Neighborhood Effects in Migration*

Rubén Irvin Rojas Valdés, C.-Y. Cynthia Lin Lawell, and J. Edward Taylor

Abstract

Given the economic significance of migration and its relevance for policy, it is important to understand the factors that cause people to migrate. We add to the literature on the determinants of migration by examining whether neighborhood effects matter for migration decisions. We define 'neighborhood effects' as arising whenever the migration decisions of other households in the village affect a household's decision to have a member migrate. Using instrumental variables, a rich set of controls, and village fixed effects to address the endogeneity of neighbors' decisions as well as exclusion bias, we empirically examine whether there are neighborhood effects in household migration decisions in rural Mexico. We conduct several tests and analyses to make the case for the validity of our instruments and to address possible concerns about our instruments. We develop and conduct a falsification test to show that the instruments are not correlated with unobserved village factors that may affect migration decisions, thereby providing supportive evidence that the exclusion restriction is satisfied. Results show that, for households in rural Mexico, neighborhood effects have a significant positive effect on international migration to the US, but do not have a significant effect on decisions to migrate domestically within Mexico. As a consequence, policies that affect the international migration decisions of one household in a village may have a multiplier effect and induce similar decisions by neighboring households.

JEL Codes: O15, O54, C26
Keywords: migration, Mexico, neighborhood effects, strategic interactions, peer effects, falsification test, exclusion restriction
This draft: August 2020

*Rojas Valdés: Centro de Investigación y Docencia Económicas (CIDE); irvin.rojas@cide.edu. Lin Lawell: Cornell University; clinlawell@cornell.edu. Taylor: University of California at Davis; jetaylor@ucdavis.edu.
We thank Steve Boucher, Colin Carter, Marcel Fafchamps, Emilio Gutiérrez, Tom Hertel, Seema Jayachandran, Travis Lybbert, Shaun McRae, Erich Muehlegger, Barbara Petrongolo, John Rust, Yaniv Stopnitzky, and Bruce Wydick for invaluable comments and discussions. We benefited from comments from seminar participants at the University of San Francisco; and conference participants at the Agricultural and Applied Economics Association (AAEA) Annual Meeting; at the Oxford Symposium on Population, Migration, and the Environment; at the Pacific Conference for Development Economics; at the Midwest International Economic Development Conference; and at the Gianinni Agricultural and Resource Economics Student Conference. We thank Gerardo Aragon, Diane Charlton, Katrina Jessoe, Rebecca Lessem, and Dale Manning for their help with the data. We are also indebted to Antonio Yunez-Naude and the staff of PRECESAM and of Desarrollo y Agricultura Sustentable for their invaluable assistance and data support. We received financial support from the University of California Institute for Mexico and the United States (UC MEXUS). Lin Lawell is a former member and Taylor is a member of the Giannini Foundation of Agricultural Economics. All errors are our own.

1 Introduction

According to estimates from the World Bank (2010a), around 3 percent of the world population live in a country different from the one in which they were born. The US is the country with the highest immigrant population in the world, with more than 46 million people who were foreign born (United Nations, 2013), of which about 11 million are from Mexico (World Bank, 2010b).

Since the mid-1980s, migration to the US has represented an employment opportunity for Mexicans during a period of economic instability and increasing inequality in Mexico. In addition, it has represented an important source of income via remittances, especially for rural households (Esquivel and Huerta-Pineda, 2007).[1] Remittances from the US to Mexico amount to an estimated 22.8 billion US dollars per year (World Bank, 2012). Each of the nearly 11 million Mexicans living in the US sends an average of 2,115 US dollars in remittances, accounting for up to 2 percent of the Mexican GDP (D'Vera et al., 2013), 13 percent of household total income, and 16 percent of per capita income in Mexico (Taylor et al., 2008).[2] With a border 3,200 kilometers long, the largest migration flow between two countries, and a wage differential for low-skilled workers between the US and Mexico of 5 to 1 (Cornelious and Salehya, 2007), the US-Mexico migration relationship also imposes challenges to policy-makers in both countries (Rojas Valdés, Lin Lawell and Taylor, 2020b).

Given the economic significance of migration and its relevance for policy, it is important to understand the factors that cause people to migrate. We add to the literature on the determinants of migration by examining whether neighborhood effects matter for migration decisions. We define 'neighborhood effects' as arising whenever the migration decisions of other households in the village affect a household's decision to have a member migrate.[3]

Using instrumental variables, a rich set of controls, and village fixed effects to address the endogeneity of neighbors' decisions as well as exclusion bias, we empirically examine whether there are neighborhood effects in household migration decisions in rural Mexico, and whether there are nonlinearities in the neighborhood effects. The instruments we use for the migration decisions of a household's neighbors are based on characteristics of the household's neighbors, and we include as controls the same characteristics of the household itself in order to address exclusion bias (Caeyers and Fafchamps, 2019). Our IV regressions

[1]Esquivel and Huerta-Pineda (2007) find that 3 percent of urban households and up to 10 percent of rural households in Mexico receive remittances.

[2]Castelhano et al. (2020) find that migrant remittances are not associated with increases in rural investment in agricultural production in Mexico, however.

[3]We choose to use the term 'neighborhood effects' instead of 'peer effects' because the term 'peer' often connotes an individual; in contrast; the decision-makers we examine are households rather than individuals. Nevertheless, our concept of 'neighborhood effects' is very similar to that of 'peer effects'.

also include controls at the household, village, municipality, state, and national level; border crossing variables; as well as village fixed effects that control for unobserved conditions that affect all households in the village in the same fashion. As a consequence, our instruments, whose values vary across households within the same village and across years, exploit the variation in the characteristics of neighbors that affect their probability of migration. We conduct several tests and analyses to make the case for the validity of our instruments and to address possible concerns about our instruments. We develop and conduct a falsification test to show that the instruments are not correlated with unobserved village factors that may affect migration decisions, thereby providing supportive evidence that the exclusion restriction is satisfied.

An empirical assessment of neighborhood effects in migration has important implications for policy. If neighborhood effects are positive – for example through network effects or information spillovers – then policies that affect the migration decisions of one household in a village may have a multiplier effect and induce similar decisions by neighboring households. As a consequence, migration policies (to deter or induce migration) may be more cost-effective if they are targeted at the village level, and focused on directly affecting the migration decisions of a subset of households in each village. In contrast, if neighborhood effects are negative – for example when there are higher returns to migrating when there are fewer migrants – then deterring one village household from engaging in migration may induce other neighboring households to engage in migration.

Results of our IV regressions show that, for households in rural Mexico, neighborhood effects have a significant positive effect on international migration to the US, but do not have a significant effect on decisions to migrate domestically within Mexico. In our base case specification, an increase of one standard deviation in the fraction of neighbors with migration to the US increases a household's probability of engaging in migration to the US by around 12 percentage points. As a consequence, policies that affect the international migration decisions of one household in a village may have a multiplier effect and induce similar decisions by neighboring households.

The balance of the paper proceeds as follows. We review the literature in Section 2. In Section 3 we discuss sources of neighborhood effects in migration. Section 4 describes the data. Section 5 presents the empirical strategy that allows us to identify the neighborhood effects. Section 6 presents the empirical results. Section 7 concludes.

2 Literature Review

2.1 Determinants of migration

The first strand of literature upon which our paper builds is the literature on determinants of migration. The new economics of labor migration posits the household as the relevant unit of analysis. Using the household as the relevant unit of analysis addresses several observed features of migration that are ignored by individualistic models, including the enormous flows of remittances and the existence of extended families which extend beyond national borders. Most applications of the new economics of labor migration assume that the preferences of the household can be represented by an aggregate utility function, and that income is pooled and specified by the household budget constraint.

For example, Stark and Bloom (1985) assume that individuals with different preferences and income not only seek to maximize their utility but also act collectively to minimize risks and loosen constraints imposed by imperfections in credit, insurance, and labor markets. This kind of model assumes that there is an informal contract among members of a family in which members work as financial intermediaries in the form of migrants. The household acts collectively to pay the cost of migration by some of its members, and in turn migrants provide credit and liquidity (in form of remittances), and insurance (when the income of migrants is not correlated with the income generating activities of the household). In this setting, altruism is not a precondition for remittances and cooperation, but it reinforces the implicit contract among household members (Taylor and Martin, 2001). In their analysis of how migration decisions of Mexican households respond to unemployment shocks in the US, Fajardo, Gutiérrez and Larreguy (2017) emphasize the role played by the household, as opposed to individuals, as the decision-making unit at the origin. Garlick, Leibbrandt and Levinsohn (2016) provide a framework with which to analyze the economic impact of migration when individuals migrate and households pool income. Kinnan, Wang and Wang (2018) discuss and analyze several channels through which migration might affect rural households and their remaining household members.

Characteristics of individual migrants are important determinants of migration decisions, and affect the impacts that migration has on the productive activities of the remaining household. Human capital theory à la Sjaastad (1962) suggests that migrants are younger than those who stay because younger migrants would capture the returns from migration over a longer time horizon. Changes in labor demand in the United States has modified the role of migrant characteristics in determining who migrates, however. Migrants from rural Mexico, once mainly poorly educated men, more recently have included female, married, and better educated individuals relative to the average rural Mexican population (Taylor and Martin, 2001). Borjas (2008) finds evidence that Puerto Rico emigrants to the United States have lower incomes, which is consistent with the prediction of Borjas (1987) that migrants have (expected) earnings below the mean in both the host and the source economy when the source economy has low mean wages and high inequality. On the other hand, Feliciano (2001), Chiquiar and Hanson (2005), Orrenius and Zavodny (2005), McKenzie and Rapoport (2010), Cuecuecha (2005), and Rubalcaba et al. (2008) find that Mexican migrants come from the middle of the wage or education distribution. McKenzie and Rapoport (2007) show that migrants from regions with communities of moderate size in the United States come from the middle of the wealth distribution, while migrants from regions with bigger communities in the United States come from the bottom of the wealth distribution.

The financial costs of migration can be considerable relative to the income of the poorest households in Mexico.[4] Migration costs reflect in part the efforts of the host country to impede migration (Hanson, 2010). Migration costs for illegal crossing from Mexico to the United States are estimated to be 2,750 to 3,000 dollars (Mexican Migration Program, 2014). Border enforcement grew by a factor of 13 between 1986 and 2002 (Massey, 2007; Lessem 2018). Estimates reported in Hanson (2010) suggest that the cost of the "coyote" increased by 37 percent between 1996-1998 and 2002-2004, mainly due to the increase of border enforcement due to the terrorist attacks of 9/11. On the other hand, Gathmann (2008) estimates that even when the border enforcement expenditure for the Mexico-United States border almost quadrupled between 1986 and 2004, the increase in expenditure produced an increase the cost of the coyote of only 17 percent, with almost zero effect on coyote demand. Nevertheless, Lessem (2018) finds that increases in border enforcement decrease migration from Mexico to the United States. Feigenberg (2020) finds that US-Mexico border fence construction induces migrants to substitute toward alternative crossing locations, disproportionately deters low-skilled migrants, and reduces the number of undocumented Mexicans in the United States.

2.2 Strategic interactions among neighbors

In addition to the literature on migration, our paper also builds on the previous literature using reduced-form empirical models to analyze strategic interactions. Strategic interactions arise whenever the actions of others affect the payoffs of a decision-maker and therefore the decision-maker's actions. The neighborhood effects we examine in this paper are strategic interactions in which the decision-makers are households in the same village.

[4]Data from the National Council for the Evaluation of the Social Policy in Mexico (CONEVAL) show that the average income of the poorest 20 percent of rural Mexican households was only 456 dollars a year in 2012.

We build on the work of Lin (2009), who analyzes the strategic interactions among firms in offshore petroleum production. In particular, she analyzes whether a firm's exploration decisions on federal lands in the Gulf of Mexico depend on the decisions of firms owning neighboring tracts of land. To address the endogenity of neighbors' exploration decisions, she uses variables based on the timing of a neighbor's lease term as instruments for the neighbor's decision.

Pfeiffer and Lin (2012) use a unique spatial data set of groundwater users in western Kansas to empirically measure the physical and behavioral effects of groundwater pumping by neighbors on a farmer's groundwater extraction. To address the simultaneity of neighbors's pumping, they use the neighbors' permitted water allocation as an instrument for their pumping. Bollinger, Burkhardt and Gillingham (2020) identify causal peer effects in residential water conservation during the summer using variation from movers.

Robalino and Pfaff (2012) estimate neighbor interactions in deforestation in Costa Rica. To address simultaneity and the presence of spatially correlated unobservables, they instrument for neighbors' deforestation using the slopes of neighbors' and neighbors' neighbors' parcels. Similarly, Irwin and Bockstael (2002) investigate strategic interactions among neighbors in land use change using physical attributes of neighboring parcels as instruments to identify the effect of neighbors' behavior.

Morrison and Lin Lawell (2016) investigate the role of social influence in the commute to work. Using instruments to address the endogeneity of commute decisions and a dataset of US military commuters on 100 military bases over the period 2006 to 2013, they show that workplace peers positively influence one another's decisions to drive alone to work and carpool to work.

Durlauf (2004) surveys the economics literature on the role of neighborhoods in influencing socioeconomic outcomes. Topa and Zenou (2015) provide an overview of the research on neighborhoods and social networks and their role in shaping behavior and economic outcomes for crime, education, and the labor market. Graham (2018) reviews econometric methods for the identification and estimation of neighborhood effects, particularly the effects of one's neighborhood of residence on long-run life outcomes.

Munshi (2003) identifies job networks among Mexican immigrants in the United States labor market. The network of each origin community in Mexico is measured by the proportion of the sampled individuals who are located at the destination (the US) in any year. He uses rainfall in the origin community as an instrument for the size of the network at the destination.

3 Sources of Neighborhood Effects

In this paper we build upon the literature on the determinants of migration by empirically examining whether neighborhood effects are a determinant of migration decisions as well. We define 'neighborhood effects' as arising whenever the migration decisions of other households in the village affect a household's decision to have a member migrate. There are several reasons why households may take into the account the actions of other households in their village when making their migration decisions (Rojas Valdés, Lin Lawell and Taylor, 2020b).

The first source of neighborhood effects are migration networks. Migration networks may affect migration decisions because they may reduce the financial, psychological, and/or informational costs of moving out of the community. Contacts in the source economy lower financial or information costs and reduce the utility loss from living and working away from home. The role of migration networks has been studied by Du, Park and Wang (2005) on China; Munshi and Rosenzweig (2016) on India; Bauer and Gang (1998) and Wahba and Zenou (2005) on Egypt; Battisti, Peri and Romiti (2016) on Germany; Neubecker, Smolka and Steinbacher (2017) on Spain; Mahajan and Yang (2020) on migration from countries affected by hurricanes; Alcosta (2011) on El Salvador; Alcala et al. (2014) on Bolivia; Calero, Bedi and Sparro (2009) on Ecuador; Acosta et al. (2008) on Latin America; and several others on Mexico, including Massey and Espinosa (1997) and Massey, Goldring and Durand (1994). These papers find a positive effect of migration networks on the probability of migration. Elsner, Narciso and Thijssen (2018) find evidence that networks that are more integrated in the society of the host country lead to better outcomes after migration.

