Migration Decisions of Mexican Workers: Within and Across the Borders*

Gihoon Hong

Indiana University South Bend

January, 2012

Abstract

This paper develops a discrete choice dynamic structural model that explains migration decisions of Mexican workers within Mexico and across the American border. Each Mexican worker in the model makes a sequence of location decisions and has an option to apply for an immigration document when he chooses out-migration to the U.S. The model accounts for differentials in wages and location-specific amenities as determinants of migration decisions, and network effects influence the value of the U.S. location choice. The structural parameters of the model are estimated using source-country based retrospective panel data from the Mexican Migration Project. I use the results to evaluate the impact of three alternative immigration policies: an increase in the number of available visas, an increase in the number of border patrol officers, and tougher control on unauthorized residence in the U.S. Simulation results show that Mexican migrants are very responsive to changes in pecuniary and non-pecuniary utility flows associated with illegal employment, as the proportion of U.S. immigrants decreases by as much as 60% in response to a policy that reduces the utility flows related to illegal U.S. residence.

JEL classification: J24, J3, J61, F22

Keywords: Mexican immigration, Discrete choice dynamic programming model, Structural estimation, Immigration policy

*Work in Progress. I am grateful to Steven Stern and Leora Friedberg for their guidance and encouragement. I owe special thanks to Chris Woodruff for providing me with the distance to rail lines data. Financial support from the Bankard Fund for Political Economy at the University of Virginia is gratefully acknowledged. All errors are mine.
E-mail for correspondence: honggi@iusb.edu

1 Introduction

The number of Mexican workers in the U.S. has surged over the last few decades. In 1970, the Mexican-born population in the U.S. was 800,000. In 2000, it was nearly ten times as large, with about 400,000 more Mexicans arriving each year. In 2008, it was estimated that 13 million Mexicans lived in the U.S., which represented approximately 11% of Mexico's population.[1] Since Mexican immigrants represent a growing share of the nation's total foreign-born population, recent discussions about U.S. immigration policy are closely connected with Mexican immigration.[2] For example, the debate over the optimal number of working visas and green cards assigned to foreigners or the issue of U.S.-Mexico border enforcement is centered around its impact on Mexican immigration.

Many researchers have sought to understand what drives the massive immigration flow from Mexico to the U.S. Since the seminal paper by Sjaastad (1962), neoclassical economists view migration as the reallocation of workers in response to differentials in wages and other living standards across locations. Prospective migrants undertake a cost-benefit analysis and move to a new location when the net benefit of moving is positive. This simple theoretical framework has been extended to analyze migration across nations. For example, Borjas (1987) assumed that potential immigrants compare expected income in the host and source countries and make the immigration decision based on income differentials (net of moving costs). But, what is essential in studying migration is that a decision to move incurs significant fixed costs that need to be paid immediately, while the benefits from the move are uncertain and accrue over time. This implies that the migration decision has to be viewed as an "investment" within a dynamic perspective. If expectations are rational and the components that affect the costs and benefits of the U.S. immigration are time-varying, then static migration models will generate biased estimates. For example, ignoring dynamics would underestimate moving costs and the effect of intensified border controls on migration if individuals expect continued tighter border enforcement in the future. This suggests that a policy change which aims to reduce the stock of undocumented immigrants by increasing the cost of illegal migration may not be effective because it could lead more Mexicans to stay in the U.S. for the fear of higher costs associated with returning to the U.S. in the future.

[1] Using the March 2000 CPS, Camarota (2001) estimated that illegal aliens in the U.S. amount to 10 million, among which 40% are Mexican nationals.

[2] According to Hanson (2006), Mexican immigrants in the United States were 31% of the U.S. foreign-born population in 2004.

There are at least two further reasons why a dynamic model is necessary to study the migration behavior of Mexican workers. First, in addition to its large scale, a distinguishing feature of Mexican immigration is that migration events of Mexican workers have tended to be circular.[3] Many Mexican immigrants return to their place of origin after spending a few years in the U.S. This implies that Mexican immigrants make decisions about how long to move as well as whether to move. Second, the decision to apply for a legal document to work in the U.S. also requires dynamic considerations. Given that applicants wait in line for varying amounts of time after application costs are paid and that the benefit from becoming a legal immigrant will be realized only in the future, it would be difficult to justify application decisions observed in the data without a structure that allows the past sequence of choices to affect present alternatives.

In this study, I develop and estimate a dynamic structural model of migration decisions with many alternative location choices. A potential migrant in the model has the option of moving both within Mexico and to the U.S. The dynamic structure of the model sheds light on the timing of migration, return migration, and the duration of each stay. An optimal sequence of migration and application decisions are influenced by differentials in wages and amenity values across locations. An individual's network, defined as the size of a migrant's community in the U.S. who originated from the same Mexican community, also affects the value of locating in the U.S. In addition to choice-specific random shocks, uncertainty in the model arises from the stochastic nature of occupation transitions, the length of the waiting period associated with visa application decisions, and the possibility of apprehension at the border. The structural model is estimated using maximum likelihood estimation techniques and data from the Mexican Migration Project (MMP), which contains detailed information on location and occupation histories of many households in Mexico. I restrict the sample to the group of male household heads aged between 16 to 20 years old in 1965 in order to reduce the computational burden.

My contribution to the literature on Mexican immigration is twofold. First, I introduce legal immigration as a possible choice in the model – a crucial consideration, as about 6 million Mexicans are legal residents of the U.S. Combined with a measure of enforcement against illegal flows, this allows me to simulate the effect of relevant policy changes such as increases in expenditure on

[3] Massey et al. (1987) observed that "the Mexican migrants are temporary rather than permanent immigrants to the United States."

border enforcement or in visa availability. Second, I introduce unobserved heterogeneity in many dimensions. Unobserved person-specific effects have been important in explaining job durations in the labor literature, and it appears to be an important consideration in migration models, too.[4] In case individuals are different in terms of unobserved location preference, neglecting to control for unobserved heterogeneity in a dynamic migration model would result in biased estimates of structural parameters associated with duration dependence. Also, considering that domestic location choices within Mexico in the model are likely to be more similar to one another in terms of observed and unobserved attributes relative to the U.S. location, allowing for correlations between choice-specific errors is necessary to minimize potential biases that may arise from distributional assumptions. In addition, it is intuitively appealing to allow for persistent difference in unobserved person-specific tastes. For example, unobserved location preference in non-pecuniary utility helps explain why we observe individuals staying in the U.S. as illegal workers even when income prospects are not high.

The parameter estimates of the structural model indicate that the migration decisions of a Mexican worker are positively associated with the size of a migrant's community at the U.S. location. Also, an authorized foreign worker is found to enjoy higher wages and amenity values relative to undocumented immigrant workers during the U.S. residence after conditioning on qualifications. I find that the estimated coefficients of the moving cost function are significant and large in magnitude, which confirms the importance of considering dynamics in studying migration. Counterfactual experiments show that Mexican migrants are very sensitive to more stringent policies against illegal employment, as the proportion of U.S. immigrants decreases by as much as 60% in response to a policy that reduces the utility flows related to illegal U.S. residence. Also, I find that a 30% decrease in the U.S. wages associated with illegal employment leads to about 70% decrease in the percentage of undocumented immigrants among Mexican workers in the U.S. because potential migrants without immigration documents at the margin might pursue a legal immigration option or might not emigrate when such a policy change is expected to persist.

The paper proceeds as follows. Section 2 reviews related literature and provides a background on the history of Mexican immigration to the U.S. and U.S. immigration policy. Section 3 presents the dynamic migration model of Mexican workers. Section 4 illustrates the econometric model and

[4]See Berkovec and Stern (1991).

discusses the estimation approach. Section 5 describes the data used in the empirical analysis. Section 6 discusses non-structural estimation results. Section 7 presents the structural estimates and provides analysis of the main results. Section 8 presents the results from the policy experiments. Section 9 outlines future directions for this research and concludes.

2 Background

In this section, I provide background knowledge that is closely related to this research. First, I review the relevant literature on migration and Mexican immigration. Then, I discuss the history of Mexican immigration to the U.S. and U.S. immigration policy controlling the flows of immigrants. Finally, I briefly describe the legal immigration process to the U.S. and border enforcement efforts at the U.S.-Mexico border.

2.1 Related Literature

There is a growing literature on migration.[5] In particular, the increased availability of micro-level data over the last few decades that contains information on both migration and occupation histories at the individual level has made it possible for economists to analyze how migration decisions are related to other important economic behavior such as marriage/divorce (Holt 1997, Gemici 2008) or retirement (Dresher 1994, Duncombe et al. 2000). Each of the aforementioned studies offers a unique perspective on migration, answering important questions with regard to the determinants of migration/return migration, duration of stay, self-selection in the migration process, etc.

The migration literature that is most relevant to my research can be classified into two groups and is briefly summarized here. The first group of studies attempts to identify the driving forces of Mexican immigration. Hanson and Spilimbergo (1999) studied the effects of U.S.-Mexico wage differentials and U.S. border enforcement efforts on illegal immigrant flows from Mexico. Using an aggregate measure of the migration rate, they showed that a 10 percent decline in Mexican real wages is associated with a 6 to 8 percent increase in border apprehensions. They also found that border apprehensions increase as border enforcement increases, which suggested that expanded enforcement made crossing the border more costly. A recent development in the economic theories

[5]See Holt (1997) for an extensive survey on the migration literature.

on migration emphasized the importance of migrants' social capital in lowering the cost of migration. Potential immigrants who are tied to a network of friends or relatives might find a better job with less effort than those who are not. However, simply controlling for network size in an empirical model is likely to generate endogeneity due to the correlation of unobserved characteristics common to leaders and followers. Massey and Espinosa (1997) used the existence of family in the U.S. as a measure of one's network but failed to consider the correlation of unobserved tastes within the family. Hanson and Woodruff (2004) used the historic state-level migration rate to study the relationship between household migration behavior and schooling in Mexico, but this raises concerns about possible location-specific unobserved components. Munshi (2003) used the lagged rainfall in the origin village as an instrument to assess the effect of networks on migrants' labor market outcomes. But, his sample was limited to small agricultural villages because of the nature of the instrument, and the dependent variable in the estimation was occupation status instead of the actual wage.

The second group of studies used discrete choice dynamic programming to model the migration decision. Even though the dynamic programming approach using Bellman's equation seems to be well-suited to migration studies, it has not been used much until recently in the migration literature because of the computational complexity in estimation when the number of location choices is large. Also, because the determinants of costs and benefits associated with each location choice in migration models differ across population segments, the existing literature tends to focus on the group of individuals who share similar demographic characteristics.[6] For example, Kennan and Walker (2008) analyzed the migration patterns of young workers and found that income prospects are the most important determinant of migration among workers in their twenties and thirties. They addressed the computational issue by limiting the role of past migration history. Woo (2005) used a similar model to Kennan and Walker and examined the migration decisions of older workers approaching retirement. He found that retiree migration behavior is driven largely by local amenities and fiscal policy rather than income prospects. Holt (1997) employed a Nash bargaining framework to account for marriage and divorce within a dynamic migration model. Using data from the Panel Study of Income Dynamics, he showed that the cost of moving is higher when the agent

[6] Walker (2006) noted, "The theoretical framework (of migration study) is general, but its empirical implementation is context driven."

is part of a family, relative to when the agent is single. Finally, Colussi (2006) used the same data I use and developed a dynamic general equilibrium model to assess the effect of networks on the migration decision, but his sample was limited to small agricultural villages. I extend his efforts by including legal versus illegal immigration and domestic migration choices, and by including every originating community in the sample.

2.2 Mexican Immigration

Massive Mexican immigration began with the implementation of the Bracero program, as the flow of workers from other countries started to decrease in the early 20th century. The Bracero program was a large scale guest worker program that issued temporary visas to workers in Mexico, and under it the number of Mexican workers admitted to the U.S. steadily increased until the 1950's. When the program ended in the 1960's, the U.S. Congress began to control the flow of workers from Mexico by reducing the number of available visas. Reductions in quotas made in 1968, 1976, and 1980, however, failed to prevent Mexicans from crossing the border illegally.

In 1986, in an effort to reduce the benefits of illegal migration, the U.S. Congress enacted the Immigration Reform and Control Act (IRCA), which imposed sanctions on employers who knowingly hired or recruited illegal immigrants. The IRCA also awarded immigrant visas to undocumented immigrants who had been continuously and illegally present in the U.S. since 1982, and about 2 million Mexicans were awarded permanent residence through the amnesty.[7]

During the 1990s, a series of additional efforts were made to reduce illegal border crossings, but the number of illegal workers has increased drastically throughout the last three decades. In 1996, for example, President Clinton signed the Illegal Immigrant Reform and Immigrant Responsibility Act of 1996 (IIRIRA 96). The act instituted upgrades needed for border patrol enforcers, equipment, and the overall patrolling process; increased domestic enforcement and practices of the Immigration and Naturalization Service (INS), including employer sanctions and expedited removal; and reduced eligibility for legal immigration. For example, the bill required aliens to get various vaccination shots and submit greater documentation as part of immigrant visa applications. IIRIRA also created the notion of "unlawfully present" aliens, stating that if an immigrant has been unlawfully present in the United States for 180 days but less than 365 days he or she must remain

[7] See Hanson (2006).

outside the United States for three years. If the person has been in the United States for 365 days or more, he or she must stay outside the United States for ten years. This placed a greater restriction on immigrants who enter with tourist visas and then overstay the tourist visa and work illegally.[8]

In 2007, some attempts to make major changes in U.S. immigration law were made, as the STRIVE act of 2007 and Senator Arlen Specter's immigration-reform bill were introduced.[9] Both proposals aimed to increase the number of visas, especially in employment-based immigration category, but neither was ever voted on, even though a number of amendments to the bill were considered. On March 18, 2009, President Obama announced his plan to address the nation's immigration system and made it clear that the administration would push for a comprehensive immigration reform bill. The proposed legislation that followed include the "DREAM Act" which would provide certain undocumented alien students who arrived in the U.S. as minors the opportunity to earn conditional permanent residency, and the "AgJOBS Act" which identifies undocumented farm workers and allow those already working in U.S. agriculture to continue to work in the United States legally if they first pay a fine, show that they are current on their taxes, have clean criminal records, and commit to working in U.S. agriculture for the next five years.

