ELSEVIER

Journal of Development Economics 84 (2007) 1 – 24

JOURNAL OF Development ECONOMICS

www.elsevier.com/locate/econbase

Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico ☆

David Mckenzie [a,*], Hillel Rapoport [b]

[a] *Development Research Group, The World Bank, United States*
[b] *Department of Economics, Bar-Ilan University, CADRE, University of Lille II and SCID, Stanford University, United States*

Received 31 August 2005; received in revised form 13 November 2006; accepted 16 November 2006

Abstract

International migration is costly and initially only the middle class of the wealth distribution may have both the means and incentives to migrate, which can increase inequality in the sending community. However, the migration networks formed lower the costs for future migrants, which can in turn lower inequality. This paper shows both theoretically and empirically that wealth has a nonlinear effect on migration, and then examines the empirical evidence for an inverse U-shaped relationship between emigration and inequality in rural sending communities in Mexico. After instrumenting, we find that the overall impact of migration is to reduce inequality across communities with relatively high levels of past migration. We also find some suggestive evidence for an inverse U-shaped relationship among communities with a wider range of migration experiences.

JEL classification: O15; J61; D31
Keywords: Migration; Remittances; Inequality; Network effects

☆ Thanks to Nicole Hildebrandt, Cristobal Huneeus and Pascal Lévy-Garboua for excellent research assistance; suggestions and comments from Gordon Hanson, two referees, Frederic Docquier, Andrew Foster, Kaivan Munshi, Christopher Udry and seminar participants at Boston University, Brown, UC Berkeley, UC Davis, UC Santa Cruz, the World Bank DECRG and Yale; and Chris Woodruff for generously supplying his historic migration rate data. We are grateful for financial support for this project from the Rockefeller Foundation.
* Corresponding author. Mailing address: MSN MC3-301, The World Bank, 1818 H Street N.W., Washington, DC 20433, United States. Tel.: +1 202 458 9332; fax: +1 202 522 3518.
E-mail address: dmckenzie@worldbank.org (D. Mckenzie).

0304-3878/$ - see front matter
doi:10.1016/j.jdeveco.2006.11.003

1. Introduction

The United States–Mexico border is the longest between a developed and developing country in the world and there is a long history of migration between the two countries. It is estimated that Mexican emigrants now represent more than 15% of the Mexican population of working age, against only about 3% in 1970 (Mishra, 2007). This paper examines the impact of these large emigrant flows on inequality in the rural sending communities in Mexico. Income distribution in Mexico displays a high level of inequality by international standards, and there is now a large body of both theoretical and empirical research which suggests that inequality can retard growth. To the extent that emigration is non-neutral with respect to inequality, it can therefore have important political and growth consequences for rural Mexico.

The overall impact of international migration on economic inequality at origin is a priori unclear, depending upon where migrants are drawn from in the initial wealth distribution, and on the impacts of their migration decisions on other community members. If migration costs are sizeable, migrants will be initially primarily drawn from households at the upper–middle of the community wealth distribution, causing inequality to initially increase as such households get richer from income earned abroad. In contrast, if migration costs are low or liquidity constraints do not bind, the lower part of the distribution is also able to migrate, resulting in a more neutral or even inequality-reducing effect of migration income. Related research on the education-selectivity of migration from Mexico has generally found positive selection (Cuecuecha, 2003; Chiquiar and Hanson, 2005; Mishra, 2007), although Ibarraran and Lubotsky (2006) have recently challenged this finding.[1] From a dynamic perspective, however, there are a number of channels through which past migration impacts on current migration incentives. In particular, sociologists have emphasized that social networks have a strong impact on the size of migration costs. For example, Espinosa and Massey (1997) report that social networks play an important role in mitigating the hazards of crossing the border, with friends and relatives with previous migrant experience often accompanying new immigrants across the border, showing them preferred routes and techniques of clandestine entry.[2] Munshi (2003) finds that individuals with larger networks are more likely to be employed and to hold higher paying jobs upon arrival in the U.S. As a result, net migration costs become endogenous to the migration process, as modeled theoretically in Carrington, Detragiache and Vishwanath (1996)[3], and migration is therefore likely to have different effects on inequality at different levels of a village's migration history.

Previous literature has not examined the overall impact of migration on inequality, focusing instead mainly on examination of the effect of remittances alone on inequality in only a couple of communities.[4] Stark, Taylor and Yitzhaki (1986) analyzed the direct effect of remittance income in two villages in Michoacán, Mexico, by comparing the Ginis with and without remittance income, and found remittances reduced inequality. Following a similar approach in Yugoslavia, Milanovic (1987) found that remittances increase inequality among agricultural households. Noting that migrant workers would otherwise be working and earning income at home, Adams (1989) predicted what income would have been without remittances. Using a sample of three

[1] Although these studies have found Mexican migrants to be positively selected with respect to the Mexican educational distribution, over time Mexican immigrants' relative education levels have fallen compared to U.S. natives (Borjas and Katz, 2005).

[2] Note that most migrants making their first trip to the U.S. do so without documents.

[3] See also Kanbur and Rapoport (2005).

[4] See Rapoport and Docquier (2006) for a survey of the causes and consequences of migrants' remittances.

villages in Egypt, he then found that the inclusion of remittances from abroad worsens inequality. Taylor (1992) and Taylor and Wyatt (1996) noted that in addition to the direct immediate impact on income, remittances can ease credit constraints for liquidity constrained households. Using a sample of 55 households from one part of Michoacán in Mexico, they found a greater impact of remittances on inequality reduction once these indirect effects were included. Barham and Boucher (1998) used data from 3 neighborhoods in Bluefields, Nicaragua, and estimated a double-selection model to allow for the counterfactual of no migration and no remittances to impact on the participation decisions and earning outcomes of other household members. Treating remittances as exogenous would lead them to conclude that remittances reduce income inequality, whereas treating them as a substitute for home earnings results in remittances increasing inequality.

Our methods allow, and indeed force, us to examine the overall impact of migration on inequality. This is composed of the direct and indirect effects of remittances detailed above, multiplier effects of remittances through their spending on products and services produced by other community members, and other potential spillover and general equilibrium effects; this also includes the network effects of migration on the costs and benefits of migration for other community members. We begin by writing down a simple theoretical model of rural migration, show that it leads to a non-monotonic relationship between migration and wealth for a given cost of migration, and then use it to examine the consequences of changes in costs and benefits which might arise from the presence of networks. Two data sets are used for the empirical analysis: the Mexican Migration Project (MMP) survey consists of data from 57 rural communities typically located in areas of high migration, while the national demographic dynamics survey (ENADID) consists of a representative sample of 214 rural communities in Mexico. This enables us to construct data on inequality and migration for a large number of communities with a range of different migration experiences, in contrast to previous case studies which focused on only a couple of villages, typically in areas of high emigration. Since there are likely to be unobserved community level factors correlated with both household migration decisions and current inequality, we employ an instrumental variables strategy to isolate the impact of community level migration on household propensities to migrate and on inequality. The main instruments employed are historic state-level migration rates and U.S. labor market conditions.

These data are first used to empirically confirm the non-monotonic relationship between migration and household resources predicted by our theoretical model. We find support for our model's prediction that household resources and migration networks interact in determining migration propensities. In particular, among households with low networks we find evidence that a marginal increase in the network size increases the likelihood of migrating more for wealthier households, whereas poorer households benefit more from a marginal increase in network size once network size becomes larger. As a result, less wealth is needed in order to be able to migrate in communities with larger networks.

We then focus on the empirical relationship between emigration and inequality in the sending communities. In accordance with our theoretical predictions, we find that further migration reduces inequality among communities with reasonably high initial levels of migration experience, and we find small positive but insignificant effects of migration on inequality in communities with smaller migration networks. Employing panel data for a sample of communities observed in both 1992 and 1997, we do find suggestive evidence for an inverse U-shaped relationship between migration and inequality, with migration increasing inequality at first before subsequent migration lowers it. These findings suggest that migration may increase inequality in the sending regions in the short term, but that this should be viewed as a temporary phenomenon, as migration has a strong equalizing effect once migration networks are sufficiently developed.

2. The model

In this section we first discuss how the impact of wealth on the migration decision depends on migration costs. Since in practice the primary cross-section variation in migration costs comes from migration networks, we then discuss the testable implications of our model with respect to the likely effect of such networks on interhousehold inequality.