In his analysis of job networks among Mexican immigrants in the US labor market, Munshi (2003) finds that the same individual is more likely to be employed and to hold a higher paying nonagricultural job when his network is exogenously higher. Orrenius and Zavodny (2005) show that the probability of migrating for young males in Mexico increases when their father or siblings have already migrated. McKenzie and Rapoport (2010) find that the average schooling of migrants from Mexican communities with a larger presence in the United States is lower. Networks and the presence of relatives or friends in the host country are consistently found to be significant in studies such as those of Greenwood (1971) and Nelson (1976), among others.

A second source of neighborhood effects are information externalities between households in the same village that may have a positive effect on migration decisions. When a household decides to send a migrant outside the village, other households in the village may benefit from learning information from their neighbor. This information may include information about the benefits and costs of migration, as well as information that enables a household to increase the benefits and reduce the costs of migration (Massey, Goldring and Durand, 1994). Previous studies have found that information externalities and social learning from the experimentation of others play an important role in developing countries for decisions regarding technology adoption (Foster and Rosenzweig, 1995; Munshi, 2004; Bandiera and Rasul, 2006; Conley and Udry, 2010; Maertens, 2017).

A third source of neighborhood effects may be relative deprivation (Taylor, 1987; Stark and Taylor, 1989; Stark and Taylor, 1991). Models of relative deprivation consider that a household's utility is a function of its relative position in the wealth distribution of all the households in the community. The relative deprivation motive may explain why local migration is different from international migration because when a migrant moves within the same country it is more likely that she changes her relative group since it is easier to adapt in the host economy, where the same language might be spoken and the cultural differences might not be as dramatic as in the case of international migration. In contrast, with international migration, the migrants likely still consider their source country as their reference group, especially when the source and the host countries are very different. Thus, the relative deprivation concept would predict that those individuals from a household that is relatively deprived might decide to engage in international migration rather than domestic migration even though the former is more costly because by migrating locally her position in the new reference group would be even worse than the position she would have if she did not migrate (Rojas Valdés, Lin Lawell and Taylor, 2020a).

A fourth source of neighborhood effects is risk sharing. Chen, Szolnoki and Perc (2012) argue that migration can occur in a setting when individuals share collective risk. Cheng et al. (2011) show that migration might promote cooperation in the prisoner's dilemma game. Lin et al. (2011) show that when migration occurs with a certain probability if the aspired level of payoffs is greater than the own payoff, aspirations promote cooperation in the prisoner's dilemma game. Morten (2019) develops a dynamic model to understand the joint determination of migration and endogenous temporary migration in rural India, and finds that improving access to risk sharing reduces migration.

A fifth source of neighborhood effects is a negative competition effect whereby the benefits of migrating to the US or within Mexico would be reduced if others from the same village also migrate to the US or within Mexico. This negative competition effect may be compounded if there is a limited number of employers at the destination site who do not discriminate against migrants from elsewhere (Carrington, Detragiache and Vishwanath, 1996).

A sixth source of neighborhood effects is the marriage market. Marriage and migration decisions of one's own household and those of one's neighbors are often intertwined. The possibility of marriage with someone from another household in the same village may affect migration decisions, and vice versa. For example, a household may care about what its neighbors do in terms of migration since it may affect the marriage prospects of members of one's household. Riosmena (2009) finds that single people in Mexico are most likely to migrate to the US relative to those who are married in areas of recent industrialization, where the Mexican patriarchal system is weaker and where economic opportunities for both men and women make post-marital migration less attractive.

A seventh source of neighborhood effects are cultural norms that make migration a typical activity for some households and communities. Kandel and Massey (2002) find that, in some communities, migration can be seen as a normal stage of life, reflecting a transition to 'manhood' and a means of social mobility.

Owing to migration networks, information externalities, relative deprivation, risk sharing, competition effects, the marriage market, and/or cultural norms, households may take into account the migration decisions of neighboring households when making their migration decisions. Our empirical model is general enough to capture multiple possible sources of neighborhood effects, and enables us to analyze their net effect.

4 Data

We use data from the National Survey of Rural Households in Mexico (ENHRUM) in its three rounds (2002, 2007, and 2010[5]). The survey is a nationally representative sample of Mexican rural households across 80 villages and includes information on the household characteristics such as productive assets and production decisions. It also includes retrospective employment information: individuals report their job history back to 1980. With this information, we construct an annual household-level panel data set that runs from 1990 to 2010[6] and that includes household composition variables such as household size, household head age, and number of males in the household. For each individual, we have information on whether they are working in the same village, in some other state within Mexico (domestic migration), or in the United States (international migration).

In our sample, 446 households (28%) had at least one member migrate within Mexico during at least one year of our data set, but never had any member migrating to the US during the time period of our data set; 321 households (21%) had at least one member migrate to the US during at least one year of our data set, but never had any member

[5]The sample of 2010 is smaller than the sample of the two previous rounds because it was impossible to access some villages during that round due to violence and budget constraints.

[6]Since retrospective data from 1980 to 1989 included only some randomly selected individuals in each village who reported their work history, we begin our panel data set in 1990.

migrating within Mexico during the time period of our data set; and 316 households (20%) had both at least one member migrating within Mexico during at least one year and at least one member migrating to the US during at least one (possibly different) year of our data set. The remaining 533 households (33%) never engaged in either type of migration during the time period of our data set.

The survey also includes information about the plots of land owned by each household, including irrigation status and land area.[7] We reconstruct the information for the complete panel using the date at which each plot was acquired. We interact the area of land owned by the household that is irrigated for agricultural purposes with a measure of contemporaneous precipitation at the village level (Jessoe, Manning and Taylor, 2018). Precipitation data is available only for the subperiod of 1990 to 2007. Because information on the irrigation status and land area of a household's plots of land is only available for the plots owned by the household, households that do not own any plots of land have missing observations for irrigated land area and irrigated land area interacted with precipitation. We therefore also try estimating the models without using the plot-related irrigated land area variables.

We include variables that measure the level of development, schooling opportunities, and job prospects in the villages where households are located and that might affect their migration decisions. In particular, we control for the supply of schools using the number of basic schools per 10,000 inhabitants in the municipality. We also include the number of indigenous schools per 10,000 inhabitants to control for the prevalence of indigenous populations for which poverty and access to services is considerably lower than the average rural population. The level of urbanization is controlled by using the number of cars and the number of buses per 10,000 inhabitants reported by the National Statistics Institute (INEGI). These data cover the period 1990 to 2010.

We also include aggregate variables that represent the broad state of the institutional and economic environment relevant for migration. We use data from the INEGI on the fraction of the labor force employed in each of the three productive sectors (primary, secondary, and tertiary[8]) at the state level, from 1995 to 2010. We use INEGI's National Survey of Employment and the methodology used in Campos-Vazquez, Hincapie and Rojas-Valdés (2012) to calculate the hourly wage at the national level from 1990 to 2010 in each of the three productive sectors and the average wage across all three sectors.

[7]We use information on plots of land which are owned by the household because our data set does not include comparable information on plots of land that are rented or borrowed.

[8]The primary sector includes agriculture, livestock, forestry, hunting, and fisheries. The secondary includes the extraction industry and electricity, manufacturing, and construction. The tertiary sector includes commerce, restaurants and hotels, transportation, communication and storage, professional services, financial services, corporate services, social services, and government and international organizations.

We use two sets of border crossing variables that measure the costs of migration. On the Mexican side, we use INEGI's data on crime to compute the homicide rate per 10,000 inhabitants at each of the 37 the Mexican border municipalities. Criminal activity at the border might represent an additional cost for migration, which would suggest a negative sign for crime at the border. At the same time, if this criminal activity reflects an increase in the supply of all illegal business, including smuggling of people, a positive sign could also be expected on our variable of crime at the border. On the United States' side, we use data from the Border Patrol that include the number of border patrol agents, apprehensions, and deaths of migrants at each of nine border sectors,[9] and match each border sector to its corresponding Mexican municipality. We interact these border crossing variables (which are time-variant, but the same for all villages at a given point in time) with measures of distance from the villages to the border (which are time-invariant for each village, but vary for each village-border location pair).

To determine the distance from each village to each border municipality, we use a map from the International Boundary and Water Commission (2013) to obtain the location of the 26 crossing-points from Mexico to the United States. Using the Google Distance Matrix API, we obtain the shortest driving route from each of the 80 villages in the sample to each of the 26 crossing-points, and match the corresponding municipality at which these crossing-points are located. This procedure allows us to categorize the border municipalities into those less than 1,000 kilometers from the village; and those between 1,000 and 2,000 kilometers from the village.

By interacting the distances to the border crossing points with the border crossing variables, we obtain the mean of each border crossing variable at each of the three closest crossing points, and the mean of each border crossing variable within the municipalities that are in each of the two distance categories defined above. We also compute the mean of each border crossing variable among all the border municipalities.

Figure 1 presents a map of the villages in our sample (denoted with a filled black circle) and the US-Mexico border crossing points (denoted with a red X).

Table 1 presents the summary statistics for the variables in our data set. Table 2 presents the within and between variation for the migration variables. 'Within' variation is the variation in the migration variable across years for a given village. 'Between' variation is the variation in the migration variable across villages for a given year.

[9] A "border sector"' is the term the Border Patrol uses to delineate regions along the border for their administrative purposes.

5 Empirical Strategy

5.1 Econometric model

We estimate reduced-form models of a household's decision to have a member migrate to the US, and also of its decision to have a member migrate within Mexico. Our dependent variable y_{ikt} is an indicator variable of whether household i engages in migration of type $k \in [USA, Mexico]$ at time t, which is equal to 1 for household i in year t if a household has a member who is a migrant to/within k in year t.

To analyze neighborhood effects in migration, we analyze whether the fraction of neighbors who engage in migration to the US and the fraction of neighbors who engage in migration within Mexico affect a household's decision to have a member migrate to the US and/or a household's decision to have a member migrate within Mexico. In particular, we regress household i's decision y_{ikt} to engage in migration on the fraction s_{ikt} of the sampled households in the same village as household i, excluding i, that engage in migration of type k.

We estimate the following econometric model:

$$y_{ikt} = \alpha + \sum_k \beta_{sk} s_{ikt} + x'_{it}\beta_x + \mu_i + \tau_t + \varepsilon_{ikt}, \qquad (1)$$

where the vector x_{it} includes covariates at the household, village, municipality, state, and national level as well as border crossing variables; μ_i is a village fixed effect; and τ_t is a year effect. We also run a set of specifications using a time trend instead of year effects.

The regressors at the household level in x_{it} include the number of males in the household, the age of the household head; the schooling of the household head; the maximum level of schooling achieved by any of the household members; the average level of schooling, measured as the number of years of education that have been completed, of household members 15 years old and above; a dummy if the household's first born was a male; the area of land owned by the household that is irrigated for agricultural purposes, interacted with village precipitation; the lagged fraction of household members working in the US; and the lagged fraction of household members working within Mexico. The area of land owned by the household that is irrigated for agricultural purposes interacted with contemporaneous precipitation captures shocks to agricultural home production and therefore to household income that vary by household and year and that may affect migration decisions.

The regressors at the municipality level in x_{it} include the number of schools in the basic system, the number of schools in the indigenous system, the number of cars, and the number of buses. Basic schools in Mexico include preschools, elementary schools, and junior high schools. Indigenous schools are a particular type of basic schools: in particular, an indigenous schools is a public basic school that is specifically targeted to people in communities where there is a large share of people who speak indigenous languages and the curriculum is adapted to these special needs.[10] The number of indigenous schools is a proxy for the size of indigenous population, overrepresented in poverty status and with economic disadvantages in terms of health and access to services.

The state-level variables in x_{it} include employment by sector. The national variables in x_{it} are aggregate variables that represent the broad state of the institutional and economic environment relevant for migration, including the average hourly wage, and wage by sector.

The border crossing variables in x_{it} include variables that measure crime, deaths, and border enforcement at nearby border crossing points. While these border crossing variables are likely to affect migration decisions, the effect of these border crossing variables on migration can be either positive or negative, and is therefore an empirical question. On the one hand, criminal activity and enforcement at the border might represent an additional cost for migration, which would suggest a negative effect of the border crossing variables on migration. On the other hand, if criminal activity and enforcement at the border reflects an increase in the supply of all illegal business, including the smuggling of people, one may expect a positive correlation between our border crossing variables and migration.

In addition to controls at the household, village, municipality, state, and national level as well as border crossing variables, we use village fixed effects to control for unobserved conditions that affect all households in the village in the same fashion. For example, the village fixed effects absorb village characteristics such as the overall educational level of the village, infrastructure, and teaching quality.

5.2 Instruments

In order to identify neighborhood effects, it is important to address endogeneity problems and other potential sources of bias.

Measuring neighborhood effects is difficult owing to two sources of endogeneity. One source is the simultaneity of the neighborhood effect: if household i is affected by its neighbor j, then household j is affected by its neighbor i. The other arises from spatially correlated unobservable variables (Manski, 1993; Manski, 1995; Brock and Durlauf, 2001; Conley and Topa, 2002; Glaeser, Sacerdote and Scheinkman, 1996; Moffitt, 2001; Lin, 2009; Robalino

[10] Indigenous schools teach a curriculum in both Spanish and the predominant indigenous language in the community. The main features of indigenous schools are the bilingual teaching, the use of contextual content in the curriculum (for example, incorporating group-specific traditional literature), and the linkage between traditional and scientific knowledge.

and Pfaff, 2012; Pfeiffer and Lin, 2012).

An additional source of bias is an exclusion bias: because household i cannot be its own neighbor, there is a mechanical negative relationship between household i's characteristics and those of its neighbors, irrespective of whether neighbors self-select each other or are randomly assigned, and this negative correlation produces a negative bias in ordinary least squares (OLS) estimates of neighborhood effects (Guryan, Kroft and Notowidigdo, 2009; Angrist, 2014; Caeyers and Fafchamps, 2019).

To address the endogeneity of neighbors' migration decisions s_{ikt} as well as exclusion bias, we use instruments for the fraction of neighbors that engage in migration that are correlated the neighbors' migration decisions but do not affect a household's own-migration decision except through their effect on the neighbors' migration decisions. The instruments we use for the migration decisions of a household's neighbors are based on the characteristics of the household's neighbors, and exploit the variation in characteristics of neighbors that affect the neighbors' probability of migration. For each of our instrumental variables, there is variation in that instrumental variable both across households within the same village and also across years.

