NETWORKS IN THE MODERN ECONOMY: MEXICAN MIGRANTS IN THE U. S. LABOR MARKET*

KAIVAN MUNSHI

This paper attempts to identify job networks among Mexican migrants in the U. S. labor market. The empirical analysis uses data on migration patterns and labor market outcomes, based on a sample of individuals belonging to multiple origin-communities in Mexico, over a long period of time. Each community's network is measured by the proportion of the sampled individuals who are located at the destination (the United States) in any year. We verify that the same individual is more likely to be employed and to hold a higher paying nonagricultural job when his network is exogenously larger, by including individual fixed effects in the employment and occupation regressions and by using rainfall in the origin-community as an instrument for the size of the network at the destination.

I. INTRODUCTION

Economists have taken a very favorable view of nonmarket institutions in recent years. The general perception is that these institutions emerge in response to market failure, harnessing social ties to avoid information, enforcement, and coordination problems. While nonmarket institutions may be more prevalent in developing countries, where market imperfections tend to be more severe and pervasive, a strong implication of this view is that these institutions should also be observed in those sectors of the modern economy in which markets function imperfectly.

In this paper I attempt to identify network effects among Mexican migrants in the U. S. labor market. While community networks serve many roles, my specific objective is to test whether the network improves labor market outcomes for its members. There is an old and extensive literature in labor economics that documents the importance of friends and relatives in providing job referrals (see Montgomery [1991] for a review).

* This project could not have been completed without the help of Payal Gupta and Judith Alejandra Frias, who collected the Mexican rainfall data. Nolan Malone and Mariano Sana at the Mexican Migration Project at the University of Pennsylvania patiently answered all my questions. I thank Aldo Colussi and George Mailath for many helpful discussions. Abhijit Banerjee, George Borjas, Esther Duflo, Andrew Foster, Lawrence Katz, Douglas Massey, Mark Rosenzweig, two anonymous referees, and seminar participants at Brown University, Columbia University, El Colegio de México, Harvard-MIT, Instituto Tecnológico Autónomo de México, the World Bank, and the University of Pennsylvania made very helpful comments on the paper. Nauman Ilias and Chun-Seng Yip provided superb research assistance. Research support from the University Research Foundation at the University of Pennsylvania and NIH grant R01-HD37841 is gratefully acknowledged. I am responsible for any errors that may remain.

The Quarterly Journal of Economics, May 2003

Within the labor market we would expect these network effects to be stronger in migrant communities [Borjas 1992]. Migrants are by definition newcomers in the labor market, and so will be more susceptible to the information problems that generate a need for job referrals in the first place. Migrant communities also tend to be more socially cohesive. The application that I have chosen would thus seem to be ideally suited to test for the presence of network effects in the U. S. economy.[1]

The bulk of the data used in this paper comes from the Mexican Migration Project (MMP), conducted jointly by researchers based in Mexico and the United States since 1982 (see Massey et al. [1987] for details of the study). In this project a small number of Mexican communities are surveyed each year. Each community is surveyed once only, and a retrospective history of migration patterns and labor market outcomes is obtained from typically 200 randomly sampled household heads. Setting aside recall and sampling issues for the time being, this leaves the econometrician with a panel data set of individual location decisions and labor outcomes, from multiple communities, over a long period of time.

The communities in the sample are drawn from a region in Southwestern Mexico that has traditionally supplied between half and three-quarters of the Mexican migrants to the United States [Bustamante 1984; Jones 1984]. Migration from this region tends to be *recurrent*: individuals move back and forth between Mexico and the United States, and only a small fraction settle permanently abroad. If the individual's network at the destination consists of other migrants from his origin-community, then this tells us that both the size and the vintage of the network will be changing over time. I use this variation within the community over time rather than across communities to estimate the network effects in this paper.

Using variation within each origin-community's network over time to identify network effects has two major advantages. First, the network at the destination is drawn from a well-defined

1. The fact that the majority of Mexican migrants (67 percent in the data) are undocumented would only reinforce the use of such informal recruitment channels. For interesting recent studies on social interactions in the U. S. labor market, and migrant networks, see Topa [2001] and Bertrand, Luttmer, and Mullainathan [2000], respectively.

and well-established social unit: the origin-community.[2] Massey et al. [1987] use both quantitative and ethnographic data to study network relationships among migrants. They find that most relationships are based on kinship, friendship, and in particular, *paisanaje* (belonging to a common origin-community). Ties among *paisanos* actually appear to strengthen once they arrive in the United States, and this sociological change is reinforced by the emergence of community-based institutions, such as soccer clubs, which bring the migrants together.

The second advantage of my estimation strategy is that the econometrician is in a position to control for both selectivity in the migration decision, as well as for the endogeneity of the network itself, in the employment regression. The individual migrant's network is measured by the proportion of sampled individuals in his community who are located at the destination (the United States), at each point in time. The basic specification of the regression equation includes the size of the network, the individual's unobserved ability, and unobserved labor market shocks, as determinants of the migrant's labor outcome in the United States. If migration is based on both the individual's ability as well as the size of the network at the destination, then changes in the size of the network will be associated with compositional change in the pool of migrants, biasing the estimated network effects. Since we have panel data, this selection bias can be corrected by including individual fixed effects in the employment regression, under the assumption that individual ability does not vary over time.

While fixed effects control for the individual's unobserved ability, network size could also respond to unobserved shocks in the U. S. labor market. For example, positive shocks at the destination could induce additional migration, biasing the network effect upward. Alternatively, improved labor market conditions could hasten the speed at which migrants achieve their target savings, increasing the rate of departure among the more established members of the network and biasing the network effects in the opposite direction. Individual fixed effects do not solve the problem in this case. What we need, to avoid this simultaneity bias, is a statistical instrument that determines changes in the size of the network but is uncorrelated with labor

2. In contrast, previous studies based in the United States have typically used administrative or census boundaries to define social units [Case and Katz 1991; Glaeser, Sacerdote, and Scheinkman 1996; Borjas 1995; Bertrand, Luttmer, and Mullainathan 2000; Topa 2001].

market shocks in the United States. A major innovation of this paper is the use of rainfall in the *origin-community* (collected from local weather stations) as an instrument for the size of the migrant network at the *destination*.[3]

The empirical analysis begins with employment status as the outcome of interest. After controlling for individual fixed effects and year effects, we find that current (period t) employment in the United States is negatively correlated with distant-past (period $t-3$ to $t-6$) rainfall in the individual's Mexican community. In contrast, recent-past (period t to $t-2$) rainfall has little effect on employment.

Why is an individual located in the United States more likely to be employed if rainfall in his Mexican origin community was low more than three years ago? Rain-fed agriculture is the major occupation in the Mexican origin-communities, and we will find a strong negative correlation between rainfall at the origin and (immediate) migration to the United States. The results (with fixed effects) just described tell us that lower than average rainfall more than three years ago would have induced greater than average migration at that time, which translates into a greater than average pool of established migrants today. These established migrants are able to provide job referrals for other members of the network, leading to higher than average employment levels among the migrants in the community.[4] This interpretation of the results will be later borne out in the corresponding Instrumental Variable (IV) employment regression as well, with distant-past (recent-past) rainfall instrumenting for established (new) migrants at the destination, where we see that it is the number of established migrants that determines employment levels in the network.

We complete the empirical analysis by replacing employment

3. As Manski [1993, 2000] has pointed out repeatedly, the fundamental problem with much of the literature on social interactions is its inability to control for correlated unobservables within the community, which would be the labor market shocks in this application. Recently, however, a number of papers have used an experimental approach to identify social effects [Katz, Kling, and Liebman 2001; Ludwig, Duncan, and Hirschfield 2001; Sacerdote 2001; Duflo and Saez 2002; Miguel and Kremer 2002]. Taking a similar approach, I use random rainfall variation to identify the network effects in this paper.

4. Low distant-past rainfall increases the number of older migrants in the network, which could in turn increase average employment levels if individuals are independently more likely to find jobs as they gain exposure at the destination. However, employment regressions presented later show that migrants who have just arrived at the destination are also more likely to be employed when distant-past rainfall is low, ruling out this alternative explanation.

status with the individual's occupation as the outcome of interest. Migrants in nonagricultural jobs earn substantially more than the agricultural workers in our sample: Their annual income (in 2001 US dollars) is on average $12,000, versus $8,700 for the agricultural workers. My strategy to establish that the network is actively channeling its members into preferred nonagricultural jobs is to investigate whether the *same* individual is more likely to hold a nonagricultural job when his network is exogenously larger, by including individual fixed effects in the occupation regression and by using rainfall at the origin as an instrument for the size of the network at the destination. We will see that low rainfall at the origin increases the probability that the migrant will be occupied in a nonagricultural job, but once again with a lag. The network not only finds jobs for its members, it also channels them into higher paying occupations.

The empirical analysis in this paper provides us with a first glimpse of a remarkable institution. The number of Mexican migrants in the United States is difficult to estimate since so many of them are undocumented, but 2.3 million Mexicans applied for the amnesty offered by the Immigration Reform and Control Act (IRCA) in 1986 [Bean, Vernez, and Keely 1989]. We would expect the number of Mexicans living in the United States at any point in time over the past couple of decades to be at least as high as that. Migration from Mexico tends to be recurrent—the typical migrant in our sample will spend three–four years in the United States before returning home after a single migration spell. This tells us that millions of individuals in Mexico must form the pool of workers that supplies low-skill labor to the United States.

How do these workers find jobs when they arrive? Our results tell us that it is the more established members of the network that provide most of the referrals and the support. In this decentralized equilibrium there are always enough established migrants at the destination, but it is a different group of individuals that provides this support from one period to the next. Thus, the migrant will typically be matched with a completely different group of individuals from his community on each trip to the United States. A very dense web of social ties must necessarily be in place for the network to function so well without repeated interactions between individuals at the destination.

Our results tell us that the network significantly improves labor market outcomes among its members. Unemployment levels among the migrants in the sample are quite low, around 4 percent. My most conservative estimates suggest that if we were to exogenously shut down the networks, but leave migration patterns unchanged, these levels would increase substantially, up to nearly 10 percent. Similarly, nonagricultural jobs account for 51 percent of all jobs at the destination. If we were to shut down the network, this statistic would decline to 32 percent. This is just a simple thought exercise; we would never expect to see such large changes in equilibrium since migration would decline in this case. These results nevertheless tell us that network effects are economically very significant, at least in the particular segment of the economy that we are looking at. While we are accustomed to thinking of social networks as being a feature of a developing economy, our results suggest that networks could play an important role in the modern economy as well.

The paper is organized in seven sections. Section II describes the institutional setting that the migrants operate in. Section III provides a motivation for the presence of networks in the labor market, and Section IV discusses the identification of network effects. Section V presents the estimation results with employment as the outcome of interest, while Section VI studies the choice between agricultural and nonagricultural jobs. Section VII concludes.

II. The Institutional Setting and the Data

Migration from Southwestern Mexico to the United States began in 1885, when the first rail line reached the region. This period coincided with the closure of labor migration from China and Japan, and Mexican workers were actively recruited, particularly in U. S. mining and agriculture, from the turn of the century onward. This trend continued over the first half of the twentieth century, and especially during the Bracero Accord (temporary work arrangement) from 1942 to 1964 [Cardoso 1980]. Four states in Southwestern Mexico—Jalisco, Michoacan, Guanajuato, and Zacatecas—accounted for 45 percent of all *bracero* migration between 1951 and 1962 [Craig 1971], and this region continues to supply the majority of Mexican migrants to the United States today [Durand, Massey, and Charvet 2000].[5]

5. The states in this region include Jalisco, Michoacan, Zacatecas, Colima, Aguascalientes, Nayarit, San Luis Potosi, and Guanajuato. With the exception of

In this section I use the Mexican Migration Project (MMP) data to describe the setting in which migration occurs in our communities. Each community in the MMP data set is surveyed only once, and retrospective information is collected from typically 200 household heads over a long period of time. Much of the analysis in this paper restricts attention to the fifteen years prior to the survey year in each community. Communities that display no change in employment over the sample period do not contribute to the identification of the network effects, since fixed effects are included in all the regressions in this paper. Excluding these communities, as well as communities for which rainfall data are unavailable, we are left with 24 communities in seven states: Jalisco, Guanajuato, San Luis Potosi (SLP), Michoacan, Zacatecas, Nayarit, and Colima.

In the discussion that follows, I study the characteristics of these origin communities, the pattern of settlement in the United States, the nature of migrant activity, and the role of the network in providing employment in the United States, separately by state. The person-year is typically treated as the unit of observation, and we will compute descriptive statistics over the full sample period (the fifteen years prior to the survey-year in each community), for community-years in which rainfall data are available with a six-year lag, to be consistent with the regressions reported later. While we often use all the available person-years, in some cases we restrict attention to observations at home in Mexico, or abroad in the United States. Some of the descriptive statistics will also be computed with the community-year as the unit of observation. The patterns that I describe below match well with other studies, mostly by anthropologists and sociologists, which have been conducted in the area.

