On the Push and Pull Forces of Migrant Border Crossing: The Role of Networks*

Nancy H. Chau　　Filiz Garip†

This version: January 2023

Abstract: Can the locational determinants of migrant border crossings be traced back to the push-pull origins of migration? This paper studies crossing behavior along the Mexican-US border where migrants confronted competing apprehension and environmental hazards, and where US family network ties presented migrant-level variations in pull-push forces. We demonstrate theoretically and verify empirically the uneven hazardous border crossing response to a major border enforcement reform that strategically leveraged terrain hazards for migrant deterrence. Our findings shed light on family reunification-based immigration quotas as a means to mitigate against border deaths, particularly when border enforcement reforms limit safe undocumented border crossing options.

JEL Classification: J64, J31, O15.
Keywords: Border Walls; Trade Walls; and Migrant Border Crossing.

*Funding from the Institute of Social Sciences and the Atkinson Center for a Sustainable Future at Cornell University is gratefully acknowledged. We thank Simone Bertoli, Patricia Fernández-Kelly, Rachel Gisselquist, Douglas Massey, Omar McDoom, Kunal Sen, Maheshwor Shrestha, and seminar participants at Princeton, Cornell, UNU-WIDER, NEUDC, the IIFT Kolkata Development Economics Workshop, the IZA-World Bank Jobs and Development Conference, the European Society of Population Economists Annual Meeting, the University of Texas at San Antonio, and the AfD-World Bank-LISER Migration and Development Annual Conference for insightful comments and encouragement. Tanya Bakshi, Mario Molina, and Julia Zhu provided excellent research assistance. This paper is completed while Chau is visiting scholar at the UNU-WIDER. The hospitality of the institute is gratefully acknowledged. All errors are our own.

†Chau: Charles H. Dyson School of Applied Economics and Management, Cornell University, 201A Warren Hall, Ithaca, NY 14853 & Institute for the Study of Labor (IZA) (email: hyc3@cornell.edu); Garip: Department of Sociology, Princeton University, Wallace Hall, Princeton, NJ 08544 (email: fgarip@princeton.edu).

1 Introduction

For undocumented migrants, the choice of border crossing location is a matter of survival. Tragically, over 49,400 deaths and disappearances of migrants en route to their destinations have been recorded globally since 2014.[1] Conditional on cost, what markers indicate individuals / communities of migrants that seek out loosely enforced border crossings? What features characterize migrant population at risk of making perilous border crossing journeys? Answers to these questions are of first order importance if we are to gain better understanding about the distributional consequences of border enforcement efforts at origin communities beyond migrant displacement along the border, as well as the sources of any disproportionate harm potentially inflicted on individuals compelled to brave hazardous conditions to cross the border.

There is general consensus that large scale migrant displacements along the Mexican-US border have been a response to a policy shift towards a "prevention through deterrence" stance over enforcement operations in the country's interior since the mid 1990's (Congressional Research Service, 1997; Cornelius, 2001). The objectives of the policy are to deter undocumented entry by disrupting traditional crossing traffic and smuggling routes through large increases in border enforcement spending, with the expectation that any remaining entry attempts will have to take place over hostile terrains (United States Border Patrol, 1994). Studies have leveraged a broad range of enforcement intensity proxies of border operations to gauge the displacement effect.[2] Among these, Gathmann (2008) presents some of the earliest evidence pointing to enforcement-induced migrant displacement along the US-Mexican border. The study estimates a likelihood function predicting the probability that a repeat migrant will switch crossing locations depending on enforcement intensity at the previous crossing, after controlling for individual characteristics. Feigenberg (2020) shows that individual-level crossing intentions at the border municipality level are associated with border fencing construction at and adjacent to that location, controlling for

[1]The International Organization of Migration's Missing Migrant Project collects data from a variety of sources including staff reports by United Nations High Commissioner for Refugees based on survivor surveys (Mediterranean region), medical examiners and sheriff offices (US), as well as media and year-end government reports.

[2]These include border patrol budgets (Gathmann, 2008), personnel (Gathmann, 2008; Lessem, 2012), and fencing (Allen et al., 2018; Feigenberg, 2020), or a combination of both Cornelius (2001).

lagged border municipality characteristics. Allen et al. (2018) frames the impact of border wall construction in a general equilibrium setting of trade and migration. The study also shows crossing probabilities to be systematically associated with border fence expansion after controlling for distance, and a full set of fixed effects.

The collective message of this growing literature about the migrant displacement effects of border enforcement in the aggregate is unambiguous. But what remain unexplored are the mechanisms driving the incidence of the displacement effect at the migrant- and community-levels, as well as the triggers that result in a migrant's resolve to risk undertaking hazardous border crossing journeys. In this context, this paper furnishes a first study on the locational choice of border crossings via the lens of the push and pull forces of migration. We write a simple theoretical model of a migrant's choice between multiple crossing locations, each distinctive in terms of the likelihood of successful crossing, the likelihood of accidents / death while crossing the border, and the cost of crossing. The objective is to clarify any link between traditional pull and push forces of migration, and the border location choice calculus at the migrant level.

The model's findings are simple and intuitive: pull factors of migration attract border crossers to favor loosely enforced locations with higher probability of crossing success to take advantage of the gains from reaching the destination, while push factors of migration additionally encourage migrants to discount the hazards at the border as their fallback (origin) option worsens. This conceptual model guides our empirical work in three ways. First, we seek proxies for the pull-push mediators of migration at the individual level to understand crossing location choices. Second, to help rule out confounders, we choose a study period during which border enforcement regime changes are well known to have altered the relative ranking of the enforcement intensities across crossing locations. Finally, our empirical specification is designed to account for the time varying interaction effect between changes in border enforcement and changes in the push and pull forces, controlling for different types of coyote price discrimination conduct.

Since any discussion about the determinants of migration is incomplete without its social context (Massey et al., 1993; DiMaggio and Garip, 2012; Munshi, 2020)), we leverage the ubiquitous role of networks as a push-pull force of migration. Munshi (2020) dissects a voluminous literature on the role of social ties to the destination community as a pull factor of migration by facilitating social learning and information sharing (Borjas, 1992; Chau, 1997; Munshi, 2003). Networks strengthen desire to relocate by assisting in destination job search (Basu, Chau, and Lin, 2021), providing material support for new migrants (Munshi, 2014), and offering access to credit while away from home (Massey, 1988; Orrenius, 1999; Orrenius and Zavodny, 2005; Dolfin and Genicot, 2010). DiMaggio and Garip (2012) synthesizes a rich interdisciplinary literature and expands on this theme to additionally include the normative influence of networks.[3] To wit, network ties can reinforce norms of behavior through social approval and / or sanctions. A culture of migration (Piore, 1979) adds to the allure of migration abroad, while the stigma associated with not moving and the ensuing social sanctions against non-conformers can effectively serve as push forces of migration (Kandel and Massey, 2002; Garip and Asad, 2016).

We draw lessons from a novel data set, the Mexican Migration Project (MMP), to retrieve border sector choices at the individual level stretching over a period of 26 years (1980 - 2005) when undocumented migration across the Mexico-US border was at its peak. The MMP is a repeated cross-sectional survey conducted annually since 1982. Importantly for our study, the MMP provides detail crossing history records along the Mexico-US border, namely when, where, and the cost incurred along with individual level characteristics such as family ties in the U.S.

To assert family networks as the cause of border sector choices, a key challenge is the existence of confounders. For example, families with a history of migrants may simply be located closer to easy border crossings. Guided by theory, our identification strategy involves juxtaposing individual-level variations in network ties with major shifts in the nature of competing risks of enforcement and crossing hazards confronting migrants at the border. The 1,954-mile long US-Mexican border is arguably one of the world's most dangerous land borders in terms of total number of reported migrant fatalities (IOM 2019). The hazards of border crossing are not uniform, with traditional crossing locations through San Diego and El Paso interspersed with notoriously hazardous landscape and climate conditions surrounding the Tucson border sector.[4] Complementing these selective and persistent en-

[3] Notable works include Mines and de Janvry (1982); Moretti (1999); Winters et al. (2001); Munshi and Rosenzweig (2016); Garip (2008), to name a few.

[4] The causes of migrant deaths differ case-by-case, but deaths due to hypothermia, drowning, and dehydration along the border are common. Other reported man-made causes of migrant death such as

vironmental crossing hazards along the border, our study period features individual-level crossing observations before, during, and after the border enforcement policy shifts towards the prevention through deterrence stance. The policy resulted in the implementation of a series of border enforcement operations along popular crossing routes,[5] which greatly increased spending in fencing, patrols and border enforcement technologies. The result is a wholesale shift in the distribution of border apprehensions along the border from a San Diego/ El Paso majority to the emergence of Tucson as an important crossing point, consistent with the literature on migrant displacement subsequent to border enforcement (Cornelius, 2001).

We leverage this exogenous and selective change in crossing success likelihoods to test the role of networks on border sector choices. Applying the insights from our model, the prevention through deterrence policy is a recipe for a reversal in the pattern of self-selection among migrants with network ties to the US. To wit, individuals with family ties in the US – potentially both a pull and a push force of migration – should be negatively selected in the Tucson sector when enforcements in safe locations such as San Diego and El Paso were still relatively low, and positively selected in the Tucson sector after the border sector reforms, controlling for cost, year and regional fixed effects.

Informed by these potential interactive effects on crossing choice that differ by border sectors, our empirical methodology adopts the alternative specific conditional logit model (Mcfadden, 1974) – a choice model in which determinants at the migrant level (e.g. individual network ties to the US, year of migration), community level (e.g. norms of behavior, concurrent shocks), and border sector specific features (e.g. distance to sector, community level crossing history) are jointly taken into account as potential triggers of location choice among US-Mexico border sectors.[6]

We draw three broad findings from the evidence. First, our baseline specification focuses on the time-specific impact of U.S. family ties to shed light on the nuanced role

homicide, vehicular accidents, and border wall related injuries, while present, are strictly a minority (IOM 2019).

[5]These include Operation Gatekeeper (California 1994), Operation Hold the Line (Texas 1993), Operation Safeguard (Arizona 1994) and Operation Rio Grande (Texas 1997).

[6]There are nine US-Mexico border sectors: San Diego (CA), El Centro (CA), Yuma (NV), Tucson (AZ), El Pas (NM), Big Bend (TX), Del Rio (TX), Laredo (TX) and Rio Grande (TX). Yuma and Big Bend have extremely small number of reported crossings. Our study thus work with the remaining seven border sectors, covering 99% of all reported border crossings in our data set.

of migration push-pull forces on crossing location choices.[7] Consistent with our model's prediction, we find a reversal in the pattern of self-selection unfolding over time – migrants with US family ties were negatively selected at the Tucson border sector prior to the prevention through deterrence strategy starting in 1995. This pattern then shifted and reversed over time to exhibit positive selection later on, and particularly so after 2000. Echoing studies on mediators that strengthen network impacts on migration motives (e.g., Curran et al., 2005; McKenzie and Rapoport, 2007), we find that this pattern of self-selection to be especially salient among low skilled, and / or blue collar workers in agriculture and manufacturing, and to work through direct family networks rather than community migrant networks.

Building on this baseline, we evaluate alternative pathways through which networks can function beyond traditional push and pull forces, different mechanisms that may have been derivatives of family network ties, as well as other measures of push and pull forces.[8] Notably, community-level migrant networks can introduce a commons effect on coyote price (DiMaggio and Garip, 2012; Garip and Asad, 2016). For example, the cost of finding and hiring a smuggler to cross the border through a particular border sector can be lower when smugglers can count on a reliable history of migrants through the community. Smugglers may also charge different prices depending on the average characteristics of the typical migrant, such as age, when border sectors present different risks to individuals across age groups. The practice of second and third degree price discrimination by smugglers depending on community-level network characteristics (e.g. community-level lagged shares of border crossing locational choices, and average age of border crossers) thus constitute alternative mechanisms that can direct border crossing traffic. We include these alternative network measures into our crossing choice regressions, and find that the pattern of self-selection reversal at the Tucson border sector among individuals with family connections remains robust.

Beyond network externalities, we additionally conduct checks to examine how our

[7] The baseline specification additionally includes cost controls (e.g. distance to (destination) and from (origin) the border sector in question, fencing and time period varying border sector fixed effects), time fixed effects to capture common policy trends, and regional fixed effects.

[8] DiMaggio and Garip (2012) presents a tripartite typology of the impact of networks on migration incentives, namely through social learning; normative influence, and network externality. The first two are what we have included already under the banners of push and pull forces.

findings may have been susceptible to the possibility that (i) ties to the US convey better information about border conditions, (ii) the share of migrants with US family ties has changed as border enforcement intensified, and (iii) other migrant characteristics (e.g. age, education) may also have changed in response to enforcement regime change.