2.3 Legal Immigration to the U.S.

The majority of new Mexican immigrants in the U.S. are undocumented. A significant number of qualified Mexicans, however, immigrate to the U.S. with legal documents. According to historical data from the Office of Immigration Statistics, an average of about 175,000 Mexicans obtained legal immigrant status annually between 2000 to 2007, which exceeds the number of new legal immigrant flows from three other major source countries (China, The Philippines, and India) combined.[10]

The U.S. legal immigration system is complex, but there are two common ways of obtaining an immigrant visa among Mexicans: family-based immigration and employment-based immigration.[11] Under family-based immigration provisions, immigrants obtain visas through the sponsorship of a

[8]The law also authorized immigration officers to deny immigrant visas to persons deemed likely to become a public charge by considering factors such as age, health, family status, financial resources, education, and skills.

[9]The Security Through Regularized Immigration and a Vibrant Economy Act of 2007 was a bipartisan comprehensive immigration reform bill, introduced by Representative Luis Gutierrez (D-IL) and Jeff Flake (R-AZ).

[10]Source: 2008 Yearbook of Immigration Statistics

[11]Other ways to obtain immigrant visas include asylum and refugee status, the diversity lottery, and amnesty.

close family member who is either a U.S. citizen or permanent resident.[12] Waiting periods before obtaining visas vary depending on the closeness of the family relationship, ranging from 0-2 years for a spouse of a U.S. citizen to more than 10 years for a sibling, due to the limited number of available visas each year. Employment-based immigration, denoted by "EB" hereafter, refers to the process in which employers sponsor a potential employee for immigration. Before a potential migrant worker submits a visa application based on the EB provision to the United States Citizenship and Immigration Services (USCIS), an employer sponsor must demonstrate that there are insufficient number of qualified workers available in the U.S. to fill the position in need.[13] The waiting time for obtaining visas through the EB process also varies, depending on the applicant's qualifications in terms of education and work experience. The wait may last 1-2 years for migrants with master's or baccalaureate degrees with at least 5 years of experience in the profession (EB-2), but it can be more than 7 or 8 years for an unskilled workers (EB-3).

Prospective legal immigrants face different application costs depending on the category under which each application falls. For example, while the cost of pursuing family-based immigration only includes the fees paid to the USCIS, employment-based immigration involves additional costs such as obtaining labor certificates from the Department of Labor and advertising. In addition, the fees that immigration attorneys charge are much higher for EB cases.[14]

2.4 U.S. Border Enforcement Efforts

The Office of Immigration Statistics Annual Report (2009) showed that the Department of Homeland Security (DHS) apprehended 792,000 foreign nationals in 2008, among which 88% were Mexican nationals.[15] Given the scale of illegal immigrants flows from Mexico, the DHS uses considerable

[12]To be a sponsor of an immigrant relative, one must show that his household income is equal to or higher than 125 percent of the U.S. poverty level adjusted for the size of the household.

[13]In rare cases, extraordinarily-qualified immigrants who possess special skills may immigrate without a sponsor (EB-1 category).

[14]As of 2010, the application fee is $355 for I-130 (Petition for Alien Relative), $475 for I-140 (Immigrant Petition for Alien Worker), and $1,010 for I-485 (Application to Register Permanent Residence).

[15]This number does not include undocumented immigrants who were apprehended, detained, or deported from within the U.S. by U.S. Immigration and Customs Enforcement (ICE), the primary agency responsible for the detection of undocumented immigrants inside of the U.S. Even though apprehensions have increased greatly over the last decade due to the increased efforts enforcing immigration laws within the U.S., I do not model the possibility of apprehensions and removals once in the U.S. for two reasons. First, apprehensions at the border are still far more common than within the U.S. According to the 2009 Yearbook of Immigration Statistics (http://www.dhs.gov/files/statistics/publications/archive.shtm#1), only about 9.2% of "deportable aliens" were located by ICE in 2009, which had increased from about 4.5% in 1992. It is likely that the percentage was even lower during the earlier period (1965-1995) I am focusing on in this paper. Second, information on deportations within the

resources to undertake immigration enforcement at the U.S.-Mexico Border. The U.S. Border Patrol, which is part of U.S. Customs and Border Protection (CBP) within the DHS, is the primary agency responsible for the deterrence or apprehension of illegal immigrants between ports of entry. Figures 2-4 display time series plots of a variety of inputs the U.S. Border Patrol devoted for border enforcement between 1965 and 2003.[16] Figures 2 and 3 show that the border patrol enforcement budget and the number of linewatch hours spent patrolling the U.S.-Mexico border have nearly tripled since 1990. Figure 4 also shows that the number of border patrol officers has increased dramatically over the last four decades.

The massive increase in the resources devoted to border enforcement, however, has not effectively deterred influxes of illegal immigrants. Even though the number of apprehensions has risen over the period (Figure 5), the rate of increase is lower than that of increase of the enforcement budget. In addition, the number of Mexican immigrants in the U.S. still grew to be about 8 million in 2000 from 4.3 million in 1990. The MMP124 has information on the likelihood of arrest while attempting to cross the border with false or no documents calculated with the MMP114, and it also shows that the effect of expanded border enforcement is modest (Figure 6).[17]

3 Model

Migration events are common among the Mexican population. Table 1 shows that approximately 67 percent of the sample in the MMP124 data set (described later) moved at least once over the course of their lifetime, either domestically or internationally. The dynamic migration model presented in this section seeks to explain why an individual chooses to move and what affects the values associated with each location choice. First, I describe a potential migrant's choices each period, then the utility gained from each possible choice, and finally the solution approach to the lifetime utility maximization problem.

A prospective immigrant in Mexico solves a finite horizon, discrete time, dynamic programming problem. He enters the model at some initial time $t = 1$ and starts making location decisions every

U.S. is not available in the MMP.

[16]Source: MMP124

[17]However, calculation of the apprehension probability in the MMP did not account for person-specific characteristics. This is problematic because a significant number of Mexican migrants are observed to make multiple attempts within a given year when they were caught at the border.

year until $t = T$ to maximize the discounted sum of expected lifetime utility. An individual's choice set at time t is denoted by $D_t = [l_t, a_t]$, where l_t represents the location decision and a_t is the visa application decision. I allow up to 3 domestic location choices within Mexico, *North* ($l_t = 1$), *West* ($l_t = 2$), *South* ($l_t = 3$), and the individual has the option to move to the U.S. ($l_t = 4$). In addition, an individual makes a decision regarding whether to apply for an immigration document.[18] A potential migrant who chooses to apply for the document is required to pay application costs, which include the possible cost of finding a sponsor, as well as monetary costs associated with preparing and filing an application.[19] I denote a_t as a dummy that takes a value of 1 if an individual chooses to apply for the document at time t. Once he applies at time t, he waits until the application is approved and cannot apply again in the future.

Thus, the individual chooses from up to 8 alternatives in every period, $d_t = 1, ..., 8$, each of which corresponds to the unique combination of $(l_t,\, a_t)$ pair, where the last four choices are associated with the individual applying for the immigration document at time t. Assuming that the immigrant is allowed to apply for an immigration document only once, anyone who has applied in the past, $a_s = 1,\, for\, s < t$, makes the location decision only at time t.[20] Furthermore, the choice set is location-specific in that a Mexican immigrant in the U.S. is not allowed to change the document status. It means that choice combinations associated with the application decision ($a_t = 1$) are excluded from the U.S. choice set.[21] The choices of illegal border crossing and legal immigration, as discussed later, introduce additional uncertainty because of the possibility of apprehension at the border and randomness in waiting time involved in legal immigration applications.[22]

[18] A potential migrant may be eligible for an immigration document if he has someone in the U.S. who may agree to be a sponsor as an employer or as a close relative who is either a U.S. citizen or permanent resident. Given that eligibility for legal immigration is not observed in the data, I deal with heterogeneity in eligibility within the sample using variation in waiting time in the visa application process as a function of person-specific qualification measures such as education, work experience, and network. In addition, I allow waiting periods in visa application to vary with unobserved heterogeneity.

[19] In principle, application costs as well as the degree of split in costs between employer and employee vary significantly across cases. However, because neither employer decisions about whom to hire nor employee decisions about which visa category to choose is included in the model, I do not allow for variation in application costs.

[20] I assume that every visa application is approved. In the MMP124, the denial rate is less than 1 percent.

[21] This restriction is made for two reasons. First, the application process for immigration documents within the U.S., called "Adjustment of Status (AOS)", is generally different from the process outside of the U.S., called "Consular Processing (CP), as the former tends to involve additional steps to change the non-immigrant visa status from one that does not allow "dual intent" (e.g. B or F visas) to another that does (e.g. H-1B). "Dual intent" is said to be allowed if foreigners possessing certain non-immigrant U.S. visas are allowed to show immigrant intent. Second, given the small number of AOS cases observed in the data, I choose to ignore the possibility of changing document status within the U.S. However, to the degree that Mexican immigrants strategically enter the U.S. illegally and seek an opportunity to become legalized through either marriage with a citizen or amnesty, the model would be misspecified.

[22] I assume that the probability of apprehension is zero once in the U.S.

3.1 Utility Flows

Individual i gets a per-period utility flow $U_{ijt}(d_t)$ consisting of wages and non-pecuniary benefits when choosing d_t in location j at time t. I assume that each location choice differs in two dimensions: the means of the wage distribution and of the non-pecuniary amenity value, both of which are known to agents at time 0.[23]

Wages

Wages depend on observable fixed and time-varying person-specific characteristics, location, and a random shock. Then the wage offer a Mexican worker i receives in location j when employed in occupation q at time t is written as

$$W_{ijt}^{q} \quad = \quad w_q(S_{it}^{w}) + \eta_{1j} + e_{ijt}. \tag{1}$$

The term $w_q(S_{it}^{w})$ denotes the deterministic component of the wage that is a function of a vector S_{it}^{w}, which includes human capital acquired in both countries, education, duration of residence, legal status, and a worker's social network.[24] The next term η_{1j} is a coefficient on the location dummy that shifts the mean of wages for all workers at location j relative to the base Mexican location – *South*. True randomness in the wage is captured by e_{ijt}. A worker's occupation evolves according to an exogenous stochastic function. While this is an obvious simplification, it is necessary because it is much less computationally expensive than treating occupation as endogenous.[25] I assume that a probability distribution governs the likelihood of getting job offers from each occupation: *unemployed*, *agriculture*, *blue collar*, *white collar*.[26] The location-specific occupation transition probability varies by last period's occupation (q_{t-1}) and location (j_{t-1}), age, education (*educ*), age at migration (age_0) and the last occupation held in Mexico (q_0) for an immigrant worker

[23]Alternatively, it is possible to allow for learning about the wage distribution in alternative locations. In this case, a worker can draw a wage only by visiting a location, thereby incurring a moving cost. See Kennan and Walker (2008) for details.

[24]Education is treated as exogenous.

[25]To the extent that agents are self-selected into occupations in which their innate ability is high, parameter estimates of the occupation transition functions are biased.

[26]I assume that at most one job offer arrives within a single period and that a worker accepts any job offer with probability of 1.

in the U.S., and it is written as

$$q_t = f(q_{t-1}, j_{t-1}, j_t, age_t, age_0, educ, q_0, \epsilon_t). \tag{2}$$

I assume that $\epsilon_t \sim$ i.i.d. extreme value, which implies that the occupation transition equations are specified as a multinomial logit process.

Non-pecuniary Utility Flows

The choice-specific non-pecuniary component of utility corresponding to the choice combination that person i receives in location j at time t, denoted by $H_{ijt}(d_t)$, is composed of a deterministic function of observable characteristics, a person-specific preference for location, a location-specific effect, and a choice-specific random shock to non-pecuniary benefits. The non-pecuniary utility flow equation is

$$H_{ijt}(d_t) = h_j(S^h_{it}) + \varphi^j_i + \eta_{2j} + \varepsilon_{ijt}(d_t). \tag{3}$$

The first term $h_j(S^h_{it})$ represents the deterministic component of the non-pecuniary utility, which includes observable characteristics that may influence the amenity value or psychic cost of living in location j. The endogenous state variables in S^h_{it} include education, duration of residence, and the indicator of document status in case the U.S. location is chosen, as legal status confers numerous non-pecuniary benefits.[27] Also, I allow Mexican migrants to have a bias for their home location. Person i's unobserved preference for location j is denoted by φ^j_i. The next component of the non-pecuniary utility flows η_{2j} represents location-specific amenity values of location j relative to the base location that is common across all individuals in location j. The error $\varepsilon_{ijt}(d_t)$ captures the effect of idiosyncratic shocks on the utility flow associated with the choice combination d_t that are not explained by other variables in the model.

[27]For example, legal immigrants in the U.S. are eligible for government or state-sponsored welfare programs such as TANF (Temporary Assistance for Needy Families).