II.A. Economic Conditions at the Origin

We begin in Table I, Panel A, with the basic characteristics of the individuals in our sample. Note that at this point we are using person-years in the United States and in Mexico. We see that the household heads tend to be in their forties, over the sample period. Most are married, and fertility rates appear to be fairly high. Notice that education levels are very low, just five years of

Aguascalientes, all the other states are represented in our sample of communities, which I describe below.

TABLE I
DESCRIPTIVE STATISTICS AT THE ORIGIN (IN MEXICO)

Origin state:	Full sample	Jalisco	Guanajuato	SLP	Michoacan	Zacatecas	Nayarit	Colima
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: individual characteristics								
Age	41.95	40.30	43.64	47.57	41.94	40.97	40.58	39.28
	(16.55)	(16.08)	(16.53)	(15.69)	(15.84)	(16.74)	(17.20)	(16.36)
Education	5.06	3.89	5.12	6.88	3.57	5.40	5.39	5.47
	(4.69)	(3.85)	(4.57)	(5.21)	(4.45)	(4.81)	(5.07)	(4.57)
Marital status	0.76	0.75	0.79	0.79	0.86	0.74	0.67	0.67
	(0.43)	(0.43)	(0.41)	(0.40)	(0.34)	(0.44)	(0.47)	(0.47)
Number of children	4.15	4.41	4.44	4.10	4.21	3.97	3.60	3.67
	(3.47)	(3.82)	(3.38)	(3.05)	(3.34)	(3.39)	(3.30)	(3.49)
Number of observations	48,386	11,221	8,600	5,030	3,560	14,776	3,000	2,199
Panel B: occupational choice (%)								
Professional	8.26	3.04	8.82	15.97	4.30	11.06	6.95	7.08
Sales person	13.43	12.61	16.21	16.49	13.16	11.94	15.11	7.08
Skilled manual	17.90	18.33	26.79	18.76	12.61	12.45	20.52	16.34
Unskilled manual	11.48	13.06	13.87	8.55	6.49	12.25	9.80	6.08
Agriculture	30.88	35.80	17.72	23.56	48.34	31.88	28.44	42.77
Service	12.72	11.72	12.25	12.00	11.56	13.67	11.83	18.53
Other	5.33	5.44	4.34	4.67	3.54	6.75	7.35	2.12
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Panel C: fraction of land irrigated								
Fraction	0.18	0.11	0.36	0.16	0.49	0.17	0.11	0.00
	(0.36)	(0.26)	(0.46)	(0.35)	(0.48)	(0.37)	(0.28)	—
Number of observations	6538	1857	567	696	396	2540	206	276

Standard deviations are in parentheses.
The sample covers a fifteen-year period prior to the survey-year in each Mexican community.
Each observation is a person-year.
Panel A covers all available person-years. Panel B is restricted to observations in which the individual locates at the origin in a given year.
Panel C applies to landowners only.
Marital status = 1 if married, 0 otherwise.

schooling on average, which suggest immediately that employment opportunities in the United States will be limited to low-skill jobs. All of these patterns appear to be uniform across the sending states.

Turning next to the occupational patterns at the origin in Panel B, based on person-years in which individuals are located at home over the sample period, we see that agriculture is the main occupation in all the states except Guanajuato. That state has a tradition of silver craftsmanship and leatherwork (around the city of Leon), which may also explain the importance of "Skilled Manual" in column (3).[6] Southwestern Mexico is relatively undeveloped, and given the low education levels that we saw above, it is not surprising that agriculture and manual labor are the dominant activities in the origin communities. Notice also, from Panel C, that the fraction of irrigated land tends to be very low in these communities, which suggests that there has been very limited investment in agriculture. For the purpose of our statistical analysis, the observed dominance of rain-fed agriculture in the local economy is fortunate, since this suggests that migration is very likely to respond to rainfall shocks at the origin.[7]

II.B. Employment and Location Patterns at the Destination

We saw in Table I that education levels in the sample were very low, and that the main occupations in the origin communities were agricultural work and manual labor. Restricting attention now to person-years in which individuals are located at the destination, in Panel A of Table II, we would predict a similar occupational profile in the United States as well. As expected, agriculture is the dominant occupation (except for migrants from San Luis Potosi), followed by unskilled manual labor. These are low-skill activities associated with little human capital accumulation on the job, which supports the view that I take later in the paper that the migrant's ability in the United States is effectively constant over time.

Turning to location patterns in Table II, Panel B, we see that the migrants in our communities end up at a fairly limited num-

6. Note that the results that I report later in the paper are robust to the exclusion of Guanajuato from the sample.

7. The coefficient of variation for rainfall within a community, averaged over all 24 communities, is 0.21. We will see later that this variation is sufficient to identify the network effects of changes in the level of migration over time.

TABLE II
DESCRIPTIVE STATISTICS AT THE DESTINATION (IN THE UNITED STATES)

Origin state:	Full sample	Jalisco	Guanajuato	SLP	Michoacan	Zacatecas	Nayarit	Colima
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: occupational choice (%)								
Skilled manual	12.01	10.66	16.36	15.50	7.91	12.10	9.89	5.56
Unskilled manual	24.88	20.02	30.48	64.19	21.43	18.35	21.98	23.33
Agriculture	48.90	53.81	39.03	12.66	56.63	54.47	57.69	51.11
Service	10.61	10.66	9.48	2.40	12.76	12.01	9.89	18.89
Other	3.60	4.85	4.65	5.25	1.27	3.07	0.55	1.11
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Panel B: location patterns (%)								
San Joaquin Valley	19.90	6.28	5.81	35.52	30.40	24.61	6.77	24.73
San Francisco	5.97	8.30	4.36	2.05	21.07	3.95	2.08	1.08
Los Angeles	32.63	37.55	23.00	12.73	29.60	36.46	35.42	48.39
San Diego	7.89	26.92	5.57	1.23	6.40	1.64	6.25	0.00
Houston	3.81	1.21	5.08	15.61	2.40	2.22	0.00	0.00
Dallas	3.83	2.63	14.04	10.27	0.27	1.97	0.00	1.08
Chicago	5.15	7.49	17.68	14.58	0.53	0.43	4.69	0.00
Other urban	5.56	3.04	10.17	2.67	3.20	6.21	12.50	7.53
Other rural	10.96	2.63	7.75	4.31	3.47	17.10	29.17	4.30
Other	4.30	3.95	6.54	1.03	2.66	5.41	3.12	12.89
Total	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Number of observations	4624	988	413	487	375	2076	192	93

Standard deviations are in parentheses.
The sample covers a fifteen-year period prior to the survey-year in each Mexican community.
Each observation is a person-year.
We restrict attention to observations in which the individual is located at the destination.

ber of U. S. destinations over the sample period.[8] California is clearly the dominant destination region, and within that state, Los Angeles and to a lesser extent the San Joaquin Valley and San Diego attract the most migrants. However, notice the enormous variation *across origin states* in Panel B. As an example, take the second destination zone, San Francisco: 21 percent of the migrants from Michoacan locate there, yet the proportion of migrants from the other six states that locates there never exceeds 8 percent. To take another example, 27 percent of the migrants from Jalisco and only 1 percent of the migrants from San Luis Potosi (SLP) settle in San Diego. When it comes to locating in Houston, this pattern is reversed: 1 percent of the migrants from Jalisco and 16 percent of the migrants from SLP settle there. We saw in Table I that individual characteristics are fairly uniform across the origin states, which are all located in one region of Mexico, yet the wide variation in location patterns in the United States continues to be observed as we move down from row to row in Panel B, consistent with the view that historical accident may often play an important role in the formation of community-based migrant networks.[9]

Previous versions of the paper also described the location pattern within each community in some detail. Individual communities do not channel all their migrants to a single destination zone. Instead, it appears that the community establishes itself at a limited number of destination zones in the United States (typically three), with a tight spatial concentration within each zone (on average 90 percent of the migrants from the community locate in the same SMSA). I do not formally model the spatial distribution of the community network in this paper, but will take account of these patterns later in the estimation section.

II.C. Individual Migration Patterns

In the preceding section we studied how communities locate themselves in the United States. We now turn our attention to individual migration patterns over the sample period. I begin with the most basic migration statistics in Panel A of Table III. Roughly 12 percent of the observations in each community-year

8. The number of observations in Panel B is slightly lower than what we use later in the employment regressions at the destination because the exact location in the United States is missing for a few migrants.

9. Carrington, Detragiache, and Vishwanath [1996] describe similar migration patterns in the United States during the Great Black Migration.

TABLE III
INDIVIDUAL MIGRATION PATTERNS

Origin state:	Full sample	Jalisco	Guanajuato	SLP	Michoacan	Zacatecas	Nayarit	Colima
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: migration and employment								
% migrants	11.47	10.51	6.79	11.90	12.37	17.71	6.92	4.64
	(8.28)	(7.65)	(3.69)	(7.58)	(4.20)	(10.18)	(1.33)	(1.10)
% new migrants	5.17	5.49	3.09	4.31	5.87	7.12	4.18	2.82
	(3.78)	(3.97)	(2.00)	(3.25)	(2.77)	(4.61)	(1.46)	(0.86)
% established migrants	6.31	5.02	3.70	7.59	6.50	10.58	2.75	1.83
	(5.19)	(4.27)	(2.10)	(4.57)	(1.93)	(6.57)	(0.64)	(0.51)
% employment in the United States	95.66	96.38	96.35	92.95	97.40	95.66	92.39	95.83
% employment in Mexico	86.48	90.48	87.07	82.01	90.17	84.23	82.50	88.01
Panel B: individual migration patterns over the sample period								
Avg. number of trips	1.35	1.45	1.22	1.30	1.29	1.38	1.34	1.21
	(0.69)	(0.77)	(0.54)	(0.56)	(0.59)	(0.73)	(0.69)	(0.63)
Avg. duration at destination (years)	3.57	3.36	2.59	4.16	3.10	4.08	3.18	2.98
	(3.57)	(3.53)	(2.41)	(3.47)	(2.92)	(4.04)	(3.35)	(3.09)
% with 1 trip	74.50	69.23	84.21	74.26	75.25	72.92	76.00	85.71
% with 2 trips	17.85	19.66	9.87	22.77	21.78	18.29	16.00	10.71
% with 3 trips	5.83	8.12	5.92	1.98	0.99	6.94	6.00	3.57
% with 4 trips	1.55	2.99	0.00	0.99	1.98	1.16	2.00	0.00
% with 5 trips	0.27	0.00	0.00	0.00	0.00	0.69	0.00	0.00
Number of observations	1098	234	152	101	101	432	50	28

Standard deviations are in parentheses.
Panel A begins by computing the fraction of the community that is located at the destination in each community-year.
An individual is located at the destination if he spent more than one month in the United States in a given year.
A new migrant refers to a person-year in which the individual was located continuously at the destination for less than three years.
Established migrants are located continuously at the destination for three or more years.
Subsequently, the employment rate is computed for person-years in which individuals are located at the destination and the origin, respectively.
The individual is employed if he works more than one month in the year.
Panel B focuses on the 1098 individuals who migrate at some point during the sample period. There are 4450 individuals in the full sample.
Note that the individual is now the unit of observation.

over the sample period are located in the United States. The MMP data are coded so that the United States is listed as the location for any person-year in which the individual spent longer than one month at the destination. Thus, a migrant with seasonal employment who returns home for a few months each year is still treated as being at the destination continuously. About 55 percent of the migrants are "established" migrants, where an established migrant is defined as a worker who has located continuously at the destination for three or more years (a justification for the three-year cut-off is provided later in Section V). Finally, the MMP data are coded so that an individual listed as being located at the destination in a given year, is also listed as being employed if he held a job for at least one month in the United States. The unemployment rate, for person-years in which the individuals are located at the destination, is just over 4 percent over the sample period.[10] In contrast, the corresponding unemployment rate in Mexico is nearly 14 percent. Looking across columns in Panel A, notice that there is considerable variation across origin states in these statistics.

Turning next to Panel B, we first focus on individuals who migrate at some point during the sample period. The average number of trips is well over one, and the average duration at the destination is roughly 3.5 years. This tells us immediately that there must be considerable movement back and forth between the origin and the destination, despite the fact that most of the migrants are undocumented (67 percent of the person-years in our sample). Looking at these migration patterns more closely, while the majority of the migrants make a single trip to the destination over the sample period, a substantial fraction make two trips, and three, four, and even five trips (over a fifteen-year period) are seen in the data.[11]

10. Unemployment rates among the migrants in our sample appear to be very low, perhaps because they travel to the United States specifically to work. They are also drawn from a region in Mexico that has supplied short-term workers to the United States for nearly a century, so labor market networks in these communities are likely to be well established, with correspondingly favorable employment outcomes.