This study complements a rich and vibrant migration literature that has to date given limited attention to how migrants make decisions about the border crossing journey. The canonical theoretical treatment of migration has a longstanding tradition dating from the expected utility framework of Sjaastad (1962). Recent studies present important supporting evidence and extensions by assessing the role of migration distance (Friebel et al., 2018), spatial general equilibrium features such as commuting costs (Monte et al., 2018; Bryan and Morten, 2019; Caliendo et al., 2019; Tombe and Zhu, 2019), credit constraints (Mckenzie and Rapoport, 2007; Munshi and Rosenzweig, 2016), and social networks to explain group-based heterogeneity in migration destination choices (Chau, 1997; Munshi, 2003; DiMaggio and Garip, 2012).The implicit assumption in these analyses is that conditional on the decision to migrate and the choice of destination, migrants select the least cost route of migration. This paper makes the point that the best route need not be the same for all migrants from the same origin, or even the same migrant over time, when the border is porous and enforcement is selective and changing. This suggests that modeling the impact of border enforcement on migration incentives should pay close attention to individual incentives as a starting point.

Our findings also provide new insights on the design of border enforcement policy. Notable studies include work on border enforcement as means of migrant deterrence, and / or migrant displacement Ethier (1986); Chau (2001, 2003); Epstein and Weiss (2011); Facchini and Testa (2021).[9] Specifically, we offer two related perspectives. First, border enforcement strategies have distributional consequences. In the context of the prevention through deterrence policy along the Mexcian-US border, individuals with strong desire

[9]Empirical studies in this area is understandably scant, and the handful of studies available unanimously feature the US-Mexican border as an excellent illustration of how border enforcement deters migrants and how a combination of geo-climatic diversity and enforcement heterogeneity can jointly contribute to wholesale displacements in migrant crossing locations. Hanson and Spilimbergo (1999) demonstrate that strengthening border enforcement at the US-Mexican border, as measured by overall border patrol manpower, is an effective migrant deterrent device. Other studies include (Cornelius, 2001; Gathmann, 2008; Angelucci, 2012; Lessem, 2012; Allen et al., 2018; Feigenberg, 2020; Bazzi et al., 2021).

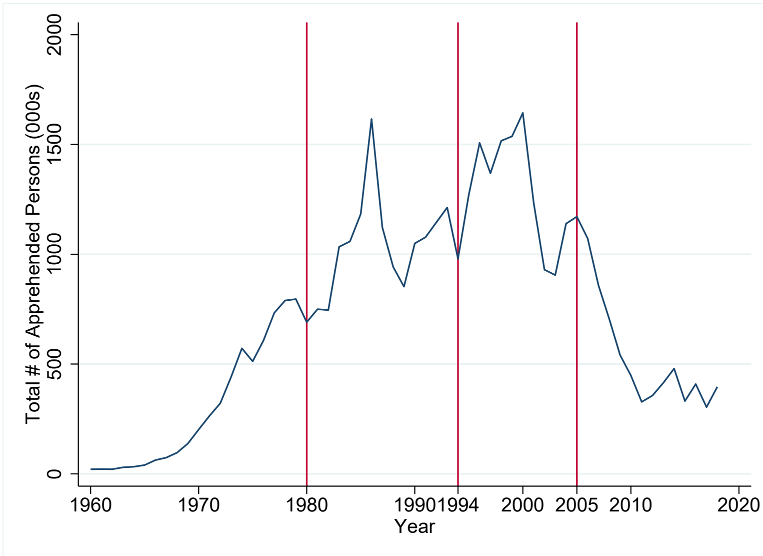
to migrate were diverted to take on perilous border crossing journeys, risking exposure to hazardous climate conditions and environmental terrains. In this paper, we find that this policy disproportionately affected low skilled, blue-collar individuals who were well connected to the US through family network ties. Second, our findings also provide novel humanitarian justification for border enforcement policies to go hand-in-and with family-based immigration policy – arguably the bedrock of U.S. immigration policy since the Immigration and Nationality Act in 1965. Giving preference for family reunification in a priority-weighted system of immigration quotas not only makes sense for immigrant families, it also helps steer family members of existing migrants away from border hazards, particularly since these individuals with direct family ties are most at risk of choosing to confront dangerous border crossing when alternative and safer routes are made inaccessible through border enforcement reforms.

The rest of this paper is organized as follows. In Section 2, we provide an overview of the institutional and policy landscapes guiding the flow of Mexican-US migration. In Section 3, we outline a simple theoretical model of migrants' choice of border sector in the presence of heterogeneous rewards and risks profiles among the border sector choices, as well as how this analysis motivated our empirical specification. In Section 4, we provide an overview of the data we assembled for this study, and Section 5 presents our main empirical specifications and results, in addition to a series of robustness checks. Section 6 concludes.

2 Mexican-US Border Crossings: Migrant Networks and Border Enforcement

There are nine border patrol sectors (San Diego (CA), El Centro (CA), Yuma (NV), Tucson (AZ), El Paso (NM), Big Bend (TX), Del Rio (TX), Laredo (TX) and Rio Grande (TX)) along the Southwestern border of the US. Figure 1 plots the total number of apprehended individual undocumented immigrants over time as reported by the US Customs and Border Protection (CBP). These numbers were low and relatively stable during the 1960s at around a total of 27,000 apprehensions per year. The reason for the increase is the end of the so-called Bracero Program, the labor agreement between Mexico and the United States that gave temporary work visas to agricultural workers. The end of the program stopped the

Figure 1: Total Number of Apprehensions at the Mexican-US Border (000s)

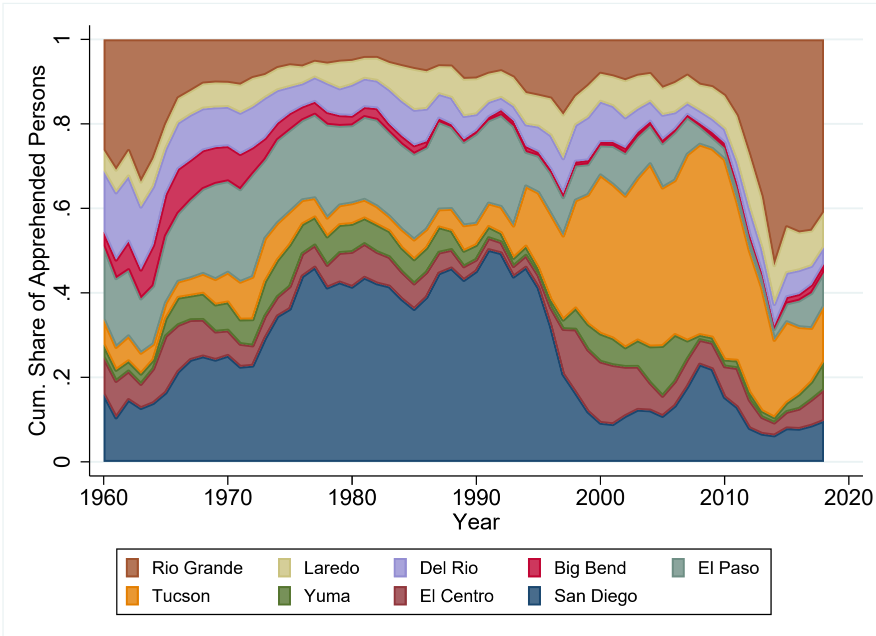


Source: United States Customs and Border Patrol

visas but not the demand for agricultural labor in the United States. Migrants, as a result, turned to undocumented crossing. By the 1970s and early 1980s the number rose at a significantly faster pace, and between 1970 and 1985, total apprehension increased at the rate of about 60,000 *additional* apprehensions per year.

These migrant movements prompted multiple border enforcement initiatives. In 1990, a 14-miles stretch of border fencing was approved to be built along the Tijuana-San Diego border. Its completion in 1993 became prelude to the prevention through deterrence border enforcement policy formally articulated in a national border patrol strategic plan (United States Border Patrol, 1994). A series of border operations to control the flow of illegal drugs as well as illegal migration were authorized during this period: Operation Hold the Line (1993) in Texas, Operation Gatekeeper (1994) in California, Operation Safeguard (1994) in Arizona, and Operation Rio Grande (1997) in Texas. The locations of these operations were designed to stem the flow of undocumented immigrants at the most commonly taken pathways along the border. These operations led to sharp increases in border patrol funding in the 1990s, in the number of border patrol officers, equipment and sensors, as well as the construction of additional border barriers. In 1996, the Illegal Immigration Reform and Immigration Responsibility Act was signed into law, which authorized further

Figure 2: Apprehension Shares Across Border Sectors by Year



Source: United States Customs and Border Patrol

fortification of fencing starting in 1990. Between 1993 and 2005, total border patrol budget tripled from \$500 million to \$1.5 billion.[10] An iconic feature of the border enforcement lesson throughout this time period is that migrants readily respond to new border barriers by switching to alternative crossing locations (Cornelius, 2001; Fernández-Kelly and Massey, 2007; Massey et al., 2016; Gathmann, 2008; Allen et al., 2018; Feigenberg, 2020). Figure 2 shows the share of apprehended undocumented migrants at each border sector from 196.[11] The pattern of migrant displacement away from San Diego and El Paso, to a Tucson majority starting in the mid 1990's can be easily discerned.[12] Confronting these

[10]This fortification of the border continued even after the end of our study period in 2005. President George W. Bush signed the Secure Fence Act into law in 2006, which approved another 700 miles of border walls to be erected from California to Texas. Total border patrol budget more than doubled again from \$1,500 million to over \$3,500 million by 2010.

[11]Of course, apprehension is a function of both enforcement intensity, and the actual flow of migrants. To the extent that the prevention through deterrence policy strengthened enforcement in previously popular crossing locations potentially make apprehension more likely there, the pattern shown in Figure 1b is likely an undercount of the true extent of the shift in migrant displacement.

[12]By siphoning migrants to cross the border through increasingly hazardous terrains, studies have alluded to the possibility that the prevention through deterrence policy may have inadvertantly made organized human smuggling and transborder crime an even more lucrative business as migrants have no choice but pay coyotes to overcome hostile border crossing conditions (Massey et al., 2016). Rising smuggler fees and ever riskier journeys to evade border patrols have been observed (Cornelius, 2001), in addition to extortion

shifts in border apprehension likelihoods and crossing hazards, how do migrants self-select into border sectors? Can these changes in crossing strategies be traced back to the pull and push origins of migration? We now turn to a model of migrant border crossing behavior in which these questions can be addressed.

3 Modeling the Choice of Border Crossing

Consider a large population of migrants of size N. Each migrant i must choose a border crossing location out of K feasible options, $k = 1, ..., K$. Let p_k^s denote the likelihood of successful border crossing at k, and p_k^a the likelihood of encountering an accident, which we take to mean any event that deters a migrant from work or employment at home or abroad. With complementary probability $1 - p_k^s - p_k^a$, the migrant returns to the origin after a failure to cross the border.[13] Furthermore, let $V_i^d \geq 0$, V_i^a, and $V_i^o \geq 0$ respectively denote the expected discounted lifetime utilities, henceforth expected values, associated with reaching the destination after a successful border crossing, accident / death at the border, and an unsuccessful migration attempt. We henceforth normalize the expected value associated with an accident at the border at zero ($V_i^a = 0$). Let C_{ik} denote the cost of border crossing at k, which depends on the cost of hiring a coyote, plus the total distance required to travel to crossing k and then again from crossing k to the destination location for example.

We model the push and pull forces embodied in V_i^d and V_i^o as functions of location-specific factors ($w^j,\ \ j = \{d, o\}$) and network ties characteristics n_i:[14]

$$V_i^d = w^d + \kappa^d n_i, \ \ V_i^o = w^o - \kappa^o n_i. \tag{1}$$

where w^j denotes baseline living standard in location $j = \{d, o\}$. Destination network contact provides a social support structure, job search assistance, and mitigates against the cost of staying abroad. $\kappa^d \geq 0$ parameterizes these features of migrant networks n_i as

and kidnapping (Amnesty International, 2010). We discuss the implications of these changes in crossing risks in Section 4.

[13]The implicit assumption is thus that credit constraints prevent migrants from taking unlimited repeated migration attempts.

[14]In Section 5, we expand on this and include individual characteristics, such as skills, as in McKenzie and Rapoport (2012).

a pull force of migration. Networks can serve as a push force as well, $\kappa^o \geq 0$ parameterizes this particular role of networks when, for example, a negative stigma may be attached to individuals well-connected to the US who nonetheless decide to stay behind (DiMaggio and Garip 2012, Garip and Assad 2016). n_i is a generic placeholder for direct family ties to the US, and / or community level network ties to the US.

