a certain destination relative to other locations, we present define the flow variable in relative terms.

In an uncertain environment, networks provide information about the host locations. Although knowledge of current labor market conditions may deteriorate over time, migrants who have returned several years ago may still provide key links and support for new migrants, such as

---

captured by our *flow* measure, while our *Mexican share of the total population* variable captures the generalized impact of having other Mexicans around in a US location.

arranging a coyote to smuggle them across the border, provide information about alternative locations, or simply telling stories about their experiences and passing on knowledge. More recent migrants with current first hand information about job opportunities are likely to help their community members find jobs. Others from the broader ethnic group set the tone and atmosphere of living in locations away from home.

Our village migration experience variable and our flow variable are scaled by the village's experience in the United States, making them relative measures reflecting the influence of the village network on location choice. The spread of migrants across the US has an important impact on the utility a migrant obtains from the network; there are both positive and negative network effects. As the concentration of migrants' increase, their wages decrease; however, as geographic mobility is high and US labor markets are highly integrated, wages and network effects are relative. A similar argument can be put forward when considering the attitude of the local population towards immigrants.

To control for other factors that may affect the utility levels associated with a US location, we include several variables capturing the economic and social characteristics of a location in the multivariate analysis. These factors include, for each US location, population size, the consumer price index to capture cost of living differences, and the unemployment rate in those locations. Though the unemployment rate is sometimes problematic in migration studies, the literature often assumes that the probability of choosing a particular location decreases with the unemployment rate in this location (see the discussion in Jaeger (2000)).[4] A detailed description of the variables used in the empirical analysis is given in Appendix A.

Migration costs affect location choice. Most Mexican migrants have a very low income in their home village. Therefore, the cost of migrating may be an important factor in determining the specific location to which to migrate. To control for these costs we include road mileage from the migrant's origin village in Mexico to the alternative US locations.[5] We also examine

---

[3] We thank Julie Phillips for making this variable available to us.

[4] Our empirical analysis treats the network for each Mexican community as exogenous to the individual migration decision. We feel this is the appropriate specification. However, one could argue that unobserved autocorrelated fluctuations in US local labor markets may draw migrants together to one location even if previously migrants from that community had gone elsewhere. Local unemployment rates in the US receiving communities also control for this.

[5] In addition to road mileage, we also examined hours by car and the actual migration costs expressed by the migrant himself. All three cost variable yield similar results in our estimations.

US location specific fixed effects in order to control for time constant determinants of the location choice.

The covariates just discussed are US location specific, as dictated by our desire to examine determinants of migrants' location choice, and by the conditional logit model we discuss in the next section. In addition we use several individual specific variables and examine how these individual dimensions interact with our stock and flow effect variables. We look at the interaction of the location specific variables with skill level, legal status and whether it is someone's first trip to the US or their last trip (as recorded in the data). Migrants with six or fewer years of schooling are assumed to be unskilled; those with more than six years are considered skilled. Migrants report themselves whether they migrated legally (documented) or illegally (undocumented). We expect the migrant's use of the information provided by the stock of previous migrants or their inclination to follow the flow will vary depending on these factors. In particular, we expect the impacts of stocks and flows to vary between the first-time an individual migrates to the US and repeat movers.

Table 1 presents a description of the data we use in our analysis. For the first migration, we have information on 1739 individuals from 47 Mexican villages who migrated to 43 different locations in the US The geographic unit in the US varies – some are cities, some are parts of a county, and some are counties – but they are generally recognizable as sensible divisions (See Appendices B and C for a list of the locations). We assume that each person has the possibility of going to each of these 43 locations, but does not consider other locations.[6] This generates 74,777 observations – each person may or may not go to each of the 43 locations. For the last migration, we have 1561 individuals from 47 Mexican villages going to 46 US locations, resulting in 71,806 observations. Unskilled migrants dominate, comprising 67 percent of first time migrants and 74 percent of last time migrants. On the other hand, 88 percent of first time migrants are undocumented, while only 46 percent of repeat migrants are undocumented, indicating that Mexicans obtain US residence permits over time.

Table 1 further shows that Mexicans make up about 5.5 percent of the population of the US locations in our sample. The highest concentration could be observed in Laredo, Texas, where 24.2 percent of the residents are of Mexican origin (Appendix B). Laredo has the highest

[6] Under the conditional logit formulation we apply, it is assumed the potential availability of other location choices will not affect the coefficient estimates.

unemployment rate in our sample (more than 16 percent), a very small local population and is very close to Mexico. Though the city is small and has a high unemployment rate, many appear to migrate there, because it is close to the border. The *Village Migration Experience* variable averages 1.9 percent. It reaches a maximum of 29.2 percent in Los Angeles, followed by Chicago with 9.2 percent (Appendix B). The migration flow appears to be about twice as large for first time migrants than for repeat migrants. Each of our locations has, on average, an unemployment rate of 7.1 percent, a population of 1.35 million, and is approximately 1460 miles away from the sending village in Mexico.

Figures 1 and 2 describe some typical patterns of our two stock variables. In Figure 1 we plot the Herfindahl index of the concentration of the US migration experience of nine typical Mexican villages for the time period covered in our sample. The index is given by

$$HERF_{mt} = \sum_{j=1}^{J} \left(\frac{VME_{mjt}}{100} \right)^2 \tag{4}$$

with $0 \leq HERF_{mt} \leq 1$. Higher values of $HERF_{mt}$ indicate a higher concentration of the migration experience of a Mexican village. The villages differ in the concentration of their migration experience. Compared to the other villages depicted in Figure 1, concentration is relatively low in communities 36 and 38 in the Mexican State S.L.P., community 46 in Zacatecas, and community 33 in Colima.[7] In most of the villages the concentration of the migration experience is increasing over time and flattens out at the end of the sample, indicating some kind of quadratic pattern, though most of the villages do not reach a turning point. Only in community 36 do we observe the concentration of the migration experience increasing at the very beginning of the sample period, reaching a maximum and then decreasing. In contrast to all other communities we observe a U-shaped pattern in community 52 in Oaxaca. Note that we find such a pattern only in two communities. In terms of US locations, Los Angeles County is the location with the highest average value of migration experience.

Figure 2 shows the development of our second stock variable, the share of the Mexican population, in six US locations for the period covered in our sample. We display the Imperial

[7] The data set does not provide names for the Mexican villages.

Valley, Chicago, Houston, and Miami for their geographical dispersion and generic interest. In these five US locations the share of the Mexican population is increasing. The sixth US location is Laredo, Texas, which has the highest average share of Mexican population in our sample. In Laredo, the share of the Mexican population shows a U-shaped pattern over time; it decreases until 1982 and then increases.

3. MULTIVARIATE ANALYSIS

Econometric Approach

To analyze the determinants of the location choice of Mexican migrants to the US, we estimate a conditional logit model (McFadden, 1984).[8] Each Mexican migrant *i*, who is assumed to maximize his utility, faces a choice among *J* alternative US communities. Assume that the utility of choosing location j is given by

$$U_{ij} = X_j \beta + \varepsilon_{ij} , \qquad (5)$$

where X_j is a vector of the characteristics of the US location *j*, including stock and flow effects, and ε_{ij} is an error term that is assumed to be independent and identically distributed. The probability that an individual *i* chooses location *j* is given by

$$\Pr(U_{ij} > U_{ik}) \quad \text{for all} \quad k \neq j. \qquad (6)$$

Let Y_i be a random variable that takes the values 0 and 1, indicating the location choice made by the migrant. The probability that individual *i* chooses the US location *j* can then be written as

$$\Pr(Y_i = j) = \frac{exp(X_j \beta)}{\sum_{j=1}^{J} exp(X_j \beta)}, \qquad (7)$$

where X_j is a vector of characteristics of the US communities in our sample and β is a parameter vector to be estimated. Equation (7) can be estimated using maximum likelihood. Note that our sample is restricted to individuals who actually migrated at some point in time to the US. The

[8] Bartel (1989) and Jaeger (2000) also use this model to study the location choice of migrants in

analysis does not consider migration within Mexico.