11. In previous versions of the paper we also reported how individual migrants locate within the United States over multiple migration spells. Only about 54 percent of these migrants return to the same destination zone on each spell over the sample period. Individuals do not appear to form lasting ties directly with their employers in the United States. Instead, return migrants seem to take full advantage of the multiple locations that their communities establish at the destination to improve their employment prospects.

II.D. Job Search at the Destination

The literature in labor economics and sociology is replete with references to the importance of friends and relatives in finding employment in the U. S. labor market, across occupational categories and ethnic groups. For example, Rees [1966], in an early study set in Chicago, found that informal sources account for about 50 percent of all hires in four white-collar occupations, and 80 percent of all hires in eight blue-collar occupations. Similarly, Holzer [1988] found that friends and relatives were the two most frequently used methods for finding employment in the 1981 panel of the National Longitudinal Survey of Youth (NLSY). The same job search patterns have been obtained, with remarkable regularity, in study after study of the U. S. labor market (Montgomery [1991], provides a summary).

Turning to migrant communities, we would expect the importance of social ties in the job search process to be even stronger in these groups. Certainly, the received evidence overwhelmingly supports the view that friends and relatives, and particularly those who belong to a common origin-community, are the main source of information about jobs. Chavez [1992, p. 136], for instance, tells the story of an undocumented Mexican migrant: "Leonardo shared an apartment with seven other friends, all *paisanos* from Sinaloa. Seven of the eight friends worked as gardeners. The first two friends had been in the area for five years, and provided referrals for employers for each of the subsequent migrants, the last of whom migrated two years earlier." Over 70 percent of the undocumented Mexicans, and a slightly higher proportion of the Central Americans, that Chavez interviewed in 1986 found work through referrals from friends and relatives. Similar patterns have been found in contemporary studies of Salvadoran immigrants [Menjivar 2000], Guatemalan immigrants [Hagan 1994], Chinese immigrants [Nee 1972; Zhou 1992], as well as historically during the Great Black Migration [Gottlieb 1991; Grossman 1989; Marks 1989].

Direct evidence from the MMP accords perfectly with this referral-based view of the job search process. The household heads in our sample were asked how they obtained employment on their last visit to the United States; individual search (23 percent), relatives (35 percent), and friends or *paisanos* (35 percent), account for the bulk of the jobs that were obtained. If we include relatives, friends, and *paisanos* in the network, then it is clear that social ties play a significant role in obtaining employment among the migrants in the sample.

III. Networks in the Labor Market

My main objective in this section is to discuss conditions under which networks emerge in the labor market, and to suggest ways in which these networks function. Some simple testable implications of network effects emerge from this discussion, which also leads naturally to the discussion on the identification of network effects that follows in Section IV. This section is based for the most part on a model of labor market networks that was laid out in some detail in previous versions of the paper (available from the author). Only one type of job is available to workers in that model: the individual is either employed or unemployed. I will relax this assumption at the end of this section since occupational choice plays such an important role in the empirical analysis.

III.A. Why Do Networks Emerge?

To generate a role for social networks in the labor market, we must begin with a positive level of unemployment in equilibrium, which could for instance be generated by exogenous job turnover. The type of activities that our migrants are employed in, such as agriculture and manual labor, are associated with frequent shifts in demand, so job turnover is likely to be fairly high in this setting.

While job turnover will generate a positive level of unemployment, it does not by itself motivate the emergence of a community-based network. For that, we must introduce some sort of information problem in the labor market. Here one way to proceed would be to consider a model of costly search, in which unemployed workers benefit from information about newly available jobs that they receive from the employed members of their network [Carrington, Detragiache, and Vishwanath 1996].

Alternatively, we could shift the information problem to the firm. Suppose that the firm is unable to identify a freshly hired worker's ability. If we make the usual assumption that the firm is unable to specify a performance-contingent wage contract, then it would always prefer to hire a high ability worker when a new position becomes available.[12] The firm could choose to enlist the help of one of its incumbent workers in this case, to recruit able workers from his network (as in Montgomery [1991]). The discussion that follows will restrict attention to this adverse selection model, since unobserved ability plays such an important role in the identification of network effects.

III.B. How Do Networks Function?

The simplest model of labor market networks with adverse selection treats the composition of the network as exogenously given. Assuming that ability is positively correlated within a network, the proportion of high ability workers will be higher on average in the incumbent high ability worker's network, as compared with the corresponding proportion in the market as a whole. At least some firms will use referrals in this case, drawing randomly from the unemployed members of the incumbent worker's network, instead of drawing from the pool of (all) unemployed workers in the market.

We could imagine instead that the incumbent worker has better information than the firm about the ability of individuals in his network. This information asymmetry would also generate a role for referrals, with the incumbent worker searching purposefully for high ability workers from his network. We could relax the assumption that the composition of the network is exogenous in this case, although this would be a more complicated model to solve. In addition, we would need to ensure that the incumbent worker has an incentive to refer the ablest individual from his network's unemployment pool to the firm (see Saloner [1985] for an analysis of this problem).

III.C. Who Contributes to the Network?

Focusing now on migrant networks, we would expect that it is the older migrants, those who have been at the destination longer, who contribute disproportionately to the network.

If migrants arrive at the destination without a job, then

12. Piece-rate contracts are rarely used in the U. S. economy, and among the occupations that our migrants are employed in only agriculture is associated with the use of such incentive schemes. Data from the 1997–98 National Agricultural Workers Survey [U. S. Department of Labor 2000] suggest that only 20 percent of agricultural workers are paid piece rates, with a slightly higher figure (25 percent) for certain crops such as fruits, nuts, and vegetables. With about 50 percent of our migrants engaged in agriculture, these statistics tell us that only about 10 percent will face piece-rate contracts.

employment levels will be increasing in their duration at the destination, as they gradually escape from the unemployment pool (a detailed characterization of these employment dynamics was provided in an earlier version of the paper). Older migrants provide more referrals in this case simply because they are more likely to be employed.

Further, among the employed migrants, older migrants will on average have been employed longer by the firms that hired them. These workers will presumably have risen within the organizational hierarchy, or accumulated a firm-specific reputation over time, and so have more to lose if they are separated from their firms. The threat of separation, which helps ensure that the incumbent worker only refers the ablest available workers from his network, consequently has greater bite for the older workers. This tells us in turn that the firm will be more likely to use referrals from such workers in equilibrium. Older workers contribute more to the network in this case not necessarily because they are more likely to be employed, but rather because they are employed longer on average.

III.D. Who Benefits from the Network?

Evidently, it is individuals who would otherwise be unemployed who benefit most from the network.[13] When firms draw randomly from the incumbent worker's network, it is low ability workers in networks with a large proportion of high ability workers that benefit most from the referrals.

When incumbent workers search purposefully for high ability recruits, only high ability workers will be referred in equilibrium. Now, it is individuals with unfavorable *observed* characteristics, competent older migrants and women for instance, who will benefit most from the network.

III.E. Introducing Multiple Occupations

Up to this point, we have assumed that there are only two labor market outcomes: the individual is either employed or un-

13. This need not be true if ability and network effects are complements, or if individuals could self-select into networks. In that case, high ability workers could end up benefiting more from their network. The origin-community exogenously determines the boundaries of the network in this application, and the low-skill jobs that the migrants are employed in would seem to rule out the complementarity assumption.

employed. However, all of the preceding discussion would still apply if multiple occupations were available in the labor market.

For example, suppose that two occupations—higher paying nonagricultural jobs and lower paying agricultural labor—are available. The network would now try to channel its members into the higher paying nonagricultural jobs. Individuals are more likely to occupy these coveted positions as they gain exposure at the destination, so the more established members of the network would also be better positioned to provide nonagricultural referrals and channel individuals into preferred occupations. Here again it would be individuals less likely to find nonagricultural jobs on their own, those who are less educated for example, who would benefit most from the network.

Once we allow for multiple occupations, individuals might wait to receive a preferred job, and larger networks could in principle be associated with *lower* levels of employment. However, as long as switching jobs is sufficiently easy (which would seem to be the case for the kinds of jobs that our migrants hold), a larger network should improve employment outcomes *and* channel individuals into preferred occupations.

IV. Identifying Network Effects

My objective in this section is to discuss the biases that arise with the estimation of network effects. I make three assumptions to simplify the exposition, all of which will be relaxed later. First, there are only two possible labor market outcomes: the individual is employed or unemployed. Second, each individual works for two periods only. Third, he makes an irreversible location decision at the beginning of his working life: he must choose between the "origin" (his Mexican community) and the "destination" (the United States). This location decision will depend on the returns at the origin and the destination over the next two periods, so our first task will be to describe these returns.

Begin with the employment outcome at the destination, which is in general determined by the migrant's ability, his duration at the destination, the network effect, and employment shocks at the destination. Leaving aside the individual's duration at the destination for the time being, the employment outcome for individual i in period t can be expressed as

$$\Pr(E_{it} = 1 | X_{it} = 1) = \beta X_{t-1} + \omega_i + C_t, \tag{1}$$

where $E_{it} = 1$ if the individual is employed, $E_{it} = 0$ otherwise. $X_{it} = 1$ if the individual chooses to work at the destination, $X_{it} = 0$ otherwise. X_{t-1} is the measure of migrants from his origin community who moved to the destination in the previous period.[14] Individuals work for two periods, and we take it that they provide referrals only in the second year of their working life, so a single cohort provides referrals in each period. ω_i is an idiosyncratic ability term which does not vary over time. C_t is an employment shock that is common across individuals in the community but varies over time. Both ω_i and C_t are unobserved by the econometrician, and we will see below that it is these terms that create problems for consistent estimation of the network effects, in the employment regression, by being correlated with X_{t-1}.

The corresponding expression for the individual's employment outcome in period $t + 1$, $\Pr(E_{it+1} = 1 | X_{it+1} = 1)$ is obtained by replacing X_{t-1} with X_t, and C_t with C_{t+1}. X_t, C_{t+1} are unobserved by the individual when he chooses his location at the beginning of period t, and we will see in a moment that this will complicate his migration decision slightly.

Turning to the returns at the origin, we assume that the individual will be employed in the traditional activity (farming). Returns from farming depend on the weather, but not on the individual's ability:

$$\Pi_{it} = \Pi(Z_t), \tag{2}$$

where Π_{it} is the economic return at the origin and Z_t is the rainfall in period t. Introducing other determinants of Π_{it} in equation (2) would not affect the discussion that follows in any way. All that we require is a single variable that determines the location decision exclusively through the returns at the origin, to later use as an instrument for the size of the network. As before, the expression for the returns in period $t + 1, \Pi_{it+1}$, is obtained by simply replacing Z_t with Z_{t+1}. When computing these returns,

14. If we took the model laid out in the previous section seriously, then it is only employed individuals who can provide referrals, and so the relevant network size should be the measure of *employed* migrants at the destination. However, we will see later that the network provides other support, such as financial assistance and housing, as well. So it would seem more appropriate to use the measure of migrants, regardless of their employment status, as the size of the network. The discussion on identification would follow through with either network measure, and I will later verify that the estimated network effects are robust to the method used to measure the network.

the individual must account for the fact that Z_{t+1} is unobserved at the beginning of period t.

Normalizing so that the wage at the destination is unity, the individual will locate at the destination if the expected return there over the next two periods, net of moving costs, is higher than the expected return at home.[15] The complication that arises immediately in this case is that the returns at the destination in period $t + 1$ depend on the measure of migrants that moves with the individual in period t, X_t, so there is a strategic element to the individual's location decision. In this case it is easy to verify that a migration equilibrium for the cohort that starts working in period t is characterized by a threshold ability $\underline{\omega}$, such that all individuals with ability greater than $\underline{\omega}$ will choose to locate at the destination.

Under what conditions will a unique interior solution for $\underline{\omega}$, in which a positive fraction of the cohort locates at the destination in each period, be obtained? To begin with, we must consider the coordination problem that could arise when a sufficiently large fraction of migrants is required to sustain a viable network at the destination: everyone in the cohort could choose to remain at home in that case. To rule out this possibility, we need to assume that a few of the highest ability individuals in each cohort will always migrate, regardless of the (expected) size of the network at the destination.