Finally, we present a parsimonious specification of the cost of border crossing C_{ik}, and take C_{ik} to be a function of border characteristics (e.g. cost of hiring a coyote at the border) c_k, crossing-individual characteristics (e.g. distance from origin community to crossing location) d_{ik}, and other individual costs associated with migration not specific to the sector in question $\bar{c_i}$. For example, the psychic cost of migration can be subsumed under $\bar{c_i}$. We address additional complications such as individual-specific differences in border crossing cost c_k in Section 4:

$$C_{ik} = \bar{c_i} + c_k + \tau d_{ik}.$$

The expected value of crossing k for migrant i is thus

$$V_{ik} \equiv p_k^s V_i^s + (1 - p_k^s - p_k^a) V_i^o - C_{ik} + \epsilon_{ik}. \tag{2}$$

Purely in terms of relative merits, crossing location k dominates k' if and only if

$$V_{ik} - V_{ik'} > 0. \tag{3}$$

Assume that the $\epsilon_{ik}'s$ are Type I extreme value distributed with density function:

$$f(\epsilon_{ik}) = exp(-\epsilon_{ik} - exp(-\epsilon_{ik})).$$

Crossing k offers the highest value with probability (Maddala 1983 pp. 60-61) :

$$P_{i,k} \equiv Pr(V_{ik} = \max\{V_{i1}, ..., V_{iK}\}) = \frac{exp(p_k^s V_i^s + (1 - p_k^s - p_k^a) V_i^o - C_{ik})}{\sum_{j=1}^{K} exp(p_j^s V_i^s + (1 - p_j^s - p_j^a) V_i^o - C_{ij})}. \tag{4}$$

The relative log odds that migrant i will choose crossing location k relative to 1 can thus be simply expressed as:

$$\log \frac{P_{i,k}}{P_{i,1}} = (p_k^s - p_1^s)(V_i^d - V_i^o) - (p_k^a - p_1^a)V_i^o$$
$$-(c_k - c_1) - \tau(d_{ik} - d_{i1}) \tag{5}$$

Strong pull forces of migration (a high V_i^d) disproportionately favor crossing locations with a high probability of successful crossing which, in this context, are sectors with looser enforcement (i.e. sector k, if $p_k^s - p_1^s > 0$). While strong push forces (a low V_i^o) likewise favor easy crossings, and additionally put less weight on the risks associated with accident-prone crossings (k, if $p_k^a - p_1^a > 0$) – unlike V_i^d, a lower V_i^o worsens the fallback option when a migrant returns home. These migrant have little to lose even when attempting hazardous crossings. Cost and distance considerations ($c_{ik} - c_{i1}$ and $d_{ik} - d_{i1}$) enter into (5) in an intuitive way, favoring low cost, and short distance crossing prospects.

More specifically on the role of networks, consider the incremental effect of network ties to the US via n_i:

$$\frac{\partial \log P_{i,k}/P_{i,1}}{\partial n_i} = (p_k^s - p_1^s)(\kappa^d + \kappa^o) + (p_k^a - p_1^a)\kappa^o \tag{6}$$

Note that the effects of network ties on crossing probabilities via either pull $\kappa^d > 0$ and / or push κ^o forces are mediated by the relative ease of crossing success ($p_k^s - p_1^s$), and relative crossing hazards ($p_k^a - p_1^a$). This suggests that the effect of network ties on relative crossing likelihoods will differ between border sector pairs with different relative likelihood of crossing success, $p_k^s - p_1^s$. Furthermore, the effect of network ties on relative crossing likelihoods between any two sectors k and 1 can change over time, whenever border enforcement intensities change in ways that disproportionately impact one sector over the other.[15]

In order to tease out the impact of networks as a push-pull force on migration ($\kappa^d + \kappa^o > 0$) on crossing preference rankings P_{ik}/P_{i1}, we leverage the prevention through deterrence policy shift in the mid 1990s's, where the stated objective is to turn the tide

[15]Whereas the reversal in the sign of $p^s_{Tucson} - p^s_{SanDiego}$ is the implicit driver of the many important studies on enforcement drive displacement of migrants along the border (Gathmann, 2008; Allen et al., 2018; Feigenberg, 2020, e.g.), the corresponding changes in the environmental risks ($p_k^a - p_1^a$) over the time period of our study is not well-studied. In Section 5.1, discuss this issue.

of undocumented immigration by reversing the relative ease of a successful crossing, in our notations, the sign of $(p^s_k - p^s_1)$, to favor hazardous terrains by increasing border enforcement along traditional traffic routes. An implication of our model, one which we will explore further in this paper, is that such a policy driven change in $p^s_{tucson} - p^s_1$ can in fact trigger a reversal in the self-selection pattern in border crossing choices by individuals with network ties.

4 Data

We employ data from the Mexican Migration Project (MMP).[16] The MMP is a repeated cross-sectional dataset documenting the life and migration experiences of members of over 27,000 households surveyed between 1982 and 2018. The survey covers the full migration history of household heads and spouses, along with the migration experiences of family members. The survey also documents household as well as community characteristics, covering employment, and environmental variables.[17]

Border Crossers and Networks

Importantly for this study, the dataset collects detailed border crossing information for undocumented migrants, including information on when, how and where they crossed the border for each trip they took. In order to rule out possible path dependencies, we limit our analysis to first-time border crossing decisions. Furthermore, even though the survey began in 1982, recorded crossings based on recall go as far back as the 1920s. Observations prior to 1980 are sparse and can dip under single digits even for the most popular crossing locations. Unlike earlier border enforcement attempts, such as Operation Gatekeeper and Operation Hold the Line, where border fencing and patrol efforts were strengthened in the

[16] See <https://mmp.opr.princeton.edu/research/studydesign-en.aspx> for details on survey methodology, sample selection, as well as survey questionnaire.

[17] While the MMP provides a rich source of information on migrants and their families, it is not without caveats. MMP surveys are conducted in typically rural areas known to have higher concentration of migrants. An overwhelming majority of the migrants in our sample period ($> 99\%$) are undocumented at the time of first crossing, and as Lessem (2012) notes, most surveys are conducted in Mexico, and as such temporary and circular migrants are the focus, while permanent migrants without any remaining household members behind are not captured. For our study, the focus on the undocumented migrant population presents a strength rather than a limitation for naturally the effect of border enforcement applies to clandestine crossings only. In addition, temporary and circular migrants as a group are particularly vulnerable to the hazards of border crossings.

historically popular crossing places such as San Diego and El Paso, the Secure Fence Act of 2006 put in place significant increases in border fencing near the Tucson-Mexico border.

In order to avoid conflating the impact of distinctly different border enforcement regimes, and thus potentially differential rankings of the risks and rewards of crossing locations, we limit the time frame of our analysis to between 1980 and 2005. Finally, we consider only migrants 18 years of age or older to account for any agency concerns that may arise with decision-making for young migrants. This leaves us with 2,447 observations of individuals in 153 communities distributed in 24 Mexican states with migrants bound for 38 US states.

Table 1 presents an overview of migrant characteristics in terms of years of education, age at first crossing, gender, and family connections in the US. These are organized according to the location of border crossing (Tucson, or not), and the year of border crossing depending on whether migration occurred pre- (1980-94) or post- (1995-2005) prevention through deterrence policy. The average migrant in our sample received around 7 years of education, and was 28 years old at the time of first migration. Migrants are overwhelmingly male. Comparing the time periods before and after the prevention through deterrence policy in 1994, we see that there are very minor changes in the education ($<$ 1 year more schooling after) and age ($<$ 3 years older after) profiles of the average border crosser. The share of females fell slightly (1% lower after). These figures are quite uniform for both Tucson and non-Tucson crossers.

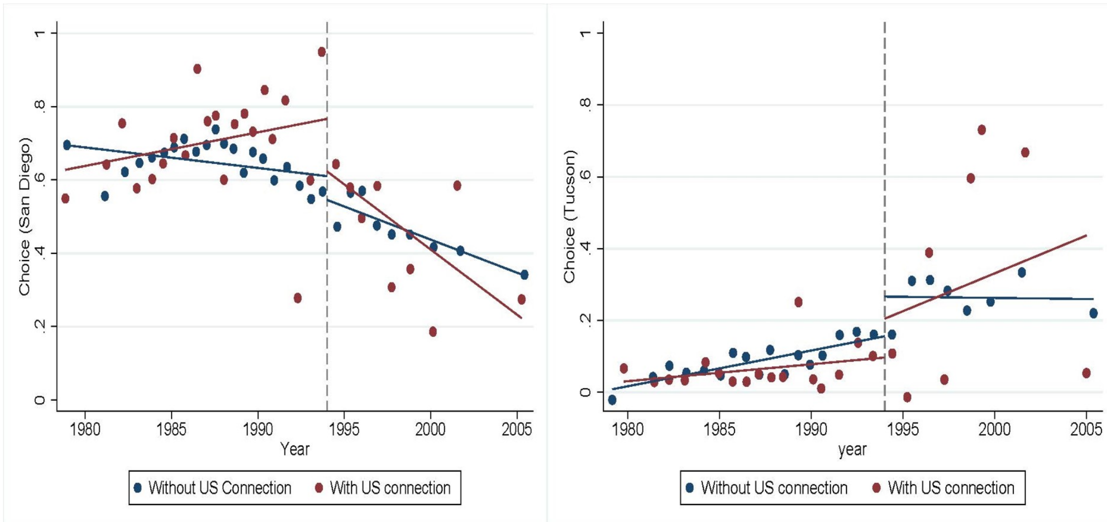
The pattern of family connections for Tucson and non-Tucson crossers are of particular interest. From Table 1, prior to the prevention through deterrence policy, non-Tucson crossers are better connected to the U.S. – the share of migrants with at least one family member who has previously migrated to the US is close to 10 percentage points higher for non-Tucson than Tucson crossers. After the policy, the pattern has reversed, where the share of crossers with US family contacts is higher among the Tucson crossers relative to non-Tucson crossers by about 1 percentage point (Table 1).

Before endeavoring to uncover the push and pull implications of networks on crossing behavior, it is useful to evaluate two arguably more straightforward claims: (i) only connected migrants are privy to up-to-date information about enforcement condition at the border; (ii) the overall shift in border crossings favoring Tucson may have been due to a mechanical increase over time in the share of networked first time migrants. We can establish that both of these claims lack validity based on the data.

To assess the former claim, we visualize the data with 1994 as the cutoff point. In Figure 3 we provide a binscatter plot of individual-level choices in favor of Tucson (1 if Tucson border crossing and 0 otherwise) by year of crossing for individuals with and without family US network ties (either or both parent having had migration experience in the US). We repeat the same binscatter plot to cover individual-level choices in favor of San Diego (1 if San Diego border crossing and 0 otherwise) as well. Both are residualized plots with distance to destination and distance to origin from Tucson and San Diego respectively as controls, and 1994 as the regression discontinuity threshold year. Figure 3 reveals strikingly that after controlling for distance, migrants regardless of family ties seemed to have quite drastically revised their crossing preferences away from San Diego over time in favor of Tuscon since the 1994 mark. These observations are relevant for our analysis because it would be inconsistent with these pictures for our analysis to henceforth assume that network operates purely via an informational route, where only networked individuals are privy to the information that border enforcement intensities have changed. Quite the contrary, regardless of network status, migrants responded to the policy in the expected way. Networked migrants do seem to have been more deeply impacted by the policy as their choice pattern reversals appear to be stronger over time post 1994. In Section 4, we turn to this question by exploring the push-pull role of networks as a determinant of border crossing decisions.

Turning next to the possibility that a rise in the share of networked migrants may have contributed to the rise in Tucson border crossing, we note from Table 1 that the overall share of migrants with US family contact has in fact decreased from 44% to 37% respectively before and after 1994. This may be attributed to a shift in favor of longer stays in the US in response to the prevention through deterrence policy (Fernández-Kelly and Massey, 2007). Thus, the shift in favor of Tucson cannot solely be attributed to a change in the share of networked migrants with up-to-date information about enforcement intensities at the border.

Figure 3: Binscatter Plots of Border Crossing Choices by Network Type



Cost of Crossing: Distance

To provide a sense of the extent to which migrants take on long distance journeys in order to reach the border crossing of their choice, we measure the shortest road distance traveled by each migrant i from their origin community to their stated destination in the US (m) and their stated crossing location choice using data from Google Map ($actualdist_{i,m}$)).[18] We find that migrants travel long distances to reach the border, and then again to reach the destination. The mean distance traveled from migrant community to the chosen border sector is 1327.57 miles, and the mean distance traveled from the chosen border sector to the destination state is 740.97 miles.