2.1. Migration costs and the wealth–migration relationship

Consider a family of size N making its living from agriculture, with initial illiquid household wealth A, such as land holdings. Family members are assumed to live for two periods. Since our prime interest lies in the study of interhousehold inequality, we assume for simplicity that income is equally shared between members of the same family. In the first period, all members are in Mexico, and each household member inelastically supplies one unit of labor to household production. Total farm production with L workers is $AL-\frac{bL^2}{2}$. The marginal product of farm labor is linearly increasing in wealth and decreasing in the number of workers.[5] A household member can migrate to the U.S. and earn the foreign wage w by incurring a fixed migration cost c, which is initially assumed to be fixed and exogenous. To account for credit market imperfections we assume that migration costs cannot be financed by borrowing, and so no household member can migrate in the first period. In the second period, households must instead use savings from the first period to finance migration, after having met the first period subsistence needs of I per member. We assume $w>I$ and that $A-\frac{bN}{2}\geq I$.

The household's problem is to choose the proportion of members who migrate, m. We assume no discounting, so the household makes this decision to maximize second period household income net of migration costs, subject to the subsistence constraint:

$$\max_{\{m\}} \mathrm{AN}(1-m)-\frac{bN^2(1-m)^2}{2}+Nm(w-c) \quad \text{s.t. } A-\frac{bN}{2}-mc\geq I. \tag{1}$$

Solving the first-order condition for the optimal household migration rate, m^*, gives:

$$m^* = 1-\frac{[A-(w-c)]}{bN}-\frac{\lambda c}{bN^2} \tag{2}$$

The Lagrange multiplier $\lambda=0$ unless (1) binds.[6] When (1) binds, we can solve for the constrained migration rate:

$$\tilde{m} = \frac{1}{c}\left(A-\frac{bN}{2}-I\right). \tag{3}$$

[5] Although the model is written in terms of farm production, it can also be more generally applied to other home production and family businesses, in which labor is a complement to capital in production.

[6] When (1) binds, we have $\lambda=\frac{N}{2c^2}(2bNc-2[A-w+c]c-2bNA+b^2N^2+2bNI)$.

From (3) we can solve to find the highest level of assets at which a household is constrained by subsistence needs to not have any migrants, $\underline{A}=\frac{bN}{2}+I$. Let us now see how the rate of migration changes with the level of wealth, A. From (2) we have:

$$\frac{\partial m^*}{\partial A}=\begin{cases}-\frac{1}{bN} & \text{when} \quad \lambda=0\\ \frac{1}{c} & \text{when} \quad \lambda>0\end{cases}. \tag{4}$$

Interpreting (4), we see that when subsistence constraints bind, increasing wealth increases migration, the extent to which depends on migration costs c. When subsistence constraints no longer bind, an increase in wealth merely causes the opportunity cost of migrating to increase in terms of lost household production, and so households will reduce migration, the extent to which depends on productivity. Using (2) and (3) to find the level of A at which $m^*=\tilde{m}$, gives a level of assets A_1 above which households are no longer bound by the subsistence constraint:

$$A_1=\frac{1}{2}\frac{b^2N^2+2bNI+2bNc+2cw-2c^2}{c+bN}. \tag{5}$$

Finally, we see in (4) that m^* is decreasing in A for $A>A_1(\lambda=0)$. Also, we can then find the lowest asset level at which unconstrained households will optimally choose no migration, $\bar{A}=bN+(w-c)$. Note that this is increasing in the net benefit from migration, $(w-c)$. So putting this altogether we have:

$$m^*=\begin{cases}0 & \text{when} \quad A\leq\underline{A}\\ \frac{1}{c}\left(A-\frac{bN}{2}-I\right) & \text{when} \quad \underline{A}\leq A\leq A_1\\ 1-\frac{[A-(w-c)]}{bN} & \text{when} \quad A_1\leq A\leq\overline{A}\\ 0 & \text{when} \quad A\geq\overline{A}\end{cases}. \tag{6}$$

That is, a household's migration rate will be a triangular function of assets, with migration increasing with wealth at low levels, and decreasing with wealth at higher levels (see the solid line in Fig. 1, which shows this relationship).

In our model, reducing costs has two effects on the desired level of household migration. First, for a given unconstrained level of desired migration, a reduction in migration costs makes it less

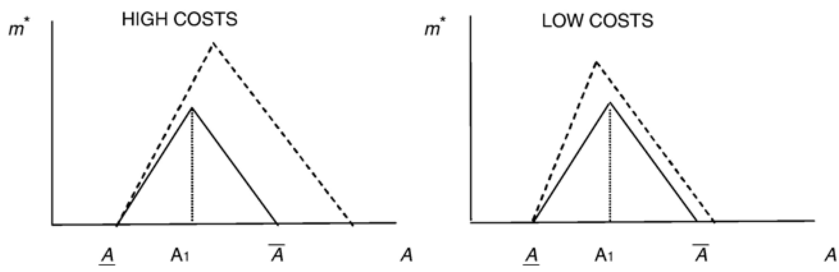

Fig. 1. Relationship between migration rate (m^*) and Asset Wealth (A), and effect of a reduction in migration cost on optimal household migration rate according to initial level of costs. Solid lines are at original costs, dotted lines at the new lower cost.

likely that subsistence concerns will prevent migration from reaching this desired level. And second, a reduction in migration costs also increases the net benefits of migrating, making households want to migrate more, and thereby increasing their likelihood of being constrained. While the first effect tends to reduce A_1, the asset level at which households are no longer constrained, the second effect tends to increase it. One can show that which effect dominates depends on whether migration costs are high to begin with, in which case the second effect dominates, or low to begin with, in which case the first effect dominates.[7] Fig. 1 plots the effect of a reduction in migration costs for initial situations of high and low costs. In both cases we see that lowering migration costs increases desired household migration rates at any asset level at which there was initially some migration, and also induces additional individuals to migrate.

2.2. The role of migrants' networks

In practice the primary cross-section variation in migration costs comes from migration networks. Indeed, as migration proceeds, migrant networks form gradually and then act to lower the cost of migration for other migrants.[8] What is then the likely impact of migration on inequality? Sociologists such as Massey, Goldring and Durand (1994) have observed that the first migrants from a community usually come from the middle class, and are the individuals who have enough resources to absorb the costs of the trip, but are not so affluent that foreign labor is unattractive. It therefore seems reasonable to interpret the level of wealth A_1, at which the initial turning point occurs, as being middle class. Initially migration will therefore occur from the middle of the distribution. These households will become richer due to remittances. When migration costs are high to begin with, the first network effects then will increase migration opportunities more for the middle and upper–middle classes, as is seen in the High Cost scenario in Fig. 1. Migration may therefore increase inequality at first as the upper middle of the distribution gets wealthier from remittances.[9] However, as the network grows larger and migration costs continue to fall, we see from the Low Cost scenario that further reductions in migration costs will benefit primarily the lower and lower–middle classes in the village, which will reduce inequality.[10]

This gives rise to the three chief predictions to be tested in this paper: (1) initial network formation raises migration incentives and propensities, but more so for relatively richer households; (2) further migration will reduce inequality when migration networks are reasonably high to begin with; and (3) there is a possibility of an inverted U-shaped relation between migration and inequality when migration networks are initially low. Note that only some of these predictions could be obtained through general equilibrium effects such as migrant household using remittance income to purchase labor or goods offered by non-migrant households. In particular, the effect of networks is to increase both the ability and desire to migrate and is likely, as explained, to impact more on relatively affluent households when networks are initially small and,

[7] In terms of our notation,

$$\frac{\partial A_1}{\partial c} \underset{>}{\leqq} 0 \Leftrightarrow c \underset{<}{\geqq} -bN + \frac{1}{2}\sqrt{6b^2N^2 + 4(w-I)bN}$$

[8] There is now an important literature demonstrating that this is indeed the case. See for example Massey et al. (1994), Carrington et al. (1996), Orrenius (1999), Bauer et al. (2002), Gathmann (2004), or Dolfin and Genicot (2006).

[9] Recall here the assumption is that the household shares all income equally, so total household income increases by $w-c$.