Once migration has been initiated, we must then consider the possibility that the entire cohort could "tip over" and locate at the destination (as in Carrington, Detragiache, and Vishwanath's [1996] characterization of migration with endogenous moving costs). Starting with the highest ability migrant in a cohort, each additional migrant will trade off the improvement in the performance of the network, as a consequence of his own migration decision, with his lower ability (relative to the migrant before him). An interior solution for the threshold ability $\underline{\omega}$ will be obtained as long as the decline in ability for each successive migrant sufficiently dominates the improvement in the performance of the network as it expands.

Finally, we need to rule out "bumps" in the distribution,

15. We could think of another decision rule in which the individual waits for a job opening at the destination (obtained through his network), before migrating. But it is difficult to imagine that the individual would be able to get from his home to the border, cross the border (most likely illegally), and then get to the job destination in time to fill the position.

which could give rise to multiple equilibria. A uniform ability distribution, together with the conditions described above, ensures that a unique interior solution for the threshold ability $\underline{\omega}$ will be obtained with each cohort.

Once we have characterized the migration equilibrium, each individual's location decision is relatively easy to describe:

(3)

$$X_{it} = 1 \text{ if } \omega_i \geq \underline{\omega}(X_{t-1}, C_t, E_t(C_{t+1}), Z_t, E_t(Z_{t+1})),\ X_{it} = 0 \text{ otherwise},$$

where $E_t(C_{t+1})$ is the predicted employment shock in period t + 1, and $E_t(Z_{t+1})$ is the predicted rainfall at the origin in period t + 1. $E_t(C_{t+1}), E_t(Z_{t+1})$ will in general be determined by the entire history of employment shocks and rainfall shocks, up to period t. Favorable conditions at the destination can support lower ability migrants, so $X_{t-1}, C_t, E_t(C_{t+1})$ will be negatively correlated with $\underline{\omega}$. In contrast, only high ability individuals migrate when rains are plentiful at the origin, so $Z_t, E_t(Z_{t+1})$ will be positively correlated with $\underline{\omega}$.

Let the distribution of ability in any cohort be characterized by the function F. In that case, the measure of migrants in period t is given by the expression,

$$X_t = 1 - F(\underline{\omega}(X_{t-1}, C_t, E_t(C_{t+1}), Z_t, E_t(Z_{t+1}))). \tag{4}$$

Working back one period, we can derive the corresponding expression for X_{t-1}:

$$X_{t-1} = 1 - F(\underline{\omega}(X_{t-2}, C_{t-1}, E_{t-1}(C_t), Z_{t-1}, E_{t-1}(Z_t))). \tag{5}$$

We are now in a position to discuss the bias in the estimated network effects in equation (1) that arises due to the unobserved ω_i, C_t terms. Starting with the employment shock, we noted above that favorable conditions at the destination are associated with a lower ability threshold $\underline{\omega}$. Thus, high C_{t-1} in equation (5) is associated with a lower $\underline{\omega}$, and hence more migration X_{t-1}. If the employment shocks are (positively) serially correlated, then X_{t-1} will be positively correlated with C_t in equation (1). This is a standard simultaneity problem that plagues the identification of social effects in general, biasing the β estimate upward.

My solution to the simultaneity problem is to instrument for X_{t-1} in the employment regression. A valid instrument in this setting would determine X_{t-1}, while remaining uncorrelated with C_t or other direct determinants of employment. A natural candidate that would appear to satisfy this condition, from equation (5), is rainfall at the origin. Low Z_{t-1} reduces returns at the origin, and by extension $\underline{\omega}$, which increases X_{t-1}. We would expect local rainfall shocks at the origin and employment shocks at the destination to be uncorrelated since the origin communities in Mexico are located very far from their U. S. destinations. Rainfall shocks at the origin and the destination, for each community, are completely uncorrelated (the correlation coefficient is 0.01). Each community is also too small to affect the level of employment at the destination through changes in its migration patterns. Z_{t-1} thus appears to be a valid instrument for X_{t-1}.

Changes in employment at the destination induce changes in location patterns, and we could in principle have estimated a migration regression, corresponding to equation (1), with the migration decision rather than the employment outcome as the dependent variable. The hypothesis in this case would be that a larger network at the destination induces additional migration from the origin. The problem with estimating this alternative regression is that while lagged rainfall Z_{t-1} may determine the size of the network at the destination, it could also directly determine the individual's migration decision by affecting current employment outcomes at the origin (for example, if local institutions that determine access to credit and other production inputs respond slowly to past rainfall shocks). We will see later that lagged rainfall does in fact directly determine current employment outcomes at the origin, and hence the individual's migration decision, which rules out its use as an instrument for the network in the alternative migration regression. Rainfall at the *origin* is a valid instrument for X_{t-1} in the employment regression precisely because we are restricting attention to activity at the *destination*.

While the use of rainfall as a statistical instrument may solve the simultaneity problem in this setting, we must still account for selectivity bias associated with the unobserved ability term ω_i in equation (1). $E(\omega_i|X_{it} = 1) = \phi(\underline{\omega}(X_{t-1},C_t,E_t(C_{t+1}),Z_t, E_t(Z_{t+1})))$, where ϕ is an increasing function of the threshold ability $\underline{\omega}$. We noted earlier that an increase in X_{t-1} improves conditions at the destination, lowering $\underline{\omega}$. Thus, ω_i will in general be *negatively* correlated with X_{t-1}. Intuitively, more favorable conditions at the destination lower the (unobserved) quality of the migrants, biasing $\hat{\beta}$ *downward.*[16] Instrumenting for X_{t-1} does not solve the selection problem since $\underline{\omega}$ is correlated with Z_{t-1} (through X_{t-1}).

My solution to the selection problem is to treat ω_i as an individual fixed effect in the employment regression. The implicit assumption here is that the individual's ability does not vary over the sample period. This seems to be reasonable, given the low-skill occupations that the individuals are engaged in, both in Mexico and the United States.

Before turning to the estimation results, I close this section with some extensions to the discussion on identification.

1. **Multiple work periods and return migration.** Up to this point we have assumed that each individual works for two periods, and makes an irreversible location decision at the beginning of his working life. We will now proceed to relax both these assumptions.

The most obvious change in the employment regression, equation (1), is that multiple cohorts will now provide referrals in each period. Following the discussion in Section III, it is the older cohorts (who have been at the destination longer) who will contribute more to the network.

While the size of a cohort continues to be determined by rainfall at the origin, each cohort no longer responds exclusively to a single rainfall lag. For example, consider a network with two cohorts, X_{t-1} and X_{t-2}. Z_{t-1} directly determines X_{t-1}, as we saw earlier.[17] However, Z_{t-2} now also determines X_{t-1}, through its effect on X_{t-2} which in turn determines migration in period $t-1$. By the same sort of argument, while Z_{t-2} directly determines X_{t-2}, Z_{t-1} also plays a role in determining the size of this cohort

16. This result does not necessarily hold once we allow returns at the origin to depend on the individual's ability at the origin, which we denote by η_i. The individual continues to make his location decision based on the returns at the origin, equation (1), and the returns at the destination, where equation (2) now includes an additive η_i term. But now it is the ability *differential* $\omega_i - \eta_i$ that determines which individuals migrate. As Borjas [1987] points out, the nature of the selection bias is ambiguous in this case. If the ability differential is systematically larger for low-ω individuals, then it is the low-ω individuals who would be the first to migrate, and unobserved selectivity would bias the network effects upward rather than downward as previously described. None of this matters, of course, as long as the individual fixed effects account for the unobserved selectivity in the employment regression.

17. The individual's location decision, equation (3), is essentially unchanged, except that he makes this decision at each point in his working life once we allow for return migration. While this decision continues to be forward looking, the worker must now account for the possibility that he could return to the origin in the future.

by affecting the level of return migration in period $t-1$. Although these cross-period rainfall effects will complicate the interpretation of the rainfall coefficients in the first-stage migration regressions reported later, notice that we continue to have a sufficient number of instruments for the employment regression, with one rainfall lag for each cohort.

The preceding discussion can also be easily extended to the case, as in our data, where a fixed number of individuals in each community make location decisions over time. Migrants who have located recently at the destination in any period t, correspond to the younger cohorts. Similarly, migrants who are well established at the destination and better positioned to provide referrals, correspond to the older cohorts. The rainfall instruments continue to apply in this case, with recent-past rainfall directly determining the number of new migrants in the network, and distant-past rainfall determining the level of established migrants. We continue to have a sufficient number of instruments for the employment regression, as above.

Individuals typically migrate to save up for a house, or to invest in a small business [Massey et al. 1987]. Once their target savings level is achieved, they will return home, with the duration of their stay depending on economic conditions at the origin and the destination, as discussed above. Favorable conditions at the destination increase the speed at which the savings target will be achieved, increasing the rate of return migration. Such attrition in the network would be most pronounced among the established migrants, since recent arrivals must spend at least a few years at the destination before they return. The estimated network effect, particularly the effect generated by the established migrants, could thus be biased *downward* once we allow for return migration. The rainfall instruments will, however, control for this additional source of bias as well since they are uncorrelated with the employment shocks that induce the return migration.

2. **The individual's duration at the destination.** Our interpretation of the reduced-form results described in the Introduction is that low rainfall at the origin, four to six years ago, increases the measure of established migrants today, improving the quality of the network and increasing employment rates among its members. An alternative interpretation of this result, following the discussion in Section III, is based on the idea that the probability of employment could be independently increasing

in the individual's duration at the destination. The negative correlation between lagged rainfall and current employment could then simply reflect compositional change in the network: employment levels are higher on average because there are more established migrants around.

One strategy to control for this confounding effect would be to estimate the employment regression with fresh arrivals only—those migrants who arrived in period t or period $t-1$. If such a migrant is more likely to be employed if rainfall in his community was low in periods $t-4$ to $t-6$, then this would provide strong evidence that he is benefiting from the relatively large number of established migrants, who arrived long before he did.[18] As noted in Section III, employment levels are likely to be particularly low for the fresh arrivals, who would then benefit disproportionately from the network. A strong implication of this discussion is that the estimated network effects should actually be larger for the fresh arrivals, as compared with the effects obtained with the full sample.

3. **Individual determinants of the employment outcome.** Notice that equation (1) contained no individual determinants of the employment at the destination, apart from ability. Many of these determinants, such as education, are time invariant and would be controlled for by the individual fixed effects. Even if individual characteristics do change over time, omitting them from the employment regression creates no problems for consistent estimation of the network effects unless they are correlated with the rainfall instrument. We noted in the Introduction that employment depends on the number of established migrants in the network, which is in turn determined by "distant-past" rainfall (more than three years ago). Thus, we need to rule out the possibility that distant-past rainfall shocks affect current individual determinants of the employment outcome in this case. About the only omitted characteristics that appear to be plausible in this environment are associated with changes in the structure of the family. Changes in marital status or the number of children could translate into an increased incentive to seek employment with a lag. However, results not reported here show that both

18. Once we allow for return migration, the individual's duration at the destination will also respond to unobserved labor market shocks. This test relies on the idea that fresh arrivals will stay at least a couple of years before they return to the origin. We can thus restrict the sample to fresh migrants without biasing the estimated network effects.

marriage and fertility in the community, unlike migration, are completely unaffected by rainfall shocks at the origin.

While individual characteristics might not change, the migrant's reservation wage or search intensity could respond to rainfall at the origin. For example, low rainfall could worsen the migrant's family's economic condition at home, lowering his reservation wage and increasing his search intensity, to the extent that he is tied financially to them.[19] While this alternative explanation generates higher employment among the migrants following a negative rainfall shock, it does not explain the long four-year delay before employment starts to rise. We will later see that low rainfall lowers employment at the origin immediately (in the same year), and we would expect information to flow fairly smoothly within the community even across national borders, so this alternative explanation would predict an employment response at the destination as early as the next year. Moreover, we will later see that low rainfall leads to improved occupational outcomes—a shift into nonagricultural jobs—among the migrants. A lowering of the reservation wage cannot explain this feature of the data.

4. **Data problems.** There are essentially three data problems that we must deal with in this paper: measurement error in the network variable, recall bias due to the retrospective nature of the data, and missing migrants on account of the fact that some of the migrants might not have returned at Christmas time in the year of the survey. I will deal separately with each of these potential sources of bias below.

Begin with the measurement error in the network variable. Remember that the econometrician's measure of the size of the network at the destination is based on a random sample of individuals drawn from the community, so this variable will certainly be measured with error if we were to treat the entire origin-community as the social unit. Measurement error attenuates the network effect down toward zero. However, the rainfall instrument will avoid measurement error as well in this case, since rainfall shocks at the origin determine the level of migration in the community, but provide no information about deviations from the true level of migration.

19. Labor contractors might also tend to visit communities which have just received poor rains, to recruit cheap labor. This effect is equivalent to an increase in the individual's search intensity.