One instrument we use for the fraction of neighbors that engage in migration is the neighbors' average household head schooling, which we define as the average of the household head schooling of other households in the village of i, but not including i, at time t. This is a good instrument because the average household head schooling of other households j may affect the migration decisions of other households j, but does not affect the decisions of household i except through its effect on the decisions of neighboring households j. For instance, all else equal, the household head schooling of a neighboring household j may affect the likelihood that neighboring household j engages in migration because the household head schooling of neighboring household j affects the decisions of neighboring household j regarding migration, schooling, and employment for its households members; how these decisions are made; absolute and relative productivity levels of its members; and/or the feasibility and desirability of having a member migrate.

A second instrument we use for the fraction of neighbors that engage in migration is the neighbors' average household average schooling, which we define as the average of the mean household schooling of other households in the village of i, but not including i, at time t. The mean schooling in each household is constructed using the years of schooling of individuals over 15 years old. This is a good instrument because the average mean household schooling of neighbor households j may affect the migration decisions of other households j, but does not affect the decisions of household i except through its effect on the decisions of neighboring households j. For instance, all else equal, the mean schooling of a neighbor household j

may affect the likelihood that neighboring household j engages in migration because the mean schooling of neighboring household j affects the decisions of neighboring household j regarding migration, schooling, and employment for its households members; how these decisions are made; absolute and relative productivity levels of its members; and/or the feasibility and desirability of having a member migrate.

We use both the neighbors' average household head schooling and the neighbors' average household average schooling as instruments in all our specifications. One may worry that the exclusion restriction may not hold for these two instruments if one household's decision to migrate depends on its relative education compared to its neighbors. For instance, one example would be if those that are the most educated within the village will not migrate as they can reap the benefits locally, whereas those that are the least educated may not migrate because of income or resource constraints that limit their ability to do so. In such a model, both one's household schooling level and the schooling level of one's neighbors matter for migration.

We address this concern that the exclusion restriction may not hold if one household's decision to migrate depends on its relative education compared to its neighbors, in several ways. First, we control for a household's own schooling. Second, we control for village fixed effects, which absorb village characteristics such as the overall educational level of the village. Because we include village fixed effects, the within-village variation in a household's own schooling captures relative differences in household schooling within a village. As a consequence, we control for the possibility that one household's decision to migrate may depend on its relative education compared to its neighbors.

A third way in which we address possible non-monotonic effects of schooling on migration is that we control for and distinguish between two measures of schooling: household head schooling and the household average schooling. Once again, because we include village fixed effects, the within-village variation in household head schooling and the household average schooling captures relative differences in household head schooling and household average schooling, respectively, within a village. As a consequence, we control for the possibility that one household's decision to migrate may depend on relative education of its household head compared to that of its neighbors, and also for the possibility that one household's decision to migrate may depend on its relative household average schooling compared to that of its neighbors; and we allow these two effects to potentially differ in sign. Fourth, as explained and presented in more detail below, we conduct several tests and analyses to make the case for the validity of our instruments, including a falsification test to provide evidence that the exclusion restriction is satisfied.

For robustness, in some specifications we use an additional third instrument for the fraction of neighbors that engage in migration. This additional third instrument is the fraction of neighbors whose first born is male, which we define as the fraction of other households in the village of i, not including i, at time t in which the first born was a male. Whether the first born was a male is exogenous. The fraction of neighboring households j in which the first born is a male is a good instrument because it is likely to affect the migration decisions of neighboring households j, but it does not affect the decisions of household i except through its effect on the decisions of neighboring households j.

The use of whether the first born is male as a source for an instrument is motivated by previous literature on migration in Mexico that shows that males are more likely to engage in migration to the US than females (Taylor and Martin, 2001), as well as previous literature on migration that finds that individuals who are men, younger, and the eldest child have larger probabilities of migrating abroad (Chort and Senne, 2018). Having the first born be a male will affect not only the first born's own probability of migration, but the probability of migration by subsequent siblings. Thus, migration is affected by the whether the first-born is male.

In our data, the mean of the fraction of neighbors whose first born is male is 0.5, as expected, and there is furthermore no evidence of systematic gender selection in Mexico. Nevertheless, there is variation in the fraction of neighbors whose first born is male (with a standard deviation of 0.13), and, as explained in more detail below, substantial variation within a village and within village-years in whether neighbors' first born is male. Thus, although there is no evidence of systematic gender selection in Mexico, there is variation in this instrument from two channels. First, the value of this instrument varies within a village owing to the exclusion of the own household in the calculation of the fraction of neighbors whose first born is male. In particular, we define the fraction of neighbors whose first born is male as the fraction of other households in the village of i, not including i, at time t in which the first born was a male. Since we exclude the own household from our calculation of the fraction of neighbors whose first born is male, a different set of households is used to calculate the value of this instrument for each household in the village. Since households in a village vary in whether their first born was male, the value of this instrument varies among households in a village. The second source of variation in this instrument is from the variation in the fraction of households with a first born being a male across villages.[11]

When instrumenting the average outcome of neighbors with the neighbors' average of certain characteristics, the same characteristics at the household level must also be included

[11] In their analysis of the effects of low skilled immigrants on the labor market outcomes of low skilled natives, Foged and Peri (2016) exploit the exogeneity of a Danish dispersal policy that distributed migrants from refugee counties to Denmark across municipalities to construct their instrument. Unfortunately, we do not have any similar exogenous policy in our context to exploit as an additional source of instruments.

as additional explanatory variables; failing to do so leads to biased and inconsistent estimates of neighborhood effects (von Hinke, Leckie and Nicoletti, 2019). Using the neighbors' average of certain characteristics as instruments for endogenous neighborhood effects while also controlling for the same characteristics at the household level also addresses the negative exclusion bias (Caeyers and Fafchamps, 2019). As our instruments – the neighbors' average household head schooling, the neighbors' average household average schooling, and the fraction of neighbors whose first born is male – are based on characteristics of a household's neighbors, we include as controls the same characteristics of the household itself – household head schooling, the household average schooling, and a dummy for whether the first born is male.

Our IV regressions also include controls at the household, village, municipality, state, and national level; border crossing variables; and village fixed effects. As explained above, the municipality controls include the number of schools in the basic system and the number of schools in the indigenous system. The village fixed effects control for unobserved conditions that affect all households in the village in the same fashion. For example, the village fixed effects absorb village characteristics such as the overall educational level of the village, infrastructure, and teaching quality. As a consequence, our instruments, whose values vary across households within the same village and across years, exploit the variation in the characteristics of neighbors that affect their probability of migration.

In the next subsection, we conduct several tests and analyses to make the case for the validity of our instruments and to address possible concerns about our instruments. As we show in detail below, there is variation within villages in the household characteristics on which the instruments are based so that the instruments are not too highly correlated with a household's own control variables; and moreover the correlations between the instruments and a household's own respective control variables are low. Thus, because the characteristics of a household are not perfectly collinear with the characteristics of its neighbors, the parameters in our endogenous effects model are identified (Manski, 1995). We also develop and conduct a falsification test to show that the instruments are not correlated with unobserved village factors that may affect migration decisions, thereby providing supportive evidence that the exclusion restriction is satisfied.

5.3 Validity of the instruments

We conduct several tests and analyses to make the case for the validity of our instruments for the fraction of neighbors that engage in migration, and to address possible concerns about our instruments.

We use both the neighbors' average household head schooling and the neighbors' average household average schooling as instruments in all our specifications. For robustness, in some specifications we also use the fraction of neighbors whose first born is male as an additional third instrument.

Table 3 presents the results of our first-stage regressions. In Specification (1), we use two instruments: the neighbors' average household head schooling and the neighbors' average household average schooling. In Specifications (2) and (3), we use all three instruments: the neighbors' average household head schooling, the neighbors' average household average schooling, and the fraction of neighbors whose first born is male. The first stage F-statistics are greater than 10 and their corresponding p-values (Pr>F) are less than 0.001 for all three specifications. To test the validity of our instruments, we also conduct an under-identification test and a Sargan-Hansen test of over-identifying restrictions. In all our specifications, we reject under-identification. We also pass the Sargan-Hansen test of over-identifying restrictions in the two overidentified specifications (Specifications (2) and (3)), since in each of these regressions we fail to reject the joint null hypothesis that the instruments are uncorrelated with the error term and that the excluded instruments are correctly excluded from the estimated equation.

One may worry that our instruments, which are based on characteristics of a household's neighbors, may be highly correlated with the same characteristics of the household itself. For example, if household head schooling does not vary much within a village, then the neighbors' average household head schooling (one of our instruments) may be highly correlated with a household's own household head schooling (one of our controls).

To address this concern and provide further evidence supporting the exclusion restriction, we examine the variation within villages in the values of household characteristics on which the instruments are based. For each village-year, we calculate the mean and standard deviation of the household characteristics on which the instruments are based across all households in that village-year. Panels A and B of Table 4 present summary statistics of the mean and standard deviation, respectively, over village-years of the household characteristics in each village-year. We then divide the mean standard deviation by the mean mean; the summary statistics for the ratio of the standard deviation over the mean over village-years are presented in Panel C of Table 4.

As seen in Panel C of Table 4, we find that the ratio of the mean standard deviation over the mean mean is non-zero for all of the covariates on which the instruments are based. For example, for household head schooling, the ratio of the mean (over village-years) of the standard deviation within a village-year, over the mean (over village-years) of the mean within a village-year is 0.7322, which provides evidence that there is considerable variation in household head schooling within a village-year, and therefore that the neighbors' average household head schooling (one of our instruments) is not highly correlated with a household's own household head schooling (one of our controls). Similarly, for the dummy variable for whether the first born is male, the ratio of the mean (over village-years) of the standard deviation within a village-year, over the mean (over village-years) of the mean within a village-year is 0.9728, which provides evidence that there is considerable variation in whether the first born is male within a village-year, and therefore that the fraction of neighbors whose first born is male (one of our instruments) is not highly correlated with whether a household's own first born is male (one of our controls). Thus, there is variation within villages in the household characteristics on which the instruments are based so that the instruments are not too highly correlated with a household's own respective control variables.

To provide further evidence that the instruments are not too highly correlated with a household's own respective control variables, we calculate the correlations between each instrument measuring neighbors' characteristics and its respective own household characteristic. As seen in Table 5, the correlations between the instruments and their corresponding own variables are low.

There is therefore variation within villages in the household characteristics on which the instruments are based so that the instruments are not too highly correlated with a household's own control variables; and moreover the correlations between the instruments and a household's own respective control variables are low. Thus, because the characteristics of a household are not perfectly collinear with the characteristics of its neighbors, the parameters in our endogenous effects model are identified (Manski, 1995).

Another concern is that the exclusion restriction may be violated because the instruments are correlated with unobserved village factors that may affect migration decisions. While we already include controls at the household, village, municipality, state, and national level; border crossing variables; and village fixed effects; and while the municipality controls include the number of schools in the basic system and the number of schools in the indigenous system, it is possible that there may be unobserved village factors such as returns to schooling or local governance that may affect migration decisions.

We develop and conduct a falsification test of the first-stage regression to examine whether the instruments are correlated with unobserved village factors that may affect migration decisions. For this test, we directly regress the indicator variable y_{ikt} of whether household i has migration of type $k \in [USA, Mexico]$ at time t on the instruments and all the regressors. Building on a technique applied by Morrison and Lin Lawell (2016), this falsification test of the first-stage regression uses pseudo-endogenous variables as dependent variables rather than our actual endogenous variables. In particular, instead of the fraction of neighbors that engage in migration, we use the following pseudo-endogenous variable: the indicator variable y_{ikt} of whether household i has migration of type $k \in [USA, Mexico]$ at time t.

Our falsification test of the first-stage regression enables us to examine whether the instruments are correlated with unobserved village factors that may affect migration decisions and therefore provides evidence regarding whether the exclusion restriction is satisfied. If the exclusion restriction is satisfied and the instruments for the fraction of neighbors that engage in migration are not correlated with unobserved village factors that affect household migration decisions, then we would expect that our instruments – the neighbors' average household head schooling, the neighbors' average household average schooling, and the fraction of neighbors whose first born is male – should not be strong predictors of whether household i engages in migration. In other words, if the exclusion restriction is satisfied, the characteristics of neighbors j that we use as instruments should not explain the migration decisions of household i.

While our test is akin to that in Morrison and Lin Lawell (2016), it is not the same technique. The pseudo-endogenous variables Morrison and Lin Lawell (2016) use are variables that would be affected by omitted variables that affect the dependent variable in the second-stage regression. In particular, their pseudo-endogenous variables are based on the behavior of individuals in the neighborhood who are not members of the subgroup of the sample they analyze. If the exclusion restriction is satisfied, then their instruments should not explain these pseudo-endogenous variables. In particular, their first-stage falsification test enables them to test whether their instruments are correlated with omitted variables that affect the dependent variable in the second-stage regression.

In contrast, for our first-stage falsification test, the pseudo-endogenous variables we use are the dependent variables in the second-stage regression themselves. If the exclusion restriction is satisfied, the characteristics of neighbors j that we use as instruments should not be strong predictors of the migration decisions of household i. Since we model the behavior of all sampled households in a village, rather than a subset of all sampled households in a village, we are not able to use the behavior of households in the village who are not members of the subset being analyzed as a source of pseudo-endogenous variables. We therefore do not implement the same test as Morrison and Lin Lawell (2016).

Since our instruments must not only satisfy the exclusion restriction but also be correlated with the endogenous variables, and since the pseudo-endogenous variables we use are the dependent variables in the second-stage regression themselves, if the instruments are correlated with the endogenous variables, then a non-zero correlation between the instruments and the pseudo-endogenous variables in our first-stage falsification test is still possible even if the exclusion restriction is satisfied. Nevertheless, if the exclusion restriction is satisfied, the characteristics of neighbors j that we use as instruments should not be strong predictors of the migration decisions of household i. Thus, unlike in Morrison and Lin Lawell (2016), our test is not whether the instruments explain the pseudo-endogenous variables, but instead whether the instruments are strong predictors of the pseudo-endogenous variables. In particular, instead of examining whether the coefficients on the instruments in the first-stage falsification test are statistically significant, we instead examine whether the first-stage F-statistics in the falsification test of the first-stage regressions are greater than 10.

Thus, if the exclusion restriction is satisfied and the instruments for the fraction of neighbors that engage in migration are not correlated with unobserved village factors that affect household migration decisions, then we would expect that the characteristics of neighbors j that we use as instruments should not be strong predictors of the migration decisions of household i, and therefore that the first-stage F-statistics in the first-stage falsification test are small.