As discussed in Section 2, our regressors include two measures of the effect of the stock of migrants, i.e., the Mexican share of the total population in US location *j* and the migration experience of a Mexican village *m* in the US location *j*, VME_{mjt}, a measure of the flow of migrants, $FLOW_{mjt}$, the population size, the consumer price index (cpi), and the unemployment rate in US location *j*, as well as the cost of migration proxied by the road mileage distance between Mexican village *m* and US location *j*. The existing theory (see, for example, Epstein, 2002) shows that we should expect the stock variables, *Village Migration Experience* and the *Mexican share of the population* to have an inverted U-shape relationship with respect to the probability of migrating to a certain location. As the stock of migrants in a location increases, the probability of a new migrant moving to that location increases at a decreasing rate, because positive network effects decrease and negative network externalities increase as the number of immigrants increase. Eventually, a turning point is reached after which a further increase in the stock of migrants will decrease the probability of a new migrant moving to that location. Hence, our specification of equation (7) includes both a linear and a squared term for the two stock variables. All other variables enter linearly.

We analyze the determinants of location choice both with and without US location fixed effects. Accounting for location fixed effects controls for the influence of time invariant heterogeneity. For example, one might argue that climate is an important determinant of location choice – especially in a study of the migration of persons from Mexico to US locations as climatically diverse as Laredo, TX, and New York City, NY.

In our empirical analysis we consider several specifications of equation (7). As individuals may have migrated more than once to the US, we divide our analysis into two parts: first and last migration. In the former we consider only the location decision made by the Mexican migrants at his/her first time migrating to the US while the latter consider only the location decisions made at his/her last time migrating to the US, conditional that he/she migrated to the US at least once before. For both specifications we estimate an overall (constrained) equation and an unconstrained equation. In the latter all variables considered in the basic specification are fully interacted with four dummy variables, one for unskilled illegal migrants, one for unskilled legal migrants, one for skilled illegal migrants, and one for skilled illegal

---

the United States.

migrants.

Estimation Results – without US location fixed effects

The second and seventh columns in Table 2 (first migration) and Table 3 (last migration) present the results for the constrained model (i.e., where we do not account for variation due to skill or legal status); columns 3-6 and 8-11 present the results for the unconstrained model; columns 2-6 do not include US location fixed effects, while columns 7-11 do. Consider the results for the constrained specification for the first migration decision. The Mexican share in the population of a US location appears to have an inverted U-shaped effect on the probability of choosing a particular location. Evaluated at the sample mean of a Mexican population share of 5.51 percent, the average marginal effect of an increase of the population share by one percent is 0.15.[9] Simulations we performed show the predicted effect of the share of Mexicans in the population of an average US location on the probability of choosing that location peaks at a population share of about 10 percent.[10]

Our other stock variable, the migration experience of a Mexican village, also follows an inverted U-shaped pattern, i.e., an increase in the share of a village's migration experience in a particular US location relative to its total Mexican migration experience increases the probability of choosing a particular US location at a decreasing rate. At a *Village Migration Experience* (VME) of approximately 63 percent the impact peaks, declining afterwards. While most cities in most times are on the uphill side of this turning point, we do observe four US locations where the value of VME_{mjt} exceeds 63 percent: Los Angeles County, Orange County and San Diego County

---

[9] The marginal effects of a change in the characteristics X_j of a US location j on the probability that a Mexican migrant will choose location j are given by the derivative of equation (15) with respect to the characteristics X_j., Note that these marginal effects will vary with the characteristics of a US location j, which leads to a very large number of marginal effects to interpret. Therefore, we follow the simplifying approach chosen by Jaeger (2000) and calculate average effects of a change in the characteristics X on $Pr(Y_i=j)$, i.e. $\partial Pr(Y_i = j)/\partial X_j = [(1/J)(1-(1/J))]\hat{\beta}$, where J=43 for the first migration decision and J=47 for the last migration decision. Hence, to obtain average marginal effects, the coefficients reported in Table 3 have to be multiplied by 0.0227 and those in Table 4 by 0.0208.

[10] In particular, we calculated $Pr(Y_i = 1) = \dfrac{exp(\beta' X_j)}{1 + exp(\beta' X_J)}$ using sample means for X_j for all variables except the variable of interest and assuming that the location specific fixed effects are zero.

in California and Chicago.[11] The coefficient on the variable capturing flow effects is significantly positive. The average marginal effect for this variable is calculated to be 0.0053, indicating that a 1 percent increase in the flow of migrants to a specific US location in the last year increases the probability that a migrant chooses this location on average by 0.53 percent.

For the constrained model without US location fixed effects, and for the four subgroups considered in the unconstrained model, the Mexican share in the population of a US location has an inverted U-shaped pattern. It appears that the Mexican stock in a US location is more important for unskilled as compared to skilled workers. Whereas the probability of choosing a US location peaks at a Mexican population share of approximately 10 percent for the latter, it reaches a maximum for skilled workers at a population share of 8 percent. Comparing legal and illegal migrants, however, no clear pattern emerges.

As in the constrained model, the estimated inverted U-shaped pattern for the village experience variable is much flatter than the respective pattern for the Mexican population share. However, in contrast to the Mexican population share, important differences between legal and illegal migrants appear. For illegal migrants, the effect of village migration increases the average probability of choosing a US location up to a share of 61 percent for unskilled, and 71 percent for skilled. For legal migrants this variable reaches its maximum effect at a share of 48 percent for unskilled and a share of 53 percent for skilled migrants.

The flow of migrants significantly affects all sub-groups considered. It further appears that there are no significant differences of the estimated flow effect between the different groups. Finally, the response of illegal migrants is more sensitive to changes in the migration flow before their migration decision as compared to legal migrants. However, as already noted above, these differences are not statistically significant.

Overall, these results indicate that legal and skilled migrants are less dependent on the stock of migrants when deciding on the location. The results further suggest that village-specific links, captured by the migration experience of a village, are on average relatively more important for the location choice of a migrant than ethnic goods, captured by the Mexican population share.

The estimation results for the last migration decision, not accounting for the possibility of US location fixed effects, are reported in Table 3. As for the first migration decision, both stock variables appear to have an inverted U-shaped pattern on the probability of choosing a US

[11] This only happened in certain years and does not show up in the Appendix tables.

location and the pattern of the effect is much flatter for the village migration experience as compared to the share of the Mexican population in a US location. Comparing the different groups differentiated in the unconstrained model does not give a significantly different picture than the one obtained in Table 2. Comparing the first and last migration decision, however, it appears that both flow and stock effects are slightly more important for the last migration decision: the peaks are at a higher probability level and a higher share for the two stock variables. The effect of the flow variable on the probability of choosing a US location is steeper for the last as compared to the first migration decision.

Let us now consider what effects US location characteristics have on migrant location choice, still not accounting for the possibility of US location fixed effects. In the constrained model, the unemployment rate in a US location has a negative effect on the probability of choosing a location. However, only for the migrants' first trip is this effect statistically significant. In the unconstrained model, the effect of the unemployment rate on the location decision of a migrant is unclear for his/her first trip. According to the results reported in Table 2, the unemployment rate has a significant negative impact on the location decision of skilled illegal migrants and an unexpected significant positive impact on unskilled legal migrants. For the last trip of a migrant, the unemployment rate in the US location j affects only the location choices of skilled migrants on a statistically significant level; an increase in the unemployment rate in a US location decreases the probability that a skilled Mexican migrates there by 0.4 percent for illegal migrants and by 0.2 percent for legal migrants. Cost of living differences as captured by the consumer price index do not seem to drive location choice. Where the CPI is significant, it lowers the probability of moving to a location. We also examined a specification omitting the CPI., which left our other coefficient estimates essentially unchanged.