State Space

The location decision at time t depends on age, the current location (j_t), tenure in the current location (z_t), current occupation (q_t), work experience in the current occupation both in Mexico and in the U.S. (hc_{it}), years of U.S. experience in total (x_t), current visa status (v_t), and years since submitting an application for a document (y_t). In addition, each agent in the model is endowed with a set of permanent unobserved heterogeneity terms in the non-pecuniary flow, the moving cost function, and the waiting time function, all of which are known to them at time 0 but not to the econometrician. To reduce the computational burden arising from having to evaluate value functions in each state point in the state space, I make simplifying assumptions. First, I assume that only the most recent occupation determines the occupation-specific work experience and that duration and tenure effects are constant after 5 years. Second, I consider the baseline hazard of exiting the waiting spell as constant 2 years after submitting a visa application, which implies that I have to keep track of only three values for y_t. Third, I assume that individuals in the model have perfect foresight regarding the future values of two policy variables: intensity of border enforcement and number of available visas.[28] This reduces the size of state space significantly because I do not need to keep track of these additional state variables. However, there are five different age groups in any given period, each of which faces a different series of policy variables as they age. This implies that I still need to solve for value functions separately for each cohort. Given these restrictions on state variables and setting the number of different values that human capital can take at 3, and the length of time horizon at 30, the size of state space is $4(j) \times 4(q) \times 5(x) \times 3(hc_{mx}) \times 3(hc_{us}) \times 5(z) \times 2(v) \times 3(ysa) \times 30 = 648,000$, which is manageable.[29]

3.2 Value Functions

Given the specification of per-period utility flows as a sum of the wage and non-pecuniary benefits and an individual's discount factor, β, I can write down the optimization problem each individual

[28]The model can incorporate uncertainty about the time-varying aggregate variables by assuming that they follow some joint stochastic path, such as a VAR process. In this case, agents in the model will form an expectation about the future values with a distribution. The computational burden, however, would be considerably larger depending on the number of possible realizations of the new state variables.

[29]I approximate unobserved occupation-specific human capital using work experience in the current occupation.

solves at time t as

$$V\left(\Omega_t, t\right) = \underset{\{d_s\}_{s=t}^{T}}{max}\ E\left[\sum_{s=t}^{T} \beta^{s-t} U_s|\, \Omega_t\right] = U_t\left(\Omega_t, t\right) + \beta \cdot E\left[\underset{d_t \in D_t}{max} V_{d_t}(\Omega_{t+1},\, t+1)|\, \Omega_t\right]. \tag{4}$$

Let V_{d_t} be the value of choosing d_t at time t. The expectations are taken over those components that are random to the agent in the model, which includes shocks to the wages and the choice-specific utility flows, stochastic occupation transitions, and uncertainty in the visa waiting time. Following Berkovec and Stern (1991), the discount factor β is set equal to .95 instead of being estimated.[30] Ω_t is the information set that contains observed state variables both from the wage (S_t^w) and non-pecuniary flows (S_t^h) equations, person-specific heterogeneity, demographic characteristics, and the realization of idiosyncratic shocks at time t. The set of observed state variables includes location, occupation, duration in the current location, years of U.S. experience, work experience, document status, and years since submitting an application in case one has applied for a visa in the past.

Expectations are taken over the distributions of idiosyncratic shocks. Besides the sources of uncertainty in the utility flows, individuals face uncertainty about whether they can complete their migration plans. Since the choice set D_t depends on the location in the previous period, it is useful to represent the maximization problem in terms of location-specific value functions.

First, suppressing Ω_t and letting $V_t^{MX_m}$ denote the value function of an individual located in region m in Mexico at time t, the value function is

$$V_t^{MX_m} = U_t^{MX_m} + \beta \cdot E\left[\underset{d_t \in D_t^{MX}}{max} V_{t+1}|\, d_t\right],\ for\, m = North,\, West,\, South. \tag{5}$$

When the individual chooses to move to the U.S. without a document, the value function next period is

$$EV_{t+1} = (1 - A_{it}) \cdot EV_{t+1}^{US_{illegal}} + A_{it} \cdot EV_{t+1}^{MX_m} - C_{illegal}, \tag{6}$$

where $EV_{t+1}^{US_{illegal}}$ is the value function associated with illegal residence in the U.S., and $C_{illegal}$ denotes the cost of moving to the U.S. illegally, which includes a start-up cost as well as a trans-

[30] Sullivan (2010) also fixed the discount factor to be 0.95. Rust and Phelan (1997) estimated the discount factor and found that the estimate was very close to 1. In Keane and Wolpin (1997), the discount factor was estimated to be .936.

portation cost.[31] A_{it} denotes the probability of being apprehended, given that an attempt to cross the border illegally is made.[32] I model the apprehension technology as

$$A_{it} = \frac{1}{1 + \exp\left(\delta_0 + \delta_1 B_t + \delta_2 I_{it}\right)}. \tag{7}$$

I assume that the apprehension technology is a function of B_t, a measure of intensity of border enforcement at time t. For example, line watch hours (total person hours spent patrolling border), the INS and border patrol enforcement budget, and other efforts made by the U.S. government to control the U.S. border may affect the probability of apprehension (I will discuss later how I measure B_t). I allow A_{it} to vary over individuals by including individual specific characteristics I_{it} with the idea that more previous experience of illegal crossing or of simply living in the U.S. may increase familiarity with the process of illegal entry and, therefore, the probability of being successful on a border crossing attempt.

Next, denoting $EV_{t+1}^{US_{legal}}$ as the value function associated with legal residence in the U.S., the expected value to the potential immigrant in Mexico of choosing to immigrate to the U.S. with a document at time t can be written as

$$\begin{aligned} EV_{t+1} = &\left(EV_{t+1}^{US_{legal}} - C_{legal}\right) \cdot 1(v_t = 1) + \\ &\left\{\left(EV_{t+1}^{US_{legal}} - C_{legal} - C_a\right) \cdot Pr(w_t = 0) + \left(EV_{t+1}^{MX_m} - C_a\right) \cdot Pr(w_t > 0)\right\} \cdot 1(v_t = 0). \end{aligned} \tag{8}$$

The first term relates to the value of legal entry for an individual with an immigration document, as v_t is a dummy variable that indicates whether the individual possesses a legal document to enter the U.S. at time t. The next term is associated with values of pursuing legal migration options when one is not documented. The waiting time variable w_t indicates how long it takes

[31] In the MMP, even though the apprehension probability ranges from 20% to 40% during the sample period (1965-2006), the probability of failing to move to the U.S., given an illegal migration attempt was made is observed to be less than 3%. This implies that migrants make multiple attempts when deportations occur at the border. As a result, I choose to directly estimate the number of expected attempts as a function of border enforcement variables and individual characteristics, instead of estimating the apprehension probability itself.

[32] Note that there is no other penalty imposed when migrants failed to move to U.S. illegally. According to Hanson (2006), those apprehended who agree to voluntary deportation are not processed by the U.S. justice system, and 95% of those apprehended by the Border Patrol between 1990 and 2003 agreed to this.

for an application submitted at time t to be processed.[33] Due to uncertainty in the application process, I assume that w_t is a random variable that is determined by a discrete hazard function, $\rho(s|\, X_t, \upsilon) = Pr(S = s|\, X,\, \upsilon,\, S \geq s)$, where S denotes the waiting time. I use the Cox proportional hazard model and specify $\rho(\cdot)$ as

$$\rho(s|\, X_s, \upsilon) \quad = \quad h(\gamma_s)\exp(X_s\beta + \kappa \cdot n_s + \nu), \tag{9}$$

where $h(\gamma_s)$ is the baseline hazard. The hazard rate $\rho(\cdot)$ is assumed to be a function of observable characteristics (X_s) that affect the length of the waiting period when an individual applies for a travel document.[34] For example, network, education, work experience in both locations, and the existence of a U.S. citizen in the family are all relevant here. The number of available visas at discrete time s (n_s) and time-invariant unobserved heterogeneity in the application process (ν) are also allowed to affect the arrival rate of immigration documents.[35] Lastly, C_{legal} and C_a denote the cost of legal migration and the cost of applying for a visa, respectively.

The value function next period when an individual chooses to move within Mexico is given by

$$EV_{t+1} = EV_{t+1}^{MX_k} - C_{mk},\, for\ k \neq m \tag{10}$$

where C_{mk} is the cost of moving domestically from location m in period t to location k in period $t+1$. This will depend only on the wage and non-wage amenities available at location k, as described earlier.

Finally, the value of choosing to stay in the current location m in Mexico is

$$EV_{t+1} = EV_{t+1}^{MX_m}. \tag{11}$$

Likewise, I can write the value function of someone residing in the U.S. legally at time t as

$$V_t^{US_{legal}} = U_t^{US_{legal}} + \beta \cdot E\left[\max_{d_t \in D_t^{US}} V_{t+1}|\, d_t\right]. \tag{12}$$

[33] I observe waiting times for completed spells in the MMP.

[34] Waiting time to get the document is usually less than a few months for temporary visas, but, for immigrants visas, it can take more than 10 years depending on the category (employment-based or family-based) the application falls under.

[35] Note that s inside the hazard function denotes the spell length after applying rather than a calendar year.

As an authorized immigrant in the U.S., he has only two options: *stay in the U.S.* or *move to location m in Mexico*. The value functions of choosing each alternative are,

$$EV_{t+1} = EV_{t+1}^{US_{legal}}, \, if \, stay \tag{13}$$

$$EV_{t+1} = EV_{t+1}^{MX_m} - C_m \, , \, if \, move \, to \, m \, in \, Mexico, \; for \, m = North, West, South, \tag{14}$$

where C_m is the cost of moving from the U.S. to location m in Mexico.

The value functions for someone in the U.S. illegally are analogous. In order to account for inherent stayers (movers), who are more likely to stay (move) after conditioning on observable characteristics, I allow all the moving costs in the model to vary with unobserved heterogeneity.

3.3 Solving the Location Decision Problem

Before discussing how to solve the optimization problem, it is useful to specify what agents know at time t as well as the distributions of the random components in the model. Agents in the model know the distribution of the idiosyncratic errors, ε_{ijt} and e_{ijt}, and the value of all other errors. I assume that $e_{ijt} \sim N(0, \sigma_e^2)$. The joint distribution (Ψ_i) of person-specific unobserved heterogeneity in the waiting function (ν), in moving costs (ω), and person-specific preferences for each location choice (φ_i^j) is specified as a discrete multinomial distribution, following Heckman and Singer (1984). Finally, the choice-specific time shocks to non-pecuniary utility (ε_{ijt}), unknown to the econometrician but known to the individual at time t, are distributed i.i.d extreme value. Using the extreme value distribution has two advantages. First, conditional on unobserved heterogeneity and wage errors, the expected value of the best choice next period has a closed form solution:

$$E\left[\max_{d_t \in D_t} V_{t+1}(d_t)|\, e,\, \Psi_i\right] = \gamma + \ln\left[\sum_{d_t \in D_t} \exp\left(\overline{V}(d_t|\, e,\, \Psi_i)\right)\right], \tag{15}$$

where $\overline{V}(d_t) = V_t(d_t) - \varepsilon_t(d_t)$, and γ is Euler's constant. Second, the choice probabilities conditional on unobserved heterogeneity are multinomial logit.[36]

Given the state space and distributional assumptions, I can solve the individual's optimization problem by backward recursion. Assuming that no choices are made after some period T, it is straightforward to evaluate the value functions at $t = T$. Then the value functions can be evaluated recursively for all $t < T$ using the value function specifications given above.[37] An issue arises, however, in solving the dynamic programming problem, as it requires obtaining the value functions for each sample person at each point in the state space for a given guess of parameters. Given the degree of heterogeneity in the sample in terms

[36] See Rust (1987) and Berkovec and Stern (1991).

[37] The time horizon is set to 30 years.

of exogenous characteristics, one of which*, distance to the railroad*, is continuous, this procedure quickly becomes too demanding computationally. In order to deal with this "curse of dimensionality," I use the weighting scheme suggested in Brien et al. (2006). Specifically, instead of evaluating value functions for every person in the sample, I group them into cells based on having similar exogenous traits and solve value functions only once for the representative person in each cell, given a set of parameters.[38] Then, I compute the conditional choice probability of a sample person i being a representative j, conditional on unobserved heterogeneity, as

$$P_{ijt} = \frac{\exp\left(\overline{V}_{jt}(d_t = k \mid \Phi_i)\right)}{\sum\limits_{k' \in D_t} \exp\left(\overline{V}_{jt}(d_t = k' \mid \Phi_i)\right)}. \tag{16}$$

The unconditional choice probability of the sample person i is obtained as a weighted average of conditional probabilities across representative people, $P_{it} = \sum\limits_j \omega_{ij} P_{ijt}$, where the weights (ω_{ij}) are inversely proportional to the distance between the sample person i and the representative person j. This interpolation method reduces the computational burden substantially while still preserving heterogeneity across people in the sample as value functions need to be evaluated only for a (relatively) small number of representative people.[39] I further increase computation speed by employing parallel processing.[40]

4 Estimation of the Structural Model

4.1 Econometric Specification

The deterministic components of the wage and non-pecuniary amenity value functions are specified in this section. Again, $w_q(S_{it}^w)$ and $h_j(S_{it}^w)$ have different arguments depending on location. First, the deterministic components of the log wage equations when employed in occupation q are assumed to take the form

$$\begin{aligned} w_q^{MX}(S_{it}^w) \;&=\; \beta_{0q}^{MX} + \beta_{1q}^{MX} \cdot hc_{it}^{MX}(q) + (1-\gamma_q^{MX}) \cdot \beta_{1q}^{MX} \cdot hc_{it}^{US}(q) + \beta_{2q}^{MX} \cdot educ_i + \beta_{3q}^{MX} \cdot age_{it} \\ &\quad + \beta_{4q}^{MX} \cdot age_{it}^2 + \beta_{5q}^{MX} \cdot x_{it} + \beta_{6q}^{MX} \cdot 1(j = north) + \beta_{7q}^{MX} \cdot 1(j = west), \end{aligned} \tag{17}$$

$$\begin{aligned} w_q^{US}(S_{it}^w) \;&=\; \beta_{0q}^{US} + (1-\gamma_q^{US}) \cdot \beta_{1q}^{US} \cdot hc_{it}^{MX}(q) + \beta_{1q}^{US} \cdot hc_{it}^{US}(q) + \beta_{2q}^{US} \cdot educ_i + \beta_{3q}^{US} \cdot age_{it} \\ &\quad + \beta_{4q}^{US} \cdot age_{it}^2 + \beta_{5q}^{US} \cdot x_{it} + \beta_{6q}^{US} \cdot v_{it} + \beta_{7q}^{US} \cdot network_i, \end{aligned} \tag{18}$$

[38] The continuous variable is discretized.

[39] Given five cohorts, two unobserved type, two education, three hometown categories, and discretizing the network instrument with three points, the number of representative people used in estimation is 180.