We present the results of our falsification test of the first-stage regressions in Table 6. While each of our regressions has two endogenous variables, corresponding to the fraction of neighbors that engage in migration to the US and within Mexico, only at most one instrument is significant in each of them. Moreover, the first-stage F-statistics are all small and much lower than 10 in all instances, and the corresponding p-values indicate that the null hypothesis of instruments being jointly non-significant cannot be rejected in all except for one case. The characteristics of neighbors j that we use as instruments therefore do not explain the migration decisions of household i. Thus, the instruments are not correlated with unobserved village factors that affect migration decisions, which provides supportive evidence that the exclusion restriction is satisfied.

6 Results

In this section we present the results of our empirical analysis of whether the migration decisions of other households in the village affect a household's decision to have a member migrate. We run two sets of IV regressions: one for the decision to migrate to the US, and the other for the decision to migrate within Mexico. The parameters of interest are the coefficient on the fraction of households in the same village as household i, excluding i, with migration to the US; and the coefficient on the fraction of households in the same village as i, excluding i, with migration within Mexico. We empirically examine whether there are neighborhood effects in household migration decisions in rural Mexico, and whether there are nonlinearities in the neighborhood effects.

6.1 Neighborhood effects

Tables 7 and 8 present the main results of our paper for migration to the US and migration within Mexico, respectively. In both Tables 7 and 8, Specification (1) is our preferred specification. Specification (2) includes households' plot level area interacted with precipitation. Specification (3) includes year effects instead of a time trend.

According to the results for the US migration regressions in Table 7, there is a positive and significant own-migration neighborhood effect that is robust across specifications. In the base case Specification (1), an increase of one standard deviation (which is 0.21, as seen in Table 1) in the fraction of neighbors with migration to the US increases a household's probability of migration to the US by 11.92 percentage points. The result is robust across specifications; an increase of one standard deviation in the fraction of neighbors with migration to the US increase a household's probability of migration to the US by 11.33 to 11.92 percentage points across Specifications (1), (2), and (3). Additionally, the other-migration neighborhood effect is negative and significant at a 5% level in Specifications (2) and (3): the fraction of neighbors with migration within Mexico has a significant negative effect on a household's probability of migration to the US in some, but not all, specifications. An increase of one standard deviation (which is 0.17, as seen in Table 1) in the fraction of household with migration within Mexico reduces the probability of migration to the US by 14.91 to 15.24 percentage points.

Table 8 presents the results for the migration to other states within Mexico. Neither the coefficient on the the fraction of neighbors with migration within Mexico nor the coefficient on the fraction of neighbors with migration to the US is significant at a 5% level in any of the specifications. We therefore do not find evidence of neighborhood effects in migration within Mexico.

6.2 Other determinants of migration

Table 7 also presents the results of other characteristics that may explain migration to the US. The number of males in the household has a positive effect that is significant in all specifications, which is consistent with the previous literature on the determinants of migration, which finds that migrants from Mexico to the US are more likely to be male than female. The household head age has a significant positive effect in all specifications.

Household head schooling has a significant negative effect on the probability of migration to the US, which could provide evidence in favor of the negative selection hypothesis. In contrast, the household's maximum level of schooling has a significant positive effect on the probability of migration to the US. Thus, while the household maximum schooling increases the probability of migration to the US, having a highly educated household head decreases the probability of migration to US.

To control for persistence in migration, we include the lag of the fractions of household members migrating to the US and within Mexico, respectively, as regressors. The lagged fraction of household members migrating to the US has a significant positive effect on the probability of migration to the US. The lagged fraction of household members migrating to other states within Mexico has has no effect on the probability of migration to the US.

Table 8 also presents the results of other characteristics that may explain migration to other states within Mexico. The number of males in the household has a significant positive effect on migration to other states within Mexico. The age of the household head has a significant positive effect on the probability of migrating within Mexico. Whether the first born in the household is a male has a significant negative effect, meaning that the first male of the family tends to stay at the home locality, likely to become the next generation's household head.

Household head schooling has a significant negative effect on the probability of migrating within Mexico. Household maximum schooling has a significant positive effect. Thus, while the household maximum schooling may increase the probability of migration within Mexico, having a highly educated household head decreases the probability of migration within Mexico.

The lagged fraction of household members migrating within Mexico has a significant positive effect on the probability of migration within Mexico, which provides evidence for a high degree of persistence. In contrast, the lagged fraction of household members migrating to the US has no significant effect on the probability of migration within Mexico.

6.3 Robustness

We examine the robustness of our results by running alternative specifications that are variants of the base case Specification (1) from Tables 7 and 8 for migration to the US and migration within Mexico, respectively. The results of our robustness checks are in Tables A.1 and A.3 in the Appendix for migration to the US and within Mexico, respectively. In Specification (4) we divide the municipality-level variables by municipality population. In Specification (5) we do not divide the municipality-level variables by municipality population, and we include an alternative set of border crossing variables. In Specification (6) we drop all the wage and border crossing variables. In Specification (7) we use year fixed effects instead of a time trend.

We also estimate a set of specifications that include the lagged dependent variable (i.e., lagged migration to the US and lagged migration within Mexico) using the augmented version of the Arellano-Bond (1991) estimator outlined in Arellano and Bover (1995) and fully developed in Blundell and Bond (1998). This dynamic panel estimator for lagged dependent variables is a Generalized Method of Moments estimator that treats the model as a system of equations, one for each time period, and which addresses the issue with the original Arellano-Bond estimator that lagged levels are often poor instruments for first differences, especially for variables that are close to a random walk. These results are presented in Tables A.2 and A.4 in the Appendix for migration to the US and within Mexico, respectively.

The main results from our base case specification are robust across the alternative specifications. In the US migration results for the alternative specifications in Table A.1 in the Appendix, an increase of one standard deviation in the fraction of neighbors with migrants to the US increases a household's probability of migration by 10.43 to 12.73 percentage points. As before, the number of males in the household, the household head age, the maximum level of schooling in the household, and the lagged fraction of household members with migration to the US have significant positive effects on the probability of having a migrant to the US. Household head schooling has a significant negative effect on the probability of having a household member migrate to the US.

In Table A.2 in the Appendix, we present the US migration results of specifications that include the lagged dependent variable (i.e., lagged migration to the US and lagged migration within Mexico) using the augmented version of the Arellano-Bond (1991) estimator outlined in Arellano and Bover (1995) and fully developed in Blundell and Bond (1998). As before, we find a positive and significant own-migration neighborhood effect for US migration. We also find a positive and statistically significant effect of the number of males in the household, and a negative statistically significant effect of the household head schooling, on the probability of migration to the US. The other-migration neighborhood effect is not significant at a 5% level in any of the specifications in Table A.2 in the Appendix: when using the Arellando-Bond estimator, the fraction of neighbors with migration within Mexico does not have a significant effect on a household's probability of migration to the US.

In Figure 2 we summarize our findings on the own-migration neighborhood effect for migration to the US across Specifications (1)-(3) of Table 7, Specifications (4)-(7) of Table A.1 in the Appendix, and Specifications (8)-(12) of Table A.2 in the Appendix. Across all specifications, we find a positive and significant own-migration neighborhood effect for US migration.

For migration within Mexico, Table A.3 in the Appendix shows that, just as in the main results from our base case specification, we do not find evidence of a statistically significant own-migration neighborhood effects in migration within Mexico. The schooling of the household head and whether the first born is male have a negative effect on the probability of migration. The number of males in the household, the household head age, and the lagged fraction of household members with migration within Mexico have a significant positive effect on the probability of migration to other states within Mexico.

In Table A.4 in the Appendix, we present the US migration results of specifications that include the lagged dependent variable (i.e., lagged migration to the US and lagged migration within Mexico) using the augmented version of the Arellano-Bond (1991) estimator outlined in Arellano and Bover (1995) and fully developed in Blundell and Bond (1998). Results show that, when we use the Arellando-Bond estimator, we find a statistically significant own-migration neighborhood effect. Nevertheless, the effect is still small: an increase of one standard deviation (which is 0.17, as seen in Table 1) in the fraction of household with migration within Mexico increases the probability of migration within Mexico by 2.12 to 4.26 percentage points. As before, we do not find evidence of a statistically significant neighborhood effect of migration to the US on migration within Mexico. We also find that the number of males in the household and the household head age have a statistically significant and positive effect on the probability of migration within Mexico, while the household head schooling has a negative and statistically significant effect on this probability.

In Figure 3 we summarize the findings of our results on the own-migration neighborhood effect for migration within Mexico across Specifications (1)-(3) of Table 8, Specifications (4)-(7) of Table A.3 in the Appendix, and Specifications (8)-(12) of Table A.4 in the Appendix. We find that own-migration neighborhood effects on migration within Mexico are either statistically insignificant or small in magnitude.

6.4 Nonlinearities in the neighborhood effect

We also examine whether the effect of the fraction of neighbor households with migration on a household's migration decision may be a nonlinear function of the fraction of neighbor households with migrants.

To allow for nonlinearities in the neighborhood effect in the most flexible way, we estimate a semiparametric partially linear IV model using Robinson's (1988) double residual estimator to semi-parametrically identify nonlinearities in the own-migration neighborhood effect. We use a two-step procedure to account for the endogeneity of the strategic variables with a control function approach (Verardi, 2013). In the first stage, we fit two panel fixed effects models regressing the fraction of households with migrants to the US and within Mexico, respectively, on the same instruments and controls as we used in our base case specification 1a from Tables 7 and 8 respectively, and estimate the residuals. In the second stage, we run a semi-parametric estimation where the dependent variable is the migration dummy, controlling for the same controls as in our base case specification 1a, and adding the first-stage residuals as additional regressors. The controls, the residual estimated in the first stage, and village and year fixed effects enter the second stage parametrically, while the endogenous regressor – the fraction of households engaged in migration to the US (within Mexico) – is allowed to vary non-parametrically.

To examine any nonlinearities in the own-migration neighborhood effect for migration to the US, Figure 4 plots the partialled-out residual in the analysis of the decision to migrate to the US as a function of the fraction of neighbors with migration to the US. In general we seem to find a linear effect.

Similarly, to examine any nonlinearities in the own-migration neighborhood effect for migration within Mexico, Figure 5 plots the partialled-out residual in the analysis of the decision to migrate within Mexico as a function of the fraction of neighbors with migration to within Mexico. The relationship again seems to be linear.

We also run two additional alternative specifications of the semiparametric partially linear IV model. In one alternative specification, we run the semiparametric partially linear IV model without instrumenting for the other-migration neighborhood effect. In the other alternative specification, we run the semiparametric partially linear IV model without including the other-migration neighborhood effect as a regressor. Our results are robust across all specifications. Thus, our test for nonlinearities in the migration neighborhood effects do not reject the hypothesis that the neighborhood effects are linear in the fraction of neighbor households with migrants.

7 Conclusion

We contribute to the literature on the determinants of migration by examining whether neighborhood effects matter for migration decisions. There are several reasons why a household's migration decisions may depend on the migration decisions of its neighbors, including including migration networks, information externalities, relative deprivation, risk sharing, competition effects, the marriage market, and cultural norms. Using instrumental variables, a rich set of controls, and village fixed effects to address the endogeneity of neighbors' decisions as well as exclusion bias, we empirically examine whether there are neighborhood effects in household migration decisions in rural Mexico, and whether there are nonlinearities in the neighborhood effects.

We conduct several tests and analyses to make the case for the validity of our instruments and to address possible concerns about our instruments. We develop and conduct a falsification test to show that the instruments are not correlated with unobserved village factors that may affect migration decisions, thereby providing supportive evidence that the exclusion restriction is satisfied.

Results of our IV regressions show that there is a significant and positive own-migration neighborhood effect on migration to the US. In our base case specification, an increase of one standard deviation in the fraction of neighbors with migration to the US increases a household's probability of engaging in migration to the US by around 12 percentage points. We find that this neighborhood effect on migration to the US is linear in the fraction of neighbors that migrate to the US in the village.

In contrast, for migration within Mexico, we find that own-migration neighborhood effects on migration within Mexico are either statistically insignificant or small in magnitude. The fraction of neighbors with migration within Mexico has either a statistically insignificant or small effect on a household's probability of engaging in migration within Mexico.

Our results therefore suggest that neighborhood effects among households in a village have an important role in household migration decisions, and that neighborhood effects are different for international migration and domestic migration. This is consistent with the previous theoretical and empirical literature incorporating the different costs associated with migration and analyzing how these costs may have different effects on migration if the migration were international rather than domestic, for example in the form of cultural barriers or strong preferences for establishing close to home. Our results suggest that the neighborhood effects that arise from decisions to migrate to other countries are different than those generated by local migration. In particular, we find that, for households in rural Mexico, neighborhood effects are important for international migration to the US, but do not have a significant effect on decisions to migrate domestically within Mexico.

An empirical assessment of neighborhood effects in migration has important implications for policy. If neighborhood effects are positive – for example through network effects or information spillovers – then policies that affect the migration decisions of one household in a village may have a multiplier effect and induce similar decisions by neighboring households. As a consequence, migration policies (to deter or induce migration) may be more cost-effective if they are targeted at the village level, and focused on directly affecting the migration decisions of a subset of households in each village. In contrast, if neighborhood effects are negative – for example when there are higher returns to migrating when there are fewer migrants – then deterring one village household from engaging in migration may induce other neighboring households to engage in migration.

Our result that there is a significant and positive own-migration neighborhood effect on migration to the US suggests that policies that affect the international migration decisions of one household in a village may have a multiplier effect and induce similar decisions by neighboring households. As a consequence, international migration policies to deter or induce international migration from Mexico to the US may be more cost-effective if they are targeted at the village level, and focused on directly affecting the international migration decisions of a subset of households in each village.

There are several potential avenues for future research. First, the analysis in this paper was based on reduced-form regressions of migration. In Rojas Valdés, Lin Lawell and Taylor (2018) and Rojas Valdés, Lin Lawell and Taylor (2020a), we further analyze neighborhood effects in migration by developing and estimating a structural econometric model of the dynamic game between households deciding whether or not to engage in migration to the US and within Mexico.

Second, our empirical model is general enough to capture multiple possible sources of neighborhood effects, including migration networks, information externalities, relative deprivation, risk sharing, competition effects, the marriage market, and cultural norms, and enables us to analyze their net effect. In future work we hope to further distinguish among different possible sources of neighborhood effects.

The focus of our paper is on neighborhood effects. Having demonstrated the importance of neighborhood effects, a third potential avenue for future research is to analyze the effects of other individual, household, and village characteristics on migration decisions after properly accounting for and controlling for these neighborhood effects. Altonji and Mansfield (2018) show that controlling for group averages of observed individual characteristics potentially absorbs all the across-group variation in unobservable individual characteristics, including peer effects and neighborhood effects, and use this insight to bound the treatment effect variance of school systems and associated neighborhoods for various outcomes. In future work we hope to build on their insights and empirical techniques to analyze the effects of other individual, household, and village characteristics on migration decisions after properly accounting for and controlling for neighborhood effects.