[10] The migrant network may also provide loans to potential migrants, which information costs prevent other potential lenders from making. The result is that even households with wealth below $\underline{A}$ may eventually be able to migrate, further lowering inequality.

conversely, to impact more on less affluent households once networks are sufficiently large. In contrast, general equilibrium effects (e.g., if initial migration causes wages at home to rise in response to the lower labor supply and higher labor demand of some households) would increase both the ability to migrate and the opportunity cost of migration, thus decreasing the desired level of migration. Hence, these effects can contribute to reduce inequality but are unable to explain the initial increase in migration incentives and propensities among households at the middle–upper end of the wealth distribution. On the other hand, it is possible that remittances are partly used to increase households' assets and initially serve to relax credit constraints for richer households only. This in turn can explain the initial rise in inequality, but not its subsequent decrease.[11]

This simple model illustrates the likely effects of increasing network size on the selectivity of migration and on inequality. This effect has received the most attention in the sociological literature, and we believe it to be of crucial importance in explaining how the impact of migration on inequality varies across communities. This motivates our focus on migration networks in the empirical analysis. The presence of migrant networks makes the migration process dynamic, and in practice migration will also be accompanied by other dynamic effects such as the general equilibrium effects discussed above. While our empirics will consider the overall impact of previous migration on inequality, using the insights of and showing results consistent with the model exposited here, data constraints prevent us from implementing a fully dynamic empirical analysis and so we will examine the migration–inequality relationship using a cross-section of communities with a wide range of migration histories.[12]

3. Data

Mexico has some of the most comprehensive surveys of migration available for any developing country. In order to examine the effect of migration on inequality in the sending communities, one would ideally like to have individual and community level panel data on both assets and migration. While no single survey fits this criterion, we use two surveys which approach it: the Mexican Migration Project (MMP) data and the Encuesta Nacional de Dinámica Demográfica (ENADID).

The Mexican Migration Project is a collaborative research project based at the Princeton University and the University of Guadalajara.[13] The MMP surveyed five communities in 1982, between two and five communities each year from 1987–97, and fourteen communities in 1998. In general, 200 households were surveyed in each community, with smaller samples taken in communities with less than 500 residents. We use the MMP71 database, which contains data on 71 communities. Since our theoretical model applies best to rural communities and small towns, we restrict our analysis to the 57 communities which had a population below 100,000 in 1990. Each community is surveyed only once, but household heads are asked entire life retrospective migration histories. In addition, the survey asks for each individual in the household whether they have ever been to the United States, and if so, in what year was their first migrant experience. This

[11] As shown by Docquier and Rapoport (2003), an inverse U-shape relationship between migration and income inequality can theoretically be obtained with constant migration costs by combining remittances invested in households' assets and wage effects. The required assumptions appear very unrealistic in the context we study. Notably, the richer classes must be initially credit-constrained, and initial inequality must be sufficiently low.

[12] We do not have a long panel of data on inequality and migration (no survey in the world that we know of does), and so we have to estimate a dynamic relationship off of the cross-section.

[13] Full details of the methodology, the data, and the questionnaires are available at http://mmp.opr.princeton.edu.

enables the construction of a time-series of the stock of current residents in a community who had migrant experience in a given year.[14] In addition to questions about migration, households are asked about their current household infrastructure and durable asset ownership. The dataset also contains community level variables taken from past years of the Mexican Census. The survey is typically taken in December and January, which is when traditionally most migrants return to their communities, but if initial fieldwork suggests migrants tend to return during other months instead, a portion of the interviews are conducted then.

The MMP surveys have the advantage of containing the most detailed migration data, allowing construction of both community and household head panel data on migration and migration networks. However, since data on assets are collected only for the survey year, we only have cross-sectional data on inequality for each community. Moreover, although migration history itself is not an explicit criteria in selection of communities, the survey contains data from only 13 of Mexico's 32 states, with many of the surveyed communities coming from the traditional migrant-sending states in West-Central Mexico. Hanson (in press) and Orrenius and Zavodny (2005) examine the selectivity of the MMP, and conclude that it is reasonably representative for seasonal and agricultural migrants, not the general population. For these reasons we also carry out some estimation using data from the ENADID.

The ENADID is a national demographic survey intended to provide information on fertility, infant and general mortality, national and international migration, births, deaths and contraceptive practices. It was taken in 1992 and again in 1997 by Mexico's national statistical agency, the Instituto Nacional de Estadística, Geografía e Informática (INEGI, 1994, 2001). Approximately 2000 households were surveyed in each state, with a total sample size of 57,017 households in 1992 and 73,412 households in 1997. We use the 1997 survey for the majority of our analysis since it contains more information on the variables of interest.

The ENADID asks whether household members have ever been to the U.S. in search of work. This question is asked of all household members who normally live in the household, even if they are temporarily studying or working elsewhere, and an additional question asks whether any household members have gone to live in another country in the past five years. Thus U.S. migrants are recorded as long as they return to Mexico or have family members remaining in the community. We restrict our analysis to households sampled in municipalities which are outside of cities of population 100,000 or more, and in which 50 or more households were sampled in order to measure the community migration network, and use the number of households in the municipality as weights in our analysis. This gives a sample of 214 rural municipalities in the 1997 survey from all 32 Mexican states. Although the same households were not sampled in both years, some of the same municipalities were. We are able to match 33 municipalities in which 100 or more households were surveyed in both 1992 and 1997.[15] This matched data forms a short panel on inequality and migration which will be used to examine how changes in migration relate to changes in inequality.

[14] Retrospective migration histories are likely to be subject to recall bias (see e.g. Smith and Thomas, 2003). However, our main measure, migration prevalence, measures whether an individual has *ever* migrated, which is a salient event that should be less subject to recall bias. The other variable we use, migration in the past two years, measures a recent event, which again should not be subject to large recall bias.

[15] One concern is that since we observe a different sample of individuals from a given municipality in 1997 than we did in 1992, some of the changes in inequality and migration prevalence found in the data may just be the result of small sample measurement-error. Verbeek and Nijman (1992) find that with sample sizes of 100–200 households, pseudo-panel data provides a good approximation to genuine panel data and hence we restrict our panel analysis to communities with at least 100 households.

The MMP and ENADID surveys provide the most comprehensive information available about Mexican migration to the United States. However, neither survey contains data on consumption or complete information on income. The surveys do collect a variety of measures of household infrastructure, such as whether the household has a dirt or tile floor, and access to running water, electricity and sewerage facilities. The MMP additionally asks whether households own certain durable assets, such as a car, radio, television, stove and fridge. McKenzie (2005) shows how this information can be used together with the national income and expenditure surveys (ENIGH), which contain data on both consumption and asset indicators, in order to predict non-durable consumption and inequality for each community. The inclusion of durable asset indicators in the MMP survey additionally allows us to form a measure of relative asset inequality, defined in McKenzie (2005) as the standard deviation within a community of the first principal component of these assets, divided by the standard deviation within all communities.

Both the MMP and ENADID surveys only contain data on migrants who have either returned to Mexico, or who have at least one household member remaining in Mexico, excluding households which have migrated to the United States in their entirety. This tends to underrepresent permanent migrants (Hanson, in press), who are more likely to take their whole household. By comparing the U.S. and Mexico Censuses, Ibarraran and Lubotsky (2006) conclude that males from these excluded households are likely to be more educated. As a result, omitting these individuals will lead us to understate the degree of overall positive selection in migration. However, since our main outcome of interest is the effect of migration on inequality in the sending communities, looking only at the selectivity among migrants from households in which a household member remains in these communities will still be informative for what we should expect to see in terms of changes in inequality.

Table 1 provides summary statistics for both the MMP and ENADID for the key variables used in this paper. Our main measure of migration at the community level is the migration prevalence ratio, defined by Massey, Goldring and Durand (1994) to be the proportion of all individuals aged 15 and over in a given community who have ever migrated in the reference year. Fig. 2 plots the estimated density functions of migration prevalence for the two surveys. We see that the MMP survey largely omits communities with low levels of migration prevalence. This difference is not due to differences in measurement across the two surveys since Massey and Zenteno (2000) find that the two surveys match well along a number of dimensions, including migration prevalence, when compared over geographically similar communities. As a result of these differences, according to our theory we should find migration reducing inequality by more for the MMP sample than for the ENADID sample since the former has larger migrant networks, and hence lower migration costs on average.

4. Instrumental variables strategy

Our empirical work will examine the impact of the migration prevalence in the community a household lives in on that household's own propensity to migrate and on inequality in the community. A potential concern is that there are unobserved community characteristics which are correlated with both migration prevalence and current inequality. For example, Munshi (2003) finds a strong negative correlation between rainfall and migration to the United States. Rainfall and other community shocks would be expected to affect community inequality and migration rates. To account for this possibility we follow Woodruff and Zenteno (2007) in using historic state-level migration flows as instruments for current migration.