[40] Swann (2002) describes the use of parallelization (MPI) in solving a maximum likelihood problem.

where hc_{it}^{j} denotes work experience acquired in location j as a proxy for the level of human capital.[41] γ^q measures the degree of discount of human capital in occupation q accumulated from foreign country in the domestic labor market. For example, if $\gamma_A^{US} < \gamma_B^{US} < \gamma_W^{US}$, then human capital accumulated in the white collar occupation in Mexico is discounted the most severely of any Mexican occupations in the U.S. labor market. Another interesting parameter in the wage equations is β_{5q}^{MX}, which measures the effect of general U.S. experience on the wage in Mexico. Along with γ_q^{MX}, this coefficient allows me to test the transferability of human capital acquired in the U.S. Given that there is little variation in educational attainment among Mexicans in the MMP, edu_i is replaced by a dummy for having a high school diploma in the estimation. Note that legal status and network have occupation-specific effects in the U.S. but not in Mexico. Also, location within Mexico has occupation-specific effects.

Likewise, the deterministic components of non-pecuniary utility flow in both locations are specified as

$$
\begin{aligned}
h_j^{MX}(S_{it}^h) &= \alpha_{0j}^{MX} + \alpha_{1j}^{MX} \cdot \chi_{ij} + \alpha_{2j}^{MX} \cdot z_{it},\ for\, j = North, West, South && (19)\\
h^{US}(S_{it}^h) &= \alpha_0^{US} + \alpha_1^{US} \cdot age_{it} + \alpha_2^{US} \cdot age_{it}^2 + \alpha_3^{US} \cdot edu_i + \alpha_4^{US} \cdot network_i \\
&\quad + \alpha_5^{US} \cdot z_{it} + \alpha_6^{US} \cdot v_{it}, && (20)
\end{aligned}
$$

where χ_{ij} is a dummy variable that takes a value of 1 if individual $i's$ hometown is in location j. This term allows agents to have a preference for their native location.[42] The parameters of the duration term, z_{it}, measures the degree to which one gets additional utility or disutility from being more familiar with the location choice as duration of residence increases. Finally, α_1^{US}, α_2^{US}, and α_3^{US} capture the effect of age and education on the non-pecuniary value of residing in the U.S. For instance, the younger and the more educated one is, the less likely one suffers from living in a foreign country.[43]

4.2 Estimation Approach

The estimation proceeds in two stages. First, I estimate exogenous stochastic processes outside the model such as the occupation transition function, the hazard function associated with the application process, and the expected number of attempts conditional on the illegal border crossing option. I also estimate the wage equation ahead of time in order to reduce the computational cost. Despite the loss of efficiency involved in this

[41] Wage earned in the U.S. is measured in 1990 dollars and is converted to pesos measured in 1990 pesos using the 1990 real exchange rate. Annual wages are computed by multiplying hourly wages reported in the MMP by 2080 (52 weeks X 40 hrs/wk) hours.

[42] See Kennan and Walker (2008) for an example of a migration model that allows for a bias in favor of the home location.

[43] The deterministic component associated with the *South* location choice, $h_{South}^{MX}(S_{it}^h)$, is set equal to zero because the non-pecuniary utility flow coefficients are identified only relative to a base choice (*South*).

two-step estimation, it is still worthwhile given that it significantly reduces the number of parameters I need to estimate inside the full structural model.[44] To correct for the endogeneity associated with estimating wage equations separately from the choice process, I estimate a simple model that incorporates wages, worker's document status, and selection into the labor market at each location.[45]

In the second stage, the parameters associated with the choice process are estimated using maximum likelihood estimation. The solution to the dynamic programming problem presented above provides the choice-specific value functions which are used in the likelihood function. Given unobserved individual "type" k, and the observed outcome O_{it}, his likelihood contribution is the joint probability of observing the choice made by the individual. The conditional likelihood function is,

$$L_i(\Theta|\, type\, k) = \iint \left(\prod_{t=0}^{T_i} Pr(O_{it} = d_t|\, e,\, \varepsilon,\, \Theta,\, S_{it},\, type\, k) \right) dF(e)dF(\varepsilon) \tag{21}$$

where Θ is the vector of parameters in the model and T_i is the number of periods that individual i is observed in the sample.

Using the fact that the randomness in the non-pecuniary flow is distributed extreme value, I can rewrite the conditional likelihood as

$$L_i(\Theta|\, type\, k) = \int \left(\prod_{t=0}^{T_i} \frac{\exp\left[\bar{V}(d_t|\, e,\, \Theta, S_{it},\, type\, k)/\tau\right]}{\sum\limits_{d_t' \in D_t} \exp\left[\bar{V}(d_t'|\, e,\, \Theta, S_{it},\, type\, k)/\tau\right]} \right) dF(e) \tag{22}$$

where $\bar{V}(d_t) = V(d_t) - \varepsilon_t(d_t)$.

Unobserved heterogeneity enters the estimation as follows. I assume that Mexican workers can be divided into K groups or "types," each of which share the same unobserved heterogeneity in terms of four characteristics: location preferences, moving costs, the effect of an additional year of staying on non-pecuniary utility in the U.S., and the baseline hazard in waiting time after applying for a visa.[46] In other words, estimation of the model reveals what proportion of the population is each type and determines the value of unobserved heterogeneity for each type in each of the four characteristics just noted. Unobserved location preference parameters are identified by time-persistent differences in location choices across types who are otherwise observably similar. Permanent unobserved heterogeneity in moving costs is identified by variation in the frequency of moving across types conditional on observed characteristics. The unobserved heterogeneity in duration is identified by variation across types in the speed of returning to Mexico conditional on being in the U.S and observable characteristics. Finally, the unobserved heterogeneity in waiting time

[44] There are 7 parameters of the wage equation for each location-occupation pair.

[45] Marital status is included in the estimation to instrument for the participation equation.

[46] K is set to 2 in estimation.

is separately identified from the baseline hazard as long as there are observable individual characteristics included in the Cox proportional hazard specification.[47] Then, the unconditional likelihood contribution of person i is the weighted average of type-specific likelihood contributions, where the weights are type probabilities.[48]

5 Data

To estimate the model, I use 3 different data sets. The primary source of data for this study is the Mexican Migration Project (MMP), a retrospective panel that contains data at the individual level on occupation and migration histories of Mexican workers.[49] Each year since 1982, the MMP has randomly selected households in communities throughout Mexico and collected socioeconomic and demographic information as well as life histories of a sample of household heads. The MMP124 database offers a unique, source-country-based data set that contains retrospective migration and occupation histories of 20,621 households from 124 communities in Mexico. Due to the sampling method, I drop 981 U.S. households in the sample, which leaves 19,640 households, including 7,952 (40%) household heads with work experience in the U.S.[50]

However, the MMP is subject to some important sample selection problems. First, because the project initially focused on rural villages located in Western-Central Mexico, which was traditionally known as the heartland for out-migration to the U.S., it is not likely to be representative of the general population in Mexico.[51] Comparing the 1990 U.S. and Mexico censuses with the MMP, Hanson (2006) confirmed that the U.S. migrants in the MMP are disproportionately male, uneducated, and agricultural. Second, because surveys take place in Mexican villages, only households with at least one member remaining in Mexico are included in the sample. This implies that the sample is missing information on immigrants who have established their households in the U.S. The MMP attempts to deal with this issue to some degree by tracking small numbers (10 to 20) of the U.S. households from each Mexican community and adding them to the sample.[52] But, the "snow-ball" sampling method used to survey the U.S.-based sample introduces an additional selection problem regarding the referring process.[53] Finally, there is a potential issue of

[47] See Elbers and Ridder (1982) for details.

[48] See Heckman and Singer (1984).

[49] The Mexican Migration Project (MMP) is a collaborative research project based at the Princeton University and the University of Guadalajara. The MMP is publicly available at http://mmp.opr.princeton.edu. For more information on the MMP, see Massey and Espinosa (1997).

[50] 981 U.S. households were surveyed in the U.S. through the "snow-ball" sampling method. While it is often cost-effective in collecting data, the estimation of choice probabilities using data from a choice-based sampling method requires prior knowledge of marginal distributions of attributes and population shares of choices. See Manski and McFadden (1981) for details.

[51] See Orrenius and Zavodny (2005) and McKenzie and Rapoport (2007).

[52] The U.S.-based sample was compiled using snow-ball sampling (also known as the chain-referral or reputational sampling method) with which future subjects are recruited through recommendations or referrals within networks.

[53] For example, to the degree that a U.S. migrant refers someone who share similar characteristics with himself, the

misreporting given the retrospective nature of the survey. This raises concerns about the accuracy of data on "seasonal migrants," because the migration history of the household head who was absent at the time of survey is collected by interviewing other remaining members within the household.

Despite the apparent problems regarding the sampling method and the statistical problems it implies, the MMP124 is still well-suited to my study for the following reasons. First, since the survey takes place in the source country, I can address the selection problem that would arise if considering immigrants only in the host country. Second, having information on both migrants and non-migrants enables me to investigate the determinants of migration behavior. Third, it includes undocumented migrants, who are likely to be absent in the Current Population Survey (CPS) or any other host country based survey. Finally, the vast amount of quantitative information on location and occupation histories of many households allows for esimation of a complex structural migration model as shown in this paper.

For tractability, I focus on a group of male household heads in the MMP124 whose age was between 16 and 20 years old in 1965.[54] This implies that I am restricted to using only a small portion of this rich data set, but it still involves solving the structural model over a 30-year horizon. Solving the full dynamic model with additional overlapping cohorts would increase the computational burden severely due to the existence of time-specific aggregate state variables: border control index variables and number of available visas. In constructing the location state variable, 32 states in Mexico are aggregated into 3 locations. *North* includes Baja California del Norte, Baja California del Sur, Coahuila, Chihuahua, Nuevo Leon, Sonora, Tamaulipas. *West* includes Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacan, Nayarit, Queretaro, San Luis Potosi, Sinaloa, Zacatecas. The rest of the states correspond to *South* (Figure 1). After dropping 52 communities missing document status information which were surveyed in earlier years, the final sample that is used in the estimation includes 673 male household heads. I keep track of each person from 1965 until 1994, which results in a total of 18,720 person-year observations.[55]

Table 4 presents summary statistics for the 673 household heads who are included in the estimation. Only about 5% of Mexican workers in the sample have a high school degree. An average person in my sample is observed for 27.8 years, lives in Western states in Mexico, and is employed in low-skilled occupation. About 33% of those have some U.S. migration experience, though only 6.1% of person-year observations take place in the U.S. Panel C summarizes the U.S. experience of 223 individuals who emigrated at some point. On average, they spent about 7 years in the U.S. between 1965 and 1994. 80% did not have an employment authorization document to work in the U.S., and the majority were employed in blue-collar occupations.

The second data set I use has distance to rail lines. In their study on the effect of migration networks

estimation of the choice probability requires information about identity of who referred whom.

[54] I choose this particular year, since the Bracero program, a guest worker program which started in 1942, ended in 1964.

[55] Sample selection criteria used to construct the final sample is provided in Table 3.

in Mexico on the development of micro-enterprises in the country, Woodruff and Zenteno (2007) used the distance from the capital of the Mexican state in which an individual was born to the nearest station on the north/south rail lines as they existed in the early 1900s. Since these rail lines were the main means of transportation when the massive migration from Mexico began through the guest worker program during the 1910s, the distance-to-rail-line data provides plausibly exogenous variation that identifies the effect of the network on individual migration decision.[56]

Lastly, I use border apprehensions and enforcement data from Hanson and Spilimbergo (1999). Specifically, I use the number of border patrol officers to construct the measure of intensity of border enforcement. Also, the total number of available visas to Mexicans are computed as a sum of MXQUOTA (Mexicans Admitted Under Numerically Limited Quotas) and MXRELAT (Mexicans Admitted as Relatives of U.S. Citizens) from the MMP124.[57] Time series plots of the two policy variables are shown in Figures 4 and 7.

6 Non-Structural Estimation

Before presenting estimation results of the structural model, this section discusses a non-structural estimation approach. These exercises are useful in revealing the relationships in the data and identifying which variables should be included in the structural model without specifying the full dynamic structure. In addition, the results from the non-structural estimation methods serve as a basis for comparison to the structural estimates. First, I describe a simple model of Mexican workers' labor market performance that incorporates wages, network effects, and worker's immigration document status. The estimated coefficients of wage equations will be used as inputs in the structural estimation. Next, I estimate a logit model with random effects to analyze the effect of various explanatory variables on the probability of choosing the U.S. location.

6.1 Multivariate Probit Estimation of Wage Equations

6.1.1 Non-Structural Model of Labor Market Performance

I estimate wage equations for Mexican workers both in Mexico and in the U.S. In order to explicitly account for selection into the U.S. labor market and the endogeneity associated with document status, the probability of working and of being "*documented*" are estimated inside of the model, along with parameters of the wage equation. I denote L^*_{ij} and Q^*_{ij} as latent variables indexing the propensity for worker i from village j in Mexico to work in the U.S. and to have a valid employment authorization document. Let L_{ij} be a dummy variable equal to 1 when $L^*_{ij} > 0$, which indicates that an immigrant is observed to be working, while

[56] The rail lines were no longer used when the guest worker program ended in the 1960s.

[57] The number of visas granted is used to measure variation in visa availability over time. In most years, the number of available visas coincides with the number of visas granted because there are more applicants than available visas.