To gauge the extent to which migrants deviate from distance minimizing choices of border crossing, we first measure the shortest road distance between each of the 153 possible MMP communities that migrant i belongs to and each of the 9 border sectors (k) ($Dist_{o,k}$).[19] We also measure the shortest road distance from the border sector k to the stated destination m of each migrant ($Dist_{k,d}$), and the shortest road distance from the community i to destination k, ($\min_{k=1,\dots,9}(Dist_{o,k} + Dist_{k,d})$) calculated based on the road distance required for each of the 9 border sectors. The deviation of actual distance traveled

[18] We use city destination if the information is recorded, otherwise, if only destination state information is recorded, we select the state capital as the destination city.

[19] Where there are multiple crossing places within a sector, we select the most popular crossing place.

($ActualDist_{o,d}$) and the minimum distance possible is denoted

$$DevDist_{i,m} = ActualDist_{o,d} - \min_{k=1,\dots,9}(Dist_{o,k} + Dist_{k,d}).$$

Table 2 summarizes the matrix of $DevDist_{o,d}$ aggregated across the main origin regions in Mexico and the destination regions in the US. The data is further divided into two periods (pre-1994 and post-1994).[20] Evidently, Mexican migrants are not distance minimizers, and the patterns of deviation are not uniform. Migrants' border crossing choices often meant many additional miles traveled. Pre-1994, the range of average deviation from the shortest ran from 8 miles (Border (Mexico) to Plains (US)) to 1770 miles (Central (Mexico) to Northeast (US)). Post-1994, the range of $DevDist_{o,d}$ has changed, where migrants bound for the Great Lakes, the Northeast, and Northwest saw a reduction, while the rest saw even lengthier journeys. Of course, distance traveled during migration does not capture the changes in conditions that may have occurred at migrant origins, or at border crossings. We turn to these next.

Cost of Crossing: Border Enforcement

To capture the extent of border enforcement by sector, we use the cumulative mileage of border fencing at each of the 9 border sectors reported in Guerrero and Castaňeda (2017) based on a Freedom of Information Act request. Prior to 1990, border walls were nonexistent, and migrants often simply walked across the border in the cover of darkness at night. The 1990s saw the first wave of border wall / fencing construction. Every year between 1990 and 2005 with the exception of 2001 and 2003, new sections of border wall were being built as part of initiatives to curb undocumented migration, and by 2005, six out of the nine border sectors had wall contructions, resulting in a total of 84 miles

[20] We use the following regional classification. In Mexico, the regions are Historical (Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Liu Potosi), Central (Distrito Federal, Guerrero, Hidalgo, México, Morelos, Oaxaca, Puebla, Queretar, Tlaxcala), Border (Baja California, Chihuahua, Coahuila, Nuevo León, Sinaloa, Sonora, Tamaulipas) and Southeast (Campeche, Chiapas, Quintana Roo, Tabasco, Vercruz, Yucatán). In the US, the regions are Borderlands (Arizona, California, New Mexico, Texas,), Northwest (Idaho, Nevada, Oregon, Utah, Washington), Great Lakes (Illinois, Indiana, Michigan, Ohio, Wisconsin), Northeast, Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, Pennsylvania, Rhode Island, Vermont, Wyoming), Southeast (District of Columbia, Florida, Georgia, Maryland, North Caroline, South Carolina, Virginia, West Virginia), Deep South (Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Tennessee), Plains (Colorado, Iowa, Kansas, Minnesota, Missouri, Montana, Nebraska, North Dakota, Oklahoma, South Dakota). See for example Massey et al. (2016).

of border wall. One concern here is that sector characteristics that encourage migrant border crossings, such as elevation, rainfall and temperature for example, also impact where border walls are built, thus biasing the estimate upwards. Another possibility is to use border personnel as an alternative gauge of the intensity of border enforcement. However, border patrol personnel data are only available from 1992 at the sector level and hence unsuitable for our analysis. Acknowledging these challenges, we incorporate time fixed effects and allow these effects to vary by border sector. The idea is thus to encapsulate time-specific enforcement differences by sector within these fixed effects without disentangling the individual impacts of each border enforcement tool.

5 Estimation

Informed by theory, we adopt the alternative specific conditional logit model, or McFadden's choice model (McFadden 1974), which estimates (4) with simultaneous inclusion of variables that are border sector (and border sector × individual) specific, as well as individual migrant specific. In this setup, individual-specific controls are interacted with indicator variables of each of the choice alternatives in a conditional logit model to yield border sector specific responses. In all specifications, San Diego serves as the baseline alternative, and standard errors are clustered at the origin community level.

To recall, the log choice likelihood ratio is

$$\begin{aligned}\log\frac{P_{i,k}}{P_{i,1}} &= (p_k^s - p_1^s)(w^d - w^o) - (p_k^a - p_1^a)w^o - (c_k - c_1)\\ &\quad + \left[(p_k^s - p_1^s)(\kappa^d + \kappa^o) + (p_k^a - p_1^a)\kappa^o\right] n_i\\ &\quad - \tau(d_{ik} - d_{i1})\end{aligned}$$

Starting with the last term, $\tau(d_{ik} - d_{i1})$, the alternative specific conditional logit model produces a common estimated coefficient for border sector × individual (community) specific variables (e.g. distance to and from the border d_{ik} that differ from sector to sector at the individual level), τ. For individual specific variables (e.g presence of network ties n_i that remains the same regardless of sector choice), the model produces an $N-1$ vector of coefficients corresponding to $(p_k^s - p_1^s)(\kappa^d + \kappa^o) + (p_k^a - p_1^a)\kappa^o$, $k \neq 1$. In our estimation, we work with seven border sectors (San Diego, El Centro, Tucson, El Paso, Laredo, Del Rio, Rio Grande) covering 99% of all cases / individuals in our data.[21]

The next expression in the log relative likelihood ratio, $\log P_{i,k}/P_{i,1}$, relates to the impact of the pull $\kappa^d n_i$ and the push $\kappa^o n_i$ forces of networks on crossing preferences.These effects are mediated by border sector specific enforcement and environmental hazards, $(p_k^s - p_1^s)$ and $(p_k^a - p_1^a)$, that can naturally change as policy and crossing hazards evolve over time. To account for these potential variations in border risk landscapes, before and after the roll out of the prevention through deterrence policy, we interact n_i with five five-year dummies (1980-84, 1985-89, 1990-94, 1995-99, 2000-05) indicating the year of crossing, to allow the impact of network ties to change with crossing time intervals.[22] Thus we estimate 5 × (7-1) = 30 different coefficients to assess the impact of network ties on crossing choice likelihoods at each of the 5 time intervals across 6 border sector-base alternative pairings.

Finally, to complete our specification, we additionally include stand-alone border sector specific five-year dummies, with 1980-84 as base period, and border sector constants to capture any changes in network-free relative returns on migration through border sector k. This addresses the remaining term in the predicted log relative likelihood ratio $(p_k^s - p_1^s)(w^d - w^o) - (p_k^a - p_1^a)w^o - (c_k - c_1)$.[23]

In our first specification, we employ a minimalist approach and only use the border-sector specific five-year dummies indicating the year of migration (with 1980-84 as base) as individual characteristic controls. For sector-specific characteristics, we use distance to origin community, and distance to destination ($Dist_{o,k}$ and $Dist_{k,d}$) as controls. Column 1 of Table 3 displays the estimated coefficients for each border sector relative to the base (San Diego). The estimated marginal effects associated with these determinants are available in the Appendix, and the interpretation of the marginal effects will be discussed in Section 5.

Distance to ($Dist_{o,k}$) and from ($Dist_{k,d}$) a border sector are negatively associated

[21] Yuma and Big Bend are rarely chosen border sectors with less than a handful of entries in many years. In the Appendix, we document results using the full set of 9 border sectors whenever feasible. The qualitative results reported here do not change upon including these two additional border sectors.

[22] The period-interaction approach is agnostic about whether any change in border risk is due to enforcement $(p_k^s - p_1^s)$ or environmental hazards $(p_k^a - p_1^a)$, and does not place specific assumptions about the signs and magnitudes of these diverse risk factors.

[23] In the sequel, we introduce additional individual characteristics control to model $w^d - w^o$ as well as c_k ad c_1.

with the likelihood of border sector choice, suggesting that long distance is indeed a deterrence to border crossing choices. Interestingly, controlling for time-invariant distance considerations, Table 3 showcases shifts in the crossing sector preferences over time via the year interval dummies. What is singularly striking about these results is that in odds ratio terms, individuals whose first crossing took place after 1995 are $exp(1.915) - 1 = 579\%$ (1995-99) and $exp(2.356) - 1 = 955\%$ (2000-05) respectively more likely to cross via Tucson relative to San Diego. These increases are statistically significant at the 1 percent level. In all other border sectors, the likelihood of crossing either strictly fell (e.g. Laredo), or saw no statistically significant change (e.g. El Paso). These findings are consistent with studies that confirmed migrant replacement responses to selective border enforcement reforms (e.g. Cornelius (2001); Gathmann (2008); Lessem (2012); Allen et al. (2018); Feigenberg (2020)).

5.1 Family and Community Ties

We now turn to specifications that include individual ties to the US as proxies for networks, n_i. In Table 4, we show only the results pertaining to the Tucson border sector. The full set of estimation results for every sector is relegated to the Appendix. To recall, direct family network n_i takes on a value of 1 if the migrant's father or mother had migration experience in the US prior to the year of crossing, and 0 otherwise. In Column 1 of Table 4, we look at the full sample results. In Column 2, we ascertain the role of direct family networks for individuals with no more than 12 years of education (upper secondary education), and in Column 3, individuals with no more than 9 years of education (middle / junior secondary education) are included. These regressions continue to include distance to and from the border as migration cost proxy, and a full set of five-year dummies and border-specific dummies to capture network-free motivations to cross a border sector.

In an influential study, McKenzie and Rapoport (2010) present evidence in the context of Mexican-US migration that low skill migrants are the largest beneficiaries of family ties to the US, motivating migration due to a variety of reasons such as credits constraints both at home and abroad, the need for job search assistance, and language proficiency (Borjas (1992); Chiswick and Miller (2002); Bauer et al. (2005)). To assess in our setting the role of skills on the salience of network effects as push and pull forces, Column

1 uses the full sample, while Columns 2 and 3 respectively include individuals with less than upper secondary school education, and lower secondary school education respectively. These reflect a uniformly stronger role of networks, our κ^d and / or κ^o, on the pattern of self-selection among lower skilled individuals, consistent with prior evidence suggesting a stronger role of networks as a push-pull force of migration among low skilled workers (e.g. Beine et al. (2015)). In Table 4 Column 4, we furthermore examine border crossing preference by occupation to specifically examine individuals in agricultural or manufacturing occupations among those with lower secondary education or less.[24] Here too, we see a similar pattern of self-selection reversal.

In Columns 1-4 of Table 4, we observe a pattern of crossing choice that is repeated throughout our specifications to follow: in the earlier years prior to 1995, networked migrants are less likely to choose Tucson relative to the base San Diego. This pattern is reversed in later years (after 1999 in column 1, and after 1995 in columns 2-4), where networked migrants are more likely to choose Tucson relative to base. Specifically, in odds ratio terms for Column 1 (full sample), a migrant whose father or mother had been to the US is less likely to choose Tucson over San Diego between in 1980-84 (odds ratio $exp(-1.681) = 0.187 < 1$), and more likely to do so in 2000-05 (odds ratio $exp(0.224) = 1.125 > 1$), though the latter coefficient is statistically insignificant. Going from Column 1 to 4, we see that the sign reversal pattern becomes more stark, and the coefficients are statistically significant and negative during the earliest years, statistically no different from zero in the intermediate years, and positive and statistically significant in the latest years covered in our study.

In Table 5 (Columns 1-3), we branch out to examine more indirect ties to the US as a push and pull force of migration. In Column 1, we leverage our network proxy $n_i(Family)$, which takes on a value of 1 if the migrant's own immediate family (father, mother, siblings) had migrated to the US prior to the time of first crossing. We also separately include an alternative network proxy $n_i(Community)$, which measures the share of first time crossers in a migrant's community with a direct US family connection (with $n_i(Direct) = 1$). Comparing these results with those in Table 4, we find that indirect ties are less influential

[24]We choose agriculture and manufacturing to better rule out high paying professions requiring specialized skills not attained in traditional schools.

than direct parental ties to the US both in terms of the magnitude of the corresponding odds ratio, as well as the statistical significance of the interaction terms. These findings resonate well with studies that have found that proximate ties have more influence on migration motivations involving uncertain and hazardous prospects (Davis et al., 2002; Curran et al., 2005; DiMaggio and Garip, 2012). In our context, the same migration influences are shown to impact border crossing choices through the terms κ^d and κ^o. Column 3 reiterates this message, where the network variable is the share of first time migrants in a community with a family member who had been to the US. All interaction terms, while preserving the same sign pattern, are statistically insignificant.