Table 1
Summary statistics of key variables by community

Variable	Mean	Std. Dev.	Percentiles		
			25th	50th	75th
MMP COMMUNITIES					
Household variables					
Proportion of non-migrant heads making first trip in last two years	0.026	0.159			
Log non-durable consumption	8.802	0.554	8.418	8.769	9.175
Community variables					
Community migration prevalence	0.259	0.142	0.158	0.243	0.344
State migration rate in 1924	0.013	0.008	0.007	0.013	0.020
State migration rate in 1955–59	0.025	0.016	0.013	0.025	0.032
Proportion of heads aged under 30	0.115	0.046	0.080	0.110	0.150
Proportion of heads aged over 60	0.241	0.067	0.198	0.240	0.285
Proportion of heads with education <6 years	0.630	0.191	0.500	0.602	0.750
Proportion of heads with education ≥9 years	0.161	0.112	0.065	0.140	0.245
Relative asset inequality (I_c)	0.871	0.171	0.763	0.882	0.954
Gini of non-durable consumption	0.406	0.025	0.389	0.404	0.420
ENADID 1997 COMMUNITIES					
Household variables					
Proportion of non-migrant heads making first trip in last two years	0.018	0.132			
Log non-durable consumption	8.660	0.390	8.385	8.655	8.971
Community variables					
Community migration prevalence	0.213	0.202	0.038	0.159	0.356
State migration rate in 1924	0.006	0.008	0.000	0.001	0.012
State migration rate in 1955–59	0.014	0.015	0.003	0.008	0.025
Proportion of heads aged under 30	0.167	0.056	0.128	0.165	0.200
Proportion of heads aged over 60	0.212	0.077	0.162	0.207	0.255
Proportion of heads with education <6 years	0.533	0.177	0.426	0.533	0.652
Proportion of heads with education ≥9 years	0.239	0.149	0.135	0.224	0.323
Gini of non-durable consumption	0.400	0.016	0.389	0.398	0.411

Note: household variables are for male household heads aged 15–49 who had no migrant experience two years prior to the survey.

The first instrument we use is the U.S. migration rate for 1924 for the state in which each community is located, taken from Foerster (1925). These state-level historic migration flows may be argued to be a result of largely historic demand-side factors coupled with the pattern of arrival of railroads into Mexico. Massey, Durand, and Malone (2002) outline how restrictions on immigration from Asia coupled with a booming economy in the Southwest of the United States lead U.S. employers to hire “enganchadores” (contractors) to obtain as many workers as possible. These enganchadores followed the railroads south into Mexico, stopping in the first sizeable population centers they encountered to hire workers, which were in the West-Central Mexican states. The arrival and lay-out of the railroad system thereby led to some states having different migration rates than others. Then, following the model of Carrington, Detragiache and Vishwanath (1996), one can argue that the development of these migration networks lowered migration costs for subsequent migrants and so these differences in initial networks will result in different levels of current migration prevalence.

In addition to the 1924 state migration rates, we also follow Woodruff and Zenteno (2007) in using migration rates over the 1955–59 period by state, taken from González Navarro (1974).

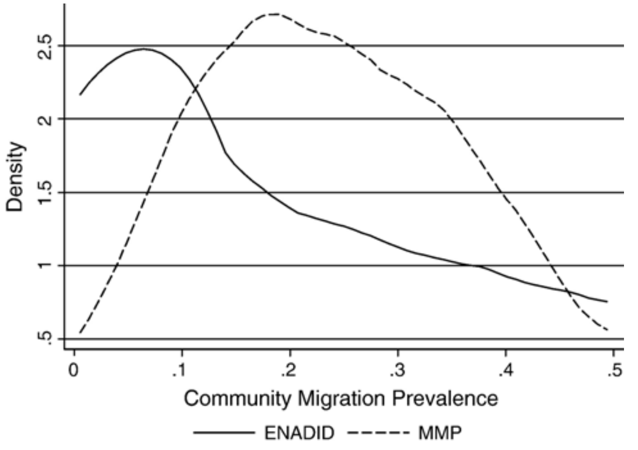

Fig. 2. Density of community migration prevalence ratios in MMP and ENADID samples.

These rates are from the peak period of the 1942–64 bracero program. This program allowed for the legal entry of temporary farm workers, providing up to 450,000 work visas annually to Mexicans during the peak years, and allowed for the immigration of around 5 million Mexicans into the United States (Massey, Durand and Malone (2002)). A sharp break in U.S. immigration policy in 1965 ended this program, and undocumented migration came to greatly outnumber legal migration in the subsequent period. In the MMP sample, on average 89% of first-time migrants between 1970 and 1990 was undocumented and a further 7% entered on tourist visas.[16] Migration networks formed during this bracero period are expected to contribute to current community migration prevalence rates, but we do not believe the bracero program to have additional impacts on current community levels of inequality, given the period of over thirty years since its peak.

The main threat to the validity of these historic migration rates as instruments for current migration is the possibility that persistent inequality was a factor determining migration both historically and at present.[17] Table 2 examines the plausibility of the exogeneity restriction needed for the validity of these instruments. The earliest inequality measures we can calculate use the 1.5% sample from the 1960 Census to obtain Gini coefficients and coefficients of variation for labor income by state. A Spearman test of independence can not reject the null hypothesis of no relationship between inequality in 1960 and migration rates in 1924 and 1955–59. The only data available at the state-level prior to 1960 are socioeconomic measures which are the results of a combination of inequality and poverty rates. We can not reject that the historic migration rates are independent of infant mortality and school attendance rates in 1930. Finally, we consider the relationship between the percentage of rural households owning land in 1910 and subsequent migration rates in 1924 and 1955. There is a significant positive relationship with the 1924 migration rate, so that households from more equal states are more likely to migrate. However, by 1955–59 this relationship is insignificant.

[16] First-time illegal migration is also the norm amongst migrants from the upper socioeconomic classes in the MMP. Between 1970 and 1990, 72% of first-time migrants with 12 years or more education migrated illegally.

[17] The model of Borjas (1987) provides one rationale for this possibility.

A second potential threat to the validity of the instrument is that the development of the railroad system in Mexico, and the corresponding historic migration rates may affect current outcomes through channels other than current migration rates. For example, historic migration may have influenced the level of current development in a state, the industrial structure, governance, and other infrastructure, all which may also influence inequality. To investigate this, Table 2 also looks at the correlations of the historic migration rates with current levels of state GDP per capita, the agricultural share of production in a state, the level of corruption in a state measured by the Mexican division of Transparency International[18], and measures of health infrastructure. We can not reject the null hypothesis of independence between the historic migration rates and any of these variables.[19] Taken together, the results from Table 2 therefore provide support for the exogeneity restriction required for our instruments.[20]

The MMP survey contains more detailed data on migration patterns enabling us to form a second set of instruments, consisting of demand-side variables from the United States. These variables affect the costs and benefits of migrating, but have no other direct impact on rural Mexican communities. For each MMP community, one can identify the most common U.S. city destination for migrants on their first trip to the U.S. Differences in geographic proximity and historic migration patterns will mean that different communities will tend to cluster at different U.S. destinations. The unemployment rate[21] in the U.S. state in which this destination city is located will then affect migration from that community to the U.S. However, since we need to instrument migration stocks rather than flows, we aggregate up unemployment in each of the ten years prior to the year in which migration prevalence is measured, and weight by the proportion of current household heads who were of prime migrant age, 20–30 years, in that year. Similarly, we also use the real depreciation[22] of the peso against the U.S. dollar, weighted by prime age population in each of the ten years prior to the year at which migration prevalence is measured as an instrument. Different communities were surveyed in different years in the MMP, and have different cohort sizes of prime migration age in the years in which large depreciations were realized, resulting in differences in the effective exposure to depreciation faced by each community. These additional instruments will be used together with the historic migration network variables to form a test of overidentification for the instruments.

5. Networks and the migration–wealth relationship

The theoretical model presented in Section 2 provides two key testable predictions for the cross-sectional relationship between migration and household resources. The first is that provided migration costs are sufficiently large that subsistence and liquidity constraints bind for

[18] Data are from the 2003 Survey on Corruption, www.transparenciamexicana.org.mx.

[19] This may strike the reader as surprising, given that we will demonstrate that migration affects the level of inequality in a community, and that we would expect household wealth and inequality to in turn affect the development process. However, at the state level over long periods of time, the effect of migration may be small relative to other economic transformations occurring, resulting in a weak cross-state relationship between historic migration and levels of development.

[20] A third possible threat to using this instrument is that historic migration affects inequality today through the migration of previous generations, rather than of the current generation. The MMP allows us to broaden the definition of migration prevalence to also include households whose parents or parents-in-law were migrants, even if they themselves are not migrants. Our results are robust to using this broader definition of migration prevalence, suggesting that this possible threat is not a concern in our data.

[21] Source: U.S. Bureau of Lab or Statistics.

[22] Source: annual average exchange rate and CPI series from the Bank of Mexico, www.banxico.gob.mx.