$Q_{ij} = 1(Q^*_{ij} > 0)$ indicates that a worker is "*documented.*" Now I can write a system of equations that determines participation, wage, and document status as

$$L^*_{ij} = X_{1ij}\theta_1 + Q_{ij}\theta_{1Q} + \varepsilon_{1i} \quad (23)$$

$$W_{ij} = X_{2ij}\theta_2 + Q_{ij}\theta_{2Q} + \varepsilon_{2i} \quad (24)$$

$$Q^*_{ij} = X_{3ij}\theta_3 + N_{ij}\theta_{3N} + \varepsilon_{3i}, \quad (25)$$

$$\text{where} \begin{pmatrix} \varepsilon_{1i} \\ \varepsilon_{2i} \\ \varepsilon_{3i} \end{pmatrix} \sim N(0, \sum), \text{ and } \sum = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 \end{bmatrix}.$$

The term X_{ij} represents a vector of demographic characteristics, such as age and education, that is common across all three equations. The individual characteristics in the participation equation X_{1ij} also includes marital status and number of children in addition to X_{ij}.[58] Worker i's wage, denoted W_{ij}, is specified as a function of X_{2ij} which includes a vector of individual characteristics that affect worker's labor market performance such as work experience in both locations, interacted with each of the occupation category (agricultural and blue collar) separately.[59] The variables in X_{3ij} include age, education, and duration of residence in the U.S. Finally, the size of the network, denoted by N_{ij}, is allowed to affect the wage and is instrumented with *log(distance to rail-line)*.[60] All the θ's in all three equations are estimated along with parameters in $\sum$ using maximum likelihood estimation.[61]

6.1.2 Wage Estimates

Table 5 and 6 show the estimation results for the wage equations. I begin describing them by focusing on the equation for wages in Mexico.[62] Table 5 shows that wage offers differ considerably across locations; the mean of the wage offer distribution is consistently higher in the *West* region of Mexico. Also, it is interesting to note that work experience acquired in the U.S. labor market has a strong positive effect on the wage in Mexico, which considerably exceeds the effect of experience acquired in Mexico, as additional year of U.S. work experience is associated with 17%-30% higher wages in Mexican labor market across occupations. The opposite does not hold, as can be seen in Table 6, which shows that work experience acquired in Mexico

[58] It is assumed that marital status and number of children affect the probability of working, but not the wage directly.

[59] The subscript for occupation is suppressed in the wage equation.

[60] The document status equation is estimated only for workers located in the U.S.

[61] σ_1 and σ_3 are restricted to be 1 in the estimation because they cannot be separately identified.

[62] Hong (2009) describes the empirical model and estimation method in more detail.

generally has little effect on the wage in the U.S, even though Mexican agricultural experience seems to raise agricultural wages in the U.S. Thus, the degree of human capital transferability is not symmetric across locations. This indicates the importance of considering dynamics in studying migration behavior of foreign workers, because, for example, the same person with 10 years of work experience in the blue-collar sector, 5 years of which was acquired in the U.S., is expected to earn higher wages in Mexico when he spent the first 5 years in his career in the U.S.

Table 6 displays the parameter estimates of the U.S wage equation. Overall, most parameter estimates have the expected sign and reasonable magnitudes. Network effects are found to be present both in the document status equation and in the wage equation. The negative and significant estimate on the network instrument (-0.18) implies that a one standard deviation increase from the mean in the value of the instrument lowers the probability of being "*documented*" by 10%.[63] This is explained by the role of networks in finding a sponsor in visa application process. As described in Section 2, a potential immigrant may be eligible for an immigration document through the sponsorship of a close family member or an U.S. employer. The estimate indicates that, holding an applicant's other characteristics constant, one is more likely to be matched with a sponsor when he has friends or relatives who have established themselves in the U.S. Turning to the wage estimates, the estimated coefficient on *Distance to Rail* is negative in both occupations, although it is statistically significant only in the agricultural sector. This is consistent with Munshi (2003) who showed that the same individual is more likely to be employed in a favorable non-agricultural job when his network is exogenously larger. I find here that a migrant is likely to earn more even within the agricultural industry, when he belongs to a larger labor market network. In general, network effects seems to be stronger in the agricultural sector, regardless of a worker's document status.[64]

6.2 Random Effects Logit Estimation of U.S. Out-migration Decisions

This section discusses estimation results from a reduced form specification of Mexican workers' U.S. out-migration decisions. The non-structural estimation exercise provides a computationally inexpensive way to examine the importance of each state variable included in the structural model. A logit model is used to investigate the relationship between observable state variables and the probability of moving to the U.S.

[63] Recall that a negative estimated coefficient on the instrument is associated with positive network effects because it measures distances from the city of birth to the nearest railroad.

[64] Several theoretical studies have investigated the formation of networks and their importance in the labor market (see Ioannides and Loury (2004) for a survey of the literature). For example, Carrington et al. (1996) considered a migration model in which moving costs decrease with the number of migrants already settled in the destination. In the presence of informational problem in the labor market, which is likely to be quite severe among newly-arrived foreign workers in the U.S., potential migrant workers benefit from information transmission within networks by learning about new job opportunities at the destination. In addition, employers may rely on networks and job referrals through its incumbent workers in recruiting new employees in order to minimize search costs when worker's productivity cannot be accurately assessed (Montgomery 1991).

In order to account for time-invariant unobserved heterogeneity, I include person-specific random effects in the regression and estimate the panel-level variance along with other parameters associated with the choice process. The same set of explanatory variables that constitute the state variables in the full structural model are included in the regression so that I can closely match the structure of the utility function in the dynamic model. In particular, expected wages in both locations are imputed using the estimated coefficients of the wage equations described above.

Table 7 presents estimation results from the random effects logit model with the base outcome being in Mexico. The estimates suggest that the probability of choosing the U.S. location significantly decreases with duration of residence in the current location and with the Mexican wage but significantly increases with the U.S. wage. Once again, the negatively significant estimate on *Distance to Rail* provides support for social capital theory in the U.S. immigration decision. One interesting parameter estimate is the coefficient on the education variable. There is abundant theoretical and empirical literature studying the selection of Mexican immigrants in the U.S. with somewhat mixed results. For example, Borjas (1987) proposed a negative selection hypothesis, arguing that immigrants with low schooling levels in poor countries have stronger incentives to immigrate to rich countries where low-skill labor is scarce, as their returns from migration are high relative to highly educated workers.[65] On the other hand, using the Mexico and U.S. population censuses, Chiquiar and Hanson (2005) provided evidence of positive selection among immigrants and explained that migration costs may decline with schooling if highly educated migrants are more motivated or may be less subject to credit constraints in financing migration compared to less skilled migrants. Table 6 shows that the education coefficient is negative, but the estimate is not statistically significant, which suggests intermediate selection of Mexican immigrants in the U.S.[66] Finally, I find that person-specific random effects are significant, as the proportion of the total variance contributed by the panel-level variance, denoted by ρ, is estimated to be over 50%. This implies that unobserved person-specific preferences are an important consideration in the structural model as well.

While the non-structural estimation provides useful reference points that can be compared to the current literature, this approach has clear limitations. First, the indicator of visa status is likely to be endogenous because the decision to apply for an immigration document may be correlated with the propensity to immigrate to the U.S., in which case the estimated coefficient on *Visa* should be interpreted as the upper limit of the effect of document status on the probability of U.S. migration. Second, the lack of dynamics and consideration of moving costs may also result in biased estimates. For example, to the degree that agents are

[65] Borjas' predictions were supported by other empirical studies such as Cobb-Clark (1993) and Durand et al. (2001).

[66] This is inconsistent with what Massey et al. (1994) and Durand et al. (2001) found using the same data I use. One explanation may be that it matters qualitatively when one corrects for endogeneity arising from unobserved heterogeneity in location preference. The other explanation may be that there are simply not enough educated workers in the MMP which makes it difficult to obtain a reliable estimate on schooling.

forward-looking and expect that their U.S. wages will grow as they accumulate U.S.-specific human capital, the estimated coefficients on *U.S. wage* is biased upward.

7 Structural Estimation Results

In this section, I present the estimates from the full model. I begin by discussing estimates from the first stage estimation, which includes the occupation transition functions, waiting time in the visa application process, and expected number of illegal border crossing attempts.[67] Then, the structural parameter estimates of the choice process from the maximum likelihood estimation are presented.

7.1 Parameter Estimates

Occupation Transition Function

Table 8 presents the estimated parameters of the occupation transition functions from equation (2), estimated using a multinomial logit specification, with the base outcome being *Blue-Collar*. I used the full sample of MMP124 here, instead of the restricted sample focused on men who were entering the labor market around 1965, to aid identification of parameters related to occupation transitions in the U.S. As expected, lagged occupation is the most significant determinant of occupation status across all the occupation categories and locations. What is interesting is that, when an immigrant worker decides to move back to Mexico, he tends to return to the occupation he had in Mexico before coming to the U.S. This suggests that location-specific human capital does not depreciate quickly when out-migration occurs. For simplicity, however, I specify the occupation transition to be a first-order Markov process in the full model, and I do not allow the last job held in Mexico to affect the occupation transition upon returning to Mexico, so that I do not have to track an additional state variable. Also, given the small number of Mexican workers in the U.S. white-collar sector, I allow an individual's occupation status in the U.S to evolve over only 3 states, *Unemployed, Agriculture,* and *Blue-Collar* and set the probability of entering the *white-collar* occupation to zero in the second stage estimation.

Waiting Time in the Application Process

Potential immigrants may apply for a legal document to work in the U.S. In such cases, they usually have to wait in line until the application is approved, and the time it takes for applicants to receive visas depends on the specific visa category for which one is applying. MMP124 provides information about the sponsor when

[67]The first stage estimates are obtained outside the structural model and used as inputs within the structural model in the second stage estimation.

a Mexican applies for legal entry to the U.S. Using this information combined with the years of application and the time to receive the document, I estimated a Cox proportional hazard model of waiting time for the pooled sample of people who apply for a visa and also separately for the two main immigration categories: Employment-Based (EB) and Family-Based (FB).

Column 1 in Table 9 presents results from the pooled sample. The parameter estimates suggest that a worker's family network, measured in terms of existence of family member at the destination, or social network, measured by the instrument, as the group of migrant's community in the U.S. who originated from the same village in Mexico, has a strong positive impact on the hazard of exiting the waiting spell, while indicators of worker qualifications such as education or previous work experience have no explanatory power.[68] When the sample is divided based on the type of sponsor, however, a distinct pattern across the two groups emerges. Column 2 shows that the coefficient on family is even larger for the FB applicants, while it becomes insignificant for the EB applicants. At the same time, an applicant's education is an important factor in EB cases, as shown in Column 3. Even though I do not consider visa categories explicitly in the structural model, this exercise provides me with useful information as to how prospective migrants sort themselves into different application categories based on individual characteristics.[69]

Expected Number of Attempts of Illegal Crossing

In the MMP, it is observed that those apprehended after failure to cross the border illegally make subsequent attempts to move to the U.S. within the same calendar year. Therefore, I directly estimate the number of expected attempts as a function of border enforcement variables and individual characteristics instead of estimating the apprehension probability. I use a semiparametric method based on Ichimura and Lee (1991) and Stern (1996).[70] Let *offcr* and *USexp* denote number of border patrol officers and years of previous U.S. migration experience, respectively. The estimated equation is

$$E(Number\, of\, Attempts) = \underset{(0.145)}{0.505} + \underset{(0.00021)}{0.00034} * offcr - \underset{(0.0051)}{0.0114} * USexp, \tag{26}$$

[68]It is not completely clear why the coefficient on U.S. work experience is negative and significant. Given that most immigrants apply for immigration documents in Mexico, those applicants with positive U.S. work experience may have been out of status while working in the U.S., which is likely to delay the legalization process.

[69]Existence of family in the U.S. is also dropped in the main structural model due to endogeneity.

[70]A total of 10,468 person-year observations in the MMP124 were included in the estimation. Each observation records the number of attempts made to cross the border and previous U.S. experience, which is merged with information on border enforcement each year.

where standard errors are in parenthesis. The estimated coefficients imply that a potential migrant with no previous U.S. experience who is considering crossing the border illegally when, for example, the number of border patrol officers is 2000 makes 1.2 attempts on average. The estimates confirm the prior hypothesis that the cost of illegal migration increases with the number of border patrol officers and that previous U.S. experience lowers the expected number of attempts necessary to cross the border successfully. The esimates in Equation (26) enters the structural model through the moving cost function associated with illegal border crossing.

Structural Estimates of the Choice Process

Having estimated all the stochastic processes in the model in the first stage, I proceed in the second stage to estimate the parameters associated with the choice process: the non-pecuniary utility flow and cost functions corresponding to each location choice. The resulting maximum likelihood estimates based on the full solution of the dynamic model are presented in Table 10.

Parameter estimates have the predicted sign, and most coefficients are significant at the 5% level. These parameter estimates are interpreted as the effect of each state variable on the non-pecuniary utility flows relative to the value of the base choice (*South*).[71] Given that utility is in term of log wages, parameter estimates are also measured in log monthly wage units.[72] For example, Table 10 shows that living in the *West* of Mexico, holding other variables constant, generates additional utility gains of 3.631 log monthly wage units relative to the base location. The estimated value of the U.S. dummy is positive and significant, which suggests that non-economic factors such as the cultural or institutional atmosphere make the U.S. attractive relative to the *South* of Mexico, even though it is less attractive than the *West* of Mexico. Additional years of residence in the U.S. incur significant disutility (-1.804), equivalent to the effect of a 88% decrease in the monthly wage on the value of U.S. location choice. This explains the frequent return migration behavior and short duration of stays in the U.S. observed in this cohort of the MMP; this will have important implications for a policy that tightens border enforcement, perhaps leading to longer U.S. stays.

There are two interesting parameters in the non-pecuniary utility component related to the U.S. location. First, the estimated effect of being documented on non-pecuniary utility is positive and large in magnitude (13.244). This can be explained by various restrictions on undocumented immigrants in the U.S. For example, illegal immigrants are not allowed to receive any welfare benefits. In addition, no restriction on travelling outside the U.S. is likely to generate a significantly positive consumption value of being a legal resident

[71] Wages are reported in terms of the log monthly wage in 1990 pesos.