The probability that migrants choose a particular US location increases with the total population in that location for the first trip. For the last trip the total population has a positive effect on the location choice of unskilled illegal and skilled migrants, and a negative effect on unskilled legal migrants. This result reflects preferences for moving to regions with relatively large labor markets. The distance between the home community and the US location has a negative impact on illegal migrants and a positive impact on documented migrants on their first trip; the estimated coefficients are, however, not statistically significant at the 5 percent-level. For the last migration decision the distance to the US location shows an unexpected pattern. For the constrained model and for unskilled workers in the unconstrained model the coefficient of the distance variable is significantly positive indicating that a higher distance increases the probability of choosing a US location. It might be that this variable captures some other effects of characteristics of the US locations we did not control for in our specification.

Estimation Results – with US location fixed effects[12]

A difficulty in our analysis is that there are probably unobserved region specific factors that determine migrants' location choice. For example, there may be variations among US locations with respect to resource endowments, cultural influences on legal and political arrangements, climate, and so on. To the extent that these factors are time invariant, we can control for these time invariant factors by including location specific fixed effects. Our results when we do this are seen in Tables 3 and 4, columns 7 to 11.

With a notable exception the inclusion of US location fixed effects does not change our results. Comparing the estimations for first migration in Table 3 without (columns 2 to 6) and with (columns 7 to 11) fixed effects, *Village Migration Experience* and our flow variable continue to be significant with approximately the same impact. The *Mexican Population Share*, however, is now not significant, small, and generally has a negative impact. Perhaps this share is slow to change and its effect is now absorbed by the location dummies. In Table 4 we observe the same phenomenon. On the bright side, the inclusion of US location fixed effects helps clear up some anomalies we noticed. In particular, distance is now negative and significant in its impact on location choice.

Our empirical results show that both of our stock measures and the flow of immigrants have significant effects on the migrant's decision about where to migrate. We should and cannot neglect these effects when analyzing location choice. These results confirm and extend other results on the importance of networks in location choice (for example, Jaeger (2000), and Winters, de Janvry and Sadoulet (2001)). However, the choice of network variable can make a difference in ones conclusions. The Mexican share of the population becomes insignificant when US location fixed effects are included. Without accounting for these fixed effects, the Mexican share of the population is significant and portrays an inverse U shape. The village stock

[12] We thank a referee for leading us to this analysis.

externality effect is larger and more robust, and also exhibits an inverse U shape, and not the simple positive linear effect as often presented in the literature.

4. A POSSIBLE EXPLANATION: HERD EFFECTS AND MIGRATION NETWORKS

One possible explanation for the results presented above is the relative importance and interaction of herd behavior (flows) and network externalities (stocks) in determining migration behavior. Both motivations give rise to immigrant clustering, a phenomenon observed in a wide variety of migration destinations. The theory we develop below builds on the work of Epstein (2002).

Let us first consider network externalities. Consider individual *j's* utility from migrating to a certain location, $U_j(.)$. $U_j(.)$ is a function of two variables: (i) the wage that the migrant will receive by migrating to the new location; and (ii) the stock of immigrants from the same origin who previously migrated to the new location, *N*. From the above discussion, the migrant's utility increases with the migrant's wage and increases with network externalities:

$$\frac{\partial U_j(w_j, N)}{\partial w_j} > 0 \text{ and } \frac{\partial U_j(w_j, N)}{\partial N} > 0 .$$

Assume a normal downward-sloping demand function for workers in the host location, $q^d(w_f)$ and an upward-sloping supply function workers, $q^s(N_L, N)$. In equilibrium demand equals supply: $q^d(w_f) = q^s(N_L, N)$. In equilibrium wages are given by $w_f^*(N)$. Note that the equilibrium wage decreases as the stock of immigrants increases. The stock of migrants (the network effect) affects utility in two ways: directly via positive externalities and indirectly via negative externalities on the wages. The "old" migrants (the stock of immigrants) who are already in the host location prefer that the maximum number of migrants coming to this location will be such that their utility is maximized. That is, the marginal increase in the migrants' utility from externalities equals the marginal effect of the decrease in wages because of the additional migrant.

Denote by N_1 the optimal stock of immigrants in the sense that this is the preferred stock of migrants who have previously migrated to this host location. Thus if the stock of immigrants exceeds N_1, further increasing the stock of immigrants raises non-wage network benefits, however wages also decrease. The effect of the increase in non-wage benefits is smaller than the effect of the decrease in wages and the utility of the immigrants who had previously migrated to this location decreases.

We may still observe migrants deciding to migrate to a location in which the stock of migrants has already exceeded N_1. Thus, the probability that an individual chooses to migrate to a location where the stock of immigrants already exceeds N_1 is positive. This probability however, decreases as the stock of immigrants already in the host location increases. We conclude,

Given network externalities, the probability an individual migrates to a certain location has an inverse U shape relationship with regard to the stock of immigrants already in the host location.

Now let us consider *herd behavior.* Following Epstein (2002) migration decisions are made sequentially, with people contemplating emigration at a given stage in their lives. Individuals respond to signals or information packets about host location possibilities. An individual receives a signal with probability p and with probability q this signal is true. The individual also observes the behavior of previous migrants. Potential migrants cannot, however, observe the information signal that was the basis for previous migrants' decisions. Given the information available, each individual chooses a location to which to migrate. The structure of the game and Bayesian rationality are common knowledge. Three assumptions govern individuals' actions: (a) An individual, who does not receive a signal and observes that everybody else has chosen to stay home, will also choose not to migrate. (b) An individual who is indifferent between following his or her own signal and copying someone else's choice will follow his or her own signal. (c) An individual who is indifferent between following more than one of the previous migrants' decisions will choose to randomize his or her decision with equal probabilities assigned to the different alternatives.

Under this framework it can be shown[13] that if an individual receives a signal to migrate to a specific place, he will follow this signal. If a second individual receives the same signal he will follow *individual 1* and if he does not receive a signal he will also follow *individual ,1* since the first migrant decided to migrate only because he had a signal and *individual 2* can see it as if

[13] See Epstein (2002) for the formal proof and generalization.

he himself received this signal. Now consider the case where *individuals 1* and *2* migrated to one location, *location 1*, and *individual 3* receives a signal to immigrate to *location 2*. Since the first individual migrated it is clear he had a signal. Therefore, *individuals 1* and *3* had different signals. Since *individual 2* also migrated to the same location as *individual 1*, he may have received a signal to migrate to *location 1* or he may not have received a signal at all. Using a Bayesian Rule it can be shown that the probability of immigrating to *location 1* is higher than that of *location 2* and thus *individual 3* will follow the first two migrants instead of following the signal he received. On average there are *1+p* signals for *location 1* while there is only 1 signal for *location 2*. Epstein (2002) shows that

As the number of individuals that have already migrated to some certain location increases, the probability of a new individual migrating to the same location increases. Individuals will migrate following the herd (flow) while disregarding their own private information.

As we can see from the empirical results, the theory of herd verses network externality may explain the behavior of the Mexican migrants in their location choices in the US.