Table 2
Exogeneity of historic migration rates

State-level variables	Spearman rank-order correlation with state migration rate from	
	1924	1955–59
Proxies for historic inequality		
Proportion of rural households owning land in 1910	0.420*	0.155
Proportion of 6–10 year olds attending school 1930	0.041	−0.053
Infant mortality rate 1930	0.227	0.329
Gini of labor income 1960	−0.274	0.104
Coefficient of variation of labor income 1960	−0.239	0.141
Current economic and political factors		
State GDP per capita	0.311	0.069
Agricultural share of production	0.095	0.144
Transparency international state-level corruption index	−0.333	−0.274
Hospitals per 100,000	−0.046	−0.276
Doctors per 100,000	0.138	−0.095
Nurses per 100,000	0.258	0.003

* and **indicates that the Spearman test of independence rejects the null of no independence at the 5% and 1% levels sources: landholdings are from the 1910 census, reported in McBride (1928); school attendance in 1930 is from the 1930 census, reported in DGE (1941); infant mortality rates are from INEGI (2001); income inequality measures from 1960 are own calculations from the 1% sample of the 1960 Census.
Current variables are from INEGI and Transparency International (Mexico).

some households, migration rates should first increase and then decrease with wealth. The second prediction is that provided migration costs are not too high to begin with, as migration costs become lower, migration propensities should increase more for poorer households. Since detailed wealth data are not available, we use the log of non-durable consumption (log ndc) as our measure of household resources, A. The community migration prevalence is used to measure migration networks, n, which are assumed to be inversely related to migration costs. These results in the following equation for the probability of migration, p:

$$p = \beta_0 + \beta_1 A + \beta_2 A^2 + \beta_3 n + \beta_4 (A^* n) + \varepsilon \tag{7}$$

where the theory predicts that $\beta_1 > 0$, $\beta_2 < 0$, $\beta_3 > 0$ and that $\beta_4 < 0$ when migration costs are relatively low.[23]

Since only current non-durable consumption is available, which is likely to reflect in part the result of any previous migration trips, we focus here on first-time migrants and estimate the probability that a household head migrated for the first-time within the last two years, conditional on not having previously migrated. We restrict our analysis to male household heads aged 15–49 years at the time of the survey, since more data is available on the migration history of heads, and individuals are highly unlikely to migrate for the first-time for work outside of this age range. As noted in the data section, estimation will be informative about the degree of selection among individuals who have a household member remaining in the community, which is what we are interested in for looking at the effect on inequality in a community.

[23] In contrast, the effect of migration through increasing wages would predict $\beta_4 > 0$.

Table 3
Network size and probability of migration in the MMP sample

Probability of household head first migrating in survey year or year prior to survey year	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	IV	IV	IV	IV	IV	IV
Log non-durable consumption (log NDC)	0.2090	0.1804	0.1809	0.1833	0.0280	0.2078	0.0069
	(2.32)**	(1.72)*	(1.62)	(1.78)*	(1.55)	(2.30)**	(0.02)
Log NDC squared	−0.0119	−0.0087	−0.0081	−0.0089		−0.0121	0.0010
	(2.34)**	(1.45)	(1.18)	(1.51)		(2.39)**	(0.05)
Migration prevalence	0.2882	1.4564	1.8395	1.4183	1.5476	0.1171	−2.1594
	(0.79)	(1.78)*	(1.39)	(1.87)*	(1.92)*	(2.23)**	(0.27)
Migration prevalence * log NDC	−0.0235	−0.1508	−0.2014	−0.1484	−0.1608		0.6625
	(0.58)	(1.71)*	(1.38)	(1.80)*	(1.85)*		(0.37)
Migration prevalence * log NDC squared							−0.0455
							(0.46)
Instrument set		A	B	C	A	A	A
Overidentification test p-value				0.43			
First-stage F-stat for migration prevalence		8.93	4.01	3.99	8.91	14.44	11.45
First-stage F-stat for migration prevalence*log ndc		9.25	4.06	4.19	9.20		11.48
First-stage F-stat for migration prevalence*log ndc squared							11.46
Observations	3116	3116	3116	3116	3116	3116	3116
Number of communities	57	57	57	57	57	57	57

T-statistics in parentheses with standard errors clustered at the community level.
*Significant at 10%; ** significant at 5%; *** significant at 1%.
For male household heads aged 15–49 who have not previously migrated.
Instrument set A consists of 1955–59 state migration rate and the interaction of this rate with log NDC, Instrument set B uses the 1924 state migration rate and its interaction with log NDC. Instrument set C combines A and B and adds the average unemployment rate in the principal U.S. destination over the past 10 years weighted by the prime age population in the sending community and the weighted depreciation rate.

We employ a linear probability model to estimate Eq. (7), which enables us to use standard two-stage least squares (2SLS) estimation once we instrument migration networks.[24] Table 3 presents the results from estimating Eq. (7) for the MMP sample. The signs of the estimated parameters concur with the predictions of the theory. The probability of migrating is found to first increase and then decrease with household resources, be higher in communities with larger networks, and be less likely to be increasing in wealth as network size increases. The interaction between the community migration network and log ndc is significant at the 10% level when the 1955–59 migration rates are used as instruments, and when a full set of instruments which also includes unemployment rates in the U.S. and depreciation rates are used. A test of overidentification fails to reject the overidentifying restrictions, lending further credence to our identification assumptions.

The last three columns of Table 3 carry out robustness checks. Column 5 shows that when only a linear term is included in log ndc, the term is positive, but insignificant, highlighting the importance of including the quadratic term, as the theory suggested. Column 6 shows further that

[24] Similar results were obtained using maximum-likelihood estimation of Amemiya's generalized least squares estimator (see Newey, 1987), sometimes referred to as IV-Probit estimation.

the quadratic term is significant at the 5% level if we drop the interaction term. Finally, column 7 shows that the interaction between migration prevalence and log ndc squared is small and insignificant, confirming that only a linear interaction term is needed.

Fig. 3 graphs the estimated probability of migration from column 2 of Table 3 at different deciles of the MMP community migration prevalence distribution. The probability of migrating is estimated to be mostly increasing in household resources for households in communities with relatively low migration networks, quadratic for households with middle levels of the network, and decreasing in wealth for households in the 60th decile and above of community migration prevalence. The shape of the migration profile for these households is that predicted by theory when migration costs are close to zero, so that no households are constrained.

The coefficients from estimating Eq. (7) for the ENADID sample are provided in Table 4. The full ENADID sample has almost four times as many communities as the MMP, over a wider range of community migration prevalence rates. The resulting parameter estimates are again in accordance with the signs predicted by theory, and are all significant at the 5% level for the 2SLS estimation. Again we find a significant negative interaction term between household resources and community migration networks, but no significant interaction with squared resources. As discussed in Section 2, at high levels of migration costs, corresponding to low migration networks, a larger network may actually benefit the upper–middle of the wealth distribution more than the lower part of the distribution. As a result, one should find β_4 to be less negative, or possibly positive, when Eq. (7) is estimated only for communities with lower levels of migration. The last three columns of Table 4 restrict estimation to the 90 ENADID communities with migration prevalence less than 0.10, that is to communities in which less than 10% of adults have ever migrated to the U.S. Over 40% of the ENADID communities satisfy this restriction, compared to just 7 of the 57 MMP communities. The magnitude of the interaction between network size and household resources falls for this subsample, and is insignificantly different from zero.

Fig. 3 also plots the estimated relationship between migration and log non-durable consumption for different deciles of the ENADID community migration prevalence distribution using column 6 of Table 4 for the bottom 4 deciles and column 2 of Table 4 for the upper deciles. In accordance with the high migration costs scenario presented in Section 2, when migration networks are low, increasing the network size shifts the turning point to the right, and so the upper–middle of the distribution benefits more from increasing the network. Once the network gains sufficient size the turning point begins to move left and at high levels of network, we have the same pattern of declining propensities of migration with wealth seen in the MMP.