A fourth potential avenue for future research is to analyze migration decisions at an individual level. In this paper, following the previous literature on the new economics of labor migration, we have posited the household as the relevant unit of analysis. In future work we hope to analyze individual migration decisions; intra-household decision-making; and the relationships among individual and household migration decisions.

References

Acosta, P. (2011). School attendance, child labour, and remittances from international migration in El Salvador. *Journal of Development Studies*, 47(6): 913-936.

Acosta, P., C. Caldron, P. Fajnzylber, and H. Lopez. (2008). What is the impact of international remittances on poverty and inequality in Latin America? *World Development*, 36(1): 89-114.

Altonji, J.G., and R.K. Mansfield. (2018). Estimating group effects using averages of observables to control for sorting on unobservables: School and neighborhood effects. *American Economic Review*, 108(10): 2902-2946.

Alcala, N.R., L.C. Adkins, B. Lahin, and A. Savvides. (2014). Remittances and income diversification in Bolivia's rural sector. *Applied Economics*, 46(8): 848-858.

Angrist, J.D. (2014). The perils of peer effects. *Labour Economics*, 30, 98-108.

Arellano, M., and S. Bond. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies*, 58: 277-297.

Arellano, M., and O. Bover. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics*, 68: 29-51.

Bandiera, O., and I. Rasul. (2006). Social networks and technology adoption in Northern Mozambique. *Economic Journal*, 116(514): 869-902.

Battisti, M., G. Peri, and A. Romiti. (2016). Dynamic effects of co-ethnic networks on immigrants' economic success. NBER Working Paper No. 22389.

Bauer, T., and I.N. Gang. (1998). Temporary migrants from Egypt: How long do they stay abroad? IZA Discussion Paper No. 3, April.

Blundell, R., and S. Bond. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87: 115-143.

Bollinger, B., J. Burkhardt, and K.T. Gillingham. (2020). Peer effects in residential water conservation: Evidence from migration. *American Economic Journal: Economic Policy*, 12(3): 107-133.

Borjas, G.J. (1987). Self-Selection and the Earnings of Immigrants. *American Economic Review*, 77(4): 531-553.

Borjas, G.J. (2008). Labor Outflows and Labor Inflows in Puerto Rico. *Journal of Human Capital*, 2(1): 32-68.

Brock, W.A., and S.N. Durlauf. (2001). Discrete choice with social interactions. *Review of Economic Studies*, 68: 235–260.

Caeyers, B., and M. Fafchamps. (2019). Exclusion bias in the estimation of peer effects. Working paper, Institute for Fiscal Studies, Norwegian School of Economics, and Stanford University.

Calero, C., A.S. Bedi, and R. Sparrow. (2009). Remittances, liquidity constraints and human capital investments in Ecuador. *World Development*, 37(6): 1143-1154.

Campos-Vazquez, R.M., A. Hincapie, and R.I. Rojas-Valdés. (2012). Family income inequality and the role of married females' earnings in Mexico: 1988-2010. *Latin American Journal of Economics*, 49(1): 67-98.

Carrington, W.J., E. Detragiache, and T. Vishwanath (1996). Migration with endogenous moving costs. *American Economic Review*, 86(4): 909-930.

Castelhano, M., C.-Y.C. Lin Lawell, D.A. Sumner, and J.E. Taylor. (2020). Migration, remittances, and productive investment in rural Mexico. Working paper, Cornell University.

Chen, X., A. Szolnoki, and M. Perc. (2012). Risk-driven migration and the collective-risk social dilemma. *Physical Review E*, 86(3): 1-8.

Chiquiar, D., and G.H. Hanson. (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.

Cheng, H., Dai, Q., Li, H., Zhu, Y., Zhang, M., and Yang, J. (2011). Payoff-related migration enhances cooperation in the prisoner's dilemma game. *New Journal of Physics*, 13: 043032.

Chort, I., and J.-N. Senne. (2018). You'll be a migrant, my son: Accounting for migrant selection within the household. *Economic Development and Cultural Change*, 66(2): 217-263.

Conley, T.G., and G. Topa. (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics*, 17: 303–327.

Conley, T.G., and C. Udry. (2010). Learning about a new technology: Pineapple in Ghana. *American Economic Review*, 100(1): 35-69.

Cornelius, W., and I. Salehyan. (2007). Does border enforcement deter unauthorized immigration?: The case of Mexican migration to the United States of America. *Regulation & Governance* 1(2): 139-153.

Cuecuecha, A. (2005). The immigration of educated Mexicans: The role of informal social insurance and migration costs. Working paper, Instituto Tecnologico Autonomo de Mexico.

Du, Y., A. Park, A., and S. Wang. (2005). Migration and rural poverty in China. *Journal of Comparative Economics*, 33(4): 688-709.

Durlauf, S.N. (2004). Chapter 50 - Neighborhood Effects. In: J.V. Henderson and J.-F. Thisse (Eds), *Handbook of Regional and Urban Economics*, Volume 4 (pp. 2173-2242). Elsevier.

D'Vera, C., A. Gonzalez-Barrera, and D. Cuddington. (2013). Remittances to Latin America Recover – but Not to Mexico. Pew Research Center. `http://www.pewhispanic.org/2013/11/15/remittances-to-latin-america-recover-but-not-to-mexico/`.

Elsner, B., G. Narciso, and J. Thijssen. (2018). Migrant networks and the spread of information. *Oxford Bulletin of Economics and Statistics*, 80(3): 659-688.

Esquivel, G., and A. Huerta-Pineda. (2007). Remittances and poverty in Mexico: A propensity score matching approach. *Integration & Trade*, 27: 45-71.

Fajardo, G., E. Gutiérrez, and H.A. Larreguy. (2017). Taking one for the team: Shocks at destination and household's supply of migrants. CAF Working paper No. 2017/10.

Feigenberg, B. (2020). Fenced out: The impact of border construction on US-Mexico migration. *American Economic Journal: Applied Economics*, 12(3): 106-139.

Feliciano, Z.M. (2001). The skill and economic performance of Mexican immigrants from 1910 to 1990. *Explorations in Economic History*, 38(3): 386-409.

Foged, M., and G. Peri. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2): 1-34.

Foster, A.D., and M.R. Rosenzweig. (1995). Learning by doing and learning from others: Human capital and technical change in agriculture. *Journal of Political Economy*, 103(6), 1176-1209.

Garlick, J., M. Leibbrandt, and J. Levinsohn. (2016). Individual migration and household incomes. NBER Working Paper 22326.

Gathmann, C. (2008). Effects of enforcement on illegal markets: Evidence from migrant smuggling along the southwestern border. *Journal of Public Economics*, 92(10): 1926-1941.

Glaeser, E.L., B. Sacerdote, and J.A. Scheinkman. (1996). Crime and social interactions. *Quarterly Journal of Economics*, 111 (2): 507–548.

Graham, B.S. (2018). Identifying and estimating neighborhood effects. *Journal of Economic Literature*, 56(2): 450-500.

Greenwood, M.J. (1971). An analysis of the determinants of internal labor mobility in India. *Annals of Regional Science*, 5(1): 137-151.

Guryan, J. , D. Kroft, and N. J. Notowidigdo. (2009). Peer effects in the workplace: Evidence from random groupings in professional golf tournaments. *American Economic Journal: Applied Economics*, 44 (3), 289-302.

Hanson, G.H. (2010). International Migration and the Developing World. In: D. Rodrik and M. Rosenzweig (Eds), *Handbook of Development Economics* (pp. 4363-4414).

Irwin, E.G., and N.E. Bockstael. (2002). Interacting agents, spatial externalities and the evolution of residential land use patterns. *Journal of Economic Geography*, 2(1): 31–54.

Jessoe, K., D.T. Manning, and J.E. Taylor. (2018). Climate change and labor allocation in rural Mexico: Evidence from annual fluctuations in weather. *Economic Journal*, 128(608): 230–261.

Kinnan, C., S.-Y. Wang, and Y. Wang. (2018). Access to migration for rural households. *American Economic Journal: Applied Economics*, 10 (4), 79-119.

Lessem, R. (2018). Mexico-US immigration: Effects of wages and border enforcement. *Review of Economic Studies*, 85(4): 2353-2388.

Lin, C.-Y.C. (2009). Estimating strategic interactions in petroleum exploration. *Energy Economics*, 31(4): 586-594.

Lin, Y.-T., H.-X. Yang, Z.-X. Wu, and B.-H. Wang. (2011). Promotion of cooperation by aspiration-induced migration. *Physica A: Statistical Mechanics and its Applications*, 390(1): 77-82.

Maertens, A. (2017). Who cares what others think (or do)? Social learning and social pressures in cotton farming in India. *American Journal of Agricultural Economics*, 99 (4), 988–1007.

Mahajan, P., and D. Yang. (2020). Taken by storm: Hurricanes, migrant networks, and U.S. immigration. *American Economic Journal: Applied Economics*, 12(2): 250-77.

Manski C. (1993). Identification of endogenous social effects: The reflection problem. *Review of Economic Studies*, 60(3): 531-542.

Manski C. (1995). *Identification Problems in the Social Sciences*. Harvard University Press: Cambridge, MA.

Massey, D.S. (2007). Understanding America's immigration 'crisis'. *Proceedings of the American Philosophical Society*, 151(3): 309–327.

Massey, D.S., and K. Espionsa. (1997). What's driving Mexico-U.S. migration? A theoretical, empirical, and policy analysis. *American Journal of Sociology*, 102(4): 939-999.

Massey, D.S., L. Goldring, and J. Durand. (1994). Continuities in transnational migration: An analysis of nineteen Mexican communities. *American Journal of Sociology*, 99(6): 1492-1533.

McKenzie, D., and H. Rapoport. (2007). Network effects and the dynamics of migration and inequality: Theory and evidence from Mexico. *Journal of Development Economics*, 84(1): 1-24.

McKenzie, D., and H. Rapoport. (2010). Self-selection patterns in Mexico-US migration: The role of migration networks. *Review of Economics and Statistics*, 92(4): 811-821.

Mexican Migration Program. (2014). Selected results. Mexican Migration Program. `http://mmp.opr.princeton.edu/results/results-en.aspx`.

Moffitt, R.A. (2001). Policy interventions, low-level equilibria, and social interactions. In: S.N. Durlauf and H.P. Young (Eds), *Social Dynamics* (pp. 45-79). Brookings Institution and MIT Press, Cambridge, MA.

Morrison, G.M., and C.-Y.C. Lin Lawell. (2016). Driving in force: The influence of workplace peers on commuting decisions on U.S. military bases. *Journal of Economic Behavior and Organization*, 125: 22-40.

Morten, M. (2019). Temporary migration and endogenous risk sharing in village India. *Journal of Political Economy*, 127(1): 1-46.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the U.S. labor market. *Quarterly Journal of Economics*, 118(2): 549-599.

Munshi, K. (2004). Social learning in a heterogenous population: Technology diffusion in the Indian Green Revolution. *Journal of Development Economics*, 74(1): 175-213.

Munshi, K., and M. Rosenzweig. (2016). Networks and misallocation: Insurance, migration, and the rural-urban wage gap. *American Economic Review*, 106(1): 46-98.

Nelson, D. (2015). Migration and Networks. In: P. Commendatore, S. Kayam and I. Kubin (Eds), *Complexity and Geographical Economics*. Springer.

Nelson, J.M. (1976). Sojourners versus new urbanites: Causes and consequences of temporary versus permanent cityward migration in developing countries. *Economic Development and Cultural Change*, 24(4): 721-757.

Neubecker, N., M. Smolka, and A. Steinbacher. (2017). Networks and selection in international migration to Spain. *Economic Inquiry*, 55(3): 1265-1286.

Orrenius, P.M., and M. Zavodny. (2005). Self-selection among undocumented immigrants from Mexico. *Journal of Development Economics*, 78(1): 215-240.

Pfeiffer, L., and C.-Y.C. Lin. (2012). Groundwater pumping and spatial externalities in agriculture. *Journal of Environmental Economics and Management*, 64 (1): 16-30.

Riosmena, F. (2009). Socioeconomic context and the association between marriage and Mexico-U.S. migration. *Social Science Research*, 38 (2): 324-337.

Robalino, J.A., and A. Pfaff. (2012). Contagious development: Neighbor interactions in deforestation. *Journal of Development Economics*, 97 (2): 427-436.

Robinson, P.M. (1988). Root-n-consistent semiparametric regression. *Econometrica*, 56(4): 931-954.

Rojas Valdés, Rubén Irvin, C.-Y. Cynthia Lin Lawell, and J. Edward Taylor. (2018). Migration in rural Mexico: Strategic interactions, dynamic behavior, and the environment. *Journal of Academic Perspectives*, 2017(3).

Rojas Valdés, R.I., C.-Y.C. Lin Lawell, and J.E. Taylor. (2020a). Migration dynamics, strategy, and policy. Working paper, Cornell University.

Rojas Valdés, R.I., C.-Y.C. Lin Lawell, and J.E. Taylor. (2020b). The economics of migration and migration policy in rural Mexico. Working paper, Cornell University.

Rubalcava, L.N., G.M. Teruel, D. Thomas, and N. Goldman. (2008). The healthy migrant effect: New findings from the Mexican Family Life Survey. *American Journal of Public Health*, 98(1): 80-84.

Schaffer, M.E. (2010). xtivreg2: Stata module to perform extended IV/2SLS, GMM and AC/HAC, LIML and k-class regression for panel data models. `http://ideas.repec.org/c/boc/bocode/s456501.htm`.

Sjaastad, L.A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5): 80-93.

Stark, O., and D.E. Bloom. (1985). The new economics of labor migration. *American Economic Review*, 75(2): 173-178.

Stark, O., and J.E. Taylor. (1989). Relative deprivation and international migration. *Demography*, 26(1): 1-14.

Stark, O., and J.E. Taylor. (1991). Migration incentives, migration types: The role of relative deprivation. *Economic Journal*, 101(408): 1163-1178.

Taylor, J.E. (1987). Undocumented Mexico-US migration and the returns to households in rural Mexico. *American Journal of Agricultural Economics*, 69(3): 626-638.

Taylor, J.E., and P.L. Martin. (2001). Human capital: migration and rural population change.

In: B. Gardner and G. Rausser (Eds), *Handbook of Agricultural Economics* (pp. 457-511). Elsevier.

Taylor, J.E., J. Mora, R. Adams, and A. Lopez-Feldman (2008). Remittances, inequality and poverty: Evidence from rural Mexico. In: J. DeWind and J. Holdaway (Eds), *Migration and Development Within and Across Borders: Research and Policy Perspectives on Internal and International migration*. Social Science Research Council.

Topa, G., and Y. Zenou (2015). Neighborhood and network effects. In: G. Duranton, V. Henderson and W. Strange (Eds), *Handbook of Regional and Urban Economics* (pp. 561-624), Vol. 5A. Amsterdam: Elsevier Publisher.