Next, consider the missing migrants. The surveys were conducted around Christmas time in each community, which is when the migrants traditionally return home to visit their families [Massey et al. 1987]. But it is always possible that some migrants might not have returned in the survey year.[20] In general, omission of individuals from the network will bias the estimated network effects upward, since we are attributing all of the network's effect to a fraction of its members.[21] If established migrants are less likely to return home at Christmas, then this could explain their disproportionate contribution to the network later observed in the data.

However, things are not as bad as they might seem. The survey is not conducted every year, but only at one point in time, which we will refer to as period T. There is a *single* group of individuals, who happen to have been located at the destination in period T, and who happen to have stayed away that Christmas. Only *those* individuals will be missing in all the sample years prior to period T. Since individuals are independently moving back and forth over time, it is very unlikely that all the missing individuals would have been together at the destination in any year other than period T. Individuals remain at the destination for a few years, so we would expect to see some persistence. But this should soon disappear, and thereafter only a small (random) fraction of the missing individuals will be together at the destination in any given year.[22] We will consequently experiment with different sample periods in the empirical analysis, discarding the

20. The MMP tracked down a few workers from each community in the United States, but these numbers are small, and the sampling problematic, so I restrict attention in the analysis to individuals surveyed in Mexico only.

21. Suppose that a fraction θ of the network is unobserved by the econometrician, and denote $\tilde{X}_{t-1} = (1 - \theta)X_{t-1}$ as the observed size of the network. Equation (1) can then be rewritten as

$$\Pr(E_{it} = 1 | X_{it} = 1) = \beta\tilde{X}_{t-1} + [\beta\theta X_{t-1} + \omega_i + C_t].$$

Ignoring selectivity and simultaneity bias, it can be easily verified that plim $\hat{\beta} = \beta/1 - \theta$. As the unobserved component of the network (θ) grows, the upward bias in the network effect grows with it.

22. The pattern over time that we have just described should hold not only for the missing group, but also for the individuals who are observed at the destination in the survey year. To empirically verify the preceding argument, I studied the location patterns over time of those individuals who were established migrants at the destination in the survey year, in a previous version of the paper. Specifically, I plotted the proportion of those individuals who continue to be established migrants as we move back in time from the survey year. As expected, there is a sharp initial decline in the proportion, followed by a flattening out thereafter.

survey year and the years just prior to that year to obtain relatively clean estimates of the network effects.

Finally, we turn to recall bias. The regression analysis uses information on where the individual was located in each year, whether or not he was employed, and the broad occupational category that he was hired in if he was working. This is fairly basic information, and we would expect accurate responses, going back many years before the survey year. In the event that there are any errors, this is only cause for concern with our instrumental variable procedure if the errors are systematic. For example, if individuals systematically report that they are at home when they are in fact at the destination, then the network size will be biased downward, and the discussion above tells us that the estimated network effects will be biased upward. If the recall error goes in the opposite direction, then the bias in the estimated network effects goes in the opposite direction as well. We would expect such errors to grow more frequent as we move further back in time, and so one way to check for such recall bias would be to experiment with longer sample lengths (twenty years prior to the survey year in each community) to verify the robustness of the estimated network effects.

5. **Occupation as the outcome of interest.** Up to this point we have assumed that there are only two labor market outcomes: the individual is either employed or unemployed. Once we allow for multiple occupations, we would also expect the network to move its members into preferred jobs. In this application, two broad classes of occupations are available to the migrants: agricultural labor and higher paying nonagricultural jobs.[23] To disentangle the effect of the network on occupational choice from its effect on employment, I will now restrict attention to individuals who are always employed in the years in which they locate at the destination ($E_i = 1$). Each of those individuals should be more likely to hold a nonagricultural job in years in which his network is larger. The occupation regression can then be specified as follows:

$$\Pr(N_{it} = 1 | E_i = 1, X_{it} = 1) = \gamma X_{t-1} + \omega_i + D_t, \tag{6}$$

where $N_{it} = 1$ if the individual holds a nonagricultural job, $N_{it} = 0$ if he has an agricultural job. D_t is an unobserved labor market

23. The MMP data set lists 81 occupations, which are further classified into broader categories, one of which is agricultural jobs.

shock, which reflects the demand for nonagricultural labor relative to agricultural labor.

As before, we control for the unobserved ω_i term with individual fixed effects. X_{t-1} is correlated with D_t for two reasons. First, higher D_t implies that more high-wage nonagricultural jobs are available, which induces additional migration and biases the $\hat{\gamma}$ estimate upward. Second, greater access to nonagricultural jobs increases the speed at which migrants achieve their target savings, hastening return migration and biasing $\hat{\gamma}$ in the opposite direction. As with the employment regression, the rainfall instrument continues to be orthogonal to D_t, providing us with an unbiased estimate of the effect of the network on occupation choice.

V. Empirical Analysis: Employment at the Destination

The network is seen to influence two labor market outcomes in this paper: the migrant's employment outcome and, conditional on being employed, the type of job that is obtained. We begin the empirical analysis with employment as the outcome of interest in this section. Subsequently, we turn to the occupation as the dependent variable in Section VI.

The individual's employment is a binary variable, which takes on a value of one if he is employed, zero otherwise. A fixed number of individuals (typically 200) were interviewed in each community, so our measure of the size of the network will be the proportion of the community that is located at the destination at each point in time.[24] The Linear Probability model, with fixed effects and year dummies, is utilized for all the employment regressions that we estimate in this paper.[25]

This section begins with the reduced-form employment regressions in subsections V.A and V.B. Subsequently, we turn to the instrumental variable (IV) estimates in subsection V.C. Following the suggestion of a referee, women who account for 3 percent of the migrants are dropped from most of the regressions that I report in this paper. We will later see in Table VI that the

24. We are implicitly ignoring the effect of community size in this case, and we will return to this point in subsection V.C.

25. While the sample period covers fifteen years prior to the survey year in most communities, the data span a much longer time period (1973–1995) since our communities were surveyed at different points over the 1982–1995 period. We use less than fifteen years for the analysis in the very earliest communities because the rainfall series does not go that far back.

TABLE IV
REDUCED-FORM REGRESSIONS: FINE PARTITION OF RAINFALL LAGS

Dependent variable:	Employment at the destination	Employment at the origin
	(1)	(2)
rain (t)	−0.003	**0.027**
	(0.013)	(0.009)
rain (t − 1)	−0.007	**0.027**
	(0.015)	(0.009)
rain (t − 2)	−0.016	**0.035**
	(0.014)	(0.009)
rain (t − 3)	−0.027	**0.024**
	(0.016)	(0.009)
rain (t − 4)	**−0.033**	0.008
	(0.014)	(0.008)
rain (t − 5)	**−0.032**	0.008
	(0.013)	(0.008)
rain (t − 6)	**−0.032**	0.009
	(0.013)	(0.010)
Individual fixed effects	Yes	Yes
Year dummies	Yes	Yes
R^2	0.705	0.812
Box-Pearson Q statistic	0.042	2.813
Number of observations	4,546	41,120

Standard errors are in parentheses.
Standard errors are robust to heteroskedasticity and clustered residuals within each community-year.
$Q \sim X_1^2$ under H_0: no serial correlation.
The critical value above which the null is rejected at the 5 percent level is 3.84.
Employment is a binary variable that measures the individual's labor market outcome in a given year. The individual is employed if he worked for more than one month in that year.
The individual is located at the destination in a given year if he spent more than one month in the United States.
Lagged rainfall at the origin as regressors in column (1) and column (2).
Rainfall coefficients in boldface are significant at the 5 percent level.
Column (1): employment at the destination as the dependent variable.
Column (2): employment at the origin as the dependent variable.

estimated network effects actually increase when the women are included in the sample, so the results that I report are most likely conservative estimates of the network effects.

V.A. Reduced-Form Regressions: Fine Partition of Rainfall Lags

We begin in Table IV with the reduced-form specification of the model, regressing employment on lagged annual rainfall. Turning to column (1), the first empirical result of this section is that employment at the destination is negatively correlated with lagged rainfall at the origin. Further, it is the longer rainfall lags that seem to have a greater effect on employment.[26]

Why should rainfall at home, more than three years ago, affect the individual's employment outcome in the United States today? Recall from Section IV that long rainfall lags directly determine the number of established migrants in the network today, and we know from Section III that the older cohorts are better positioned to provide referrals in the network. Our interpretation of this reduced-form result, which we will subsequently verify, is that long rainfall lags have a stronger effect on employment outcomes at the destination because they determine the number of established migrants in the network, at each point in time.

The basic idea behind the instrumental variable procedure in this paper is that low rainfall at the origin should adversely affect economic returns there, increasing migration to the destination. One convenient measure of the returns at the origin, which is available to the econometrician, is the individual's employment outcome. I consequently replace employment at the destination with employment at the origin as the dependent variable in column (2). Restricting attention to person-years in which the individual is located in the origin community, we see that employment is now *increasing* in lagged rainfall.[27] Moreover, it is the current and recent lags that affect employment the most, suggesting that migration should respond quickly to negative rainfall shocks at the origin. This useful result directly motivates the first-stage regressions that we will examine below.

V.B. *Reduced-Form and First-Stage Regressions: Coarse Partition of Rainfall Lags*

Roughly 12 percent of the observations in each community-year are located at the destination over the sample period. With 200 individuals in a community, this leaves us with approximately 24 migrants, spread over many years of exposure at the

26. The coefficient on lagged rainfall declines sharply after $t - 6$, which explains my choice of lag-length in these reduced-form regressions.

27. We would expect employment in communities with higher levels of irrigation to respond less to rainfall shocks. Regression results not reported here reveal that this is indeed the case. However, irrigation levels are potentially endogenous and could respond to employment shocks at the destination through remittances or savings from past migration spells. Thus, we only use uninteracted rainfall as instruments in the regressions that we report in this paper, although the IV estimates of the network effects are very similar with and without the irrigation-rainfall interaction terms.

destination, at each point in time. Clearly, there are too few migrants to estimate network effects separately for each level of exposure. My approach instead is to partition the network into new migrants and established migrants. Since lagged rainfall helps determine the pattern of migration, the estimates in Table IV provide us with a convenient cut-off to separate these categories. Recall that rainfall was insignificant for the first three lags (t to $t-2$), and significant and stable thereafter ($t-3$ to $t-6$).

Before verifying the link between rainfall and migration, which is essentially the first-stage of the instrumental variable (IV) regression, I first proceed to replicate the reduced-form regressions that we studied in Table IV, with this coarse partition of the rainfall lags. Recent-past rainfall is measured as the average rainfall over the periods t to $t-2$, while distant-past rainfall is measured as the corresponding average over $t-3$ to $t-6$. As expected, employment responds strongly in Table V, column (1), to distant-past rainfall, but is unaffected by recent-past rainfall.[28] We are implicitly restricting the coefficients on recent-past rainfall t to $t-2$ to be the same, with a corresponding restriction on the coefficients for distant-past rainfall $t-3$ to $t-6$, in this alternative specification of the reduced-form employment regression. Notice that the coefficient on distant-past rainfall is roughly four times the coefficient on the $t-3$ to $t-6$ lags in Table IV, which is what we would expect since the coefficients were fairly stable across the four years.[29]

Table V, column (2), repeats the exercise that we just described, with an alternative cut-off for the recent migrants at $t-3$, instead of $t-2$. While the coefficient on distant-past rainfall (the average over $t-4$ to $t-6$) is now slightly smaller than it was in column (1), the basic patterns that we observed earlier are unchanged.

As noted in Section IV, one simple strategy to disentangle the network effect from the individual exposure effect is to restrict attention to fresh arrivals. Restricting attention to those mi-

28. While we control for clustered residuals in each community-year when computing the standard errors in this paper, we would need to allow for clustering at a more aggregate state-year level if rainfall shocks were correlated across neighboring communities. Standard errors that allow for state-year clustering are almost identical to what we report in this paper.

29. I also experimented with the conditional (fixed-effects) logit model to check the robustness of these results. While the point estimates for the logit and the linear probability model are not directly comparable, the coefficient on distant-past rainfall continues to be negative and significant at the 5 percent level, while the coefficient on recent-past rainfall is insignificant.