Comparing the MMP and ENADID results, we see that in the MMP sample, an increase in network size always results in more of an increase in migration propensity for the poor than for the rich. The interpretation is that the majority of MMP communities have sufficiently large migration networks that we are in the low migration cost scenario discussed in Section 2. As a result, we would predict that migration should unambiguously decrease inequality in the MMP sample. The ENADID results suggest a less clear relationship between migration and inequality for the full ENADID sample. Among communities with low networks, a rise in community migration prevalence increases the probability of migration most for the middle of the distribution. As such, one would expect the middle to get relatively richer compared to both the top and the bottom. This has an uncertain effect on the Gini, but should increase measures of inequality which are more sensitive to inequality among the poor, such as the

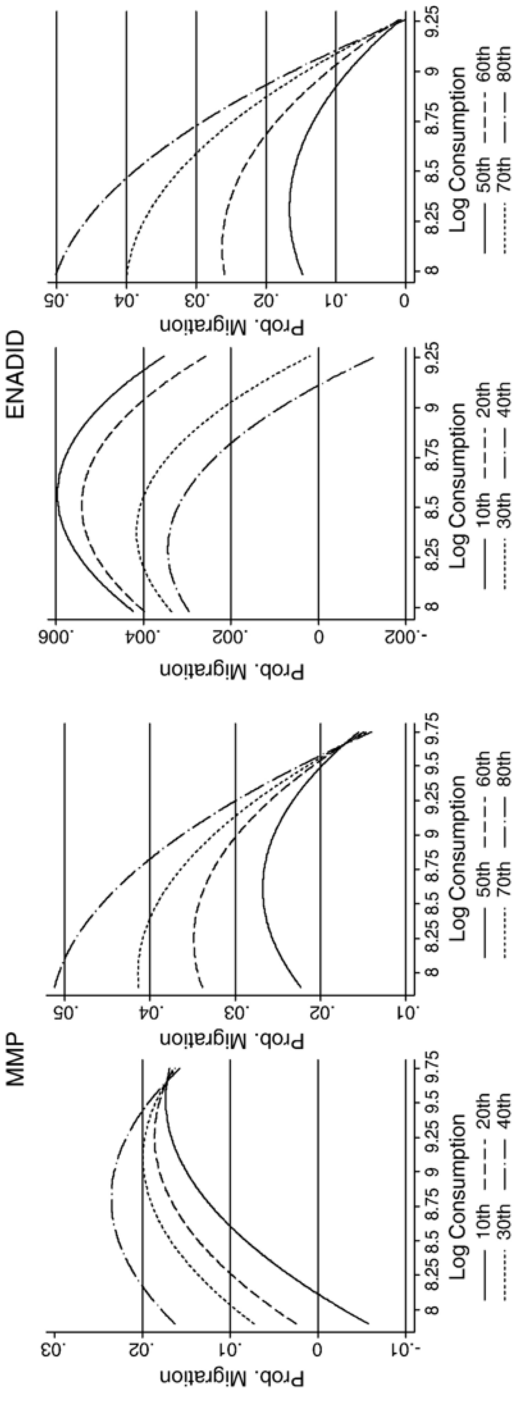

Fig. 3. Estimated probabilities of migrating at different deciles of the MMP and ENADID migration prevalence distributions.

Table 4
Network size and probability of migration in the ENADID 1997 sample

Probability of household head first migrating in survey year or year prior to survey year	Full sample				Low network sample		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	OLS	IV	IV	IV	OLS	IV	IV
Log non-durable consumption (log NDC)	0.3328	0.3318	0.3299	0.2772	0.0816	0.0926	0.0886
	(3.46)***	(3.28)***	(3.23)***	(1.40)	(1.51)	(1.72)*	(1.62)
Log NDC squared	−0.0196	−0.0189	−0.0189	−0.0158	−0.0049	−0.0054	−0.0052
	(3.48)***	(3.16)***	(3.13)***	(1.40)	(1.53)	(1.74)*	(1.61)
Migration prevalence	0.7725	1.2828	1.2169	−0.2649	0.3525	0.2316	0.2300
	(4.63)***	(3.16)***	(2.68)***	(0.04)	(0.49)	(0.47)	(0.22)
Migration prevalence * log NDC	−0.0785	−0.1386	−0.1305	0.2212	−0.0283	−0.0329	−0.0248
	(4.12)***	(2.97)***	(2.52)**	(0.14)	(0.33)	(0.59)	(0.21)
Migration prevalence * log NDC squared				−0.0209			
				(0.24)			
Instrument set		A	B	A		A	B
First-stage F-stat for migration prevalence		73.31	17.57	49.18		5.65	8.41
First-stage F-stat for migration prevalence*log ndc		72.93	17.54	48.62		5.67	8.43
First-stage F-stat for migration prevalence*log ndc squared				47.95			
Observations	11351	11351	11351	11351	5535	5535	5535

T-statistics in parentheses with standard errors clustered at the community level.
*Significant at 10%; ** significant at 5%; *** significant at 1%.
For male household heads aged 15–49 who have not previously migrated.
Instrument set A consists of 1955–59 state migration rate and the interaction of this rate with log NDC, Instrument set B uses the 1924 state migration rate and its interaction with log NDC.

Atkinson A(2) and Theil mean log deviation. For communities with migration prevalence rates above 0.10 we are in the same situation as the MMP sample, and so expect to see inequality decreasing with migration.

6. Inequality and migration

The previous Section shows that the data are supportive of our theoretical model in which household resources and community migration networks interact to influence household migration decisions. To examine the impact of community migration on inequality we would ideally use a long panel of communities, with data on both migration and inequality. However, since panel data on migration and inequality at the community level are not currently available in any survey we know of, we must use the variation across communities in migration experience to try and estimate the effect of changes in migration on inequality within a community. Using cross-sectional variation to estimate dynamic relationships is common in the cross-country growth literature, since data on many variables of interest are only available for recent years. The usual concern with such an approach is that the regressors of interest are endogenous, so that estimating the cross-sectional relationship does not uncover dynamic parameters. We attempt to overcome this concern by instrumenting for community migration prevalence. We model current inequality for community i, Ineq_i as a quadratic function of community migration prevalence, mig_i, and a vector of other community characteristics, X_i:

$$\text{Ineq}_i = \alpha_0 + \alpha_1 \text{mig}_i + \alpha_2 \text{mig}_i^2 + \gamma' X_i + \varepsilon_i. \tag{8}$$

Given the relatively small sample sizes in the MMP sample and lack of community information in the ENADID, we choose a parsimonious specification of the other community characteristics X_i, including the proportion of household heads aged under 30 in the survey year, the proportion of household heads aged over 60, the proportion of household heads with less than six years of education, and the proportion of household heads with nine or more years of education. Although Table 2 showed state-level characteristics were individually uncorrelated with historic migration rates, we also include current state conditions here as an additional check. We then use OLS and 2SLS to estimate Eq. (8). The analysis of the previous section suggests that we should find a negative effect of migration on inequality for the MMP sample and for the ENADID subsample with migration prevalence above 0.10, and perhaps evidence of migration increasing inequality among the ENADID communities with small networks.

Table 5 presents the first-stage regressions for the MMP and ENADID samples. Both the 1924 and 1955 historic migration rates are seen to be strongly significant in predicting community migration prevalence, especially in the larger ENADID sample. When these two measures are included together with the U.S. labor market conditions for the MMP sample, none of the instruments is individually significant, although the instruments are jointly significant. An overidentification test applied to this set of instruments has a p-value of 0.598, so that we can not reject exogeneity of the other instruments conditional on one instrument being exogenous. However, including the U.S. labor market conditions makes the instrument weaker. For this reason and the fact that the U.S. labor market conditions are not available for the ENADID, we will concentrate our discussions on the use of the historic migration rates as an instrument. The 1955–59 rate for the MMP sample has a first-stage F-statistic of 9.5, while the first-stage F-statistics for both historic rates are above 40 in the ENADID sample. It therefore appears that the historic migration rates are not subject to weak instrument concerns.

Table 6 presents the results from estimating Eq. (8) for the MMP sample using both the Gini of non-durable consumption and asset inequality to measure inequality. The quadratic term is insignificant in all specifications. The linear term in migration prevalence is always negative, and is significant after instrumenting. The estimated migration effect is similar across the different instrument sets, and show that an increase in migration lowers inequality. The magnitude of the coefficient suggests a reasonably strong effect: a one standard deviation increase in migration prevalence is estimated to result in a 1.1 standard deviation reduction in the Gini and a 1.8 standard deviation reduction in asset inequality.