5. CONCLUSION

Immigrant clustering is an important phenomenon to study for a number of reasons. The process by which immigrants decide where to locate is one that is not clearly understood, though there is much research on the subject. Standard economic theory argues that there are significant externalities, or “ethnic capital,” of which immigrants wish to take advantage. They move to where members of their community, generally defined, had previously gone, planning to avail themselves of these externalities. In this paper we emphasize the different information content in different types of networks by examining two stock influences and a flow influence on the migration location decision.

Although previous studies have highlighted the role of networks on migration, no one has studied the potentially different impacts of migration stocks vs. migration flows. The paper argues that the relationship between the stock of migrants and the location choice of new migrants follows an inverted U, while because of herding the relationship between migration flows and location choice is positive.

We use data from the Mexican Migration Project to investigate the location decision of Mexican migrants in the US We distinguish between two types of stock effects, capturing general ethnic goods available in a US location (*Mexican share of the total population* in a particular US location) on the one hand, and origin village connections and the history of the migration experience of a village in different US locations (*Village Migration Experience*, the tendency of residents of a given Mexican village to migrate to a particular US location, measuring information available in the sending village of a given US location) on the other hand. These two variables help us to distinguish a generalized stock effect from village-specific links. The flow effect is measured using the flow of migrants to a particular US location during the year prior to the migration decision of an individual.

We show that both stock externalities and the flow have a significant effect on the migrant's location decision, though the "cultural goods" stock effect disappears once we control for location fixed effects. Moreover, the significance and size of the effects vary according to the legal status of the migrant and whether the migrant is a "new" or a "repeat" migrant. The estimated stock effects show an inverse U-shaped pattern, not a linear positive effect as often presented in the literature. The results indicate that village-specific links are relatively more important for the location decision of a migrant than the availability of ethnic goods. Furthermore, legal and skilled migrants appear to be less dependent on the migration stock than illegal and unskilled migrants. Flow effects have significant positive effects on the location decision of a migrant. Our estimations indicate, however, that there are no significant differences in these flows between different types of migrants.

Although a number of studies have underscored the importance of networks for location choices, the argument that immigrant clustering could be explained by herd behavior has been recently introduced to the migration literature (Epstein, 2002). Networks and herds reflect different types of information. Migrants might be motivated to choose a location to benefit from the network externalities it has to offer. However, because of herd effects, the migrant may choose a location on the supposition that recent migrants had information that he does not have. Migrants may choose to follow the flow and migrate to the location recent migrants have been observed to choose. Our empirical results indicate that network externalities and herd effects can both be present and influence emigration location decisions. The network externalities and herds' story is one of many interpretations. The fact is that both stocks and flows affect the decision of the migrant of where to go; this is the most important message of this paper.

REFERENCES

Bartel, A.P. (1989): "Where Do the New US Immigrants Live?," *Journal of Labor Economics*, 7(4), 371-391.

Bauer, Thomas, Gil S. Epstein, Ira N. Gang, (2005): “Enclaves, Language, and the Location Choice of Migrants,” *Journal of Population Economics*, 18(4), 649 – 662.

Bustamante, J.A. (1998): “Some Thoughts on Perceptions and Policies,” in *Migration between Mexico and the United States: Mexico—US Binational Migration Study*. Mexican Ministry of Foreign Affairs and US Commission on Immigration Reform.

Carrington, W.J., Detragiache, E. and Vishwanath, T. (1996): “Migration with Endogenous Moving Costs” *American Economic Review*, 86(4), 909-30.

Chiswick, B.R. and P.W. Miller (1996): “Ethnic Networks and Language Proficiency among Immigrants,” *Journal of Population Economics*, 9(1), 19-35.

Church, J. and King, I. (1983): “Bilingualism and Network Externalities”, *Canadian Journal of Economics*, 26(2), 337-45.

Epstein, G.S. (2002): "Informational Cascades and Decision to Migrate," *IZA working paper 445*.

Gottlieb, P. (1987): *Making Their Own Way: Shorthorn Blacks’ Migration to Pittsburgh, 1916-30*, Urbana: University of Illinois Press.

Grossman, J. R. (1989): *Land and Hope: Chicago, Black Southerners, and the Great Migration*, Chicago, University of Chicago Press.

Jaeger, D.A. (2000): "Local Labor Markets, Admission Categories, and Immigrant Location Choice," http://dajaeg.people.wm.edu/research/ last accessed, May 22, 2006.

Marks, C. (1989): *Farewell -We’re Good and Gone: The Great Black Migration*, Bloomington, Indiana University Press.

Massey, D.S., R. Alarcón, J. Durand, and H. Gonzáles (1987): *Return to Azlan: The Social Process of International Migration from Western Mexico*. Berkley and Los Angeles: University of California Press.

McFadden, D. (1984): "Econometric Analysis of Qualitative Choice Models," in Z. Griliches and M.D. Intriligator (eds.), *Handbook of Econometrics, Volume II*. Amsterdam: North-Holland.

Munshi, K. (2003): “Networks in the Modern Economy: Mexican Migrants in the US Labor Market.” *Quarterly Journal of Economics*, 118, 549-599.

Polachek S.W and Horvath F.W (1977): "A Life Cycle Approach to Migration: Analysis of the Perspicacious Peregrinator." In Ron Ehrenberg, editor, *Research in Labor Economics*, Vol 1, 103-149.

Winters, P., A. de Janvry, and E. Sadoulet (2001): “Family and Community Networks in Mexico-US Migration,” *Journal of Human Resources*, 36, 159-184.

Table 1:
Descriptive Statistics, Means of US Recipient Locations

		First Migration	Last Migration
Unemployment Rate (in %)		7.103	7.310
		(3.309)	(3.413)
CPI		85.203	110.821
		(31.979)	(30.624)
Total Population (in 100,000)		13.351	14.066
		(18.867)	(19.216)
Miles		1459.956	1431.984
		(527.774)	(510.941)
Mexican Share of Population (in %)		5.511	5.568
		(6.476)	(6.163)
Village Migration Experience (in %)		1.986	1.870
		(7.622)	(7.563)
Flow (in %)		0.878	0.442
		(46.054)	(26.340)
Unskilled Legal	(Observations)	3784	22908
	(Individuals)	88	498
Unskilled Illegal	(Observations)	46268	30360
	(Individuals)	1076	660
Skilled Legal	(Observations)	5289	11040
	(Individuals)	123	240
Skilled Illegal	(Observations)	19436	7498
	(Individuals)	452	163
Total	(Observations)	74777	71806
	(Individuals)	1739	1561
Number of Mexican Villages		47	47
Number of US locations		43	46

Note: Standard deviations in parentheses.

Table 2:
Conditional Logit Analysis of Mexican Migrant's Location Choices:
First Migration