[72] When marginal utility of income is not linear, as is the case with the constant relative risk aversion (CRRA) utility function, an agent in a dynamic model chooses to save at the optimum to insure against uncertainty. Although the level of assets is thought to be an important state variable for migration deicisions, I do not model saving behavior because it would increase the size of the state space rapidly. Consideration of this extension is left for future research.

in the U.S. Second, I find that the estimated coefficient on *Distance to Rail* is negative and statistically significant at the 5% level. The coefficient estimate implies that compared to a Mexican worker who was born in the state of Merida, the capital of which was 1,458 miles away from the rail line, a worker who was born in the state of Tijuana in which the rail line passed through the state receives a utility flow that is 0.364 (=0.05*ln(1458)) log monthly wage units higher relative to the base choice. This provides the first evidence that there exist non-monetary benefits from joining a exogenously larger network in addition to the benefit in the labor market outcome found in Table 6.

Turning to the estimates of the moving cost function in Panel B of Table 10, I find that estimated moving costs are substantial and that they approximate the distance between locations quite well, even though I did not condition on distance. Moving costs associated with domestic migration are less than with U.S. migration. In addition, the estimated coefficient on *Undocumented* imply that moving to the U.S. without documents incurs significantly greater cost (180.85) relative to moving with documents. Although large in magnitude, it can be justified considering that parameters associated with the moving cost function are identified by the difference in the sum of pecuniary and non-pecuniary amenity values attached to the choice of moving relative to the choice of staying. In other words, the estimated moving cost associated with illegal border crossing should be interpreted as including non-monetary psychic costs as well as monetary costs that includes transportation and additional amounts that migrants pay to "Coyotes." Large magnitudes of moving cost parameters show that individuals in the dynamic migration model realize that the benefits of moving in terms of higher wages and utility are received at the destination in the future.

The estimated application cost is 74.28 utils (or log monthly wage units), which turns out to be much less than the cost of moving to the U.S. illegally. There are two possible explanations for why migrants do not apply for a visa more often, given the sharp difference between the cost of moving with and without documents. First, the need to move to the U.S. immediately dominates the benefits of having documents at a later date. The application process takes about 1-2 years on average and up to 10 years depending on availability of visas. Second, the unobserved heterogeneity in waiting time may play a significant role here. Potential migrants have private information, unobserved to the econometrician, regarding their own qualifications in each category of immigrant visa application, which implies that those who expect shorter waiting times are likely to be over-represented in the sample of visa applicants. This self-selection hypothesis is supported by the estimated coefficient of the type 2 dummy in the waiting time function. The significant and negative estimate (-0.650) implies that, all else equal, a type 2 individual is more likely to leave the waiting spell relative to type 1's, when applying for an immigration document.

The model was fit with two unobserved types. Estimated coefficients of unobserved heterogeneity parameters suggest that type 2 individuals, which comprise 47% of the population, can be classified as inherent

"*Movers*". Relative to type 1's, type 2 individuals face lower costs when moving, and the baseline utility they get residing in the U.S. is significantly lower (-2.294). Moreover, the additional year of stay in the U.S. generates a greater disutility (-1.790) for type 2 individuals.

7.2 Goodness-of-Fit

7.2.1 Within-Sample Fit

To assess how well the model fits the data, I compare some meaningful sample moments and the corresponding simulated moments from the fitted model. Specifically, the estimated structural parameters are used to simulate a sample of 1,346 (=673*2) individuals whose location and application choices are compared to those observed in the data. I used the exogenous characteristics of 673 individuals in the sample and randomly draw type 1 extreme value errors to generate a vector of choice specific shocks in each period. Antithetic acceleration is employed to reduce the variance of the simulated moments, which implies that each draw is used twice as ε and $1 - \varepsilon$.[73] Tables 11 and 12 provide evidence on the within-sample fit of the model in two important dimensions: the proportion of U.S. migrants and the proportion of visa applicants by age. Table 11 shows that the model does a fairly good job of matching the U.S. out-migration pattern observed in the MMP. Even though the simulated migration rate under-predicts the U.S. migration rate prior to age 30, the model captures the upward trend in U.S. migration decisions. Table 12 compares the proportion of visa applicants in the simulated data and in the MMP. Due to the small number of observations, it is difficult to determine the performance of the fitted model, but the model seems to predict consistently higher number of visa applicants, which results in a lower number of undocumented migrants in the U.S. in the model relative to in the actual data as shown in Table 13. Figure 9, however, illustrates that the model is able to generate the sharp decline in the number of undocumented migrants over the 30 years as observed in the data.

7.2.2 Lagrange Multiplier Tests

Given the reasonable fit of the structural model, I further investigate the possibility of misspecification due to omitted variables. Using the MMP, Hong (2009) examined the effect of networks on the labor market performance of Mexican workers in the U.S. and provided evidence of interactions between network effects and document status in that the group of documented workers in blue-collar occupations are found to get the most benefits from the existing social connection. Given, further, that the estimated coefficients of the network instrument and document status in the non-pecuniary flows equation were significant in the structural migration model above, it seems plausible that the effects of networks on amenity values related

[73] See Geweke (1988) for details of the use of simulation methods.

to the U.S. residence may be moderated by document status. To test whether the interaction effect is statistically significant in the dynamic migration model, I perform a Lagrange multiplier test. Specifically, I modify the deterministic component of the non-pecuniary flows associated with the U.S. location as

$$\begin{aligned} h^{US}(S_{it}) &= \alpha_0^{US} + \alpha_1^{US} \cdot age_{it} + \alpha_2^{US} \cdot age_{it}^2 + \alpha_3^{US} \cdot edu_i + \alpha_4^{US} \cdot network_i \\ &\quad + \alpha_5^{US} \cdot z_{it} + \alpha_6^{US} \cdot v_t + \alpha_7^{US} \cdot network_i \cdot v_t. \end{aligned} \tag{27}$$

Then, I can construct the chi-square test statistic as

$$\left\{ \left[\frac{\partial}{\partial \theta} L_i(\hat{\theta}) \right]_{|\,\alpha_7^{US}=0} \right\}' \cdot \hat{D}^{-1} \left\{ \left[\frac{\partial}{\partial \theta} L_i(\hat{\theta}) \right]_{|\,\alpha_7^{US}=0} \right\} \cdot \left\{ \left[\frac{\partial}{\partial \theta} L_i(\hat{\theta}) \right]_{|\,\alpha_7^{US}=0} \right\} \sim \chi^2(1), \tag{28}$$

where $L_i(\hat{\theta})$ is the log-likelihood contribution of observation i, and $\hat{D}^{-1}\{\cdot\}$ is the inverse of the outer product of the score statistics under the null, $\alpha_7^{US} = 0$. The value of the chi-square test statistic under the assumption that the null is true is 1.605, and I fail to reject the null hypothesis at the 5% significance level. As a result, I conclude that the effect of the added interaction term is not statistically significant.

7.2.3 Out-of-Sample Performance

The sample selection criteria used to construct the main sample was quite restrictive in that it included only those who were young adults as of 1965 (hereafter, this sample is referred to as “1965 cohort”). It may pose a problem for interpreting the results from counterfactual experiments because policy implications inferred from the exercises may not be applied to later cohorts if migration patterns between the two cohorts are qualitatively different. For example, in case there was a structural change which was not captured in the model but influenced later cohorts, in-sample inference might not be reliable. To address this concern, it is necessary to test how the model would perform in predicting the migration probabilities using a different sample.

To do this, I construct a sample of individuals who were between 16 and 20 years old in 1975. The same sample selection rule as in Section 5 was applied to the MMP124, which resulted in 884 individuals with 23,163 person-year observations in the sample (hereafter, this sample is referred to as “1975 cohort”). Table 14 presents summary statistics of the 1975 cohort. Even though similar to the 1965 cohort in most attributes, there are notable differences in some important characteristics. For example, compared to its counterpart, an individual in the 1975 cohort is more likely to have a high school diploma (11% versus 4.7%), employed in blue-collar occupations (49.8% versus 41.7%) than in agricultural occupations (33.5% versus 44.5%). In addition, the U.S. migration rate is higher among the 1975 cohort (8.4% versus 6.1%) with longer duration once in the U.S. (7.69 years versus 6.98 years). Finally, one in the 1975 cohort is more likely to be employed in blue-collar occupations (71.8% versus 57.2%) once in the U.S.

To check the out-of-sample fit of the model, the estimated structural parameters and the model are used again to simulate migration histories of the 1975 cohort. Figure 18 compares the proportion of U.S. migrants from the simulation and in the data. It shows that the model under-predicts the U.S. migration rate prior to 1985 and over-predicts after. In particular, the sharp difference in U.S. migration decisions during the 1990s seems to provide evidence of a unmodelled structural change in terms of parameters of interests.

8 Policy Simulations

Estimation of the structural model allows me to evaluate the impact of alternative immigration policies on migration behavior of Mexican workers. Assuming that a reduction in the number of illegal immigrants in the U.S. is USCIS's main objective, I perform three different policy experiments to explore how Mexican migrants would respond: an increase in the number of available visas, an increase in the number of border patrol officers, and a reduction of pecuniary and non-pecuniary flows associated with illegal residence. The location and application decisions are simulated for each observation with a set of i.i.d. preference shocks under each counterfactual scenario.

The first policy experiment of interest is a permanent increase in the number of available visas by 100%, starting in 1975.[74] To do this, I double the number of available visas starting in 1975, which is expected to affect the value of applying for an immigrant document through the visa waiting time equation. The results of the simulations are presented in Figure 10 and 11. Figure 10 compares the proportion of U.S. immigrants by year under the baseline and the policy experiment. The policy experiment raises the total number of U.S. immigrants who are older than age 30, but the change is modest. Figure 11 reports the proportion of undocumented aliens among the U.S. immigrants, and it confirms that the policy change has almost no effect on the fraction of illegal migrants in the U.S. This is not surprising because it is observed that potential migrants often do not apply for a visa even given the relatively low cost of application. This implies that what is crucial in the decision to apply is private information regarding one's eligibility, and the increase in the number of visas has little effect on the application decisions of Mexican migrants on the extensive margin.

The second policy experiment analyzes the effect of an increase in the number of border patrol officers at the U.S.-Mexico border on Mexican immigration. This policy change effectively increases the cost of illegal immigration because the expected number of border crossing attempts is a function of USCIS's efforts to prevent illegal entry, as shown in Equation (26). I use the model and the estimates to simulate location decisions

[74] This accords with Pennsylvania Senator Arlen Specter's immigration-reform bill (2007).

of Mexican workers in the sample under the counterfactual scenario in which the number of border patrol officers doubled since 1975. As of 1975, for an individual with no previous U.S. experience, this is equivalent to about a 60% increase in the cost of illegal immigration because the policy change increases the expected number of border crossing attempts from 1.2 (=0.505+0.00034*2000) to 1.9 (=0.505+0.00034*4000). The results of this exercise are shown in Figures 12-15. Figure 12 displays the proportion of U.S. migrants under the alternative policy relative to the status quo. The policy experiment lowers the number of U.S. migrants for the first 10 years after the policy change, but the effect is small in magnitude at 1-2%. What is striking, however, is that the number of U.S. migrants under the counterfactual environment is shown to overtake its counterpart beginning in 1990. This can be explained by two different effects. First, recognizing higher costs of crossing the border illegally, undocumented workers in the U.S. respond to the policy by delaying return migration. Figure 14 shows that the return migration rate actually decreases by up to 4-5% relative to the baseline.[75] Second, the policy change influences potential migrants located in Mexico by making the legal migration option more attractive than it was in the baseline. Figure 15 confirms this prediction, showing that the number of undocumented migrants in the U.S. decreases in response to increased border crossing costs.

The third set of policy experiments investigates how Mexican workers within the U.S. would respond to a stricter control on unauthorized immigrants. Specifically, I simulate permanent 10%, 20%, and 30% reductions in illegal workers' wages in the U.S. labor market, beginning in 1975, to resemble tougher penalties on unauthorized employment.[76] The impact of these policy changes on the proportion of the U.S. migrants are shown in Figure 16.[77] This policy experiment causes interesting differences for migrants already here illegally versus potential migrants in Mexico, because it makes legal immigration more attractive, since changing visa status within the U.S. is not allowed in the model. Simulation results reveal a substantial impact of the policy change, as a 20% or a 30 % decrease in illegal workers' wages in the U.S. decreases the stock of U.S. migrants by as much as 40% or 60%, respectively. This counterfactual experiment suggests that Mexican migrants are very sensitive to the changes in the U.S. wage. But, the time series plots show that the effect of the wage cut on the total number of Mexican migrants starts reversing about 10 years after the policy has been in effect, as immigration starts to rise again, though it does not return to the

[75] This is in line with what is found in the literature. For example, solving for the equilibrium number of U.S. migrants from small agricultural villages in Mexico, Colussi (2006) showed that the average trip length increases from about 5 years for the baseline model to more than 19 years when the cost of crossing the border is increased to six months of U.S. wages.

[76] Given the partial equilibrium context of the present analysis modelling the unskilled labor market in the U.S., I cannot predict how increased control on illegal employment would map into wages. To the degree that the policy change reduces the supply of unskilled workers in the U.S., the outcomes shown in this excercise should be interpreted as the upper bound of the effects of wage reductions in the U.S. on Mexican immigration.

[77] Colussi (2006) predicted that illegal worker's wage will decrease as penalties imposed on illegal immigrants' employers in the U.S. increase, because the employers are likely to transfer part of the expected penalties on the illegal employees.

baseline levels predicted with no wage cut. This can be explained by the change, caused by the alternative policy, in the composition of migrants flow in terms of the document status, which is endogenous to the model. For example, at the margin, a potential immigrant in Mexico who might have chosen to cross the border illegally under the status quo regime may consider applying for a travel document in response to the unexpected but permanent change in the relative price of each migration choice. This prediction is supported by Figure 17, which depicts the change in the proportion of undocumented immigrant in the U.S. under the two different regimes. The figure shows that the number of undocumented workers decreases rapidly relative to the number of documented under the counterfactual scenario, which is driving the substantial decrease of the total number of U.S. migrants over the first 5-10 years after the policy change in Figure 16.