TABLE V
REDUCED-FORM AND FIRST-STAGE REGRESSIONS: COARSE PARTITION OF RAINFALL LAGS

Dependent variable:	Reduced-form					First-stage	
	Employment at the destination				Employment at the origin	New migrants	Established migrants
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Recent-past rainfall	−0.028	−0.049	−0.023	−0.047	0.085	−0.091	0.005
	(0.027)	(0.035)	(0.072)	(0.040)	(0.018)	(0.037)	(0.020)
Distant-past rainfall	−0.125	−0.092	−0.226	−0.129	0.046	0.103	−0.106
	(0.035)	(0.027)	(0.108)	(0.044)	(0.021)	(0.033)	(0.023)
Individual fixed effects	Yes	Yes	Yes	No	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.705	0.705	0.647	0.038	0.812	0.768	0.940
Q statistic	0.041	0.041	0.036	0.660	2.813	0.010	0.316
Number of observations	4546	4546	1732	4546	41,120	4546	4546

Standard errors are in parentheses.
Standard errors are robust to heteroskedasticity and clustered residuals within each community-year.
$Q \sim X_1^2$ under H_0: no serial correlation. The critical value above which the null is rejected at the 5 percent level is 3.84.
Recent-past rainfall is average rainfall at the origin over the past three years; t to $t-2$.
Distant-past rainfall is average rainfall at the origin over the preceding four years; $t-3$ to $t-6$.
New migrants measures the proportion of the community located at the destination for one–three years in period t.
Established migrants measures the proportion of the community located at the destination for four or more years in period t.
Employment was defined in Table IV.
Columns (1)–(5): reduced-form employment regressions.
Column (1) and column (5): repeat reduced-form employment regressions in Table IV with coarse partition of lagged rainfall.
Column (2): recent-past rainfall is average over the past four years, and distant-past rainfall is average over the preceding three years.
Column (3): restrict attention to person-years in which the migrant arrived in the current year or the previous year.
Column (4): reduced-form employment regression with community dummies.
Columns (6)–(7): first-stage regressions.

grants who arrived in the current year or the previous year in column (3), we see that the basic pattern observed in columns (1)–(2) is unchanged. The coefficient on the established migrants actually increases substantially, which is consistent with the view that the newcomers should be more dependent on the network.

Most of the employment regressions in this paper, including those reported up to this point, include individual fixed effects to control for unobserved selectivity in the migration decision. This implies that we are identified from the 5 percent of person-years at the destination that apply to migrants who report both employment as well as unemployment over the sample period. While

I mentioned in Section III that individuals who are less likely to be employed will benefit more from the network, I did not account for this potential variation across individuals later in the estimation section. Since most of our migrants are always employed, those individuals who report both employment and unemployment over the sample period are by definition those who would benefit most from the network. Thus, while our fixed-effects estimates would correctly measure the network effects for that vulnerable group of individuals, they could substantially overestimate the average effects for the community as a whole.[30]

One strategy to avoid this problem would be to replace individual fixed effects with community dummies in the employment regression. If we were to take the expectation of the employment outcome, across all the individuals at the destination in each community-year, the dependent variable in the equivalent community level regression would then be the proportion of migrants employed in each year. Network effects in the regression with community dummies are thus effectively estimated from changes in community-level employment over time. The reduced-form employment regression with community dummies is presented in column (4) of Table V. The point estimates are remarkably similar to what we obtained earlier in column (1), increasing our confidence in the robustness of these results.

Finally, I complete the replication of the reduced-form results, with the coarse partition of the rainfall lags, by studying employment outcomes at the origin in column (5). Employment is increasing in lagged rainfall, particularly recent-past rainfall, just as we saw in column (2) of Table IV, which leads us quite naturally to the migration regressions that follow.

Turning to the first stage of the IV regression in columns (6)–(7), we see that the numerical strength of the new (established) migrants is *negatively* correlated with recent-past (distant-past) rainfall, supporting our interpretation of the reduced-form estimates.[31] Notice also that the coefficient on distant-past

30. Goldin and Rouse [2000] must deal with the same problem in their study of gender discrimination, where they are identified from 6 percent of the individuals in the sample with individual fixed effects. While Goldin and Rouse argue that those 6 percent are drawn randomly from the full sample, their approach does not appear to be plausible in our setting; individuals who report spells of unemployment over the sample period are very likely to have lower than average ability, or observed characteristics that signal low ability.

31. Note that the additive separability in all the reduced-form and first-stage regressions in the paper is appropriate in this case. For example, this is appro-

rainfall is positively correlated with the numerical strength of the new migrants. We saw above that low distant-past rainfall reduces current employment levels at the origin, which would increase new migration. Low distant-past rainfall also improves the quality of the network by increasing the number of established migrants, which would in turn induce additional new migration. However, since there are a fixed number of individuals in each community, low distant-past rainfall also reduces the number of *potential* new migrants by increasing the number of established migrants. The positive coefficient on distant-past rainfall in column (6) suggests that the third effect must dominate. Note that the coefficient on recent-past rainfall in column (7) is insignificant. While recent-past rainfall could in principle induce return migration, this does not seem to be the case with our data.

We close this section with nonparametric reduced-form and first-stage estimates in Figure I that verify the robustness of our results. After controlling for individual fixed effects and year effects, we see that employment at the destination is negatively correlated with distant-past rainfall at the origin across the entire range of rainfall levels. Similarly, after controlling for community fixed effects and year effects, we see that established migration is negatively correlated with distant-past rainfall across the entire range of rainfall values. These results, taken together, imply a positive relationship between the number of established migrants in the network and the level of employment at any point in time, summarizing the first empirical result of this paper.[32]

V.C. OLS and Instrumental Variable Regressions

We now proceed to directly verify the relationship between the network and employment at the destination. Starting with a

priate in column (7) because recent-past rainfall changes the number of established migrants in the network at the *margin,* by shifting the ability threshold from one period to the next, instead of changing the general probability that *any* established migrant would return (in which case interaction terms would be required).

32. The nonparametric regressions are implemented in two steps. First, run a parametric regression with distant-past rainfall (measured in meters), individual fixed effects or community fixed effects, and a full set of year dummies, as regressors. The employment regression is estimated at the individual level, while the first-stage regression is estimated at the community level. Difference out the estimated fixed effects and year dummies, and then nonparametrically regress the modified dependent variable on distant-past rainfall in the second step, using the Epanechnikov kernel function.

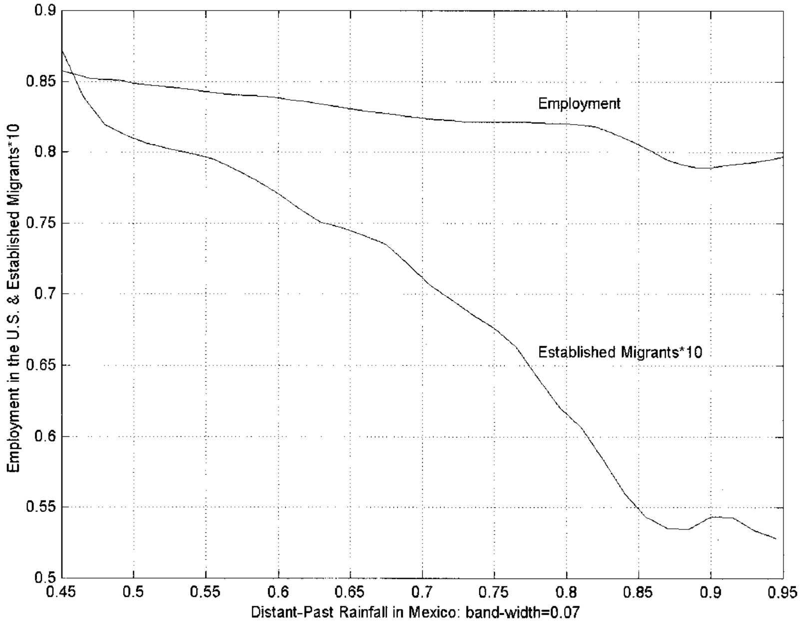

FIGURE I
Employment and Network Response to Distant-Past Rainfall—
Net Fixed Effects

preliminary OLS regression in column (1) of Table VI, we see that the individual's employment outcome depends strongly on the numerical strength of the established migrants, but is unaffected by the new migrants. Looking next at the IV estimates, in column (2) of Table VII, new migrants continue to play a negligible role, while the established migrants are even more influential than they were in the OLS regression reported in column (1).

Rainfall shocks at the origin are clearly uncorrelated with unobserved employment shocks at the destination. Rainfall is also a good instrument for the strength of the network; the rainfall coefficients in columns (6)–(7) of Table V are statistically significant, and the F-statistics are comfortably above the critical value at the 5 percent significance level in both regressions. We should thus be fairly confident about the consistency of the IV estimates. The discrepancy between the OLS and the IV estimates might arise because favorable conditions at the destination induce return migration among the established migrants who have achieved their savings target, biasing the network effects

TABLE VI
OLS AND INSTRUMENTAL VARIABLE REGRESSIONS

	Employment at the destination										
	OLS	IV									
		Basic specifications				Robustness to individual characteristics			Robustness to sample lengths		
Dependent variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
New migrants	−0.032	0.397	0.522	0.093	0.626	0.623	0.394	0.424	0.511	0.377	0.251
	(0.070)	(0.315)	(0.376)	(0.537)	(0.501)	(0.353)	(0.306)	(0.326)	(0.321)	(0.400)	(0.356)
Established migrants	0.670	1.554	1.474	2.073	1.745	2.021	1.321	1.565	1.699	1.304	1.058
	(0.154)	(0.551)	(0.545)	(1.069)	(0.661)	(0.594)	(0.534)	(0.656)	(0.526)	(0.578)	(0.491)
Individual fixed effects	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.707	—	—	—	—	—	—	—	—	—	—
Q statistic	0.042	0.041	0.041	0.036	0.660	0.110	0.0005	0.015	0.049	0.022	0.001
Number of observations	4546	4546	4546	1732	4546	4710	3371	4067	5214	3894	3614

Standard errors are in parentheses. Standard errors are robust to heteroskedasticity and clustered residuals within each community-year.
$Q \sim X_1^2$ under H_0: no serial correlation. The critical value above which the null is rejected at the 5 percent level is 3.84.
Employment was defined in Table IV. New migrants, Established migrants were defined in Table V.
Recent-past rainfall and Distant-past rainfall are used as instruments for New migrants and Established migrants.
Column (1): OLS employment regression with individual fixed effects.
Column (2): IV employment regression with individual fixed effects.
Column (3): recent-past rainfall (new migrants) defined as four lags, distant-past rainfall (established migrants) preceding three lags.
Column (4): restrict attention to person-years in which the migrant arrived at the destination in the current year or the previous year.
Column (5): IV employment regression with community dummies.
Column (6): include both men and women in the sample (the sample is restricted to male migrants in all other regressions).
Column (7): restrict sample to men less than 45 years only.
Column (8): restrict sample to individuals with less than ten years of education.
Column (9): extended twenty-year sample period in each community.
Column (10): discard survey year and previous year from the sample.
Column (11): discard survey year and two previous years from the sample.

TABLE VII
OCCUPATIONAL CHOICE AND LABOR MARKET OUTCOMES

Occupation at destination:	Agricultural	Nonagricultural
	(1)	(2)
Hourly wage (relative to minimum wage)	1.22*	1.42*
	(0.03)	(0.03)
Hourly wage (in 2001 dollars)	6.92*	8.06*
	(0.15)	(0.20)
Annual income (in 2001 dollars)	8682.57*	11995.28*
	(350.56)	(540.97)
Hours worked per week	49.47*	45.04*
	(0.84)	(0.60)
Months worked per year	5.73*	7.09*
	(0.15)	(0.17)
Number of observations	406	524

Standard errors are in parentheses.
This information is based on each migrant's last visit to the United States.
The individual is now the unit of observation.
These data are obtained from a separate file that provides information on earnings and community support on the migrant's last trip to the United States.
Migrants with wages in the top 1 percentile of the distribution are dropped.
* denotes rejection of equality of means for the two groups at 5 percent significance level.

attributable to this group of migrants downward. Measurement error in the network variable could also have attenuated the OLS estimates in column (1).[33]

Next, I consider an alternative cut-off for new migrants at $t-3$, so the established migrants are defined as those individuals who have been at the destination for four years or longer. The IV estimates with this alternative cut-off in column (3) are almost identical to what we observed in column (2).

Up to this point in the discussion we have not accounted for the possibility that employment rates could be independently increasing with the individual's exposure at the destination (on the current trip). Restricting attention to fresh arrivals in column (4), the estimated network effects are actually substantially

33. Another explanation for the difference between the OLS and the IV estimates is that low ability individuals are more responsive to rainfall shocks at the destination. In that case, the OLS estimates will apply to the migrant with average ability, while the IV estimates will apply to a migrant with lower than average ability, and we know that low-ability individuals benefit more from the network. I checked to see whether individuals with observed characteristics that signal low ability (women, older men, less educated men) are more likely to migrate when rainfall in the origin community is low, but was unable to uncover any systematic pattern in the location decisions.

larger than the network effects that we obtained with the full sample. This is exactly what we would expect, since the newcomers should be more dependent on the network for referrals. Subsequently, column (5) replaces individual fixed effects with community dummies, without changing the estimated network effects at all.