	Constrained	**Unconstrained**				**Constrained**	**Unconstrained**			
		Unskilled		**Skilled**			**Unskilled**		**Skilled**	
Variables		**Illegal**	**Legal**	**Illegal**	**Legal**		**Illegal**	**Legal**	**Illegal**	**Legal**
Unemployment Rate	-0.033**	0.0004	0.091	-0.169**	0.001	-0.032	0.003	0.094	-0.157**	-0.016
	(0.015)	(0.012)	(0.066)	(0.036)	(0.061)	(0.023)	(0.026)	(0.070)	(0.039)	-0,066
CPI	0.002	0.010	0.021	0.024	-0.154**	0.029	0.043*	0.047	0.036	-0.129**
	(0.016)	(0.022)	(0.068)	(0.032)	(0.055)	(0.018)	(0.024)	(0.067)	(0.031)	(0.056)
Total Population	0.012**	0.009**	0.021**	0.019**	0.013**	0.048**	0.041**	0.053**	0.046**	0.043**
	(0.001)	(0.001)	(0.005)	(0.003)	(0.005)	(0.012)	(0.013)	(0.013)	(0.012)	(0.013)
Distance in Miles (in 1,000)	0.044	0.044	0.196	0.362*	0.701*	-3.571**	-3.645**	-3.142**	-3.685**	-2.554**
	(0.085)	(0.103)	(0.416)	(0.187)	(0.381)	(0.623)	(0.628)	(0.772)	(0.650)	(0.749)
Share of Mexican Population (in %)	0.154**	0.125**	0.209**	0.283**	0.266**	-0.071	-0.082*	-0.063	-0.056	-0.025
	(0.019)	(0.022)	(0.083)	(0.057)	(0.093)	(0.048)	(0.049)	(0.089)	(0.067)	(0.091)
Share of Mexican Population (in %)Squared	-0.008**	-0.006**	-0.011**	-0.017**	-0.014**	0.001	0.002	0.001	-0.001	-0.0001
	(0.001)	(0.001)	(0.005)	(0.004)	(0.006)	(0.001)	(0.001)	(0.003)	(0.003)	(0.004)
Village Migration Experience (in %)	0.110**	0.113**	0.123**	0.100**	0.137**	0.092**	0.093**	0.098**	0.088**	0.120**
	(0.004)	(0.006)	(0.025)	(0.009)	(0.019)	(0.005)	(0.006)	(0.025)	(0.009)	(0.183)
Village Migration Experience (in %)Squared/ 100	-0.087**	-0.093**	-0.127**	-0.071**	-0.127**	-0.067**	-0.070**	-0.099**	-0.058**	-0.108**
	(0.006)	(0.008)	(0.038)	(0.011)	(0.028)	(0.006)	(0.008)	(0.038)	(0.011)	(0.025)
Flow	0.233**	0.227**	0.154*	0.292**	0.234**	0.212**	0.201**	0.193**	0.251**	0.222**
	(0.025)	(0.031)	(0.089)	(0.056)	(0.117)	(0.025)	(0.023)	(0.094)	(0.053)	(0.112)
US Location Fixed Effects	No	No				Yes	Yes			
Log-Likelihood	-4025.0	-3975,40				-3743.70	-3706,70			
Pseudo-R^2	0.385	0,392				0.428	0,4333			

Note: Observations: 74,777. Standard errors in parenthesis. *: Statistically significant at least at 10% level. **: Statistically significant at least at the 5% level.

Table 3:
Conditional Logit Analysis of Mexican Migrant's Location Choices:
Last Migration

	Constrained	**Unconstrained**				**Constrained**	**Unconstrained**			
		Unskilled		**Skilled**			**Unskilled**		**Skilled**	
Variables		**Illegal**	**Legal**	**Illegal**	**Legal**		**Illegal**	**Legal**	**Illegal**	**Legal**
Unemployment Rate	-0.031**	-0.001	-0.004	-0.200**	-0.092**	-0.055**	0.018	-0,045	-0.227**	-0.155**
	(0.015)	(0.021)	(0.027)	(0.064)	(0.045)	(0.027)	(0.032)	(0.038)	(0.071)	(0.055)
CPI	-0.038**	0.002	-0.063**	-0.035	-0.071**	-0.005	0.034	-0.041	-0.031	-0.072**
	(0.013)	(0.022)	(0.025)	(0.039)	(0.032)	(0.018)	(0.025)	(0.023)	(0.042)	(0.036)
Total Population	0.00003	0.004*	-0.010**	0.023**	0.009**	0.033**	0.048**	0.030**	0.061**	0.047**
	(0.001)	(0.002)	(0.002)	(0.004)	(0.004)	(0.011)	(0.013)	(0.012)	(0.013)	(0.013)
Distance in Miles (in 1,000)	0.471**	0.213	0.746**	-0.068	0.755**	-2.547**	-3.116**	-2.363**	-2.927**	-2.088**
	(0.101)	(0.140)	(0.206)	(0.341)	(0.296)	(0.725)	(0.737)	(0.762)	(0.811)	(0.802)
Share of Mexican Population (in %)	0.201**	0.119**	0.268**	0.312**	0.382**	0.012	-0.075	0.013	0.043	0.103
	(0.022)	0.029	(0.044)	(0.096)	(0.076)	(0.057)	(0.063)	(0.067)	(0.114)	(0.094)
Share of Mexican Population (in %)Squared	-0.008**	-0.005**	-0.011**	-0.017**	-0.019**	0,0004	0.003*	-0.0001	-0.003	-0.004
	(0.001)	(0.001)	(0.002)	(0.007)	(0.005)	(0.001)	(0.001)	(0.002)	(0.006)	(0.005)
Village Migration Experience (in %)	0.149**	0.129**	0.208**	0.096**	0.133**	0.135**	0.112**	0.197**	0.088**	0.124**
	(0.006)	(0.009)	(0.010)	(0.015)	(0.014)	(0.006)	(0.009)	(0.011)	(0.015)	(0.014)
Village Migration Experience (in %)Squared/ 102	-0.133**	-0.110**	-0.209**	-0.070**	-0.116**	-0.117**	-0.091**	-0.197**	-0.062**	-0.104**
	(0.008)	(0.012)	(0.015)	(0.020)	(0.018)	(0.008)	(0.011)	(0.015)	(0.019)	(0.018)
Flow	0.374**	0.357**	0.376**	0.451**	0.429**	0.383**	0.356**	0.438**	0.391**	0.414**
	(0.046)	(0.064)	(0.091)	(0.140)	(0.145)	(0.046)	(0.063)	(0.091)	(0.131)	(0.134)
US Location Fixed Effects	No	No				Yes	Yes			
Log-Likelihood	-3446.05	-3370,000				-3242.7	-3170,5			
Pseudo-R^2	0,423	0,436				0,457	0,47			

Note: Observations: 74,777. Standard errors in parenthesis. *: Statistically significant at least at 10% level. **: Statistically significant at least at the 5% level.

Appendix A: Data description

Data Basics:

The Mexican Migration Project is an ongoing collaborative research project that was originally based at the University of Pennsylvania and the University of Guadalajara. The American base is now at the Office of Population Research, Princeton University. The data are available to users at http://mmp.opr.princeton.edu/databases/dataoverview-en.aspx. The Project combines techniques of ethnographic fieldwork and representative survey sampling in its data collection. Interviews are generally conducted in December-January when sojourner US migrants often return to Mexico, supplemented with surveys of out-migrants located in the United States.

Each year since 1987, two to five additional communities in these states are surveyed, selected based on their diversity in size, ethnic composition and economic development, not because they were known to contain return migrants. Each community is surveyed only once. 200 households in each community are interviewed, though in smaller communities fewer households are chosen. We use the MMP52 version of the data, as some of the complementary data we use is not available to us for the MMP71 or the MMP93. In particular, we cannot recreate Mexican share of the US population for later years as the MMP is unable to make the necessary coding available for translating outside data into geographical areas consistent with the MMP geographical areas. Massey et. al. (1987), Massey, Goldring and Durand (1994), and Massey and Zenteno (1999) provide details and some data analysis. Massey and Zenteno (1999) show that the data are a source of reasonably representative retrospective data on documented and undocumented migration to the United States.

There are a few serious problems with the data. The interviews were free ranging, with the questioners following a semi-structured format. While the questioners tried to cover core questions, this process left many missing observations. Moreover, while the sample may be representative in a particular survey year, it will not be representative across time since it is retrospective and people are surveyed only once. To be included a migrant must have a link to a household in Mexico. It is impossible to know how important the "missing" information is for the analysis, but it may potentially severely bias the results. Also, as the data has been collected over a twenty year period there are issues with deflating wages, relative price changes, and the like.