Since border enforcement and crossing hazards jointly determine the role that push-pull forces of migration impact crossing preferences from (5), what exactly is a reversal of self-selection pattern indicative of? The coefficient associated with the period-interval×network control reflects a weighted combination of relative border enforcement and relative crossing hazard comparisons. Once again taking $k = tucson$ as an example, and San Diego as base, the effect of US family ties on the relative probability of choosing Tucson over San Diego is given by the following:

$$(p^s_k - p^s_1)(\kappa^d + \kappa^o) + (p^a_k - p^a_1)\kappa^o. \tag{7}$$

For the above expression to take on a negative sign, when U.S. family ties serve as positive pull and / or push forces $\kappa^d > 0$, $\kappa^o > 0$, and when a Tucson crossing implies greater chances of accident or death ($p^a_k > p^a_1 > 0$), a necessary condition is that migrants consider a San Diego crossing to offer a strictly higher chance of success $p^s_k - p^s_1 < 0$. Changes in migration policies – such as prevention through deterrence – that successively render crossing through San Diego increasingly difficult, can reverse the sign of the network effect as soon as $p^s_k - p^s_1$ becomes sufficiently large.

This is exactly what we see as we fast-forward to the post-1994 period. The stated objective of the prevention through deterrence policy is indeed to turn $(p^s_k - p^s_1)$ positive, by drastically increasing the likelihood of apprehension in previously popular locations such as San Diego. Here, the positive estimated marginal effect of US family ties on Tucson crossing reflects (i) a reversal in relative border enforcement risks $(p^s_k - p^s_1)$, and (ii) the disproportionate impact that this increase in risk has on the crossing location choices among individuals with US family ties.

Our discussion so far has implicitly taken the crossing hazard difference $(p^a_k - p^a_1)$ to be relatively stable over time, which is plausible if $p^a_k - p^a_1$ is only a function of time invariant environmental hazards. But by diverting migrants to cross the border through increasingly hazardous terrains, studies have alluded to the possibility that the prevention through deterrence policy may have made organized human smuggling and transborder crime a lucrative business as migrants have no choice but pay coyotes to overcome hostile border crossing conditions (Massey et al., 2016). Rising smuggler fees and ever riskier journeys to evade border patrols (Cornelius, 2001), extortion and kidnapping have also been reported (Amnesty International, 2010). These changes in border crossing hazards may render a Tucson crossing increasingly risky subsequent to the prevention through deterrence policy, raising $p^a_k - p^a_1$, and thus the network effect $(p^s_k - p^s_1)(\kappa^d + \kappa^o) + (p^a_k - p^a_1)\kappa^o$ over time if and only if networks are a push force of migration, $\kappa^o > 0$, all else equal.

Our estimates do not distinguish between whether migrant displacement operated through the border enforcement or the crossing hazard channel. In this sense, our study shares a common feature with other studies on the migration displacement effect of border enforcement, as enforcement proxies cannot in general separately identify the displacement effects due to changes in the likelihood of discovery in targeted points of crossing, versus changes in smuggler operations that move to harder to detect locations and attract those most desperate to migrate. In our context, what matters is that these channels are both predicted to mediate the influence of network on crossing patterns through its role as a push and / or pull force of migration. Thus, the findings in Table 4-5 suggest the presence of one or both of the two mutually reinforcing mechanisms driving networked migrants to be less likely and then more likely to choose Tucson crossings, respectively before and after the prevention through protection policies.

5.2 Network Externality on the Community

Our parsimonious modelling of the cost of migration so far only takes into account the sector specific coyote price c_k and the distance traveled τd_{ik}. Following the footsteps of Roberts et al. (2010) and Gathmann (2008), we allowed the smuggling price to depend systematically on border enforcement at each sector k, and other demand and supply shocks through border specific 5-year dummies and border sector fixed effects. But even this may be an incomplete picture of the determinants of border crossing price. Indeed, migrant networks can impact the smuggler price through a pure externality effect on everyone in the community whether or not they are connected to a network (DiMaggio and Garip, 2012; Garip and Asad, 2016). First, because of fixed cost considerations, smugglers can afford to charge a lower price when they know they can count on a sufficient scale of operation if there is a reliable history of migrants who wish to cross the border from a given sector. This is the definition of second degree price discrimination. We proxy for this type of network effect by introducing rolling average three-year lagged share of migrants in the community of migrant i crossing to the US for the first time through sector k, N_{ik}.

Alternatively, smugglers may also practise third degree price discrimination. This occurs when a history of migrants changes the characteristics of individuals who demand the services of smugglers in a community, such as the age of first time migrants due for example to when family reunification becomes an important source of migration, or the skill level or education of first time migrants because prior remittances were spent investing in the education of future migrants. Smugglers may charge a different price based on these community level characteristics even when a smuggler may well be uncertain about the age / education level of each and ever migrant. We proxy for this type of network effects by introducing the average age of first migrants, and / or the average education of first time migrants in the community of migrant i, M_i

These possibilities pose threats to identification of the role of direct family network ties to the US if network externality, which may be mechanically associated with family and community ties to the US, operate purely through the commons effect on the cost of migration, but not through the pull and push factors of migration. Thus, we extend our baseline setting to incorporate these network externality effects on smuggler prices, and in turn on crossing behavior. We bring together second degree price discrimination triggers, such as community's prior history of border sector choices, and third degree price discrimination triggers, such as the average age of first time border crossers. Our approach is not to rule out one or the other, but rather to control for both forces in our estimation, to see if our the push-pull role of direct family networks on crossing probabilities remain robust.

Rewriting the cost of migration making use of N_{ik} and M_i, we have

$$C_{ik} = \bar{c}_i + c_k + \tau d_{ik} - \gamma^N N_{ik} - \gamma_k^M M_i$$

where γ^N and γ_k^M respectively gauge the impact of second and third degree price discrimination as a result a network externality. The revised log likelihood ratio is

$$\begin{aligned} \log \frac{P_{i,k}}{P_{i,1}} &= (p_k^s - p_1^s)(\kappa^d + \kappa^o)n_i + (p_k^a - p_1^a)\kappa^o n_i \\ &\quad -(c_k - c_1) - \tau(d_{ik} - d_{i1}) \\ &\quad +\gamma^N(N_{ik} - N_{i1}) + (\gamma_k^M - \gamma_1^M)M_i \end{aligned} \tag{8}$$

Notably, if smugglers practise second degree price discrimination, (8) shows the tendency for cumulative causation in border crossing choices – popular border crossing begets future border crossing at the same location whenever $\gamma^N > 0$. Meanwhile, if smugglers practise third degree price discrimination, then changing the average characteristics of first time migrants in a community will displace border crossers if and only if smugglers at different border crossings price these characteristics differently ($\gamma_k^M - \gamma_1^M \neq 0$).

As a reality check, we regress individual level log coyote price from the MMP data with respect to these two network externality indicators, controlling for year fixed effects, state fixed effects via ordinary least squares (Table 6). Indeed, we find strong, albeit non-causal evidence that is consistent with second- and third-degree price discriminating conduct. To wit, coyote cost is strictly decreasing with respect to the average age of first time crossers in a community. Likewise, coyote cost at a sector is strictly decreasing in the share of migrants in a community choosing said border sector in prior years.

Given these observations, we introduce both the rolling average lagged crossing choice share, and the average age of first time migrants in our choice regression. In Column 1 of Table 7 we display the results. We find that even after controlling for the possibility of network externality, the reversal of self-selection patterns amongst individuals with family ties in the US continue to hold, and the estimated marginal effects are hardly affected.[25]

[25]We ran additional regressions interacting the five time period dummies with the average age of first time crosser variable in the choice regression. We find that this network externality effects do not change signs over time, indicating that community level characteristics, unlike direct family ties, play a small /

5.3 Network Impact on Individual Migrant Characteristics

Our simple model of the push and pull forces of migration have so far only taken into account the role of networks as an individual characteristic. Clearly, other forms of migrant characteristics, such as education and age at crossing, may also impact migration motivations, and by implications, crossing incentives. These are important for two reasons. First, family ties in the US may have a direct impact on the education and age of future first time migrants in the family, such as when remittances from former migrants are spent investing in the next generation of migrants through education (Rapoport and Docquier 2006, Bansak and Chezum 2009, Rapoport and McKenzie 2011), or when prior family links compel families to send young people abroad earlier for family reunification (OECD 2019 chap. 4). We assess the implications of this confounding influence of networks on crossing behavior, to gauge whether border crossing behavior was in fact responsive to migrant's skill and age, rather than the desirability of the destination or the origin community modified by family ties.

To more formally introduce these individual characteristics into our model, let e_i denote an individual characteristic (e.g. education and age):

$$V_i^d = w^d + \delta^d e_i + \kappa^d n_i, \;\; V_i^o = w^o + \delta^o e_i - \kappa^o n_i. \tag{9}$$

where δ^j is the returns on e_i (e.g. returns on human capital) in location j.

The relative log odds that migrant i will choose crossing location k relative to 1 is given by:

$$\begin{aligned}\log \frac{P_{ik}}{P_{i1}} \;\; = \;\; & (p_k^s - p_1^s)(w^s + \delta^s e_i + \kappa^d n_i) \\ & -[(p_k^s - p_1^s) + (p_k^a - p_1^a)](w^o + \delta^o e_i - \kappa^o n_i) \\ & -(c_k - c_1) - \tau(d_{ik} - d_{i1}) \\ & +\gamma^N (N_{ik} - N_{i1}) + (\gamma_k^M - \gamma_1^M) M_i \end{aligned} \tag{10}$$

The incremental effect of individual characteristics, such as education e_i, on crossing pref-

negligible role in the reversal in border crossing choice in response to border enforcement policies.

erence ranking can also be ascertained:

$$\frac{\partial \log P_{ik}/P_{i1}}{\partial e_i} = (p_k^s - p_1^s)(\delta^d - \delta^o) - (p_k^a - p_1^a)\delta^o. \tag{11}$$

Evidently, the marginal impact of skills conflates two effects – the relative likelihood of crossing success $(p_k^s - p_1^s)$, and the relative returns on skills $(\delta^d - \delta^o)$. Thus, while education is a pull force in the sense that $\delta^d > 0$, more highly educated individuals may nonetheless reject a crossing through k, even when it offers a higher chance of success $(p_k^s - p_1^s) > 0$, if there is negative selection of migrants in the destination country based on migrant skills $(\delta^d < \delta^o)$. Furthermore, to the extent that returns to education in the origin δ^o is positive, higher border crossing hazards at k $(p_k^a - p_1^a) > 0$ is a deterrent.

In Table 6 (columns 3-6), we perform another reality check to determine the extent to which the skill and age of first time migrants depends on the family connection in the US, via an OLS regression in which we regress years of education and age at first crossing on US network ties, distance, and a full set of time and state fixed effects. The results are shown in Table 6, in which we find that family network ties to the US has a modest association with the education level (plus less than 1 year) and age (minus between 2 - 3 years) of first time migrants.

We then re-ran the choice model by introducing individual-level years of education and age at first crossing. In Table 7 we show that the reversal in self-selection amongst individuals with family ties to the US remains even after introducing individual-level years of education and age at first crossing as additional controls. We ran additional regressions (available upon request) that include the interaction of these individual variables with time interval dummies. We find that these individual-level characteristics have negligible effects on border crossing choices.[26] We take these as evidence that the role of networks as a push-pull force of border crossing remains robust even upon introduction of individual characteristics.[27]

[26] We ran additional regressions including interaction effects between education and time period dummies. These effects are not significantly different from zero throughout.

[27] These results also echo a longstanding literature suggesting that an alternative reason why the relationship between income, proxied by years of education or age, and migration incentives are nuanced (Faini and Venturini, 1993; Mckenzie and Rapoport, 2007; Angelucci, 2012), for example when credit constraints are in play.