Table 7 estimates Eq. (8) for the ENADID sample. Again the quadratic term in migration prevalence is insignificant in the 2SLS results. In the full sample the linear specification in migration prevalence finds negative coefficients. However, the coefficient is only significant when the 1924 rate is used as an instrument. The magnitude of the coefficient is also small, with the point estimates in columns 4 and 6 reflecting a 0.1 to 0.4 standard deviation increase in inequality from a one standard deviation increase in migration prevalence — smaller than predicted with the MMP sample. We then split the ENADID sample into municipalities with migration prevalence rates below and above 0.10. In the low network municipalities one finds a positive and insignificant effect of migration on inequality, while negative effects are found in the high network municipalities. When the 1924 migration rates are used as an instrument, the effect

Table 5
First-stage regression for migration prevalence

Dependent Variable: Migration prevalence	MMP sample			ENADID sample	
	(1)	(2)	(3)	(4)	(5)
State migration rate in 1955–59	2.3162		2.7819	6.8311	
	(3.08)***		(0.99)	(8.51)***	
State migration rate in 1924		0.5845	−0.1001		10.3779
		(2.61)**	(0.13)		(6.69)***
Historic unemployment in host area			0.0157		
			(0.46)		
Historic real depreciation			−0.0070		
			(0.20)		
Proportion of heads aged <30	−0.0277	0.0290	0.0383	−0.1123	−0.1203
	(0.08)	(0.09)	(0.11)	(0.62)	(0.60)
Proportion of heads aged >60	1.0259	1.0918	1.0584	0.4904	0.5798
	(4.98)***	(5.42)***	(4.10)***	(3.51)***	(3.78)***
Proportion of heads with education <6 years	−0.2132	−0.2184	−0.2103	0.0345	−0.0038
	(2.22)**	(2.08)**	(2.09)**	(0.38)	(0.04)
Proportion of heads with education at least 9 years	−0.5517	−0.5543	−0.5466	−0.2137	−0.2630
	(3.39)***	(3.12)***	(3.16)***	(2.18)**	(2.47)**
State GDP per capita	0.0038	−0.0089	0.0061	0.0023	0.0011
	(0.65)	(1.26)	(0.35)	(0.86)	(0.34)
State agricultural share of production	0.0087	0.0058	0.0090	0.0017	0.0026
	(3.39)***	(2.01)*	(2.09)**	(0.70)	(0.96)
State-level Corruption measure	0.0055	0.0002	0.0063	−0.0057	−0.0089
	(1.08)	(0.03)	(0.83)	(1.46)	(2.11)**
State-level hospitals per 100,000 population	−0.0000	−0.0000	−0.0000	−0.0511	−0.1113
	(3.29)***	(3.42)***	(2.78)***	(2.65)***	(4.30)***
Constant	0.0083	0.1923	−0.0686	0.1144	0.2453
	(0.07)	(1.40)	(0.25)	(1.36)	(2.47)**
F-statistics on instruments	9.49	6.84	2.42	49.33	40.82
Observations	57	57	57	214	214
R-squared	0.67	0.67	0.67	0.52	0.45

Robust *T*-statistics in parentheses.
*Significant at 10%; ** significant at 5%; *** significant at 1%.

is significant at the 10% level and shows a 0.3 standard deviation reduction in inequality from a one standard deviation increase in migration networks.[25] These results therefore are in accordance with the theoretical prediction of migration lowering inequality in high migration communities, and a possible positive effect on low migration communities.

Although long panels are not available, further evidence on the relationship between migration and inequality can be obtained by examining changes in migration and changes in inequality over time in a short panel of 33 municipalities formed from the 1992 and 1997 ENADID surveys. This matched panel has lower migration prevalences on average in 1992 than the full ENADID 1997 sample used elsewhere in this paper. The median community migration prevalence is 0.05, and the 25th and 75th percentiles are 0.015 and 0.160 respectively, while the maximum is 0.29. As

[25] We obtain similar results using the Atkinson A(2) and Theil mean log deviation measures of inequality which are more sensitive to inequality among the poor: positive and insignificant coefficients for the low migration prevalence sample and negative and marginally significant coefficients for the high migration subsample.

Table 6
The impact of migration on consumption inequality in MMP communities

	(1)	(2)	(3)	(4)	(5)	(6)
	OLS	OLS	IV	IV	IV	IV
Panel A: Dependent variable — Gini of non-durable consumption						
Migration prevalence	−0.084	−0.058	0.935	−0.195	−0.201	−0.214
	(0.72)	(1.21)	(0.48)	(2.04)**	(1.73)*	(2.48)**
Migration prevalence squared	0.046		−1.822			
	(0.25)		(0.55)			
Instrument set			A	A	B	C
Overidentification test *p*-value						0.598
F-stat for first stage migration prevalence			4.62	9.49	6.84	2.42
F-stat for first stage squared migration			4.71			
Observations	57	57	57	57	57	57
Panel B: Dependent variable — asset inequality						
Migration prevalence	−1.166	−0.714	7.434	−2.205	−2.574	−2.240
	(2.17)**	(3.39)***	(0.46)	(3.33)***	(2.96)***	(3.32)***
Migration prevalence squared	0.790		−15.551			
	(1.01)		(0.57)			
Instrument set			A	A	B	C
Overidentification test *p*-value						0.354
F-stat for first stage migration prevalence			4.62	9.49	6.84	2.42
F-stat for first stage squared migration			4.71			
Observations	57	57	57	57	57	57

Notes: *T*-statistics in parentheses with robust standard errors.
*Significant at 10%; ** significant at 5%; *** significant at 1%.
Instrument set A consists of 1955–59 state migration rate (and the squared rate in column 3).
Instrument set B uses the 1924 state migration rate.
Instrument set C uses 1924 and 1955–59 state migration rates, and weighted state-level past unemployment and past real depreciation rates.
All regressions include age and education distribution in the community, and state-level GDP per capita, agricultural share of production, corruption, and hospitals per 100,000 population.

such, this matched panel of municipalities spans migration prevalence levels at which we might expect to potentially see both positive and negative impacts of migration on inequality according to our previous results.

We consider two specifications which allow for a nonlinear effect of migration changes on inequality. The first is to run the following regression across municipalities k:

$$\Delta\mathrm{Ineq}_{\mathrm{k}} = \delta_0 + \delta_1 \Delta\mathrm{mig}_{\mathrm{k}} + \delta_2 \Delta\mathrm{mig}_{\mathrm{k}} {*} \mathrm{mig}_{\mathrm{k},1992} + \delta_3 \Delta\mathrm{mig}_{\mathrm{k},1992} + u_{\mathrm{k}} \tag{9}$$

where $\Delta\mathrm{Ineq}_{\mathrm{k}}$ denotes the change in inequality in municipality k between 1992 and 1997, $\Delta\mathrm{mig}_{\mathrm{k}}$ is the change in migration prevalence over this same period, and $\mathrm{mig}_{\mathrm{k}}$, 1992 is the 1992 level of migration prevalence. If an increase in migration always results in a reduction in inequality, then we would expect $\delta_1 < 0$. The interaction term allows for the effect of the change in migration to vary according to the initial level. If there is in fact an inverse U-shaped relationship between migration and inequality, then one would expect to find $\delta_1 > 0$ and $\delta_2 < 0$, that is, an increase in migration would increase inequality at low initial levels of migration stock, and would reduce inequality at higher levels. The constant term captures any aggregate change in within-municipality inequality occurring in Mexico between 1992 and 1997.

Table 7
The impact of migration on consumption inequality in ENADID communities

Dependent variable: Gini of non-durable consumption										
	Full sample						Low network sample		High network sample	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	OLS	OLS	IV	IV	IV	IV	IV	IV	IV	IV
Migration prevalence	−0.046 (2.45)**	−0.013 (2.08)**	−0.087 (2.05)**	−0.010 (0.79)	0.026 (0.32)	−0.035 (2.22)**	0.053 (1.09)	0.031 (0.41)	−0.002 (0.13)	−0.021 (1.69)*
Migration prevalence squared	0.056 (2.03)**		0.122 (1.42)		−0.097 (0.75)					
Instrument set			A	A	B	B	A	B	A	B
F-stat for first stage migration prevalence			47.73	49.33	28.20	40.82	13.20	10.43	14.05	14.48
F-stat for first stage squared migration			28.77		20.65					
Observations	214	214	214	214	214	214	90	90	124	124

Notes: *T*-statistics in parentheses with robust standard errors.
*Significant at 10%; ** significant at 5%; *** significant at 1%.
Instrument set A consists of 1955–59 state migration rate (and the squared rate in column 3).
Instrument set B uses the 1924 state migration rate (and its square in column 5).
Low network sample are municipalities with migration prevalence less than 0.10, high network municipalities have prevalence greater than 0.10.

The second specification directly takes differences of Eq. (8):

$$\Delta \mathrm{Ineq_k} = \beta_0 + \beta_1 \Delta \mathrm{mig_k} + \beta_2 \Delta(\mathrm{mig_k^2}) + \eta_\mathrm{k}. \tag{10}$$

The theory then predicts $\beta_1 > 0$ and $\beta_2 < 0$. Since these equations are expressed in terms of differences, municipality fixed characteristics which are correlated with both inequality and migration levels are differenced out.

This short panel has the advantage of allowing us to difference out time-invariant community characteristics which are related to both inequality and migration. However, there is still the possibility that a common shock induced both the change in migration and the change in inequality over this period. One such example could be a drought. It is harder to think of explanations why such shocks would cause nonlinear effects on inequality of the type predicted here. Nevertheless, we acknowledge that although endogeneity is likely to be less of a concern here than in the cross-section, the possibility remains, and so we view this panel data evidence as further suggestive evidence only.[26]

Table 8 presents the results from the ENADID panel data. Under Eq. (9), column 1 shows no significant relationship when the interaction term is not included, and column 2 shows a significant inverse U relationship once the interaction term is included. The Gini of non-durable consumption increases with an increase in migration up to an initial migration prevalence ratio of 0.17, after which it decreases. Using the specification in Eq. (10), column 3 finds coefficients of the signs predicted by an inverse U theory, and a turning point similar to

[26] One possibility would be to instrument changes in migration with changes in labor market conditions in the U.S. cities where a community typically sends its migrants. Unfortunately the ENADID does not collect information on the location within the U.S. of migrants, preventing us from pursuing this strategy here.