We know if individuals ever migrated to the US, whether they were legal or not, how many times they worked in the US, the aggregate time spent in the US, when they made their first trip and when they made their last trip, how long was each of these trips, whether they were currently working in the US, their wages and occupations in the US, as well as information on the socioeconomic characteristics of the household members such as age, education and marital status. The MMP also contains more detailed migration information on household heads that have migrated.

In constructing our village migration experience and flow variables we make use of the migration event history file of the data. This file provides detailed information on each migratory experience of all heads of household, including detailed information on the first and last trip to the US such as year and duration of the trip, the documentation used, the state and city of residence, performed occupation, and hourly wage, as well as some basic information on each border crossing. See Donato, Durand and Massey (1992) for a more detailed description of the event-history file.

Mexican Share of Population: This variable has been obtained from the US Census Bureau for the censual years 1970, 1980 and 1990. A second-degree polynomial equation was estimated to these three data points to estimate the size of the Mexican foreign-population in each area during the inter-censual years. To estimate the Mexican foreign-born population in the years 1991-1995, it has been assumed that the annual growth rate during this period is the same as the annualized constant growth rate in each area between 1980 and 1990. The size of the Mexican foreign-born population is then divided by the *Total Population* in a US location. Source: We thank Julie A. Phillips for making this variable available to us.

Village Migration Experience and Flow:
These variables were calculated as indicated in the text from the event history file. *Source:* Source: *MMP 52*.

Unemployment Rate: The most recent information on the number unemployed and the size of the civilian labor force at the county level was obtained for the years 1974 and 1976-1996 from the Bureau of Labor Statistics, Local Area Unemployment Statistics Division. For the early 1970s, no information by county is available although information on unemployment for the censual years 1960 and 1970 is available. For the years 1971-1973, the assumption was made that unemployment rates in a county follow the same trends as that of the state. An estimate of the unemployment rate for 1975 was obtained by averaging the unemployment rates for 1974 and 1976. Source: *MMP 52.*

Total Population: Data were obtained from Census publications, e.g., the CPS and County and City Yearbook, for the following years: 1970, 1974, 1976, 1977, 1980, 1984, 1986, 1987, 1990, and 1991. The population for the intercensual years was estimated by assuming an exponential growth function. To estimate the population between 1992-1995, the constant growth rate that prevailed between 1980 and 1991 was applied. Source: *MMP 52*.

Migration Costs: We collected data on three measures of migration costs. For *Miles* and *Hours* we entered in the main town in the Mexican state in which the origin village is located and the main town in the US location into *Mapquest* (www.mapquest.com). For *Actual Costs* the data come from the *MMP 52*. Since the actual cost data was very sketchy, we decided not to use it. Trials with the *Hours* and the *Actual Costs* data yielded similar results to those when we used *Miles*.

Skilled vs. Unskilled, Legal vs. Illegal:
All migrants with less than 7 years of schooling are considered to be unskilled; those with more than 6 years of schooling are considered to be skilled. Undocumented migrants are labeled illegal, documented migrants *legal*. Source*: Mexican Migration Project 52*.

US Communities: Imperial Valley, CA; Lower San Joaquin, CA; Middle San Joaquin, CA; Upper San Joaquin, CA; Salinas-Monterey-Santa Cruz, CA; Sacramento Valley, CA; Ventura-Oxnard-Simi, CA; Santa Barbara, CA; Napa-Sonoma, CA; Los Angeles County, CA; Orange County, CA; San Francisco Urban Area, CA; San Jose Urban Area, CA; Riverside-San Bernardino, CA; San Diego County, CA; Rio Vista, CA; Abilene, TX; Austin, TX; Beaumont-Port Arthur, TX; Brownsville, TX; Bryan-College, TX; Corpus Christi, TX; Dallas-Ft.Worth, TX; El Paso, TX; Galveston, TX; Houston, TX; Laredo, TX; McAllen, TX; Odessa-Midland, TX; San Antonio, TX; Victoria, TX; Chicago, IL; Las Cruces, NM; Tucson, AZ; Phoenix, AZ; Denver-Boulder, CO; Reno, NV; Las Vegas, NV; Omaha, NE; New York City, NY; Washington D.C., WA; Miami, FL; Atlanta, GA.

Appendix A – Additional References

Donato, K.M., J. Durand, and D.S. Massey (1992): "Stemming the Tide? Assessing the Deterrent Effects of the Immigration Reform and Control Act," *Demography* 29(2), 139-157.

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

Massey, D.S., and R. Zenteno (1999): “A Validation of the Ethnosurvey: The Case of Mexico-US Migration,” *International Migration Review*, 766-793.

Appendix B: Descriptive Statistics by US Receiving County: First Migration

	Unemployment Rate	Total Population (in 100.000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Imperial Valley, CA	9.767	8.407	1828.160	8.137	1.703
	(2.491)	(2.603)	(160.834)	(1.533)	(5.863)
Lower San Joaquin, CA	9.573	4.278	1828.160	6.201	0.532
	(2.267)	(0.812)	(160.834)	(1.059)	(1.421)
Middle San Joaquin, CA	10.217	5.328	1828.160	7.957	2.496
	(2.192)	(0.857)	(160.834)	(1.773)	(3.548)
Upper San Joaquin, CA	11.752	7.940	1828.160	5.687	3.731
	(2.183)	(1.558)	(160.834)	(1.148)	(8.151)
Salinas-Monterey-Santa Cruz, CA	7.538	10.260	1996.099	7.474	2.934
	(2.031)	(1.676)	(160.467)	(0.513)	(4.492)
Sacramento Valley, CA	7.632	16.661	1996.099	3.134	2.377
	(2.162)	(3.029)	(160.467)	(0.534)	(3.661)
Ventura-Oxnard-Simi, CA	7.213	5.356	1608.970	7.465	1.875
	(1.314)	(0.957)	(160.619)	(0.858)	(3.747)
Santa Barbara, CA	6.056	3.121	1608.970	5.937	1.320
	(0.944)	(0.352)	(160.619)	(2.144)	(2.899)
Napa-Sonoma, CA	6.392	4.005	1996.099	2.256	0.893
	(1.899)	(0.676)	(160.467)	(1.050)	(2.686)
Los Angeles County, CA	6.866	77.237	1608.970	10.041	29.241
	(1.283)	(7.279)	(160.619)	(2.079)	(24.917)
Orange County, CA	4.856	19.621	1608.970	5.638	4.932
	(1.127)	(2.995)	(160.619)	(2.391)	(9.727)
San Francisco Urban Area, CA	5.586	33.490	1996.099	2.099	1.206
	(1.710)	(2.163)	(160.467)	(0.465)	(3.185)
San Jose Urban Area, CA	5.646	13.021	1996.099	3.713	2.595
	(0.966)	(1.345)	(160.467)	(0.921)	(6.343)
Riverside-San Bernardino, CA	7.241	17.184	1608.970	4.952	0.856
	(2.315)	(5.212)	(160.619)	(1.701)	(1.689)
San Diego County, CA	6.533	19.309	1608.970	5.490	5.184
	(1.488)	(3.758)	(160.619)	(1.374)	(12.782)
Rio Vista, CA	7.515	2.478	1996.099	1.737	0.067
	(1.397)	(0.594)	(160.467)	(0.245)	(0.314)
Abilene, TX	4.694	1.125	940.678	1.351	0.181
	(1.995)	(0.089)	(149.496)	(0.458)	(0.882)
Austin, TX	4.307	6.984	940.678	1.674	0.209
	(1.418)	(1.505)	(149.496)	(0.575)	(0.989)
Beaumont-Port Arthur, TX	8.069	3.671	940.678	0.561	0.091
	(3.559)	(0.136)	(149.496)	(0.102)	(0.581)
Brownsville, TX	11.093	2.125	621.961	21.255	1.313
	(2.788)	(0.440)	(134.766)	(5.922)	(2.760)
Bryan-College, TX	3.807	0.957	940.678	1.421	0.026
	(1.155)	(0.236)	(149.496)	(0.482)	(0.141)
Corpus Christi, TX	7.032	3.290	621.961	4.124	0.380
	(2.548)	(0.288)	(134.766)	(2.181)	(1.852)