Additional Checks and Alternative Drivers

We conduct additional checks and include alternative controls. Table 7 Column 3 includes cumulative miles of border fencing as additional border enforcement to account for annual changes in sector-specific border enforcement to complement the border-sector$\times$ period-interval fixed effects. Fencing gives rise to a combination of auxiliary effects, including potential changes in the availability of smugglers at each border crossing for example. Our inclusion of the fencing variable embodies both these direct and auxiliary effects. Our goal is to determine whether changes in period-interval fixed effects remain important after controlling for the added mileage of border walls. Interestingly, while cumulative border fencing does have a negative impact on the likelihood of migrant crossing, fencing alone does not eliminate changes in period-interval and border sector specific fixed effects. These observations suggest, not surprisingly, that displacement of migrants depends on a combination of policies and tools.

Column 3 in Table 7 also introduces regional fixed effects, to account for possible spatial origins of community-level crossing triggers. These spatial factors are arguably particularly important in the Mexican context, as migration experiences in the US are highly differentiated across origin communities. Specifically, we account for four Mexican regions, including Historical, Border, Central and Southeast. Interestingly, migrants from the border, central, as well as Southeastern origin communities are more likely than migrants from historical migrant origin communities to select Tucson over San Diego, even accounting for individual- and community-level triggers. These suggest that a history of migration experiences discourage migrants from embarking on a journey to the US through the high risk Tucson border crossing. Alternatively, a long history of migration that began with San Diego or El Paso as crossing location may have introduced its own network externality effects, with cumulative impact on the crossing choices of future migrants.

On Marginal Effects

In the Appendix, we provide the marginal effects associated with the choice model estimates in Tables 3 - 5 and 7. To recall, the coefficient estimates of the choice model in these tables assess the change in the log likelihood ratio relative to the base alternative in response to a change in network connections. From (11):

$$\frac{\partial \log P_{i,k}/P_{i,1}}{\partial n_i} = (p_k^s - p_1^s)(\kappa^d + \kappa^o) + (p_k^a - p_1^a)\kappa^o$$

which then depends on the difference in border enforcement related risks $(p_k^s - p_1^s)$, and the difference in environmental risks $(p_k^a - p_1^a)$ between any border sector k and the base (San Diego). The coefficient associated with destination networks is thus expected to take on an increasingly positive sign as enforcement loosens in k relative to 1 $(p_k^s - p_1^s)$, all else equal.

Now turning to the interpretation of the marginal effect, using (4), it is straightforward to verify the marginal effect of network ties to the US as:

$$\frac{\partial P_{i,k}}{\partial n_i} = P(i,k) \sum_{j=1}^{K} P(i,j) \left((p_k^s - p_j^s)(\kappa^d + \kappa^o) + (p_k^a - p_j^a)\kappa^o \right) \tag{12}$$

which is the weighted average analog of (11) over all possible base alternatives $j = 1, ...K$. Thus, a positive coefficient in Tables 3 - 5 and 7 based on estimated relative log likelihood ratio response to a change in network connections is neither necessary nor sufficient for a corresponding positive marginal effect. Rather, the marginal effect of networks on the choice probability of sector k is predicted to take on a negative sign at some time periods, and to switch to a positive sign later on, for example, if sector k's likelihood of migration success p_k^s moves from being an inferior choice relative to the average border sector, to being better than average later on.

With these additional observations, we note that Appendix Tables (A2) to (A5) continue to show a by now familiar self-selection reversal pattern for Tucson crossings – among migrants with family connections in the US, they are negatively selected at the Tucson sector prior to 1994, and positively selected thereafter. In these tables, we also complement the Tucson findings with the marginal effects associated with network connections on San Diego crossings. We find that a self-selection reversal (by 2000-05) in San Diego among individuals with US family connections is evident only among migrants for whom the effects on networks as a pull / push force of migration are most salient. These are migrants with lower education level (Column 3), and those in agricultural or manufacturing occupations (Column 4) (e.g. McKenzie and Rapoport (2010)).

6 Conclusion

Why do migrants embark on dangerous border crossing journeys? In this paper, we use the Mexican-U.S. border as a salient case in point, featuring stark differences in the competing risks juxtaposing border enforcement and crossing hazards that migrants take must take into account along the migration journey. These include the risks of injuries and death, and the risk of border apprehension. We write a simple model of border crossing choices, which shows how pull factors of migration attract border crossers to favor loosely enforced locations with higher probability of crossing success, while push factors of migration additionally encourage migrants to discount the hazards at the border. We use family network ties in the US to proxy for the strength of the push-pull forces of migration. In order to address the possibility of confounders, we use exogenous changes in immigration policy via the prevention through deterrence policy in the 1990's, which placed border enforcement resources strategically along the border in order to stem the tides of undocumented migration in previously popular crossing locations.

We draw three broad findings from the evidence. First, our baseline specification focuses on the time-specific impact of U.S. family ties on border sector crossings to shed light on the nuanced role of migration push-pull forces on crossing location choices. Consistent with the model's predictions, we find a reversal in the pattern of self-selection unfolding over time – migrants with US family ties were negatively selected at the Tucson border sector prior to the prevention through deterrence strategy before 1995. This pattern then shifted and reversed over time to exhibit positive selection later on, and particularly so after 2000. Echoing studies on mediators that strengthen network impacts on migration motives, we find that this pattern of self-selection is particularly salient among low skilled, and / or blue collar workers in agriculture and manufacturing, and via direct family networks rather than community migrant networks.

Further to these baseline findings, we evaluate alternative pathways through which networks can function beyond traditional push and pull forces, different mechanisms that may have been derivatives of family network ties, and as well as alternative measures of push and pull forces. These findings offer two migrant-centric perspectives on border enforcement reforms. First, border enforcement strategies have distributional consequences.

In the context of the prevention through deterrence policy along the Mexcian-US border, individuals with strong desire to migrate were diverted to take on perilous border crossing journeys, exposing migrants to some of the most hazardous climate conditions and environmental terrains. Second, our findings also provide novel humanitarian justification for border enforcement policies to go hand-in-and with family-based immigration policy – arguably the bedrock of U.S. immigration policy since the Immigration and Nationality Act in 1965. Giving preference for family reunification in a priority-weighted system of immigration quotas not only makes sense for immigrant families, it also help steer family members of existing migrants away from border hazards, particularly since these individuals with direct family ties are most at risk of choosing to confront border crossing hazards when alternative and safer routes are made inaccessible due to border enforcement reforms.

References

Allen, T., C. de Castro Dobbin, and M. Morten (2018, November). Border Walls. NBER Working Papers 25267, National Bureau of Economic Research, Inc.

Amnesty International (2010). Invisible victims: Migrants on the move in mexico. Technical report.

Angelucci, M. (2012). Conditional cash transfer programs, credit constraints, and migration. *LABOUR 26*(1), 124–136.

Basu, A. K., N. H. Chau, and G. C. Lin (2021, October). Migration gravity, networks and unemployment. Charles H. Dyson School of Applied Economics and Management Working Paper 2021, Cornell University.

Bauer, T., G. S. Epstein, and I. N. Gang (2005). Enclaves, language, and the location choice of migrants. *Journal of Population Economics 18*(4), 649–662.

Bazzi, S., G. Hanson, S. John, B. Roberts, and J. Whitley (2021, August). Deterring illegal entry: Migrant sanctions and recidivism in border apprehensions. *American Economic Journal: Economic Policy 13*(3), 1–27.

Beine, M., F. Docquier, and a. Özden (2015). Dissecting network externalities in international migration. *Journal of Demographic Economics 81*(4), 379–408.

Borjas, G. J. (1992). Ethnic capital and intergenerational mobility. *The Quarterly Journal of Economics 107*(1), 123–150.

Bryan, G. and M. Morten (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy 127*(5), 2229 – 2268.

Caliendo, L., M. Dvorkin, and F. Parro (2019). Trade and labor market dynamics: General equilibrium analysis of the china trade shock. *Econometrica 87*(3), 741–835.

Chau, N. H. (1997, February). The Pattern of Migration with Variable Migration Cost. *Journal of Regional Science 27*(1), 35–54.

Chau, N. H. (2001). Strategic Amnesty and Credible Immigration Reform. *Journal of Labor Economics 19*(3), 604–34.

Chau, N. H. (2003). Concessional Amnesty and the Politics of Immigration Reforms. *Economics and Politics 15*(2), 193–224.

Chiswick, B. R. and P. W. Miller (2002). Immigrant earnings: Language skills, linguistic concentrations and the business cycle. *Journal of Population Economics 15*(1), 31–57.

Congressional Research Service (1997). US Border Patrol Operations. Technical report.

Cornelius, W. A. (2001, December). Death at the Border: Efficacy and Unintended Consequences of US Immigration Control Policy. *Population and Development Review 27*(4), 661–685.

Curran, S. R., F. Garip, C. Y. Chung, and K. Tangchonlatip (2005). Gendered migrant social capital: Evidence from thailand. *Social Forces 84*(1), 225–255.

Davis, B., G. Stecklov, and P. Winters (2002). Domestic and international migration from rural mexico: Disaggregating the effects of network structure and composition. *Population Studies 56*(3), 291–309.

DiMaggio, P. and F. Garip (2012). Network effects and social inequality. *Annual Review of Sociology 38*(1), 93–118.

Dolfin, S. and G. Genicot (2010, May). What do networks do? the role of networks on migration and "coyote" use. *Review of Development Economics 14*(2), 343–359.

Epstein, G. S. and A. Weiss (2011). The why, when, and how of immigration amnesties. *Journal of Population Economics 24*(1), 285–316.

Ethier, W. J. (1986, March). Illegal Immigration: The Host-Country Problem. *American Economic Review 76*(1), 56–71.

Facchini, G. and C. Testa (2021). The rhetoric of closed borders: Quotas, lax enforcement and illegal immigration. *Journal of International Economics 129*, 103415.

Faini, R. and A. Venturini (1993, April). Trade, aid and migrations: Some basic policy issues. *European Economic Review 37*(2-3), 435–442.

Feigenberg, B. (2020, July). Fenced out: The impact of border construction on us-mexico migration. *American Economic Journal: Applied Economics 12*(3), 106–39.

Fernández-Kelly, P. and D. S. Massey (2007). Borders for whom? the role of nafta in mexico-u.s. migration. *The Annals of the American Academy of Political and Social Science 610*, 98–118.

Friebel, G., M. Manchin, M. Mendola, and G. Prarolo (2018, November). International migration intentions and illegal costs: Evidence from africa-to-europe smuggling routes. IZA Discussion Papers 11978, Institute for the Study of Labor.

Garip, F. (2008, 08). Social capital and migration: How do similar resources lead to divergent outcomes? *Demography 45*, 591–617.

Garip, F. and A. L. Asad (2016). Network effects in mexico–u.s. migration: Disentangling the underlying social mechanisms. *American Behavioral Scientist 60*(10), 1168–1193.

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.

Hanson, G. H. and A. Spilimbergo (1999). Illegal immigration, border enforcement, and relative wages: Evidence from apprehensions at the u.s.-mexico border. *The American Economic Review 89*(5), 1337–1357.

Kandel, W. and D. S. Massey (2002). The culture of mexican migration: A theoretical and empirical analysis. *Social Forces 80*(3), 981–1004.

Lessem, R. (2012). Mexico – u.s. immigration: Effects of wages and border enforcement. *Review of Economic Studies 85*(4), 2353–2388.

Massey, D. S. (1988). Economic development and international migration in comparative perspective. *Population and Development Review 14*(3), 383–413.

Massey, D. S., J. Arango, G. Hugo, A. Kouaouci, A. Pellegrino, and J. E. Taylor (1993). Theories of international migration: A review and appraisal. *Population and Development Review 19*(3), 431–466.

Massey, D. S., J. Durand, and K. A. Pren (2016). Why border enforcement backfired. *American Journal of Sociology 121*(5), 1557–1600.

Mcfadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*, pp. 105–142. Academic Press.

McKenzie, D. and H. Rapoport (2007, September). 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 (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-u.s. migration: The role of migration networks. *The Review of Economics and Statistics 92*(4), 811–821.

Mines, R. and A. de Janvry (1982). Migration to the united states and mexican rural development: A case study. *American Journal of Agricultural Economics 64*(3), 444–454.

Monte, F., S. J. Redding, and E. Rossi-Hansberg (2018, December). Commuting, migration, and local employment elasticities. *American Economic Review 108*(12), 3855–90.

Moretti, E. (1999). Social networks and migrations: Italy 1876–1913. *International Migration Review 33*(3), 640–657.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the u. s. labor market. *The Quarterly Journal of Economics 118*(2), 549–599.

Munshi, K. (2014). Community networks and the process of development. *The Journal of Economic Perspectives 28*(4), 49–76.

Munshi, K. (2020). Social networks and migration. *Annual Review of Economics 12*(1), 503–524.

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

Orrenius, P. (1999). The role of family networks, coyote prices and the rural economy in migration from western mexico: 1965-1994. Working Papers 9910, Federal Reserve Bank of Dallas.