9 Conclusion

In this paper, I develop a stochastic dynamic discrete choice model which explains the sequence of migrations of Mexican workers. The model explicitly incorporates a migrant's network at the U.S. destination as a determinant of migration decisions and endogenizes the decision to apply for legal documents to work in the U.S. Migration decisions are made so as to maximize expected life-time utility, and I allow individuals to differ in many unobserved dimensions including underlying location preferences. The structural parameters of the model are estimated using a source-country based retrospective panel data from the Mexican Migration Project.

The estimated model fits the decisions of location and application of the data reasonably well. The estimation results indicate that the size of a migrant's network at the destination affects the value of choosing the location choice as the same person is expected to earn higher wages and gains higher amenity values when he joins a exogenously larger network. The parameter estimates associated with unobserved heterogeneity are statistically significant, which suggests that it is important to allow for unobserved person-specific location preference in migration models. It also shows the importance of using a careful structural model instead of a non-structural estimation approach. The counterfactual experiments reveal that a change in immigration policy that reduces pecuniary and non-pecuniary utility flows associated with illegal employment is more effective in controlling the stock of U.S. migrants and unauthorized entry than an increase in number of annual visa quotas or an increase in the number of border patrol officers, as the proportion of U.S. immigrants decreases by as much as 60% in response to a policy that reduces the utility flows related to illegal U.S. residence.

In future work, I plan to explore the implication of saving decisions among migrants. As is well known, about 70% of Mexican immigrants remit the part of their earnings in the U.S. to the family in Mexico. The World Bank reported that more than $23 billion were sent back to Mexico in 2006, with remittances increasing at a rate of 15% per year since the 1990s. The MMP124 shows that about 80% of Mexican migrants remit an average of 40% of their labor income, which is about $300 per month. A growing literature studies the implication of remittance flows on various aspects of economics conditions in receiving countries in the context of development economics, but little is known about the link between remittances and the migration behavior of the senders. This issue is important because one distinguishing feature of Mexican immigration is that migration events tend to recur. Many Mexican immigrants are observed to return to their place of origin with substantial assets after spending a few years in the U.S., and often start operating small businesses upon return migration. This suggests that a major motivation for U.S. immigration among Mexicans is asset accumulation. By allowing for the possibility of saving, a structural model will allow me to investigate implications about the level, timing and duration of Mexican migration determined jointly with asset accumulation decisions.[78] Consideration of this extension is left for future work.

[78] Using a German panel data, Kirdar (2008) studied return migration behavior of immigrants and showed that an important motivation for foreign workers in Germany is to save, explaining return migration as a part of optimal life-cycle location decisions.

References

Berkovec, J. and S. Stern (1991): "Job Exit Behavior of Older Men," *Econometrica*, 59, 189–210. 4, 15, 18

Borjas, G. (1987): "Self-Selection and the Earnings of Immigrants," *American Economic Review*, 77, 531–553. 2, 27

Brien, M., L. Lillard, and S. Stern (2006): "Cohabitation, Marriage, and Divorce in a Model of Match Quality," *International Economic Review*, 47, 451–494. 19

Camarota, S. (2001): "Immigration From Mexico: Assessing the Impact on the United States," *Center for Immigration Studies*. 2

Carrington, W., E. Detragiache, and T. Vishwanath (1996): "Migration with Endogenous Moving Costs," *American Economic Review*, 86, 909–930. 26

Chiquiar, D. and G. Hanson (2005): "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States," *Journal of Political Economy*, 113, 239–281. 27

Cobb-Clark, D. (1993): "Immigrant Selectivity and Wages: The Evidence for Women," *American Economic Review*, 83, 986–93. 27

Colussi, A. (2006): "Migrants' Networks: An Estimable Model of Illegal Mexican Migration," Job Market Paper. 7, 35

Dresher, K. (1994): "Local Public Finance and the Residential Location Decisions of the Elderly," Ph.D. thesis, University of Wisconsin-Madison. 5

Duncombe, W., M. Robbins, and D. Wolf (2000): "Chasing the Elderly: Can State and Local Government Attract Recent Retirees?" Maxwell School of Citizenship and Public Affairs Aging Studies Program Paper. 5

Durand, J., D. S. Massey, and R. M. Zenteno (2001): "Mexican Immigration in the United States," *Latin American Research Review*, 36, 107–127. 27

Elbers, C. and G. Ridder (1982): "True and spurious duration dependence: The identifiability of the proportional hazard model," *Review of Economic Studies*, 49, 403–409. 22

Gemici, A. (2008): "Family Migration and Labor Market Outcomes," Working Paper. 5

Geweke, J. (1988): "Antithetic acceleration of Monte Carlo integration in Bayesian inference," *Journal of Econometrics*, 38, 73–89. 32

Hanson, G. (2006): "Illegal Migration from Mexico to the United States," *Journal of Economics Literature*, 44, 869–924. 2, 7, 16, 22

Hanson, G. and A. Spilimbergo (1999): "Illegal Immigration, Border Enforcement and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border," *American Economic Review*, 89, 1337–1357. 5, 24

Hanson, G. and C. Woodruff (2004): "Emigration and Educational Attainment in Mexico," Working Paper. 6

Heckman, J. and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320. 18, 22

Holt, F. M. (1997): "Family Migration Decisions: A dynamic Analysis," Ph.D. thesis, University of Virginia. 5, 6

Hong, G. (2009): "Measuring the Labor Market Performance of Mexican Workers in the Presence of Network Effects," Working Paper. 25, 32

Ichimura, H. and L. Lee (1991): "Semiparametric least squares estimation of multiple index models: single equation estimation," in *Non-parametric and semiparametric methods in econometrics and statistics*, ed. by W. Barnett, J. Powell, and G. Tauchen, Cambridge University Press. 29

Ioannides, Y. M. and L. D. Loury (2004): "Job information networks, neighborhood effects, and inequality," *Journal of Economic Literature*, 42, 1056–1093. 26

Keane, M. P. and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy*, 105, 473–522. 15

Kennan, J. and J. Walker (2008): "The Effect of Expected Income on Individual Migration Decisions," *Econometrica*, forthcoming. 6, 12, 20

Kirdar, M. (2008): "Labor Market Outcomes, Savings Accumulation and Return Migration," *Labour Economics*, forthcoming. 37

Manski, C. and D. McFadden (1981): "Alternative Estimators and Sample Designs for Discrete Choice Analysis," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. Manski and D. McFadden, The MIT Press, 2–50. 22

Massey, D., R. Alarcon, J. Durand, and H. Gonzalez (1987): *Return to Aztlan: The Social Process of International Migration from Western Mexico*, University of California Press. 3

Massey, D., J. Arango, G. Hugo, A. Kouaouci, A. Pelegrino, and J. Taylor (1994): "An Evaluation of International Migration Theory: The North American Case," *Population and Development Review*, 20, 699–751. 27

Massey, D. and K. Espinosa (1997): "What's Driving Mexico-U.S. Migration? A Theoretical, Empirical, and Policy Analysis," *American Journal of Sociology*, 102, 939–999. 6, 22

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, 1–24. 22

Montgomery, J. (1991): "Social Networks and Labor-Market Outcomes: Towards an Economic Analysis," *American Economic Review*, 81, 1408–1418. 26

Munshi, K. (2003): "Networks in the Modern Economy: Mexican Migrants in the US Labor Market," *Quarterly Journal of Economics*, 118, 549–597. 6, 26

Orrenius, P. and M. Zavodny (2005): "Self-Selection among Undocumented Immigrants from Mexico," *Journal of Development Economics*, 78, 215–540. 22

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, 999–1033. 18

Rust, J. and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 65, 781–831. 15

Sjaastad, L. (1962): "The Costs and Returns of Human Migration," *The Journal of Political Economy*, 72, 80–93. 2

Stern, S. (1996): "Semiparametric Estimates of the Supply and Demand Effects of Disability on Labor Force Participation," *Journal of Econometrics*, 71, 49–70. 29

Sullivan, P. (2010): "A Dynamic Analysis of Educational Attainment, Occupational choices and Job Search," *International Economic Review*, 51, 289–317. 15

Swann, C. A. (2002): "Maximum Likelihood Estimation Using Parallel Computing: An Introduction to MPI," *Computational Economics*, 19, 145–178. 19

Walker, J. (2006): "Economic Perspectives on Family and Migration," Working Paper. 6

Woo, S. (2005): "Tiebout Migration and Retirement of Older Workers," Job Market Paper. 6

Woodruff, C. and R. Zenteno (2007): "Migration Networks and Microenterprises in Mexico," *Journal of Development Economics*, 82, 509–528. 24

Table 1: Migration Pattern in the MMP124

	No Domestic Migration	Domestic Migration	All
No U.S. Migration	6826 (34.8%)	4862 (24.8%)	11688 (59.6%)
U.S. Migration	4644 (23.6%)	3308 (16.8%)	7952 (40.4%)
All	11470 (58.4%)	8170 (41.6%)	19640 (100%)

1. Every household head in the MMP124 except for 981 U.S. households who were surveyed in the U.S. through the "snow-ball" sampling method is included.

Table 2: Domestic Migration Pattern in the MMP124

	North	West	South	All
North	867 (7.29%)	308 (2.59%)	98 (0.82%)	1273 (10.7%)
West	1472 (12.37%)	5034 (42.81%)	1068 (8.98%)	7574 (63.66%)
South	337 (2.83%)	616 (5.18%)	2098 (17.63%)	3051 (25.64%)
All	2676 (22.49%)	5958 (50.08%)	3264 (27.43%)	11898 (100%)

1. Conditional on moving domestically, row variables indicate departing locations, and column variables indicate ending locations.

Table 3: Sample Construction

Inclusion Criterion	Number of Remaining Person-Year Observations
MMP124	948,832
Keep if male.	801,502
Drop if missing information on application decisions.	412,652
Keep if 1965≤Year≤2004.	303,383
Drop if missing information on education, occupation.	302,912
Drop if missing information on state of birth, city of birth.	295,202
Drop if missing information on document status.	295,081
Drop if legalized through amnesty or special contracts	286,449
Keep if age is between 16 and 20 in 1965.	18,720

Table 4: Descriptive Statistics

Variable	Mean	Std Dev	Min	Max
Panel A : Sample of 673 individuals				
Highschool	0.046	0.209	0	1
Hometown-North	0.153	0.360	0	1
Hometown-West	0.509	0.500	0	1
Years in the sample	27.81	1.603	12	30
log(distance to rail-line)	1.826	2.675	0	7.273
Never been to the U.S.	0.672	0.419	0	1
Panel B : Sample of 18,720 person-year observations				
Year	1978.4	8.098	1965	1994
Age	31.51	8.101	16	45
North	0.144	0.351	0	1
West	0.462	0.498	0	1
US	0.061	0.239	0	1
Agriculture	0.445	0.496	0	1
Blue-collar	0.417	0.493	0	1
White-collar	0.132	0.338	0	1
Panel C : Sample of 1,147 person-year observations in the U.S.				
Total years of U.S. experience	6.989	8.025	0.083	31
Documented	0.181	0.385	0	1
Agriculture	0.291	0.498	0	1
Blue-collar	0.572	0.422	0	1

Table 5: Parameter Estimates of the Mexico Wage Equation

	[1] Agriculture	[2] Blue-Collar	[3] White-Collar
Constant	4.497**	6.173**	6.208**
	(0.827)	(0.669)	(1.397)
Education	-1.149	-0.808	-0.033
	(1.146)	(0.808)	(0.365)
Work Experience in Mexico	0.106**	0.098**	0.130**
	(0.017)	(0.010)	(0.017)
Work Experience in the U.S.	0.301**	0.175**	0.323
	(0.095)	(0.076)	(0.480)
North	1.404**	-0.960**	0.042
	(0.501)	(0.235)	(0.512)
West	1.693**	0.819**	1.250**
	(0.434)	(0.197)	(0.327)
Age	-0.049	0.134**	0.211**
	(0.059)	(0.042)	(0.070)
*Age*Age*	-0.001	0.001	0.002
	(0.001)	(0.001)	(0.002)
N=5,461			

1. The parameters of the selection equation are omitted.
2. *,** indicate significance at the 10% and 5% levels.
3. *North* is a dummy that takes a value of 1 if an individual is employed in *North*.
4. *West* is a dummy that takes a value of 1 if an individual is employed in *West*.
4. Standard errors are in parentheses.

Table 6: Parameter Estimates of the U.S. Wage Equation

Panel A :	**Wage Equations**	
	Agriculture	Blue-Collar
Education	0.188*	0.051**
	(0.110)	(0.020)
Work experience in Mexico	0.012**	-0.001
	(0.002)	(0.001)
Work experience in the U.S.	-0.008	0.004**
	(0.005)	(0.002)
Documented	0.120**	0.015
	(0.061)	(0.015)
log(Distance to Rail)	-0.026*	-0.001
	(0.013)	(0.003)
Panel B :	**Visa Equation**	
Education	0.272**	
	(0.068)	
Years in US	0.093**	
	(0.004)	
Rail	-0.187**	
	(0.012)	
Panel C :	**Covariance Matrix**	
σ_{12}	-0.537**	
	(0.024)	
σ_{13}	0.988**	
	(0.024)	
σ_{2}	0.822**	
	(0.013)	
σ_{23}	-0.518**	
	(0.024)	
N=7,417		

1. The parameters of the selection equation are omitted.
2. *,** indicates significance at the 10%, and 5% levels.
3. Standard errors are in parentheses.

Table 7: Non-structural Random Effects Logit Regression of U.S. Out-migration Decisions

U.S. Location Choice	
Education	-0.891
	(0.677)
Duration	-0.262**
	(0.013)
Documented	4.521**
	(0.0294)
Wage in the U.S.	0.021**
	(0.007)
Wage in Mexico	-0.609**
	(0.079)
log(Distance to Rail)	-0.326**
	(0.050)
Age	0.444**
	(0.047)
*Age*Age*	-0.011**
	(0.564)
Constant	-3.328**
	(0.564)
σ	2.035**
	(0.093)
ρ	0.557**
	(0.022)
N=18,360	

1. *,** indicates significance at the 10%, and 5% levels.
2. Standard errors are in parentheses.
3. ρ indicates the proportion of the total variance contributed by the panel-level variance.