Appendix B: continued

Dallas-Ft.Worth, TX	4.366	31.542	940.678	2.216	2.705
	(1.370)	(5.828)	(149.496)	(1.028)	(6.872)
El Paso, TX	9.263	4.906	1036.457	20.713	0.074
	(2.087)	(0.794)	(154.082)	(3.998)	(0.224)
Galveston, TX	6.904	1.997	940.678	1.835	0.128
	(3.031)	(0.165)	(149.496)	(0.668)	(1.079)
Houston, TX	5.317	30.412	940.678	3.791	3.782
	(2.430)	(5.510)	(149.496)	(0.969)	(8.975)
Laredo, TX	16.013	1.448	621.961	24.189	0.037
	(4.430)	(0.330)	(134.766)	(6.569)	(0.234)
McAllen, TX	14.252	2.946	621.961	22.815	1.030
	(4.765)	(0.743)	(134.766)	(6.316)	(2.370)
Odessa-Midland, TX	5.230	2.052	1036.457	3.399	0.125
	(2.748)	(0.340)	(154.082)	(0.319)	(0.669)
San Antonio, TX	5.668	11.520	940.678	6.203	1.369
	(1.621)	(1.486)	(149.496)	(2.580)	(3.491)
Victoria, TX	5.259	1.722	940.678	1.458	0.300
	(1.577)	(0.148)	(149.496)	(0.623)	(1.123)
Chicago, IL	6.398	73.705	2033.580	2.461	9.197
	(1.896)	(1.345)	(149.848)	(0.766)	(17.617)
Las Cruces, NM	7.715	1.029	1298.042	11.449	0.080
	(1.030)	(0.237)	(152.066)	(2.955)	(0.315)
Tucson, AZ	5.462	5.320	1238.160	4.378	0.153
	(1.437)	(1.003)	(160.834)	(1.437)	(0.816)
Phoenix, AZ	5.401	15.676	1238.160	2.592	0.762
	(1.368)	(3.832)	(160.834)	(0.836)	(2.612)
Denver-Boulder, CO	5.240	7.615	1605.164	1.328	0.240
	(1.468)	(2.462)	(144.481)	(0.438)	(0.616)
Reno, NV	5.629	1.923	1524.070	1.550	0.263
	(1.230)	(0.444)	(160.925)	(1.557)	(1.770)
Las Vegas, NV	6.909	5.533	1524.070	1.421	0.365
	(1.672)	(1.822)	(160.925)	(0.606)	(1.173)
Omaha, NE	4.398	1.766	1687.938	6.046	0.095
	(1.133)	(1.787)	(149.981)	(7.250)	(0.347)
New York City, NY	7.246	73.383	2596.999	0.205	0.375
	(2.371)	(2.263)	(129.604)	(0.188)	(1.413)
Washington D.C., WA	7.344	6.581	2386.269	0.085	0.059
	(1.922)	(0.489)	(132.258)	(0.033)	(0.296)
Miami, FL	6.954	16.210	1926.681	0.324	0.066
	(1.849)	(2.128)	(132.039)	(0.121)	(0.293)
Atlanta, GA	5.073	10.963	1749.061	0.208	0.042
	(1.586)	(0.681)	(132.803)	(0.282)	(0.315)
Total	7.103	13.351	1459.956	5.511	1.986
	(3.309)	(18.867)	(527.774)	(6.476)	(7.622)

Observations per US county: 1739; Total observations: 74777.

Appendix C: Descriptive Statistics by US Receiving County: Last Migration

	Unemployment Rate	Total Population (in 100.000)	Miles	Mexican Share of Population (in %)	Village Migration Experience (in %)
Imperial Valley, CA	10.019	10.887	1805.511	9.201	1.463
	(2.265)	(3.172)	(128.499)	(1.665)	(5.367)
Lower San Joaquin, CA	10.889	4.991	1805.511	7.142	0.625
	(2.350)	(0.870)	(128.499)	(1.391)	(1.179)
Middle San Joaquin, CA	11.381	6.092	1805.511	9.139	2.618
	(2.165)	(0.940)	(128.499)	(2.066)	(3.107)
Upper San Joaquin, CA	12.188	9.324	1805.511	6.776	4.410
	(2.032)	(1.689)	(128.499)	(1.479)	(10.385)
Salinas-Monterrey-Santa Cruz, CA	7.774	11.598	1973.522	7.932	2.798
	(1.684)	(1.591)	(128.294)	(0.689)	(4.014)
Sacramento Valley, CA	7.802	19.363	1973.522	3.429	2.095
	(1.773)	(3.299)	(128.294)	(0.575)	(2.715)
Ventura-Oxnard-Simi, CA	6.951	6.126	1586.325	8.177	2.670
	(1.375)	(0.914)	(128.400)	(1.111)	(5.449)
Santa Barbara, CA	5.831	3.436	1586.325	7.992	1.890
	(1.094)	(0.386)	(128.400)	(2.866)	(3.975)
Napa-Sonoma, CA	5.901	4.563	1973.522	3.264	1.463
	(1.557)	(0.668)	(128.294)	(1.340)	(4.713)
Los Angeles County, CA	6.949	83.756	1586.325	11.916	30.545
	(1.596)	(7.899)	(128.400)	(2.315)	(24.412)
Orange County, CA	4.602	22.118	1586.325	7.887	4.562
	(1.251)	(2.998)	(128.400)	(2.882)	(8.758)
San Francisco Urban Area, CA	5.233	35.446	1973.522	2.539	1.175
	(1.348)	(2.393)	(128.294)	(0.623)	(2.580)
San Jose Urban Area, CA	5.300	14.100	1973.522	4.421	2.213
	(1.077)	(1.289)	(128.294)	(1.173)	(6.033)
Riverside-San Bernardino, CA	7.476	22.107	1586.325	6.448	0.920
	(2.032)	(6.221)	(128.400)	(2.143)	(1.453)
San Diego County, CA	5.977	22.546	1586.325	6.794	5.801
	(1.454)	(3.890)	(128.400)	(1.690)	(15.218)
Rio Vista, CA	7.018	3.014	1973.522	1.930	0.087
	(1.393)	(0.657)	(128.294)	(0.299)	(0.328)
Abilene, TX	5.543	1.174	918.260	1.731	0.252
	(1.642)	(0.068)	(133.864)	(0.635)	(1.200)
Amarillo, TX	4.847	1.847	1157.821	1.616	0.107
	(1.194)	(0.146)	(134.139)	(0.886)	(0.521)
Austin, TX	4.752	8.237	918.260	2.098	0.177
	(1.248)	(1.510)	(133.864)	(0.721)	(1.066)
Beaumont-Port Arthur, TX	9.060	3.682	918.260	0.588	0.097
	(2.799)	(0.100)	(133.864)	(0.093)	(0.601)
Brownsville, TX	12.328	2.451	604.505	20.797	0.794
	(2.155)	(0.390)	(126.000)	(4.099)	(1.978)
Bryan-College, TX	4.071	1.121	918.260	1.817	0.020
	(1.094)	(0.201)	(133.864)	(0.650)	(0.082)
Corpus Christi, TX	8.169	3.465	604.505	3.613	0.223
	(2.152)	(0.229)	(126.000)	(1.414)	(1.188)