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

Piore, M. J. (1979). *Birds of Passage: Migrant Labor and Industrial Societies*. Cambridge University Press.

Roberts, B., G. Hanson, D. Cornell, and S. Borger (2010). An Analysis of Migrant Smuggling Costs along the Southwest Border. Technical report.

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

Tombe, T. and X. Zhu (2019, May). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review 109*(5), 1843–72.

United States Border Patrol (1994). Border Patrol Strategic Plan, 1994 and Beyond. Technical report.

Winters, P., A. de Janvry, and E. Sadoulet (2001, 01). Family and community networks in mexico-u.s. migration. *Journal of Human Resources 36*, 159–184.

Table 1

Migrant Characteristics by Crossing Choice and Period

Variables	**Crossing Choice and Year** 1980-1994 All Choices	1980-1994 Not Tucson	1980-1994 Tucson	1995-2005 All Choices	1995-2005 Not Tucson	1995-2005 Tucson	All Years All Choices
Years of Education (years, avg.)	6.35	6.33	6.69	7.14	6.97	7.52	6.58
Age at First Crossing (years, avg)	27.26	27.31	26.58	30.05	29.98	30.20	28.08
% Female (Male=0, Female=1, %)	5.27	5.039	8.40	4.31	3.98	5.07	4.99
% with US Family Connections (%)	44.39	45.12	34.45	37.27	37.05	37.79	42.30
Total # of US Family Connections	0.98	0.99	0.77	0.73	0.74	0.71	0.91
N	1,728	1,609	119	719	502	217	2,447

Notes. 1. Source: Mexican Migration Project. 2. A US family connection is defined as having a parent or a sibling living in the US prior to the migrant's first crossing attempt. 3. Total US family connection is a count of the number of family members (parents and siblings) who had lived in the US prior to the migrant's first crossing attempt.

Table 2

Deviation of Actual Total Distance from Minimal Total Distance By Origins and Destinations (miles)

	Borderlands	Deep South	Great Lakes	Northeast	Northwest	Plains	Southeast
Pre-1994							
Border	52.94	.	.	.	206.00	8.00	39.30
Central	53.73	615.33	1334.49	1770.80	137.00	720.67	1131.69
Historical	78.02	624.33	886.36	1390.03	136.98	597.60	926.72
Southeast	181.49	.	1308.40	.	119.00	1113.50	239.00
Post 1994							
Border	78.69	163.00	37.50	17.20	39.00	46.64	43.33
Central	111.07	1695.17	1065.53	1443.29	78.58	737.84	1100.20
Historical	140.93	1116.67	688.69	1137.44	108.50	638.28	824.15
Southeast	167.85	1267.00	1068.26	1187.57	99.63	652.00	1337.44

Notes. 1. Source: Mexican Migration Project and Google Map. 2. Actual Total Distance: regional average of the shortest road distance from origin communities in Mexico to US destination via the chosen border crossing location. 3. Minimum Total Distance: regional average of the shortest road distance from origin communities in Mexico to US destination via the border crossing that minimizes total road distance. 4. Regional Classification for Mexico: Historical (Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Liu Potosi), Central (Distrito Federal, Guerrero, Hidalgo, México, Morelos, Oaxaca, Puebla, Queretar, Tlaxcala), Border (Baja California, Chihuahua, Coahuila, Nuevo León, Sinaloa, Sonora, Tamaulipas) and Southeast (Campeche, Chiapas, Quintana Roo, Tabasco, Vercruz, Yucatán). 5. Regional Classification for the US: Borderlands (Arizona, California, New Mexico, Texas,), Northwest (Idaho, Nevada, Oregon, Utah, Washington), Great Lakes (Illinois, Indiana, Michigan, Ohio, Wisconsin), Northeast, Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, Pennsylvania, Rhode Island, Vermont, Wyoming), Southeast (District of Columbia, Florida, Georgia, Maryland, North Caroline, South Carolina, Virginia, West Virginia), Deep South (Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Tennessee), Plains (Colorado, Iowa, Kansas, Minnesota, Missouri, Montana, Nebraska, North Dakota, Oklahoma, South Dakota). See Massey et al. (2016).

Table 3

Baseline Model Coefficients (Base alternative: San Diego)							
	Border Sector	El Centro	Tucson	El Paso	Del Rio	Laredo	Rio Grande
x_k: (Border Var.)							
$Dist(o,k)$	-0.112						
(100 miles)	(0.080)						
$Dist(k,d)$	-0.231***						
(100 miles)	(0.017)						
(x_i: Ind. Var)							
d85_89		-0.974***	-0.289	-0.694*	-1.048***	-1.723***	-0.993***
		f (0.367)	(0.329)	(0.364)	(0.353)	(0.360)	(0.345)
d90_94		-0.652*	0.770***	-0.325	-0.866**	-1.656***	-0.988***
		(0.351)	(0.260)	(0.436)	(0.416)	(0.412)	(0.362)
d95_99		-0.013	1.915***	0.667*	-0.302	-1.321***	-0.485
		(0.376)	(0.270)	(0.388)	(0.422)	443)	(0.339)
d00_05		0.374	2.356***	0.396	0.470	-1.368**	-0.358
		(0.422)	(0.369)	(0.563)	(0.464)	(0.538)	(0.544)
Constant		-2.557***	-2.609***	-2.628***	-3.475***	-2.698***	-3.178***
		(0.293)	(0.385)	(0.510)	(0.876)	(0.719)	(0.733)
Observations				16100			
No. of Individuals				2300			
Log Likelihood				-2322.283			
p-value				0.000			

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. The seven border sectors are El Centro, Tucson, El Paso, Big Bend, Del Rio, Laredo, and Rio Grande, with San Diego as the base alternative. 6. *** p<0.01, ** p<0.05, * p<0.1.

Table 4

Model Coefficients with Parental US Networks (Base alternative: San Diego)				
	(1) Full Sample Tucson	(2) Upper Secondary Tucson	(3) Lower Secondary Tucson	(4) Ag. or Manu. Tucson
(x_i: Indi. Var.)				
d85_89	-0.432	-0.329	-0.401	-0.365
	(0.337)	(0.334)	(0.370)	(0.390)
d90_94	0.622**	0.635**	0.441	0.293
	(0.268)	(0.260)	(0.297)	(0.341)
d95_99	1.742***	1.707***	1.524***	1.644***
	(0.275)	(0.270)	(0.311)	(0.361)
d00_05	2.160***	2.036***	2.003***	1.679***
	(0.377)	(0.365)	(0.394)	(0.405)
d80_84 X n_i (Direct)	-1.681*	-1.638	-15.468***	-16.127***
	(1.004)	(1.003)	(0.337)	(0.390)
d85_89 X n_i (Direct)	-0.272	-1.656	-1.414	-1.380
	(0.575)	(1.073)	(1.099)	(1.114)
d90_94 X n_i (Direct)	-0.405	-0.987	-0.506	-0.414
	(0.564)	(0.696)	(0.655)	(0.690)
d95_99 X n_i (Direct)	-0.164	0.129	1.052	0.796
	(0.585)	(0.599)	(0.700)	(0.691)
d00_05 X n_i (Direct)	0.224	0.408	16.976***	17.880***
	(0.637)	(0.687)	(0.919)	(0.904)
Constant	-2.426***	-2.394***	-2.281***	-2.079***
	(0.381)	(0.377)	(0.388)	(0.422)
Observations	16100	14350	10507	7910
No. of Individuals	2300	2050	1501	1130
Log Likelihood	-2300.380	-2043.991	-1467.618	-1158.403
p-value	0.000	0.000	0.000	0.000

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Column 1 includes the full sample; Column 2 includes individuals with upper secondary education or less; Column 3 includes individuals with lower secondary education or less; Column 4 includes individuals with lower secondary education or less and in agricultural or manufacturing occupations. 6. Only coefficients for the Tucson sector are shown. 7. *** p<0.01, ** p<0.05, * p<0.1.

Table 5

Model Coefficients with Indirect US Networks (Base alternative: San Diego)			
(1) Family Networks	Tucson	(2) Community Networks	Tucson
(x_i: Indi. Var.)			
d80_84 X n_i (Family)	-0.859	d80_84 X n_i (Avg. Comm.)	-1.428
	(0.554)		(1.201)
d85_89 X n_i (Family)	-0.815	d85_89 X n_i (Avg. Comm.)	-0.645
	(0.534)		(1.029)
d90_94 X n_i (Family)	-0.726*	d90_94 X n_i (Avg. Comm.)	-1.099
	(0.393)		(1.030)
d95_99 X n_i (Family)	0.168	d95_99 X n_i (Avg. Comm.)	0.727
	(0.361)		(0.845)
d00_05 X n_i (Family)	1.133*	d00_05 X n_i (Avg. Comm.)	2.916
	(0.608)		(3.076)
Constant	-2.242***	Constant	-2.309***
	(0.402)		(0.391)
Observations	10507		10507
No. of Individuals	1501		1501
Log Likelihood	-1473.758		-1471.079
Wald χ^2	3483.526		5495.599
p-value	0.000		0.000

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Family) is a dummy variable equalling 1 if at least one of the individual's parents or siblings has migrated to the US prior to crossing year. n_i(Avg. Community) is a dummy variable equalling the share of individuals in a community who has a direct network tie to the US. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Includes individuals with upper secondary education or less. 6. Only coefficients for the Tucson sector are shown. 7. *** p<0.01, ** p<0.05, * p<0.1.

Table 6

Network Externality, Crossing Cost and Individual Characteristics Least Square Regressions with Two-way FE						
	(1) Cost	(2) Cost	(3) Education	(4) Education	(5) Crossing Age	(6) Crossing Age
Avg Lag (3) Choice Share	-0.308***	-0.241***				
	(0.069)	(0.076)				
Avg Age at First Crossing	-0.005**	-0.000				
	(0.003)	(0.003)				
n_i (Direct)	-0.067	-0.011	1.074***	1.149***	-2.585***	-3.269***
	(0.056)	(0.058)	(0.226)	(0.232)	(0.560)	(0.583)
Observations	1323	1323	2300	2300	2300	2300
R2	0.553	0.633	0.101	0.222	0.095	0.194
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	No	Yes	No	Yes	No
Community Fixed Effects	No	Yes	No	Yes	No	Yes

Note: 1.This table displays OLS regression estimates of log individual crossing cost, year of education, and age at first crossing on network externality and direct network variables with two-way fixed effects. 2. n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Avg Lag (3) Choice Share is the rolling average three-year lagged share of migrants in the community that crossed via the same sector as the migrant. Avg Age at First Crossing is the community level average age of all first time crossers. 4. Columns 1 and 2 include distance to and from border sector, crossing year fixed effects, state / community fixed effects, border sector fixed effects. 5. Columns 3 - 6 include crossing year fixed effects, state / community fixed effects, border sector fixed effects. 6. *** p<0.01, ** p<0.05, * p<0.1.

Table 7

Model Coefficients with Network Externalities, Individual Characteristics, Fencing and Regional Fixed Effects (Base alternative: San Diego)

	(1) Tucson	(2) Tucson	(3) Tucson
(x_k: Border Var.)			
Cum. Fence			-0.047***
			(0.015)
(x_i: Ind. Var.)			
d80_84 X n_i (Direct)	-14.890***	-15.721***	-15.888***
	(0.499)	(0.341)	(0.356)
d85_89 X n_i (Direct)	-1.090	-1.416	-1.251
	(1.133)	(1.097)	(1.112)
d90_94 X n_i (Direct)	-0.012	-0.492	-0.289
	(0.558)	(0.661)	(0.613)
d95_99 X n_i (Direct)	1.236*	1.101	1.332**
	(0.678)	(0.692)	(0.677)
d00_05 X n_i (Direct)	15.678***	17.186***	16.756***
	(0.808)	(0.931)	(0.935)
Years of Educ		0.041	0.010
		(0.052)	(0.049)
Age at First Crossing		0.008	0.005
		(0.010)	(0.010)
Constant	-1.372**	-2.670***	-3.214***
	(0.551)	(0.589)	(1.047)
Observations	9345	10507	10444
No. of Individuals	1335	1501	1492
Log Likelihood	-1158.867	-1465.290	-1313.545
P-Value	0.000	0.000	0.000
Comm Avg lag(3) Choice Share	Yes	No	No
Comm Avg Age First Crossing	Yes	No	No

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Includes individuals with upper secondary education or less. 6. Only coefficients for the Tucson sector are shown. 7. Column 1 includes the rolling average three year lagged border sector choice share as a border sector control x_k, and the community-level average age of first time crossers as individual control x_i; Column 2 includes individual characteristics: year of education and age at first migration; Column 3 includes cumulative fencing (in miles) at each border sector as well as regional fixed effects. 8. Only coefficients for the Tucson sector are shown. 9. *** p<0.01, ** p<0.05, * p<0.1.