Table 8
Relation between changes in migration and changes in inequality

Dependent variable: Gini in 1997 ENADID — Gini in 1992 ENADID			
	(1)	(2)	(3)
Change in migration stock 1992–97	0.007	0.213	0.208
	(0.13)	(1.79)*	(1.53)
Change in migration stock * 1992 stock		−1.248	
		(1.98)*	
1992 Stock		−0.001	
		(0.04)	
Change in squared migration stock 1992–97			−0.525
			(1.63)
Constant	−0.002	−0.003	−0.002
	(0.87)	(0.92)	(1.01)
Observations	33	33	33
R-squared	0.00	0.12	0.08
Predicted turning point of migration stock		0.17	0.19

Absolute value of *T*-statistics in parentheses.
*Significant at 10%; ** significant at 5%; *** significant at 1%.

that in column 2. However, the coefficients are only significant at the 14% level. Given that 5 years is a relatively short time to observe changes in inequality, and that we only have panel data for 33 municipalities, we view this evidence coupled with our prior results as providing moderate support for an inverse U-shaped relationship between migration and inequality.

7. Conclusions

Migrants to the United States from Mexico are found to come from the middle of the wealth distribution when migration networks are not well developed, with the probability of migration displaying an inverse U-shaped relationship with wealth. As community migration networks grow, wealth becomes less of a constraint on individual migration, and the poor become more likely to migrate. At high levels of migration prevalence we find that this migration leads to a reduction in inequality. Large networks spread the benefits of migration to members at the lower end of the consumption and wealth distributions of the community, thereby reducing inequality. We find some evidence that migration benefits the upper–middle of the consumption distribution when networks are low and find suggestive evidence for a Kuznet relationship with migration increasing inequality at lower levels of migration stock, and then reducing inequality as one approaches the migration levels prevailing in the MMP communities. Panel data on inequality over longer time periods, and for more communities, are needed to confirm this evidence.

References

Adams, Richard H., 1989. Workers remittances and inequality in rural Egypt. Economic Development and Cultural Change 38 (1), 45–71.

Barham, Bradford, Boucher, Stephen, 1998. Migration, remittances and inequality: estimating the net effect of migration on income distribution. Journal of Development Economics 55, 307–331.

Bauer, Thomas, Epstein, Gil S., Gang, Ira N., 2002. Herd effects or migration networks? The Location Choice of Mexican Immigrants in the US. . IZA Discussion Paper, vol. 551. August.

Borjas, George J., 1987. Self-selection and the earnings of immigrants. American Economic Review 77 (4), 531–553.

Borjas, George, Katz, Lawrence, 2005. The Evolution of the Mexican-Born Workforce in the United States. NBER Working Paper, vol. 11281.
Carrington, W.J., Detragiache, E., Vishwanath, T., 1996. Migration with endogenous moving costs. American Economic Review 86 (4), 909–930.
Chiquiar, Daniel, Hanson, Gordon, 2005. International migration, self-selection, and the distribution of wages: evidence from Mexico and the United States. Journal of Political Economy 113 (2), 239–281.
Cuecuecha, Alfredo, 2003. Combining Micro Data Sets to Study Immigration from Mexico to the US", Paper presented at the NEUDC Conference, Yale University. Dirección General de Estadística (DGE) (1941) Anuario Estadístico de los Estados Unidos Mexicanos 1939. Secretaría de la Economía Nacional, DGE: Mexico City.
Docquier, Frédéric, Rapoport, Hillel, 2003. Remittances and Inequality: A Dynamic Migration Model. Stanford Center for International Development (SCID) Working Paper, vol. 167.
Dolfin, Sarah and Garance Genicot, 2006. What do networks do? The role of networks on migration and "coyote" use, Mimeo, March.
Espinosa, Kristin, Massey, Douglas, 1997. Undocumented migration and the quantity and quality of social capital". Soziale Welt 12, 141–162.
Foerster, Robert F., 1925. The Racial Problems Involved in Immigration from Latin America and the West Indies to the United States, United States Department of Labor.
Gathmann, Christina, 2004. The Effects of Enforcement on Illegal Markets: Evidence from Migrant Smuggling along the Southwestern Border. IZA Discussion Paper, vol. 1004. January.
González Navarro, Moisés, 1974. Población y Sociedad en México 1900–1970. Universidad Autónoma de México Press, Mexico City.
Hanson, Gordon, in press. Illegal migration from Mexico to the United States. Journal of Economic Literature.
Ibarraran, Pablo and Darren Lubotsky, 2006. Mexican Immigration and Self-Selection: New Evidence from the 2000 Mexican Census. Mimeo. University of Illinois at Urbana-Champaign.
Instituto Nacional de Estadística, Geografía e Informática (INEGI), 1994. ENADID Encuesta Nacional de la Dinámica Demográfica, 1992 Metodología y Tabulados, INEGI: Aguascalientes
Instituto Nacional de Estadística, Geografía e Informática (INEGI), 2001. Estadísticas Históricas de Mexico 2000 CD-ROM. INEGI: Aguascalientes.
Kanbur, R., Rapoport, H., 2005. Migration selectivity and the evolution of spatial inequality. Journal of Economic Geography 5 (1), 43–57.
McBride, George M., 1928. The Land Systems of Mexico. Octagon Books, New York.
McKenzie, David J., 2005. Measuring inequality with asset indicators. Journal of Population Economics 18 (2), 229–260.
Massey, Douglas S., Zenteno, Rene, 2000. A Validation of the ethnosurvey: the case of Mexico–U.S. migration. International Migration Review 34 (3), 766–793.
Massey, Douglas S., Goldring, Luin, Durand, Jorge, 1994. Continuities in transnational migration: an analysis of nineteen Mexican communities. American Journal of Sociology 99 (6), 1492–1533.
Massey, Douglas S., Durand, Jorge, Malone, Nolan J., 2002. Beyond Smoke and Mirrors: Mexican Immigration in an Era of Economic Integration. Russell Sage Foundation, New York.
Milanovic, Branko, 1987. Remittances and income distribution. Journal of Economic Studies 14 (5), 24–37.
Mishra, P., 2007. Emigration and wages in source countries: evidence from Mexico. Journal of Development Economics 82 (1), 180–199.
Munshi, Kaivan, 2003. Networks in the modern economy: Mexican migrants in the U.S. labor market. Quarterly Journal of Economics 118 (2), 549–599.
Newey, Whitney K., 1987. Efficient estimation of limited dependent variable models with endogenous explanatory variables. Journal of Econometrics 36 (3), 231–250.
Orrenius, Pia M., 1999. The Role of Family Networks, Coyote Prices, and the Rural Economy in Migration from Western Mexico: 1965–1994. Federal Reserve Bank of Dallas Working Paper, vol. 9910.
Orrenius, Pia M., Zavodny, Madeline, 2005. Self-selection among undocumented immigrants from Mexico. Journal of Development Economics 78, 215–240.
Rapoport, Hillel, Docquier, Frédéric, 2006. The economics of migrants' remittances. In: Kolm, S.C., Mercier Ythier, J. (Eds.), Handbook of the Economics of Reciprocity. . Giving and Altruism, vol. II. Amsterdam, North-Holland, pp. 1135–1198.
Smith, James P., Thomas, Duncan, 2003. Remembrances of things past: test–retest reliability of retrospective migration histories. Journal of the Royal Statistical Society 166 (1), 23–49.
Stark, Oded, Taylor, J. Edward, Yitzhaki, Shlomo, 1986. Remittances and inequality. Economic Journal 96 (383), 722–740.

Taylor, J. Edward, 1992. Remittances and inequality reconsidered: direct, indirect, and intertemporal effects. Journal of Policy Modeling 14 (2), 187–208.

Edward, Taylor J., Wyatt, T.J., 1996. The shadow value of migrant remittances, income and inequality in a household-farm economy. Journal of Development Studies 32 (6), 899–912.

Verbeek, Marno, Nijman, Theo E., 1992. Can cohort data be treated as genuine panel data? Empirical Economics 17, 9–23.

Woodruff, C., Zenteno, R., 2007. Remittances and microenterprises in Mexico. Journal of Development Economics 82 (2), 509–528.