Appendix C: continued

Dallas-Ft.Worth, TX	5.114	36.443	918.260	3.214	3.142
	(1.164)	(5.902)	(133.864)	(1.324)	(8.955)
El Paso, TX	10.156	5.551	1012.657	20.871	0.059
	(1.460)	(0.772)	(133.075)	(2.936)	(0.153)
Galveston, TX	7.803	2.117	918.260	1.705	0.160
	(2.234)	(0.150)	(133.864)	(0.430)	(1.345)
Houston, TX	6.120	34.397	918.260	4.741	3.309
	(1.939)	(4.840)	(133.864)	(1.258)	(8.157)
Laredo, TX	16.912	1.721	604.505	24.742	0.024
	(3.217)	(0.331)	(126.000)	(4.995)	(0.138)
McAllen, TX	17.086	3.533	604.505	22.232	0.579
	(3.884)	(0.701)	(126.000)	(4.306)	(1.445)
Odessa-Midland, TX	6.289	2.229	1012.657	3.531	0.048
	(2.363)	(0.259)	(133.075)	(0.360)	(0.332)
San Angelo, TX	5.015	1.395	1012.657	4.359	0.111
	(1.137)	(0.105)	(133.075)	(0.760)	(0.472)
San Antonio, TX	6.147	12.716	918.260	5.716	0.815
	(1.364)	(1.427)	(133.864)	(1.677)	(2.190)
Victoria, TX	5.897	1.831	918.260	2.055	0.252
	(1.338)	(0.131)	(133.864)	(0.875)	(1.174)
Chicago, IL	6.521	74.754	2011.131	3.163	8.350
	(1.456)	(1.295)	(134.142)	(0.873)	(16.917)
Tucson, AZ	5.030	6.102	1215.511	4.733	0.111
	(1.246)	(0.934)	(128.499)	(1.250)	(0.551)
Phoenix, AZ	5.084	18.915	1215.511	3.198	0.402
	(1.110)	(3.869)	(128.499)	(1.038)	(1.352)
Denver-Boulder, CO	5.456	8.437	1583.517	1.680	0.211
	(1.098)	(3.109)	(127.860)	(0.695)	(0.422)
Pueblo, CO	8.673	1.244	1583.517	0.857	0.158
	(2.370)	(0.017)	(127.860)	(0.410)	(1.058)
Reno, NV	5.403	2.289	1501.444	3.058	0.368
	(1.006)	(0.438)	(128.591)	(2.160)	(2.306)
Las Vegas, NV	6.426	7.201	1501.444	1.976	0.251
	(1.501)	(2.100)	(128.591)	(0.718)	(0.934)
St. Louis, MO	5.614	14.196	1549.915	0.069	0.063
	(1.324)	(0.434)	(134.021)	(0.017)	(0.210)
Omaha, NE	3.992	1.551	1665.505	5.543	0.157
	(1.168)	(1.679)	(134.353)	(6.134)	(0.460)
New York City, NY	7.469	73.164	2578.455	0.390	0.265
	(2.047)	(1.464)	(120.590)	(0.266)	(1.148)
Washington D.C., WA	7.281	6.264	2367.655	0.114	0.053
	(1.733)	(0.393)	(123.929)	(0.034)	(0.256)
Miami, FL	7.307	17.999	1908.071	0.436	0.056
	(1.630)	(2.156)	(123.611)	(0.142)	(0.194)
Atlanta, GA	5.388	11.564	1730.293	0.484	0.068
	(1.075)	(0.725)	(124.443)	(0.460)	0.404)
Total	7.310	14.066	1431.984	5.568	1.870
	(3.413)	(19.216)	(510.941)	(6.163)	(7.563)

Observations per US county: 1561; Total observations: 71806.

Figure 1: Concentration of Mexican Migrants in the U.S.

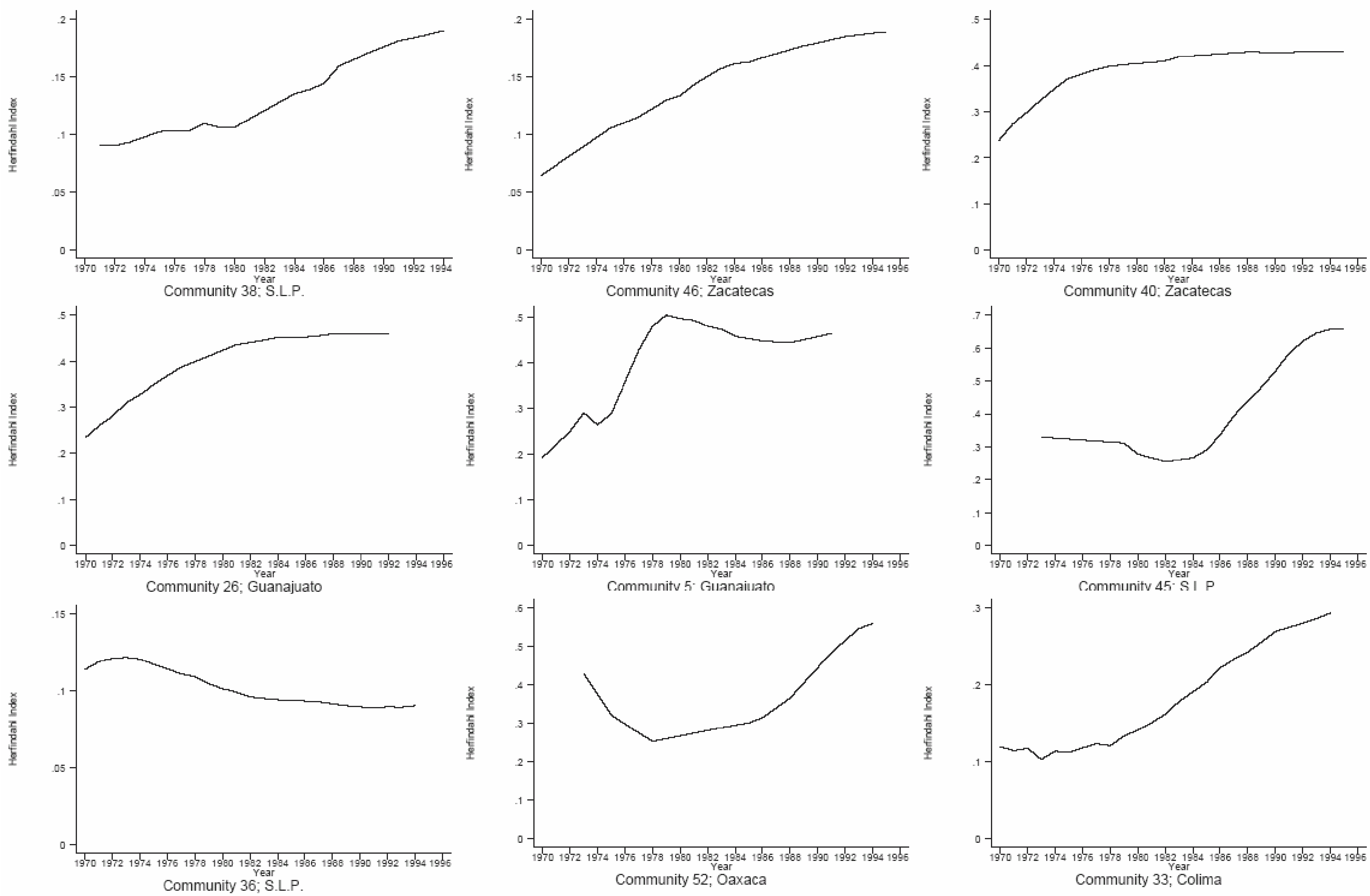

Figure 2:
Mexican Share of Population in Selected U.S. Locations