Table A1

Baseline Model Marginal Effects (dp/dx)														
	San Diego		El Centro		Tucson		El Paso		Del Rio		Laredo		Rio Grande	
x_k: (Border Var.)														
$Dist(o,k)$	-0.0226		-0.0046		-0.0125		-0.0068		-0.0017		-0.0017		-0.0014	
(100 miles)	0.0162		0.0034		0.0090		0.0054		0.0011		0.0012		0.0011	
$Dist(k,d)$	-0.0465	***	-0.0096	***	-0.0257	***	-0.0141	***	-0.0036	***	-0.0035	***	-0.0029	***
(100 miles)	0.0041		0.0016		0.0035		0.0032		0.0008		0.0006		0.0007	
(x_i: Ind. Var)														
d85_89	0.1295	***	-0.0344	***	-0.0140		-0.0336		-0.0137	**	-0.0236	***	-0.0102	**
	0.0458		0.0155		0.0355		0.0231		0.0059		0.0060		0.0048	
d90_94	0.0019		-0.0281	*	0.0985	***	-0.0211		-0.0137	**	-0.0253	***	-0.0123	***
	0.0415		0.0154		0.0303		0.0284		0.0067		0.0069		0.0047	
d95_99	-0.1845	***	-0.0117		0.2115	***	0.0269		-0.0088		-0.0241	***	-0.0093	**
	0.0438		0.0160		0.0330		0.0220		0.0068		0.0073		0.0046	
d00_05	-0.2337	***	0.0021		0.2590	***	0.0047		0.0023		-0.0259	***	-0.0085	
	0.0646		0.0176		0.0418		0.0321		0.0065		0.0081		0.0071	
Obs.							16100							
# of Cases							2300							
# of Sectors							7							
Log Likelihood							-2322.28							
Wald χ^2							507.01							
P-value							0.0000							

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. The seven border sectors are El Centro, Tucson, El Paso, Big Bend, Del Rio, Laredo, and Rio Grande, with San Diego as the base alternative. 6. *** p<0.01, ** p<0.05, * p<0.1.

Table A2

Model Marginal Effects (dp/dx) with Parental US Networks

	(1) Full Sample		(2) High School or Less		(3) Middle School or Less		4) Ag. Occupations	
	San Diego	Tucson	San Diego	Tucson	San Diego	Tucson	San Diego	Tucson
(x_i: Individual/Community-Specific Variables)								
d85_89	0.1307 **	-0.0326	0.1093 ***	-0.0229	0.1090 **	-0.0168	0.1302 ***	-0.0137
	0.0446	0.0364	0.0419	0.0353	0.0376	0.0208	0.0417	0.0238
d90_94	0.0074	0.0805 ***	-0.0050	0.0776 ***	0.0196	0.0291 *	0.0165	0.0211
	0.0415	0.0312	0.0393	0.0296	0.0331	0.0174	0.0403	0.0211
d95_99	-0.1723 ***	0.1924 ***	-0.1608 ***	0.1847 ***	-0.0808 **	0.0891 ***	-0.1033 **	0.1019 ***
	0.0432	0.0350	0.0407	0.0328	0.0361	0.0182	0.0446	0.0214
d00_05	-0.2195 ***	0.2376 ***	-0.2033 ***	0.2184 ***	-0.1130 **	0.1166 ***	-0.0626	0.1077 ***
	0.0639	0.0427	0.0593	0.0407	0.0530	0.0241	0.0619	0.0245
d80_84	0.2451 **	-0.1721 *	0.2190 **	-0.1667 *	0.8731 ***	-0.9004 ***	0.9663 ***	-1.0038 ***
$\times$ n_i (Direct)	0.1148	0.1039	0.1095	0.1022	0.1292	0.1282	0.1472	0.1505
d85_89	0.1379 *	-0.0110	0.2845 **	-0.1583	0.2250 ***	-0.0712	0.1656 **	-0.0790
$\times$ n_i (Direct)	0.0737	0.0637	0.1113	0.1159	0.0789	0.0650	0.0840	0.0704
d90_94	0.0701	-0.0396	0.6932 ***	-0.0095	1.1963 ***	0.0598	1.4568 ***	0.0939 **
$\times$ n_i (Direct)	0.0821	0.0616	0.1450	0.0758	0.1827	0.0415	0.2217	0.0468
d95_99	0.5533 ***	0.0739	0.4809 ***	0.0938	0.4301 ***	0.0986 **	0.6139 ***	0.1048 **
$\times$ n_i (Direct)	0.1053	0.0644	0.1026	0.0655	0.1110	0.0418	0.1302	0.0445
d00_05	1.0073 ***	0.2009 ***	0.8376 ***	0.1860 **	-0.7677 ***	1.0028 ***	-0.8920 ***	1.1280 ***
$\times$ n_i (Direct)	0.1750	0.0799	0.1654	0.0820	0.1623	0.1522	0.1703	0.1715
Obs.	16100		14350		10507		7910	
# of Cases	2300		2050		1501		1130	
# of Sectors	7		7		7		7	
Log Likelihood	-2300.38		-2043.99		-1467.62		-1158.40	
P-value	0.0000		0.0000		0.0000		0.0000	

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Column 1 includes the full sample; Column 2 includes individuals with upper secondary education or less; Column 3 includes individuals with lower secondary education or less; Column 4 includes individuals with lower secondary education or less and in agricultural or manufacturing occupations. 6. Only coefficients for the Tucson sector are shown. 7. *** p<0.01, ** p<0.05, * p<0.1.

Table A3

Model Marginal Effects (dp/dx) with Indirect US Networks									
(1) Family Network					(2) Community Network				
	San Diego		Tucson			San Diego		Tucson	
(x_i: Individual/Community-Specific Variables)					(x_i: Individual/Community-Specific Variables)				
d80_84	0.0370		-0.0876	*	d80_84	0.1721		-0.1454	
× n_i (Family)	0.0664		0.0525		× n_i (Comm)	0.1316		0.1222	
d85_89	0.0830		-0.0763		d85_89	0.1388		-0.0565	
× n_i (Family)	0.0578		0.0503		× n_i (Comm)	0.1240		0.1075	
d90_94	0.0873	*	-0.0660	*	d90_94	0.9408	***	0.0095	**
× n_i (Family)	0.0501		0.0364		× n_i (Comm)	0.2222		0.1125	
d95_99	-0.0347		0.0132		d95_99	0.9567	***	0.2308	**
× n_i (Family)	0.0572		0.0335		× n_i (Comm)	0.2036		0.1056	
d00_05	0.1910		0.1502	***	d00_05	1.4035	***	0.5603	
× n_i (Family)	0.1337		0.0586		× n_i (Comm)	0.5368		0.3160	
Obs.	10507				Obs.	10507			
# of Cases	1501				# of Cases	1501			
# of Sectors	7				# of Sectors	7			
Log Likelihood	-1473.76				Log Likelihood	-1471.078			
P-value	0.00				P-value	0.00			

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Family) is a dummy variable equalling 1 if at least one of the individual's parents or siblings has migrated to the US prior to crossing year. n_i(Avg. Community) is a dummy variable equalling the share of individuals in a community who has a direct network tie to the US. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Includes individuals with upper secondary education or less. 6. Only coefficients for the Tucson sector are shown. 7. *** p<0.01, ** p<0.05, * p<0.1.

Table A4

Model Marginal Effects (dp/dx) with Network Externalities, Individual Characteristics, Fencing and Regional FE

	(1)				(2)				(3)			
	San Diego		Tucson		San Diego		Tucson		San Diego		Tucson	
(x_i: Individual/Community-Specific Variables)												
d80_84	1.2299	***	-0.9362	***	0.8773	***	-0.9059	***	0.6519	***	-0.6565	*
× n_i (Direct)	0.1352		0.1242		0.1301		0.1290		0.1283		0.1282	
d85_89	0.1911	**	-0.0601		0.2243	***	-0.0707		0.2434	***	-0.0416	
× n_i (Direct)	0.0790		0.0731		0.0780		0.0642		0.0721		0.0475	
d90_94	0.8967	***	0.0739	**	1.2036	***	0.0603		1.2259	***	0.0518	*
× n_i (Direct)	0.1137		0.0389		0.1828		0.0419		0.1718		0.0307	
d95_99	0.3696	***	0.1170	***	0.4286	***	0.1004	**	0.5208	***	0.0852	**
× n_i (Direct)	0.1007		0.0437		0.1105		0.0413		0.1429		0.0337	
d00_05	-0.8970	***	1.0189	***	-0.7851	***	1.0034	***	-0.6296	***	0.6954	***
× n_i (Direct)					0.1677		0.1527		0.1611		0.1436	
Educ. (yrs)					-0.0065		0.0021		-0.0052		0.0002	
					0.0058		0.0029		0.0055		0.0020	
Crossing					-0.0014		0.0004		-0.0025	*	0.0001	
Age					0.0015		0.0005		0.0014		0.0004	
Obs.	9345				10507				10444			
# of Cases	1335				1501				1492			
# of Sectors	7				7				7			
Log Likelihood	-1158.87				-1465.29				-1313.55			
P-value	0.00				0.00				0.00			

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model. 2. dx_y denotes the interval from year x to year y, and n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Includes individuals with upper secondary education or less. 6. Only coefficients for the Tucson sector are shown. 7. Column 1 includes the rolling average three year lagged border sector choice share as a border sector control x_k, and the community-level average age of first time crossers as individual control x_i; Column 2 includes individual characteristics: year of education and age at first migration; Column 3 includes cumulative fencing (in miles) at each border sector as well as regional fixed effects. 8. Only coefficients for the Tucson sector are shown. 9. *** p<0.01, ** p<0.05, * p<0.1.

Table A5

Baseline Model Coefficients with Parental US Networks (Base alternative: San Diego) All 9 Sectors)				
	(1) Full Sample Tucson	(2) Upper Secondary Tucson	(3) Lower Secondary Tucson	(4) Ag. or Manu. Tucson
(x_i: Indi. Var.)				
d85_89	-0.321	-0.202	-0.276	-0.330
	(0.319)	(0.318)	(0.351)	(0.382)
d90_94	0.761***	0.787***	0.652**	0.345
	(0.258)	(0.250)	(0.289)	(0.332)
d95_99	1.878***	1.850***	1.681***	1.671***
	(0.274)	(0.269)	(0.307)	(0.353)
d00_05	2.214***	2.096***	2.051***	1.698***
	(0.373)	(0.360)	(0.387)	(0.403)
d80_84 X n_i (Direct)	-0.948	-0.921	-15.004***	-15.552***
	(0.705)	(0.703)	(0.323)	(0.381)
d85_89 X n_i (Direct)	-0.314	-1.713	-1.445	-1.392
	(0.559)	(1.065)	(1.091)	(1.112)
d90_94 X n_i (Direct)	-0.288	-0.704	-0.272	-0.091
	(0.488)	(0.564)	(0.535)	(0.579)
d95_99 X n_i (Direct)	-0.137	0.109	0.851*	0.578
	(0.493)	(0.492)	(0.510)	(0.548)
d00_05 X n_i (Direct)	0.239	0.411	16.704***	17.470***
	(0.642)	(0.697)	(0.918)	(0.905)
Constant	-2.575***	-2.541***	-2.451***	-2.154***
	(0.385)	(0.389)	(0.399)	(0.430)
Observations	22023	19620	14427	10503
No. of Individuals	2447	2180	1603	1167
Log Likelihood	-2577.271	-2291.566	-1664.822	-1280.422
Wald χ^2	30229.432	32433.675	3.07e+07	1.03e+08
p-value	0.000	0.000	0.000	0.000

1. This table displays the determinants of border crossing choice estimated using the alternative specific conditional logit model and data from crossings via 9 sectors. 2. dx_y denotes the interval from year x to year y, and n_i(Direct) is a dummy variable equalling 1 if at least one of the individual's parents has migrated to the US prior to crossing year. 3. Includes distance to and from border sector ($Dist(o,k)$ and $Dist(k,d)$) as alternative (border) specific variables. 4. Standard errors are clustered at the community level. 5. Column 1 includes the full sample; Column 2 includes individuals with upper secondary education or less; Column 3 includes individuals with lower secondary education or less; Column 4 includes individuals with lower secondary education or less and in agricultural or manufacturing occupations. 6. Only coefficients for the Tucson sector are shown. 7. *** p<0.01, ** p<0.05, * p<0.1.