United Nations, Department of Economic and Social Affairs, Population Division (2013). International Migration 2013 Wallchart. `http://www.un.org/en/development/desa/population/publications/pdf/migration/migration-wallchart2013.pdf`.

Verardi, V. (2013). Semiparametric regression in STATA. 2013 UK Stata Users Group meeting, London, UK, `http://www.stata.com/meeting/uk13/abstracts/materials/uk13_verardi.pdf`.

von Hinke, S., G. Leckie, C. Nicoletti. (2019). The use of instrumental variables in peer effects models. *Oxford Bulletin of Economics and Statistics*, 81(5): 1179-1191.

Wahba, J., and Y. Zenou (2005). Density, social networks and job search methods: Theory and applications to Egypt. *Journal of Development Economics*, 78: 443-473.

World Bank. (2010a). Datos sobre migración y remesas. `http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1110315015165/Factbook2011Spanish.pdf`.

World Bank. (2010b). Bilateral Migration Matrix. `http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1110315015165/T1.Estimates_of_Migrant_Stocks_2010.xls`.

World Bank. (2012). Bilateral Remittance Matrix.`http://siteresources.worldbank.org/INTPROSPECTS/Resources/334934-1110315015165/Bilateral_Remittance_Matrix_2012.xlsx`.

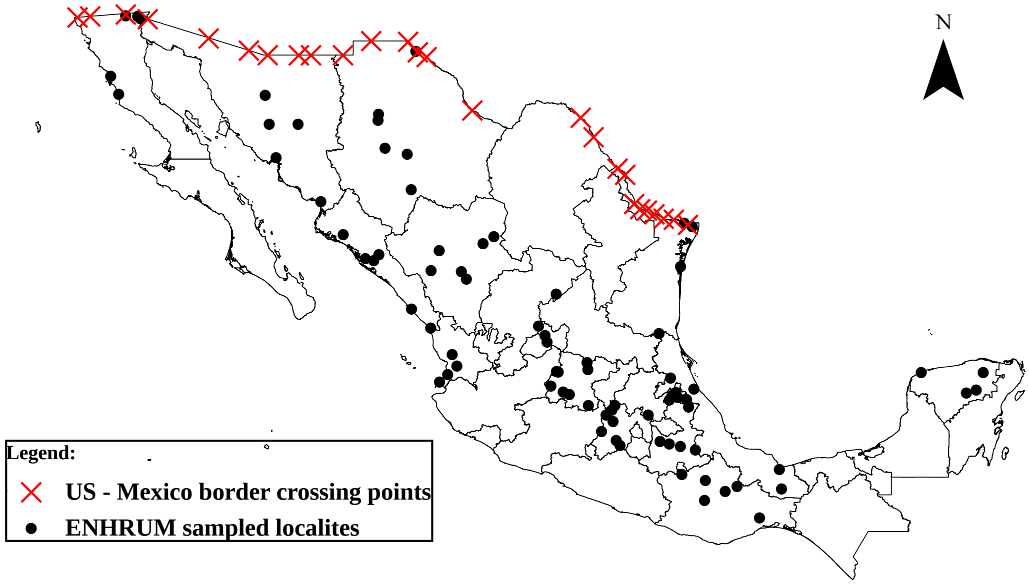

Figure 1: Location of sampled villages in the ENHRUM survey and the border crossing municipalities

Table 1: Summary statistics

	Mean	Std.Dev.	Min	Max	# Obs
Household migration variables					
Household has a migrant to the US (dummy)	0.17	0.38	0	1	25,761
Household has a migrant within Mexico (dummy)	0.20	0.40	0	1	25,761
Neighbor migration variables					
Fraction of neighbors with migrants to US	0.17	0.21	0	1	25,761
Fraction of neighbors with migrants to Mexico	0.20	0.17	0	0.89	25,761
Household characteristics					
Number of household members	5.94	3.15	1	24	25,761
Number of males in household	2.93	1.84	0	17	25,761
First born is a male (dummy)	0.5	0.5	0	1	25,761
Household head age (years)	45.15	16.26	3	100	25,725
Household head schooling (years)	4.75	3.84	0	23	25,725
Household average schooling (years)	6.21	2.97	0	20.5	25,554
Household maximum schooling (years)	8.99	3.84	0	23	25,761
Irrigated land area (hectares)	0.22	3.38	0	426	21,257
Irrigated land area (hectares) interacted with rain (millimeters)	77.40	717.45	0	30,963.9	16,743
Municipality characteristics					
Number of basic schools	284.97	332.44	0	1,762.00	22,763
Number of basic schools (per 10,000 inhabitants)	327.33	167.49	0	1,135.58	22,763
Number of indigenous schools	6.08	12.78	0	72.00	23313
Number of indigenous schools (per 10,000 inhabitants)	18.90	40.56	0	227.64	23,313
Number of cars	29,396.74	64,269.90	0	383,512.00	24,220
Number of cars (per 10,000 inhabitants)	7,626.40	7,811.72	0	41,882.23	24,220
Number of buses	371.10	841.11	0	5,355.00	24,220

Table 1: (continued)

	Mean	Std.Dev.	Min	Max	# Obs
Number of buses (per 10,000 inhabitants)	109.07	126.98	0	833.76	24,220
National variables					
Average hourly wage (2010 Mexican pesos)	35.97	3.34	29.61	41.44	33,873
Border crossing variables					
Average crime rate (murders per 10,000 inhabitants)					
... in crossing municipalities < 1000 km	11.5	8.8	1.9	83.7	12,166
... along border municipalities	14.3	2.5	9.9	18.4	17,554
... at the closest crossing point	8.7	6.6	0.0	38.2	17,554
... at the second closest crossing point	13.8	26.3	0.0	217.4	17,554
Average number of apprehensions					
... in crossing municipalities < 1000 km	139,460.2	77,498.3	44,895.0	616,346.0	21,018
... in crossing municipalities < 1000 km / border length	2,384.66	1,476.70	312.75	12,393.85	20,979
... along border municipalities	148,002.3	41,312.4	74,483.2	235,178.7	30,352
... at the closest crossing point	135,340.4	99,736.0	5,536.0	616,346.0	30,352
... at the second closest crossing point	138,760.5	99,130.0	5,536.0	616,346.0	30,352
Instruments					
Neighbors' average household head schooling (years)	4.75	1.52	0.75	11.67	25,725
Neighbors' average household average schooling (years)	6.07	1.61	1.81	12.17	25,243
Fraction of neighbors whose first born is male	0.50	0.13	0	1	25,761

Table 2: Within and between variation of migration decisions

		Mean	Std.Dev.	Min	Max	# Obs
Household has a migrant to the US (dummy)						
	Overall	0.1746	0.3796	0.0000	1.0000	25,761
	Within		0.2254	-0.7778	1.1269	
	Between		0.3095	0.0000	1.0000	
Household has a migrant within Mexico (dummy)						
	Overall	0.2000	0.4000	0.0000	1.0000	25,761
	Within		0.2477	-0.7523	1.1524	
	Between		0.3197	0.0000	1.0000	

Notes: "Within" variation is the variation in the migration variable across years for a given village. "Between" variation is the variation in the migration variable across villages for a given year.

Table 3: First-stage regressions

Dependent variable is fraction of neighbors with migration to/within:	(1)		(2)		(3)	
	US	Mexico	US	Mexico	US	Mexico
Instruments:						
Neighbors' average household head schooling	-0.0323***	-0.0350***	-0.0461***	-0.0203***	-0.0433***	-0.0177***
	(0.0023)	(0.0021)	(0.0026)	(0.0028)	(0.0027)	(0.0027)
Neighbors' average household average schooling	0.0212***	0.0495***	0.0404***	0.0211***	0.0370***	0.0255***
	(0.0025)	(0.0029)	(0.0040)	(0.0039)	(0.0041)	(0.0038)
Fraction of neighbors whose first born is male			-0.0276	-0.1314***	-0.0139	-0.1233***
			(0.0174)	(0.0170)	(0.0174)	(0.0170)
Controls	Y	Y	Y	Y	Y	Y
Irrigated land area x Precipitation	N	N	Y	Y	Y	Y
Time trend	Y	Y	Y	Y	N	N
Village fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	N	N	N	N	Y	Y
First-stage F-statistic	93.79	135.99	57.99	29.30	64.04	38.77
First-stage p-value (Pr>F)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Underidentification, Kleibergen p-value	0.0000		0.0000		0.0000	
Overidentification, J p-value			0.6553		0.8823	
# Observations	15,834	15,834	8,657	8,657	8,657	8,657

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table 4: Village-year means and standard deviations analysis

	# Village-Year Observations	**Mean (A.1)**	Std.Dev.	Min	Max	
	A: Village-year mean					
Household head schooling (years)	1,575	4.7973	1.5731	1.2000	10.2500	
Household average schooling (years)	1,575	6.1355	1.6512	2.2462	11.5026	
First born is male (dummy)	1,575	0.5110	0.1361	0.0000	0.9286	
	B: Village-year standard deviation					
	# Village-Year Observations	**Mean (B.1)**	Std.Dev.	Min	Max	
Household head schooling (years)	1,575	3.5127	1.0612	0.9003	7.3655	
Household average schooling (years)	1,575	2.6435	0.7219	0.6655	5.3905	
First born is male (dummy)	1,575	0.4971	0.0340	0.0000	0.5477	
	C: Ratio B/A					
	# Village-Year Observations	Mean	Std.Dev.	Min	Max	**Ratio B.1/A.1**
Household head schooling (years)	1,575	0.7750	0.2417	0.2715	1.7873	0.7322
Household average schooling (years)	1,575	0.4515	0.1387	0.0911	1.2056	0.4309
First born is male (dummy)	1,575	1.0501	0.2976	0.2878	2.1082	0.9728

Notes: This table examines the variation within villages in the values of household characteristics on which the instruments are based. For each village-year, we calculate the mean and standard deviation of the household characteristics on which the instruments are based across all households in that village-year. Panels A and B present summary statistics of the mean and standard deviation, respectively, over village-years of the household characteristics in each village-year. We then divide the mean standard deviation by the mean mean; the summary statistics for the ratio of the standard deviation over the mean over village-years are presented in Panel C.

Table 5: Correlations of instruments and own variables

	Correlation of own variable with corresponding instrument
Household head schooling (years)	0.2456
Household average schooling (years)	0.4167
First born is male (dummy)	0.0239

Notes: This table presents the correlations between each instrument measuring neighbors' characteristics and its respective own household characteristic.

Table 6: Falsification test of the first-stage regressions

	Dependent variable is probability of migration to/within:					
	(1)		(2)		(3)	
	US	Mexico	US	Mexico	US	Mexico
Instruments:						
Neighbors' average household head schooling	-0.0048	-0.0040	-0.0082	0.0080	-0.0074	0.0063
	(0.0057)	(0.0055)	(0.0077)	(0.0072)	(0.0079)	(0.0072)
Neighbors' average household average schooling	-0.0070	0.0058	-0.0005	-0.0174	-0.0044	-0.0150
	(0.0069)	(0.0071)	(0.0112)	(0.0109)	(0.0117)	(0.0111)
Fraction of neighbors whose first born is male			0.0936*	-0.1046**	0.0989**	-0.1098**
			(0.0493)	(0.0512)	(0.0497)	(0.0520)
Controls	Y	Y	Y	Y	Y	Y
Irrigated land area x Precipitation	N	N	Y	Y	Y	Y
Time trend	Y	Y	Y	Y	N	N
Village fixed effects	Y	Y	Y	Y	Y	Y
Year fixed effects	N	N	N	N	Y	Y
First-stage F-statistic	5.15	0.71	1.32	1.73	1.65	1.85
First-stage p-value (Pr>F)	0.0233	0.3987	0.2680	0.1771	0.1927	0.1579
# Observations	15,834	15,834	8,657	8,657	8,657	8,657

Notes: Robust standard errors in parentheses. This falsification test of the first-stage regression uses pseudo-endogenous variables as dependent variables rather than our actual endogenous variables. In particular, instead of the fraction of neighbors that engage in migration, we use the following pseudo-endogenous variable: the indicator variable y_{ikt} of whether household i has migration of type $k \in [USA, Mexico]$ at time t. Our falsification test of the first-stage regression enables us to examine whether the instruments are correlated with unobserved village factors that may affect migration decisions. If the exclusion restriction is satisfied and the instruments for the fraction of neighbors that engage in migration are not correlated with unobserved village factors that affect household migration decisions, then we would expect that our instruments – the neighbors' average household head schooling, the neighbors' average household average schooling, and the fraction of neighbors whose first born is male – should not be strong predictors of whether household i engages in migration. In other words, if the exclusion restriction is satisfied, the characteristics of neighbors j that we use as instruments should not be strong predictors of the migration decisions of household i. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table 7: IV results for migration to the US

Dependent variable is probability of migration to US			
	(1)	(2)	(3)
Fraction of neighbors with migration to US	0.5675*	0.5628*	0.5396*
	(0.2535)	(0.2837)	(0.2704)
Fraction of neighbors with migration within Mexico	-0.3854	-0.8771*	-0.8963*
	(0.2006)	(0.4273)	(0.3918)
Household characteristics			
Number of males in household	0.0283***	0.0244***	0.0243***
	(0.0017)	(0.0024)	(0.0024)
Household head age (years)	0.0011***	0.0008**	0.0008**
	(0.0002)	(0.0003)	(0.0003)
First born is male (dummy)	0.0031	-0.0059	-0.0060
	(0.0048)	(0.0064)	(0.0064)
Household head schooling (years)	-0.0045***	-0.0039**	-0.0039**
	(0.0009)	(0.0012)	(0.0012)
Household average schooling (years)	0.0031	-0.0027	-0.0026
	(0.0018)	(0.0024)	(0.0023)
Household maximum schooling (years)	0.0041***	0.0078***	0.0078***
	(0.0011)	(0.0016)	(0.0016)
Lag fraction of household members working in US	1.9275***	2.0493***	2.0472***
	(0.0437)	(0.0687)	(0.0679)
Lag fraction of household members working within Mexico	-0.0001	-0.0181	-0.0207
	(0.0234)	(0.0513)	(0.0487)
Irrigated land area (hectares) x Precipitation		-0.0000	-0.0000
		(0.0000)	(0.0000)
Municipality characteristics			
Number of basic schools	-0.0000	-0.0000	-0.0000
	(0.0001)	(0.0001)	(0.0001)
Number of indigenous schools	-0.0009	0.0011	0.0012
	(0.0017)	(0.0028)	(0.0027)
Number of cars	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
Number of buses	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
National variables			
Average hourly wage (pesos)	0.0009	0.0020	-0.0040
	(0.0010)	(0.0014)	(0.0210)
Border crossing variables			
Average crime rate (murders per 10,000 inhabitants)			
... in closest crossing border municipality	0.0002	0.0005	0.0005
	(0.0002)	(0.0005)	(0.0006)
... in second closest crossing border municipality	0.0001	-0.0001	-0.0001