The remaining regressions in Table VI verify the robustness of the IV estimates. The discussion in Section III tells us that individuals independently less likely to be employed will benefit most from the network, and we have already seen that fresh arrivals benefit disproportionately in column (4). Comparing employed and unemployed migrants (with the person-year as the unit of observation), we find that women and older men are much more likely on average to be unemployed.[34] These results are not surprising. Both women and the older men have observable characteristics that signal lower ability to perform the sort of physical jobs that the migrants engage in. The network can play a very useful role in this case, since those members of the network who are providing referrals have better information than the market about the true ability of these workers.

Following the discussion above, we would expect individuals with characteristics associated with higher unemployment—the women and the older men—to benefit more from the network. The presumption in this case is that the network cannot overcome the individual's inherent limitations in the labor market. Those individuals who would be independently less likely to be employed will benefit the most from the network, but they will continue to display lower unemployment levels in equilibrium. As expected, network effects in column (6) with both men and women in the sample are substantially larger than the corresponding point estimates in column (2), which restrict attention to male migrants. Subsequently, we return to the standard specification with men only, but restrict the sample to men less than 45 years old in column (7). The network effects do decline, and while not reported here there is a steady decline as the age ceiling is lowered from 55 years to 45 years and finally to 35 years. But the network response to age is not as dramatic as it was for women.

Table VI, column (8) continues with this exercise, restricting

34. Females account for only 2 percent of the employed person-years at the destination, but as much as 28 percent of the unemployed person-years. Similarly, the unemployed are more than thirteen years older than the employed.

attention to individuals with less than ten years of education.[35] Education levels for employed and unemployed person-years in the data are comparable, and not surprisingly the estimated network effects in column (8) are almost exactly what we observed earlier in Table VI, column (2). I experimented with alternative education cut-offs from nine to twelve years, without changing this result.

Apart from individual characteristics, we also need to verify that the network effects are robust to alternative measures of the network and community characteristics. We will see later that the network provides many services to its members: it provides financial support and housing assistance, in addition to job referrals, which is why the network was specified to be the proportion of *migrants* at the destination in Section IV, and throughout the empirical analysis. However, if we were to take the model laid out in Section III seriously, then the appropriate measure of the network is the proportion of *employed migrants* at the destination (who are in a position to provide job referrals). The empirical work this far has also not taken account of the observation in subsection II.B that each community sets up multiple centers—typically two to three large ones—in the United States. The appropriate measure of the network might then be the number of *paisanos* in the state that the migrant is located in, rather than the corresponding statistic for the United States as a whole. We would also want to include a full set of destination-state dummies as additional controls in the employment regression in this case. While not reported here, the estimated network effects with these alternative measures of the network are very similar to what we obtained in Table VI, column (2).

In a related robustness exercise I also experimented with a reduced sample of communities, by dropping small rural communities (*ranchos*) from the regression. We would expect the *ranchos* to be more cohesive, and therefore better positioned to support their members at the destination. The network effects do decline, and are no longer significant at the 5 percent level, when these communities are dropped from the sample. But the basic pattern of coefficients that we have seen throughout, with the established

35. A natural cut-off separating more and less educated migrants would be high school completion (twelve years of education). But only 5 percent of the migrants achieved this level of schooling, so I report estimates with a less stringent ten-year cut-off, which leaves us with 90 percent of the full sample.

migrants contributing disproportionately to the network, continues to be obtained.

All that remains at this point is to account for the potential data problems—recall bias and missing migrants—that we discussed in Section IV. A simple test to rule out recall bias would be to verify that the estimated network effects are robust to an increase in the sample length, say up to twenty years before the survey year. For the bias associated with the missing migrants, one solution would be to drop the survey year and the years just prior to the survey year from the sample, since we would expect the bias to fall very sharply over time.

We begin in Table VI, column (9) by increasing the sample length to twenty years preceding the survey year, in each community. The network effects are very similar to what we obtained with the fifteen-year sample length in Table VI, column (2). Subsequently, we return to fifteen years prior to the survey year as the earliest period in the sample, but discard the survey year (T) and the year before it ($T - 1$) in column (10). The network effects now decline, and this trend continues when we discard an additional year ($T - 2$) in column (11). But the network effect stabilizes thereafter—while not reported here, the point estimate remains at roughly 1.0 when we discard year $T - 3$, and then year $T - 4$.

VI. Empirical Analysis: Occupation at the Destination

The empirical analysis concludes by studying the role of the network in shifting its members into preferred nonagricultural jobs. We begin by comparing the earnings and the characteristics of agricultural and nonagricultural workers in subsection VI.A. Subsequently, we compare the community ties and the assistance received by these workers in subsection VI.B. Nonagricultural workers earn more and are more likely to receive assistance from the community, so there is some prima facie evidence that the network is channeling its members into nonagricultural jobs. Subsection VI.C. reports more robust results from occupation regressions with individual fixed effects and rainfall as the instrument for the network, where we see that the same individual is more likely to hold a nonagricultural job when his network is exogenously larger.

VI.A. Occupational Choice and Labor Market Outcomes

The MMP data set provides very limited information at *each* point in time over the individual's working life: where he was located, whether or not he was employed, and the kind of job he had if he was employed. However, much richer information is available about the migrant's earnings and the assistance that he received from the community on his *last* trip to the United States. We will use this information from the last trip to compare agricultural and nonagricultural workers in the discussion that follows.[36]

Table VII describes the wages and the duration of employment for the two types of workers. Since communities were surveyed at different points in time, and the individuals in each community would also have made their last trip in different years, the last trip of all the migrants in the sample ranges over many years. The reported nominal wages must therefore be normalized to account for this variation over time: Row 1 reports the ratio of the hourly wage to the minimum wage, while row 2 reports the real wage in 2001 dollars. Agricultural workers earn about 20 percent more than the minimum wage ($6.92 per hour) while nonagricultural workers earn about 40 percent more than the minimum wage ($8.06 per hour). All of these earnings differentials in Table VII are statistically significant at the 5 percent level, and factoring in the number of hours worked in the year by each type of worker, the nonagricultural workers earn $3,300 more than the agricultural workers, which is as much as 40 percent more than average agricultural earnings.[37]

Table VII also reports the average job durations for the two types of workers. Agricultural labor tends to be seasonal, but the work is intense during the season. Not surprisingly, the agricultural workers work longer hours per week, but for less than six months in the year; these differences between agricultural and nonagricultural workers are once more statistically significant at the 5 percent level.

36. For those cases in which the migrant held more than one job on the last trip, the MMP lists the job with the longest duration as his occupation.

37. I note below that nonagricultural workers are younger and better educated than agricultural workers, and hence may be more able. However, the income differential for agricultural and nonagricultural workers does not change appreciably when I control for individual characteristics such as sex, age, education, and the migrant's primary occupation at the origin; the income differential only declines from 3,300 to 3,200. The migrant clearly benefits a great deal by gaining access to a nonagricultural job.

The skills required for agricultural and nonagricultural work are not the same, and we saw above that the earnings from these two types of occupations differ significantly. We would thus expect the characteristics of workers in agricultural and nonagricultural occupations to differ as well. While these statistics are not reported here, nonagricultural workers are disproportionately male, and younger on average than the agricultural workers. But these differences are small when compared with the corresponding differences, along the sex and age dimension, between employed and unemployed person-years that we noted earlier. The more important difference between the agricultural and nonagricultural workers seems to be the years of schooling; the mean (with standard errors in parentheses) for the two types of workers is 4.30(0.19) and 5.71(0.18), and the difference in means is significant at the 5 percent level. We would expect the network effects to be larger for the women and the older men, and particularly so for the less educated men, in the occupation regression.[38]

VI.B. Occupational Choice and Community Support

We saw above that nonagricultural workers earn more and have different characteristics, as compared with the agricultural workers. In the discussion that follows, we will see whether the community treats these two types of workers differently as well.

While the focus of this paper is on direct assistance in the labor market, the MMP data also provide information about financial assistance and housing assistance received by the migrants. Table VIII reports the different forms of assistance received from the community on the migrant's last trip to the United States. The possible sources of assistance in each case are relatives, friends or *paisanos,* employer/labor contractor, and other. Relatives and friends or *paisanos* are also combined to construct a binary variable that describes whether or not the migrant received assistance from his community.

Looking across the columns in Table VIII, it is evident that both agricultural and nonagricultural workers benefit a great deal from the community. Roughly 40 percent of the migrants

38. As with the employment regression, the presumption here is that the network cannot overcome the individual's intrinsic limitations. Those individuals who are less likely to hold nonagricultural jobs to begin with will benefit more from the network, but will nevertheless remain less likely to hold those jobs in equilibrium.

TABLE VIII
OCCUPATIONAL CHOICE AND COMMUNITY SUPPORT

Form of assistance:	Financial		Housing		Employment	
Occupation at destination:	Agricultural	Nonagricultural	Agricultural	Nonagricultural	Agricultural	Nonagricultural
	(1)	(2)	(3)	(4)	(5)	(6)
Source of assistance (%)						
No help	55.19	50.68	4.05	3.70	25.50	21.77
Relative	20.00	30.56	36.99	61.19	31.81	37.17
Friend or paisano	16.96	14.51	23.99	26.90	31.81	36.96
Employer/labor contractor	5.06	3.66	30.92	6.37	5.16	1.44
Other	2.78	0.58	4.05	1.85	5.73	2.67
Total	100.00	100.00	100.00	100.00	100.00	100.00
Help from the community	0.37*	0.45*	0.61*	0.88*	0.64*	0.74*
	(0.02)	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)
Number of observations	349	487	349	487	349	487

Statistics are based on the last trip of the migrants in the sample. The individual is now the unit of observation.
These data are obtained from a separate file that provides information on earnings and community support on the migrant's last trip to the United States.
Migrants with wages in the top 1 percentile of the distribution are dropped.
Help from the community is a binary variable which takes the value one if relatives or paisanos provided assistance, zero otherwise.
* denotes rejection of equality of means for the two groups at 5 percent significance level.

received financial assistance from the community on their last trip to the United States, and these numbers are as high as 70 percent for housing and employment assistance. While these assistance levels are very high, notice that the nonagricultural workers receive substantially more assistance from the community in every category (these differences are all significant at the 5 percent level).

The decomposition of the various sources of assistance in Table VIII also allows us to identify the precise source of the difference between agricultural and nonagricultural workers noted above. Starting with financial assistance, we see that most of this difference is attributable to relatives. Nonagricultural workers also benefit disproportionately from their relatives when it comes to housing assistance, but this seems to substitute for the employer-provided housing that many of the agricultural workers receive. Turning finally to employment assistance, which is the immediate focus of this paper, the extra benefit that the nonagricultural workers receive from the community is attributed equally to the relatives and the friends or *paisanos.*

VI.C. Occupation Regressions

We have seen that nonagricultural workers earn significantly more than agricultural workers. We have also seen that nonagricultural workers are significantly more likely to receive financial assistance, housing assistance, and job referrals from the community. While this evidence does suggest that the network is actively channeling its members into preferred nonagricultural jobs, other explanations are also available. For example, we noted that nonagricultural workers are better educated on average. If such individuals were favored by the network for other reasons, then a spurious correlation between the network and occupational choice could be obtained. Alternatively, the information problems that generate a need for job referrals might be more severe in the nonagricultural jobs.[39] The network might be more active in the nonagricultural occupations simply because there is a greater demand for its services from individuals who target those occupations.

The test that we propose to establish an active role for the

39. As noted in Section III, piece-rate contracts are rarely used in the U. S. economy, one of the few exceptions being the agricultural sector. Once a farm can use a piece-rate contract, it is immediately less dependent on job referrals.

network in moving its members into nonagricultural jobs is to see whether the *same* individual is more likely to hold a nonagricultural job in years in which his network at the destination is exogenously larger. This test is easily implemented by replacing employment with the occupational outcome, and by including individual fixed effects and using rainfall as an instrument for network size in the occupation regression. As discussed in Section III, we run this regression with the 95 percent of the migrants who are employed in all years that they locate at the destination, allowing us to isolate the role of the network in affecting occupational choice. Unlike the employment regressions, in which there was very little variation in the dependent variable, we saw in Table II that nonagricultural jobs account for 51 percent of the person-years at the destination. Including individual fixed effects is also no longer a problem, since as many as 18 percent of the always employed migrants change their occupation over the sample period.

We begin in Table IX, column (1), with the reduced-form occupation regression. The dependent variable equals one if the individual has a nonagricultural job, zero if he is an agricultural laborer. Recent-past rainfall and distant-past rainfall are included as regressors in this fixed-effects regression, which is restricted to male migrants as before. As with the employment regression, the coefficient on recent-past rainfall is insignificant, while the coefficient on distant-past rainfall is negative and significant. We already know, from the first-stage regressions reported earlier, that low distant-past rainfall translates into a larger number of established migrants at the destination. Our interpretation of this reduced-form result is that an exogenous increase in the number of established migrants increases the likelihood that any member of the network will hold a preferred nonagricultural job.