Table 7: (continued)

	(0.0001)	(0.0001)	(0.0001)
... along border municipalities	0.0001	-0.0004	-0.0014
	(0.0003)	(0.0009)	(0.0038)
Average number of apprehensions			
... in closest crossing border municipality	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
... in second closest crossing border municipality	-0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
... along border municipalities	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
Year	-0.0023	0.0033	
	(0.0023)	(0.0025)	
Village fixed effects	Y	Y	Y
Year effects	N	N	Y
p-value (Pr>F)	0.0000	0.0000	0.0000
adjusted R-squared	0.4117	0.3532	0.3533
# Observations	15,834	8,657	8,657

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table 8: IV results for migration within Mexico

Dependent variable is probability of migration within Mexico			
	(1)	(2)	(3)
Fraction of neighbors with migration to US	-0.0077	-0.5208	-0.3770
	(0.2481)	(0.3071)	(0.2709)
Fraction of neighbors with migration within Mexico	0.1214	0.7795	0.5956
	(0.2033)	(0.4569)	(0.4002)
Household characteristics			
Number of males in household	0.0266***	0.0282***	0.0281***
	(0.0015)	(0.0024)	(0.0023)
Household head age (years)	0.0020***	0.0022***	0.0022***
	(0.0002)	(0.0003)	(0.0003)
First born is male (dummy)	-0.0020	-0.0221***	-0.0224***
	(0.0048)	(0.0066)	(0.0065)
Household head schooling (years)	-0.0030**	-0.0048***	-0.0047***
	(0.0010)	(0.0013)	(0.0013)
Household average schooling (years)	-0.0002	0.0003	-0.0000
	(0.0019)	(0.0026)	(0.0025)
Household maximum schooling (years)	0.0072***	0.0048**	0.0050**
	(0.0012)	(0.0018)	(0.0017)
Lag fraction of household members working in US	-0.0231	-0.0366	-0.0232
	(0.0236)	(0.0372)	(0.0346)
Lag fraction of household members working within Mexico	2.0681***	2.1954***	2.1781***
	(0.0432)	(0.0770)	(0.0739)
Irrigated land area (hectares) x Precipitation		-0.0000	-0.0000
		(0.0000)	(0.0000)
Municipality characteristics			
Number of basic schools	-0.0001	-0.0001	-0.0001
	(0.0001)	(0.0001)	(0.0001)
Number of indigenous schools	-0.0009	0.0056	0.0048
	(0.0018)	(0.0033)	(0.0031)
Number of cars	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
Number of buses	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
National variables			
Average hourly wage (pesos)	0.0006	0.0016	0.0052
	(0.0010)	(0.0014)	(0.0207)
Border crossing variables			
Average crime rate (murders per 10,000 inhabitants)			
... in closest crossing border municipality	0.0003	-0.0005	-0.0001
	(0.0002)	(0.0005)	(0.0006)
... in second closest crossing border municipality	0.0000	-0.0001	-0.0001

Table 8: (continued)

	(0.0001)	(0.0001)	(0.0001)
... along border municipalities	-0.0001	0.0005	-0.0007
	(0.0004)	(0.0009)	(0.0037)
Average number of apprehensions			
... in closest crossing border municipality	0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)
... in second closest crossing border municipality	-0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
... along border municipalities	0.0000	0.0000	0.0000
	(0.0000)	(0.0000)	(0.0000)
Year	-0.0010	-0.0002	
	(0.0022)	(0.0023)	
Village fixed effects	Y	Y	Y
Year effects	N	N	Y
p-value (Pr>F)	0.0000	0.0000	0.0000
adjusted R-squared	0.4881	0.4417	0.4600
# Observations	15,834	8,657	8,657

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

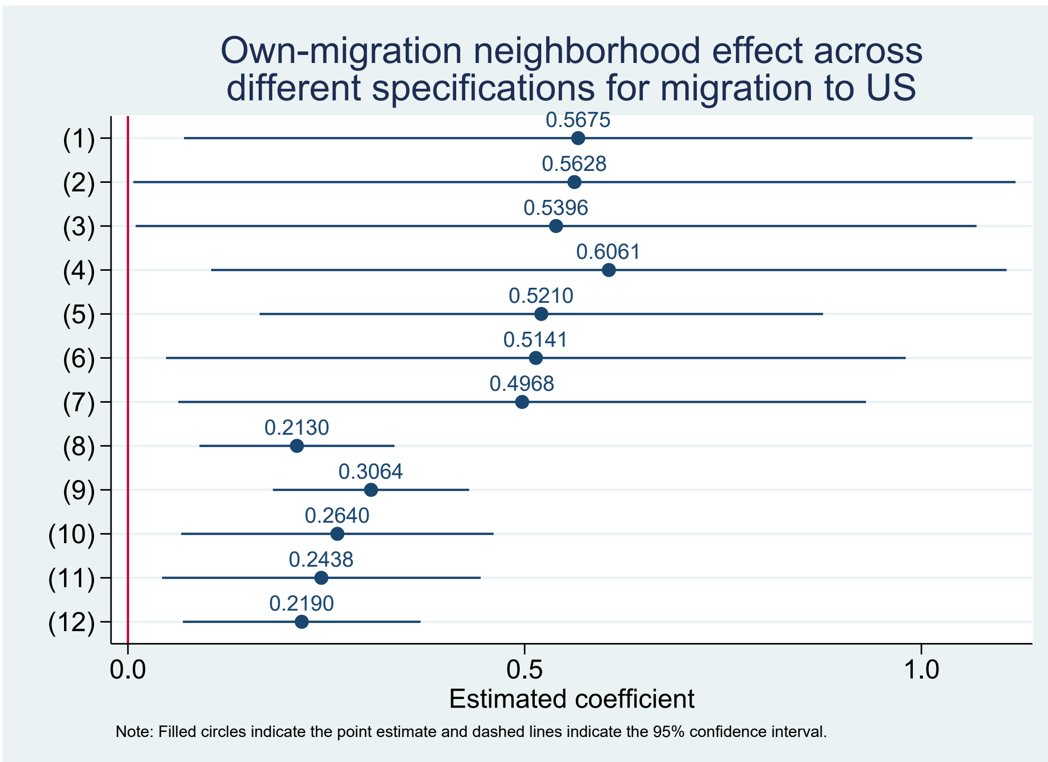

Figure 2: Summary of neighborhood effect results for migration to US

Note: Specifications (1)-(3) are from Table 7; Specifications (4)-(7) are from Table A.1 in the Appendix; and Specifications (8)-(12) are from Table A.2 in the Appendix.

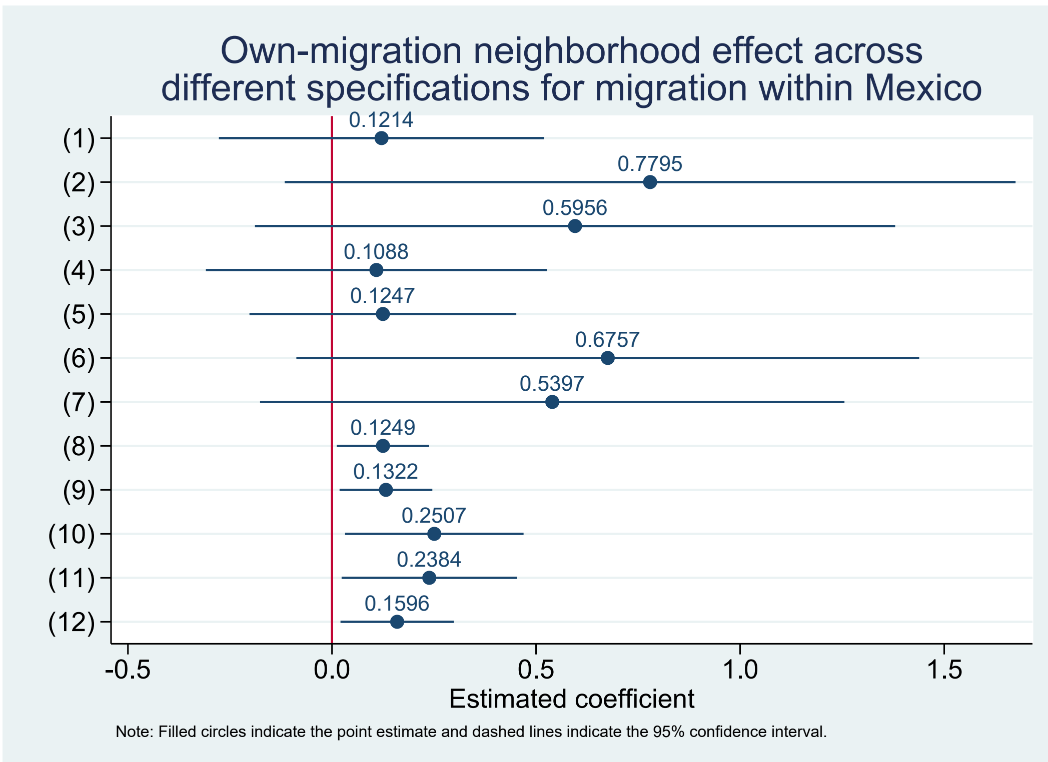

Figure 3: Summary of neighborhood effect results for migration within Mexico

Note: Specifications (1)-(3) are from Table 8; Specifications (4)-(7) are from Table A.3 in the Appendix; and Specifications (8)-(12) are from Table A.4 in the Appendix.

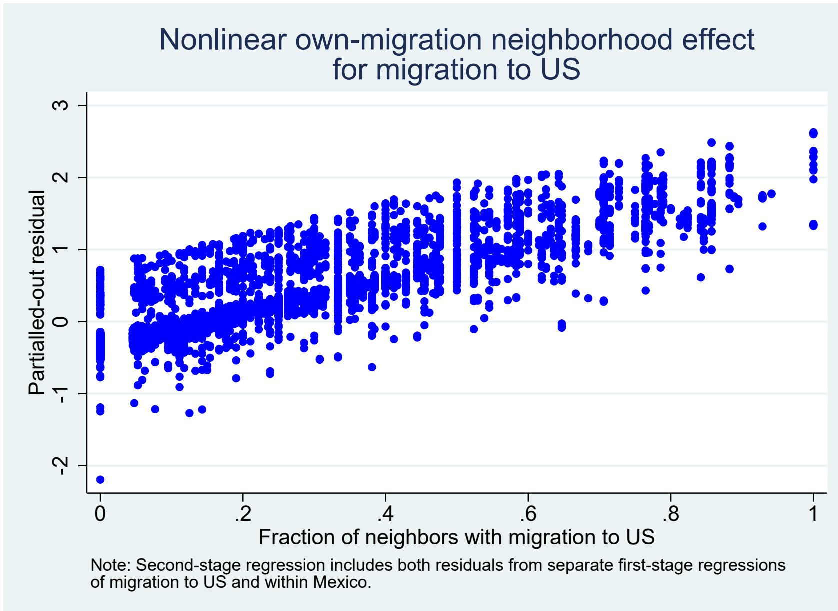

Figure 4: Nonlinearities in the neighborhood effects for migration to US

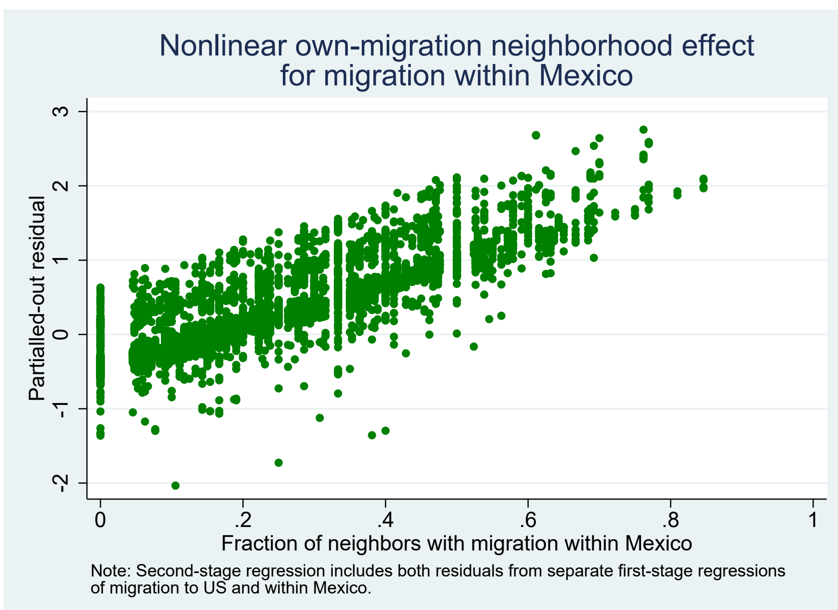

Figure 5: Nonlinearities in the neighborhood effects for migration within Mexico

Appendix

Table A.1: Robustness of IV results of migration to the US

Dependent variable is probability of migration to US				
	(4)	(5)	(6)	(7)
Fraction of neighbors with migration to US	0.6061*	0.5210**	0.5141*	0.4968*
	(0.2558)	(0.1812)	(0.2378)	(0.2211)
Fraction of neighbors with migration within Mexico	-0.4322*	-0.3427*	-0.8075*	-0.8729*
	(0.2113)	(0.1710)	(0.3722)	(0.3648)
Household characteristics				
Number of males in household	0.0283***	0.0310***	0.0244***	0.0243***
	(0.0017)	(0.0020)	(0.0024)	(0.0024)
Household head age (years)	0.0011***	0.0012***	0.0008**	0.0008**
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
First born is male (dummy)	0.0031	0.0028	-0.0059	-0.0060
	(0.0048)	(0.0055)	(0.0063)	(0.0063)
Household head schooling (years)	-0.0045***	-0.0033***	-0.0039**	-0.0038**
	(0.0009)	(0.0010)	(0.0012)	(0.0012)
Household average schooling (years)	0.0031	0.0022	-0.0026	-0.0027
	(0.0018)	(0.0020)	(0.0023)	(0.0023)
Household maximum schooling (years)	0.0041***	0.0046***	0.0078***	0.0078***
	(0.0011)	(0.0013)	(0.0016)	(0.0016)
Lag fraction of household members working in US	1.9309***	1.8999***	2.0455***	2.0438***
	(0.0441)	(0.0455)	(0.0675)	(0.0669)
Lag fraction of household members working within Mexico	-0.0031	0.0380	-0.0099	-0.0174
	(0.0239)	(0.0245)	(0.0459)	(0.0460)
Irrigated land area (hectares) x Precipitation			-0.0000	-0.0000
			(0.0000)	(0.0000)
Municipality characteristics				
Number of basic schools		-0.0000	-0.0001	-0.0000
		(0.0001)	(0.0001)	(0.0001)
Number of basic schools (per 10,000 inhabitants)	0.0001			
	(0.0001)			

Table A.1: (continued)

Number of indigenous schools		0.0009	0.0009	0.0009
		(0.0022)	(0.0027)	(0.0026)
Number of indigenous schools (per 10,000 inhabitants)	-0.0009			
	(0.0009)			
Number of cars		0.0000	0.0000	0.0000
		(0.0000)	(0.0000)	(0.0000)
Number of cars (per 10,000 inhabitants)	0.0000			
	(0.0000)			
Number of buses		-0.0000	-0.0000	-0.0000
		(0.0000)	(0.0000)	(0.0000)
Number of buses (per 10,000 inhabitants)	-0.0000			
	(0.0000)			
National variables				
Average hourly wage (pesos)	-0.0006	0.0007		
	(0.0014)	(0.0011)		
Border crossing variables				
Average crime rate (murders per 10,000 inhabitants)				
... at closest border crossing point	0.0003			
	(0.0002)			
... at second closest border crossing point	0.0001			
	(0.0001)			
... along border municipalities	-0.0001	-0.0000		
	(0.0003)	(0.0003)		
... in crossing municipalities $<$ 1000 km		0.0003		
		(0.0003)		
Average number of apprehensions				
... at closest border crossing point	0.0000			
	(0.0000)			
... at second closest border crossing point	-0.0000			
	(0.0000)			

Table A.1: (continued)

... along border municipalities		0.0000		
		(0.0000)		
... along border municipalities (normalized)	-0.0000			
	(0.0000)			
... in crossing municipalities < 1,000 km		0.0000		
		(0.0000)		
Year	-0.0031	-0.0023	0.0021	
	(0.0028)	(0.0021)	(0.0021)	
Village fixed effects	Y	Y	Y	Y
Year effects	N	N	N	Y
p-value (Pr>F)	0.0000	0.0000	0.0000	0.0000
adjusted R-squared	0.4074	0.4269	0.3605	0.3570
# Observations	15,834	11,871	8,657	8,657

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table A.2: IV results of migration to the US using the augmented Arellano-Bond estimator

Dependent variable is probability of migration to US					
	(8)	(9)	(10)	(11)	(12)
Fraction of neighbors with migration to US	0.2130***	0.3064***	0.2640**	0.2438*	0.2190**
	(0.0627)	(0.0631)	(0.1004)	(0.1025)	(0.0764)
Fraction of neighbors with migration within Mexico	0.0515	0.0258	-0.0984	-0.0661	0.0159
	(0.0511)	(0.0503)	(0.1048)	(0.1096)	(0.0749)
Household characteristics					
Number of males in household	0.0192***	0.0194***	0.0140**	0.0140**	0.0140***
	(0.0031)	(0.0031)	(0.0052)	(0.0052)	(0.0033)
Household head age (years)	0.0011**	0.0012**	0.0008	0.0008	0.0004
	(0.0004)	(0.0004)	(0.0006)	(0.0006)	(0.0004)
First born is male (dummy)	0.0121	0.0127	0.0112	0.0114	-0.0004
	(0.0080)	(0.0082)	(0.0135)	(0.0134)	(0.0083)
Household head schooling (years)	-0.0046**	-0.0048**	-0.0073**	-0.0072**	-0.0054**
	(0.0015)	(0.0016)	(0.0025)	(0.0025)	(0.0017)
Household average schooling (years)	0.0048	0.0046	0.0045	0.0045	0.0035
	(0.0029)	(0.0030)	(0.0045)	(0.0045)	(0.0031)
Household maximum schooling (years)	0.0042*	0.0042*	0.0056	0.0055	0.0038
	(0.0019)	(0.0019)	(0.0031)	(0.0031)	(0.0023)
Lag of working in US (dummy)	0.4993***	0.4943***	0.4031***	0.4051***	0.5009***
	(0.0364)	(0.0362)	(0.0672)	(0.0671)	(0.0485)
Lag of working within Mexico (dummy)	-0.0080	-0.0111	0.0139	0.0159	0.0211
	(0.0177)	(0.0179)	(0.0349)	(0.0350)	(0.0215)
Irrigated land area (hectares) x Precipitation			-0.0000	-0.0000	0.0000
			(0.0000)	(0.0000)	(0.0000)
Municipality characteristics					
Number of basic schools	-0.0000	-0.0000	-0.0001	-0.0001	
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	
Number of indigenous schools	0.0006	0.0002	0.0055	0.0057	
	(0.0013)	(0.0013)	(0.0036)	(0.0036)	

Table A.2: (continued)

Number of cars	0.0000	-0.0000	-0.0000	-0.0000	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
Number of buses	0.0000	0.0000	0.0000	0.0000	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
National variables					
Average hourly wage (pesos)	-0.0031**	0.0008	0.0010	-0.0005	-0.0031***
	(0.0010)	(0.0007)	(0.0011)	(0.0021)	(0.0009)
Border crossing variables					
Average crime rate (murders per 10,000 inhabitants)					
... at closest border crossing point	0.0001	0.0001			-0.0001
	(0.0001)	(0.0001)			(0.0002)
... at second closest border crossing point	-0.0000	0.0000			-0.0001
	(0.0001)	(0.0001)			(0.0001)
... along border municipalities	-0.0007***	-0.0002	0.0006	0.0018	0.0007
	(0.0002)	(0.0001)	(0.0007)	(0.0016)	(0.0012)
... in crossing municipalities < 1000 km			-0.0007	-0.0008	
			(0.0004)	(0.0005)	
Average number of apprehensions	0.0000	0.0000			0.0000
... at closest border crossing point	(0.0000)	(0.0000)			(0.0000)
	-0.0000	-0.0000			-0.0000
... at second closest border crossing point	(0.0000)	(0.0000)			(0.0000)
	-0.0000	0.0000	0.0000	0.0000	-0.0000
... along border municipalities	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
			-0.0000	-0.0000	
... in crossing municipalities < 1,000 km			(0.0000)	(0.0000)	
Year		-0.0018	0.0005		
		(0.0012)	(0.0020)		
Village fixed effects	Y	Y	Y	Y	Y

Table A.2: (continued)

Year effects	Y	N	N	Y	Y
p-value (Pr>F)	0.00	0.00	0.00	0.00	0.00
# Observations	15,834	15,834	6,323	6,323	12,617

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table A.3: Robustness of IV results of migration within Mexico

Dependent variable is probability of migration within Mexico				
	(4)	(5)	(6)	(7)
Fraction of neighbors with migration to US	-0.0019	0.0014	-0.4210	-0.3078
	(0.2489)	(0.1675)	(0.2451)	(0.2109)
Fraction of neighbors with migration within Mexico	0.1088	0.1247	0.6757	0.5397
	(0.2132)	(0.1668)	(0.3893)	(0.3653)
Household characteristics				
Number of males in household	0.0266***	0.0250***	0.0281***	0.0280***
	(0.0015)	(0.0017)	(0.0023)	(0.0023)
Household head age (years)	0.0020***	0.0016***	0.0022***	0.0022***
	(0.0002)	(0.0002)	(0.0003)	(0.0003)
First born is male (dummy)	-0.0020	0.0019	-0.0224***	-0.0227***
	(0.0048)	(0.0052)	(0.0065)	(0.0065)
Household head schooling (years)	-0.0030**	-0.0025*	-0.0048***	-0.0047***
	(0.0010)	(0.0010)	(0.0013)	(0.0013)
Household average schooling (years)	-0.0002	0.0007	0.0002	-0.0001
	(0.0019)	(0.0020)	(0.0026)	(0.0025)
Household maximum schooling (years)	0.0071***	0.0066***	0.0049**	0.0051**
	(0.0012)	(0.0013)	(0.0017)	(0.0017)
Lag fraction of household members working in US	2.0677***	2.0501***	-0.0287	-0.0178
	(0.0433)	(0.0509)	(0.0338)	(0.0316)
Lag fraction of household members working within Mexico	0.0004	0.0006	2.1846***	2.1723***
	(0.0014)	(0.0010)	(0.0723)	(0.0715)
Irrigated land area (hectares) x Precipitation	-0.0224	-0.0099	-0.0000	-0.0000
	(0.0241)	(0.0204)	(0.0000)	(0.0000)
Municipality characteristics				
Number of basic schools		-0.0001	-0.0001	-0.0001
		(0.0001)	(0.0001)	(0.0001)
Number of basic schools (per 10,000 inhabitants)	-0.0000			
	(0.0001)			

Table A.3: (continued)

Number of indigenous schools		0.0005	0.0047	0.0042
		(0.0021)	(0.0031)	(0.0029)
Number of indigenous schools (per 10,000 inhabitants)	-0.0001			
	(0.0009)			
Number of cars		-0.0000	-0.0000	-0.0000
		(0.0000)	(0.0000)	(0.0000)
Number of cars (per 10,000 inhabitants)	-0.0000			
	(0.0000)			
Number of buses		0.0000	0.0000	0.0000
		(0.0000)	(0.0000)	(0.0000)
Number of buses (per 10,000 inhabitants)	-0.0000			
	(0.0000)			
Border crossing variables				
Average crime rate (murders per 10,000 inhabitants)				
... at closest border crossing point	0.0003			
	(0.0002)			
... at second closest border crossing point	0.0000			
	(0.0001)			
... along border municipalities	-0.0001	-0.0002		
	(0.0003)	(0.0003)		
... in crossing municipalities $<$ 1000 km		0.0004		
		(0.0003)		
Average number of apprehensions				
... at closest border crossing point	0.0000			
	(0.0000)			
... at second closest border crossing point	-0.0000			
	(0.0000)			
... along border municipalities		0.0000		
		(0.0000)		
... along border municipalities (normalized)	-0.0000			
	(0.0000)			

Table A.3: (continued)

... in crossing municipalities < 1,000 km		0.0000		
		(0.0000)		
Year	-0.0011	-0.0011	0.0003	
	(0.0028)	(0.0018)	(0.0019)	
Village fixed effects	Y	Y	Y	Y
Year effects	N	N	N	Y
p-value (Pr>F)	0.0000	0.0000	0.0000	0.0000
adjusted R-squared	0.4884	0.4910	0.4533	0.4657
# Observations	15,834	11,871	8,657	8,657

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.

Table A.4: IV results of migration within Mexico using the augmented Arellano-Bond estimator

Dependent variable is probability of migration within Mexico					
	(8)	(9)	(10)	(11)	(12)
Fraction of neighbors with migration to US	0.0280	0.0590	-0.1122	-0.0988	-0.0192
	(0.0486)	(0.0490)	(0.0820)	(0.0823)	(0.0585)
Fraction of neighbors with migration within Mexico	0.1249*	0.1322*	0.2507*	0.2384*	0.1596*
	(0.0578)	(0.0580)	(0.1116)	(0.1096)	(0.0708)
Household characteristics					
Number of males in household	0.0142***	0.0142***	0.0166**	0.0165**	0.0121***
	(0.0028)	(0.0028)	(0.0053)	(0.0053)	(0.0035)
Household head age (years)	0.0018***	0.0018***	0.0023***	0.0023***	0.0019***
	(0.0004)	(0.0004)	(0.0007)	(0.0007)	(0.0005)
First born is male (dummy)	0.0065	0.0069	-0.0009	-0.0009	-0.0083
	(0.0077)	(0.0077)	(0.0128)	(0.0127)	(0.0092)
Household head schooling (years)	-0.0049**	-0.0050**	-0.0057*	-0.0056*	-0.0061**
	(0.0017)	(0.0017)	(0.0024)	(0.0024)	(0.0019)
Household average schooling (years)	0.0025	0.0024	0.0065	0.0064	0.0062
	(0.0030)	(0.0030)	(0.0047)	(0.0047)	(0.0035)
Household maximum schooling (years)	0.0059**	0.0059**	0.0021	0.0021	0.0025
	(0.0020)	(0.0020)	(0.0034)	(0.0033)	(0.0024)
Lag of working in US (dummy)	-0.0307	-0.0313	-0.0025	-0.0016	-0.0026
	(0.0160)	(0.0161)	(0.0213)	(0.0213)	(0.0161)
Lag of working within Mexico (dummy)	0.5527***	0.5538***	0.4484***	0.4524***	0.5145***
	(0.0366)	(0.0367)	(0.0880)	(0.0873)	(0.0575)
Irrigated land area (hectares) x Precipitation			0.0000	0.0000	-0.0000
			(0.0000)	(0.0000)	(0.0000)
Municipality characteristics					
Number of basic schools	-0.0001*	-0.0001*	-0.0001	-0.0001	
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	
Number of indigenous schools	-0.0008	-0.0009	0.0004	0.0003	
	(0.0016)	(0.0016)	(0.0034)	(0.0035)	

Table A.4: (continued)

Number of cars	0.0000	0.0000	-0.0000	-0.0000	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
Number of buses	0.0000	0.0000	-0.0000	-0.0000	
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	
National variables					
Average hourly wage (pesos)	0.0005	0.0008	0.0010	0.0006	0.0025
	(0.0013)	(0.0006)	(0.0010)	(0.0021)	(0.0014)
Border crossing variables					
Average crime rate (murders per 10,000 inhabitants)					
... at closest border crossing point	0.0002	0.0002			0.0002
	(0.0001)	(0.0001)			(0.0002)
... at second closest border crossing point	-0.0000	-0.0000			-0.0001
	(0.0000)	(0.0000)			(0.0001)
... along border municipalities	-0.0000	0.0003	0.0002	-0.0015	-0.0020
	(0.0002)	(0.0002)	(0.0007)	(0.0014)	(0.0011)
... in crossing municipalities < 1000 km			-0.0000	0.0001	
			(0.0003)	(0.0003)	
Average number of apprehensions	0.0000*	0.0000			0.0000
... at closest border crossing point	(0.0000)	(0.0000)			(0.0000)
	-0.0000	-0.0000			0.0000
... at second closest border crossing point	(0.0000)	(0.0000)			(0.0000)
	0.0000	0.0000	0.0000	-0.0000	-0.0000
... along border municipalities	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
			0.0000	0.0000	
... in crossing municipalities < 1,000 km			(0.0000)	(0.0000)	
Year		-0.0018	-0.0011	-0.0011	-0.0014
		(0.0011)	(0.0017)	(0.0017)	(0.0011)
Village fixed effects	Y	Y	Y	Y	Y

Table A.4: (continued)

Year effects	Y	N	N	N	N
p-value (Pr>F)	0.0000	0.0000	0.0000	0.0000	0.0000
# Observations	15,834	15,834	6,323	6,323	12,617

Notes: Robust standard errors in parentheses. Significance codes: * p<0.05, ** p<0.01, *** p<0.001.