Column (2) verifies the robustness of this result by restricting attention to fresh migrants (those who arrived in the current year or the previous year). As with the employment regression, the coefficient on distant-past rainfall actually increases with the reduced sample. This tells us that the new arrivals benefit disproportionately from the network, both when it comes to being employed as well as in gaining access to preferred nonagricultural jobs. We complete the robustness tests by replacing individual fixed effects with community dummies in column (3). The estimates remain very similar to what we obtained in column (1).

TABLE IX
OCCUPATION REGRESSIONS

Dependent variable:	Occupation at the destination			Occupation at the origin	Occupation at the destination				
	Reduced-form				OLS	IV	IV-Robustness tests		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Recent-past rainfall	−0.125	−0.026	−0.025	0.011	—	—	—	—	—
	(0.077)	(0.217)	(0.096)	(0.015)					
Distant-past rainfall	−0.223	−0.479	−0.185	−0.027	—	—	—	—	—
	(0.086)	(0.258)	(0.098)	(0.019)					
New migrants	—	—	—	—	0.398	1.592	1.413	1.105	1.964
					(0.231)	(0.928)	(0.954)	(1.027)	(1.081)
Established migrants	—	—	—	—	0.256	3.585	3.290	3.289	5.009
					(0.247)	(1.339)	(1.374)	(1.410)	(1.901)
Individual fixed effects	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R^2	0.825	0.734	0.143	0.898	0.825	—	—	—	—
Q statistic	0.319	0.007	1.605	1.886	0.322	0.319	0.329	0.252	0.392
Number of observations	4240	1588	4240	30,917	4240	4240	4330	3213	3817

Standard errors are in parentheses. Standard errors are robust to heteroskedasticity and clustered residuals within each community-year.
$Q \sim X_1^2$ under H_0: no serial correlation. The critical value above which the null is rejected at the 5 percent level is 3.84.
Occupation is a binary variable that takes on the value one if nonagricultural job, zero if agricultural job.
Recent-past rainfall, Distant-past rainfall, New migrants, Established migrants were defined in Table V.
Recent-past rainfall and Distant-past rainfall are used as instruments for New migrants and Established migrants in Columns (6)–(9).
Column (1): reduced-form occupation regression with individual fixed effects.
Column (2): restrict attention to person-years in which the migrant arrived at the destination in the current year or the previous year.
Column (3): reduced-form occupation regression with community dummies.
Column (4): reduced-form regression with occupation at the origin as the dependent variable.
Column (5): OLS occupation regression with individual fixed effects.
Column (6): IV occupation regression with individual fixed effects.
Column (7): include both men and women in the sample (the sample is restricted to male migrants in all other regressions).
Column (8): restrict sample to men less than 45 years only.
Column (9): restrict sample to individuals with less than ten years of education.

The regressions up to this point have restricted attention to occupational choice at the destination. Column (4) repeats this exercise at the origin. We now see that both recent-past and distant-past rainfall have no effect on the occupational outcomes at the destination. We saw earlier that rainfall at the origin had a strong and immediate effect on employment at the origin, with this effect persisting for a few years thereafter. The results just reported tell us that opportunities to switch occupations appear to be very limited at the origin, in sharp contrast to what we saw above at the destination.

Returning to the occupation regressions at the destination, column (5) presents OLS estimates, while column (6) presents the corresponding IV estimates, with individual fixed effects and the full sample of male migrants. Somewhat surprisingly, the OLS estimates are insignificant, and the coefficient on the established migrants is actually smaller than the coefficient on the new migrants. However, the familiar patterns reappear when we instrument for new (established) migration with recent-past (distant-past) rainfall. The coefficient on established migration is now very precisely estimated, and much larger than the corresponding coefficient on new migration.

The particularly severe downward bias on the established migrants in the OLS regression might be due to return migration. Migrants were seen to return when they achieved a savings target in Section III and Section IV, which would imply that the migration spells should be shorter for the better paid nonagricultural workers. As expected, migration spells are significantly shorter for the nonagricultural workers: the mean number of years (with standard errors in parentheses) is 4.52(0.24) and 3.22(0.14) for the agricultural and nonagricultural workers, respectively.[40] We would thus expect to see a higher level of return migration, particularly among the established migrants, when conditions favor nonagricultural employment, which would in turn bias the OLS estimates downward.

I conclude the empirical analysis by studying how the network effect varies by sex, age, and the level of schooling. We noted earlier that while nonagricultural workers tend to be disproportionately male, and younger on average than agricultural work-

40. These statistics are computed over the fifteen-year sample period in each community, and compare those workers who are exclusively employed in agriculture and nonagriculture, respectively, over this period.

ers, the major difference between the two types of workers seems to be in the level of schooling.

Column (7) includes both male and female migrants, while column (8) restricts the sample to men under 45 years. The network effects are very stable when compared with the corresponding estimates in column (6), in contrast with the employment regressions where we saw the women and the older men benefit disproportionately from the network.[41] However, the network effects increase sharply in Table IX when we restrict attention to male migrants with less than ten years of schooling, which tells us that less educated migrants do benefit disproportionately from the network when it comes to gaining access to nonagricultural jobs. This result is robust to alternative cut-offs ranging from nine to twelve years of schooling.

VII. Conclusion

This paper attempts to test for the presence of social networks among Mexican migrants, belonging to well-established origin-communities, in the U. S. labor market. We verify that the same individual is more likely to be employed, and to hold a preferred nonagricultural job, when his network is exogenously larger. We find that it is the more established members who contribute disproportionately to the network and that it is the disadvantaged members—women, the elderly, and the less educated—who benefit the most.

This paper provides a first glimpse of a remarkable decentralized institution that has provided a steady supply of low cost labor to the United States for nearly a century. Because migration from this region is recurrent, the individual is rarely matched with the same members of his community on different trips to the United States. However, preexisting social ties ensure that he receives various forms of assistance from those members of the community who happen to be established at the destination when he does arrive, on each trip to the United States.

While these social ties might improve the efficiency of the network, they come with a cost of their own. There is now an externality associated with the individual's migration decision, and it is very likely that members of these communities face

41. I also experimented with alternative age cutoffs at 55 years and 35 years without changing these results.

strong pressure to retain particular location patterns, and to remain in the low-skill jobs that have traditionally been chosen, to maintain the stability of the network. This observation might explain the low levels of education that we see in the data, and the prevalence of low-skill occupations, despite the long history of migration to the United States in these communities.

UNIVERSITY OF PENNSYLVANIA

REFERENCES

Bean, Frank D., Georges Vernez, and Charles B. Keely, *Opening and Closing the Doors: Evaluating Immigration Reform and Control* (Santa Monica, CA, and Washington, DC: Rand Corporation and the Urban Institute, 1989).
Bertrand, Marianne, Erzo Luttmer, and Sendhil Mullainathan, "Network Effects and Welfare Cultures," *Quarterly Journal of Economics,* CXV (2000), 1019–1056.
Borjas, George, "Self-Selection and the Earnings of Immigrants," *American Economic Review,* LXXVII (1987), 531–553.
——, "Ethnic Capital and Intergenerational Mobility," *Quarterly Journal of Economics,* CVII (1992), 123–150.
——, "Ethnicity, Neighborhoods and Human-Capital Externalities," *American Economic Review,* LXXXV (1995), 365–390.
Bustamante, Jorge, "Changing Patterns on Undocumented Migration from Mexican States in Recent Years," in Richard C. Jones, ed., *Patterns of Undocumented Migration: Mexico and the United States* (Totowa, NJ: Roman and Allanheld, 1984).
Cardoso, Lawrence, *Mexican Emigration to the United States 1897–1931* (Tucson, AZ: University of Arizona Press, 1980).
Carrington, William, Enrica Detragiache, and Tara Vishwanath, "Migration with Endogenous Moving Costs," *American Economic Review,* LXXXVI (1996), 909–930.
Case, Anne, and Lawrence Katz, "The Company You Keep: The Effects of Family and Neighborhood on Disadvantaged Youth," Working Paper No. 3705, National Bureau of Economic Research, 1991.
Chavez, Leo, *Shadowed Lives: Undocumented Immigrants in American Society* (New York: Harcourt Brace Jovanovich College Publishers, 1992).
Craig, Richard, *The Bracero Program: Interest Groups and Foreign Policy* (Austin, TX: University of Texas Press, 1971).
Duflo, Esther, and Emmanuel Saez, "The Effect of Financial Education on Savings Decisions: Evidence from a Randomized Experiment." Massachusetts Institute of Technology, mimeo, 2002.
Durand, Jorge, Douglas Massey, and Fernando Charvet, "The Changing Geography of Mexican Immigration to the United States: 1910–1996," *Social Science Quarterly,* LXXXI (2000), 1–15.
Glaeser, Edward, Bruce Sacerdote, and Jose Scheinkman, "Crime and Social Interactions," *Quarterly Journal of Economics,* CXI (1996), 507–548.
Goldin, Claudia, and Cecilia Rouse, "Orchestrating Impartiality: the Impact of Blind Auditions on Female Musicians," *American Economic Review,* LXXXX (2000), 715–742.
Gottlieb, Peter, "Rethinking the Great Migration," in Joe William Trotter, Jr., ed., *The Great Migration in Historical Perspective: New Dimensions of Race, Class, and Gender* (Bloomington, IN: Indiana University Press, 1991).
Grossman, James, *Land of Hope: Chicago, Black Southerners, and the Great Migration* (Chicago: The University of Chicago Press, 1989).
Hagan, Jacqueline Maria, *Deciding to be Legal: A Maya Community in Houston* (Philadelphia, PA: Temple University Press, 1994).

Holzer, Harry, "Search Method Used by Unemployed Youth," *Journal of Labor Economics,* VI (1988), 1–20.
Jones, Richard, "Macro-Patterns of Undocumented Migration between Mexico and the U. S.," in Richard C. Jones, ed., *Patterns of Undocumented Migration: Mexico and the United States* (Totowa, NJ: Roman and Allanheld, 1984).
Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to Opportunity in Boston: Early Results of a Randomized Mobility Experiment," *Quarterly Journal of Economics,* CVI (2001), 607–655.
Ludwig, Jens, Greg J. Duncan, and Paul Hirschfield, "Urban Poverty and Juvenile Crime: Evidence from a Randomized Housing-Mobility Experiment," *Quarterly Journal of Economics,* CVI (2001), 655–681.
Manski, Charles, "Identification of Endogenous Social Effects: The Reflection Problem," *Review of Economic Studies,* LX (1993), 531–542.
——, "Economic Analysis of Social Interactions," *Journal of Economic Perspectives,* XIV (2000), 115–137.
Marks, Carole, *Farewell—We're Good and Gone: The Great Black Migration* (Bloomington, IN: Indiana University Press, 1989).
Massey, Douglas, Rafael Alarcon, Jorge Durand, and Humberto Gonzalez, *Return to Aztlan: The Social Process of International Migration from Western Mexico* (Berkeley: University of California Press, 1987).
Menjivar, Cecilia, *Fragmented Ties: Salvadoran Immigrant Networks in America* (Berkeley: University of California Press, 2000).
Miguel, Edward, and Michael Kremer, "Why Don't People Take their Medicine? Experimental Evidence from Kenya," University of California, Berkeley, mimeo, 2002.
Montgomery, James, "Social Networks and Labor-Market Outcomes: Toward an Economic Analysis," *American Economic Review,* LXXXI (1991), 1408–1418.
Nee, Victor G., and Brett de Bary Nee, *Longtime Californ': A Documentary Study of an American Chinatown* (New York: Pantheon Books, 1972).
Rees, Albert, "Information Networks in Labor Markets," *American Economic Review (Papers and Proceedings),* LVI (1966), 559–566.
Sacerdote, Bruce, "Peer Effects with Random Assignment: Results for Dartmouth Roommates," *Quarterly Journal of Economics,* CVI (2001), 681–705.
Saloner, Garth, "The Old Boys' Network as a Screening Mechanism," *Journal of Labor Economics,* III (1985), 255–267.
Topa, Giorgio, "Social Interactions, Local Spillovers and Unemployment," *Review of Economic Studies,* LXVIII (2001), 261–295.
U. S. Department of Labor, "Findings from the National Agricultural Workers Survey (NAWS) 1997–98," Research Report No. 8, Office of the Assistant Secretary of Policy, Office of Program Economics, 2000.
Zhou, Min, *Chinatown: The Socioeconomic Potential of an Urban Enclave* (Philadelphia: Temple University Press, 1992).