Table 8: Multinomial Logit Estimates of the Occupation Transition

	[1] Within MX		[2] Within US		[3] From MX to US		[4] From US to MX	
	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.	Coef.	Std.Err.
Unemployed								
Lagged(Unemp)	8.957**	(0.045)	7.510**	(0.175)	2.062**	(0.176)	1.227**	(0.204)
Lagged(Agri)	2.555**	(0.065)	1.994**	(0.242)	-0.253	(0.164)	-0.272*	(0.125)
Lagged(White)	3.349**	(0.076)	2.904**	(0.412)	0.632**	(0.204)	0.823*	(0.390)
Age	0.046**	(0.001)	0.056**	(0.004)	0.061**	(0.005)	0.427**	(0.005)
High School	0.000	(0.000)	0.000	(0.000)	0.000*	(0.000)	0.000	(0.000)
Constant	-6.649**	(0.065)	-7.309**	(0.256)	-5.102**	(0.215)	-4.002**	(0.195)
Age at mig.	-	-	0.000	(0.000)	-	-	-	-
Last(Unemp)	-	-	-	-	-	-	2.838**	(0.174)
Last(Agri)	-	-	-	-	-	-	0.825**	(0.141)
Last(White)	-	-	-	-	-	-	0.986**	(0.243)
Agriculture								
Lagged(Unemp)	2.575**	(0.098)	1.885**	(0.432)	-0.095	(0.130)	-0.148	(0.222)
Lagged(Agri)	8.381**	(0.028)	7.295**	(0.071)	1.189**	(0.045)	-0.623**	(0.067)
Lagged(White)	2.810**	(0.082)	2.350**	(0.346)	-0.332**	(0.098)	0.230	(0.365)
Age	0.013**	(0.001)	0.005	(0.003)	-0.003	(0.002)	-0.005	(0.003)
High School	0.000	(0.000)	0.001	(0.000)	0.000	(0.000)	0.000	(0.000)
Constant	-5.016**	(0.043)	-4.286**	(0.122)	-0.476**	(0.072)	-1.789**	(0.117)
Age at mig.	-	-	-0.001	(0.000)	-	-	-	-
Last(Unemp)	-	-	-	-	-	-	1.059**	(0.206)
Last(Agri)	-	-	-	-	-	-	3.601**	(0.070)
Last(White)	-	-	-	-	-	-	0.497*	(0.214)
White-Collar								
Lagged(Unemp)	3.363**	(0.076)	3.230**	(0.037)	1.143**	(0.345)	0.304	(0.278)
Lagged(Agri)	2.633**	(0.056)	2.225**	(0.208)	-0.214	(0.246)	-0.294*	(0.119)
Lagged(White)	8.684**	(0.041)	8.468**	(0.162)	2.136**	(0.202)	1.798**	(0.294)
Age	-0.022**	(0.001)	-0.014*	(0.006)	0.005	(0.009)	-0.022**	(0.006)
High School	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)	0.000	(0.000)
Constant	-3.983**	(0.053)	-4.642**	(0.023)	-3.993**	(0.307)	-1.892**	(0.197)
Age at mig.	-	-	0.000	(0.000)	-	-	-	-
Last(Unemp)	-	-	-	-	-	-	0.703*	(0.326)
Last(Agri)	-	-	-	-	-	-	0.785**	(0.151)
Last(White)	-	-	-	-	-	-	3.822**	(0.139)
Nobs.	455312		36071		10009		8443	

1. Occupation transition functions were estimated outside the structural model.
2. *,** indicate significance at the 5% and 1% levels.
3. *Lagged(Occupation)* is the occupation held in the previous period.
4. *Last(Occupation)* is the last occupation held in Mexico before moving to the U.S.

Table 9: Cox Proportional Hazard Estimates of Waiting Time for Visas

	[1] All	[2] Family-Based Immigration	[3] Employment-Based Immigration
Highschool	0.3125	0.1393	0.9793**
	(0.2297)	(0.8709)	(0.3681)
Work Experience in Mexico	0.001	-0.0031*	0.0319*
	(0.0079)	(0.0153)	(0.0163)
Work Experience in the U.S.	-0.066**	-0.0479**	0.0044
	(0.0130)	(0.0154)	(0.0131)
Distance to rail-line	-0.001**	-0.0011	-0.002**
	(0.0003)	(0.0007)	(0.0004)
Family in the U.S.	0.518**	1.8277**	-0.237
	(0.1999)	(0.5867)	(0.1685)
Number of Available Visas	1.84e-06	6.67e-06*	4.45e-06
	(1.94e-06)	(3.37e-06)	(6.12e-06)
Nobs.	364	178	92

1. Waiting time functions were estimated outside the structural model.
2. *,** indicate significance at the 5% and 1% levels.
3. *Family in the U.S.* is a dummy that indicates existence of family or extended family in the U.S.

Table 10: Parameter Estimates of the Choice Process

Panel A :	**Non-Pecuniary flows**		
	North	West	U.S.
Constant	2.061**	3.631**	2.419**
	(0.171)	(1.005)	(0.124)
Constant (Type 2)	-1.240	2.778**	-2.174**
	(1.008)	(0.831)	(0.360)
Hometown	0.132	1.245*	-
	(0.105)	(0.732)	
Duration	-0.356	-0.733**	-1.804*
	(0.297)	(0.312)	(1.121)
Duration (Type 2)	-	-	-1.686**
			(0.512)
High School	-	-	-0.433
			(1.051)
Documented	-	-	13.244**
			(3.523)
log(Distance to Rail)	-	-	-0.051**
			(0.012)
Panel B :	**Moving Cost function**		
North-West	67.30**		
	(12.64)		
North-South	61.21**		
	(26.04)		
North-U.S.	102.59**		
	(32.49)		
West-South	65.05**		
	(14.27)		
West-U.S.	112.16**		
	(10.08)		
South-U.S.	99.00**		
	(46.21)		
Undocumented	180.85**		
	(66.44)		
Application cost	74.28*		
	(39.69)		
Type 2	-12.28*		
	(6.40)		
Panel C :	**Unobserved Heterogeneity**		
Type 2 Proportion	0.472**		
	(0.012)		
Type 2 in Waiting Time	-0.650*		
	(0.346)		

1. *,** indicate significance at the 10% and 5% levels.
2. Standard errors are in parentheses.

Table 11: Model Fit - Location Choice by Age

	U.S.		North		West	
Age	Data	Model	Data	Model	Data	Model
16-20	3.5%	5.5%	14.6%	16.7%	48.2%	46.3%
21-25	6.5%	5.0%	14.2%	17.2%	45.7%	48.7%
26-30	7.2%	6.7%	14.2%	17.8%	45.2%	49.8%
31-35	6.4%	8.4%	14.3%	18.1%	45.9%	50.0%
35-40	7.1%	9.2%	14.3%	17.6%	45.6%	50.3%
41-45	13.5%	9.4%	14.0%	18.0%	45.3%	50.1%

Table 12: Model Fit - Application Choice by Age

	Proportion of Visa Applicants	
Age	Data	Model
16-20	0.30%	0.89%
21-25	0.32%	0.59%
26-30	0.46%	0.74%
31-35	0.59%	1.63%
35-40	0.71%	1.78%
41-45	0.89%	1.48%

Table 13: Model Fit - Proportion of Undocumented Migrants by Age

	Undocumented Migrants in the U.S.	
Age	Data	Model
16-20	78.8%	86.4%
21-25	88.6%	82.0%
26-30	79.7%	71.2%
31-35	69.3%	56.4%
35-40	52.0%	38.4%
41-45	40.5%	24.8%

Table 14: Descriptive Statistics: 1975 Cohort

Variable	Mean	Std Dev	Min	Max
Panel A : Sample of 884 individuals				
Highschool	0.110	0.320	0	1
Hometown-North	0.119	0.323	0	1
Hometown-West	0.490	0.499	0	1
Years in the sample	26.37	2.048	12	30
log(distance to rail-line)	1.921	2.252	0	7.273
Never been to the U.S.	0.662	0.405	0	1
Panel B : Sample of 23,163 person-year observations				
Year	1987.7	7.706	1975	2004
Age	30.83	7.771	16	45
North	0.142	0.349	0	1
West	0.409	0.491	0	1
US	0.084	0.277	0	1
Agriculture	0.335	0.472	0	1
Blue-collar	0.498	0.500	0	1
White-collar	0.160	0.366	0	1
Panel C : Sample of 1,953 person-year observations in the U.S.				
Total years of U.S. experience	7.699	6.313	0.083	29
Documented	0.243	0.358	0	1
Agriculture	0.219	0.414	0	1
Blue-collar	0.718	0.449	0	1

Figure 1: Map of Mexico

Figure 2: Border Patrol Enforcement Budget in Nominal U.S. Dollars

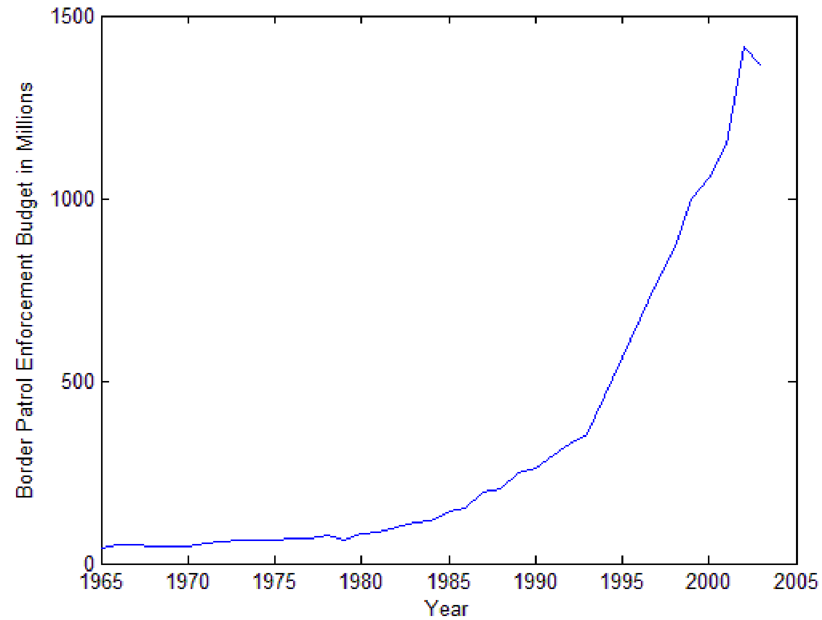

Figure 3: Linewatch Hours Spent by year

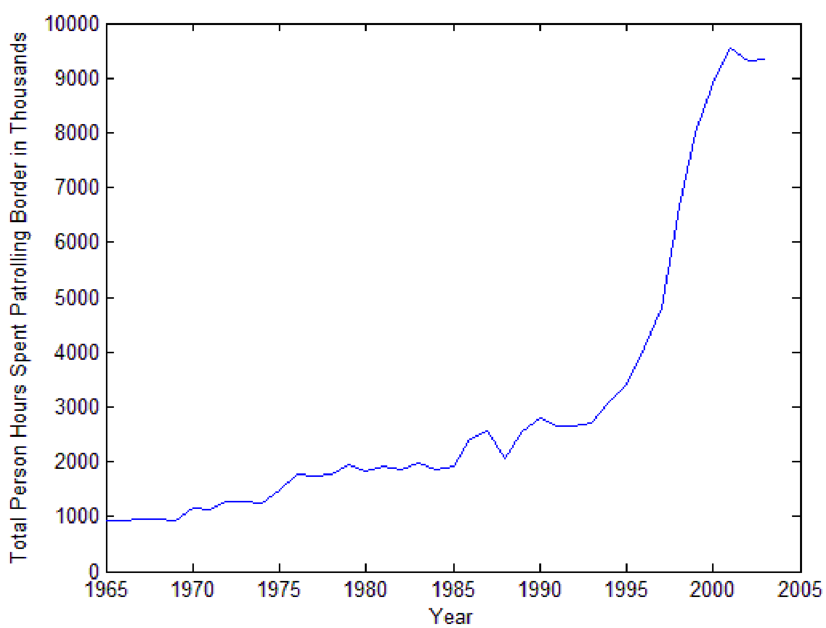

Figure 4: Number of Border Enforcement Officers by Year

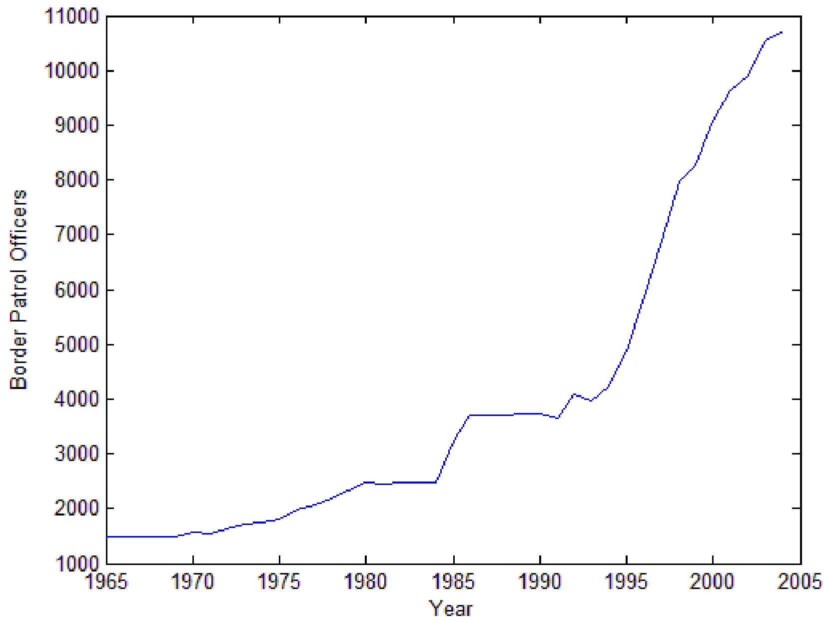

Figure 5: Number of Apprehensions of Mexicans for Illegal Border Crossing

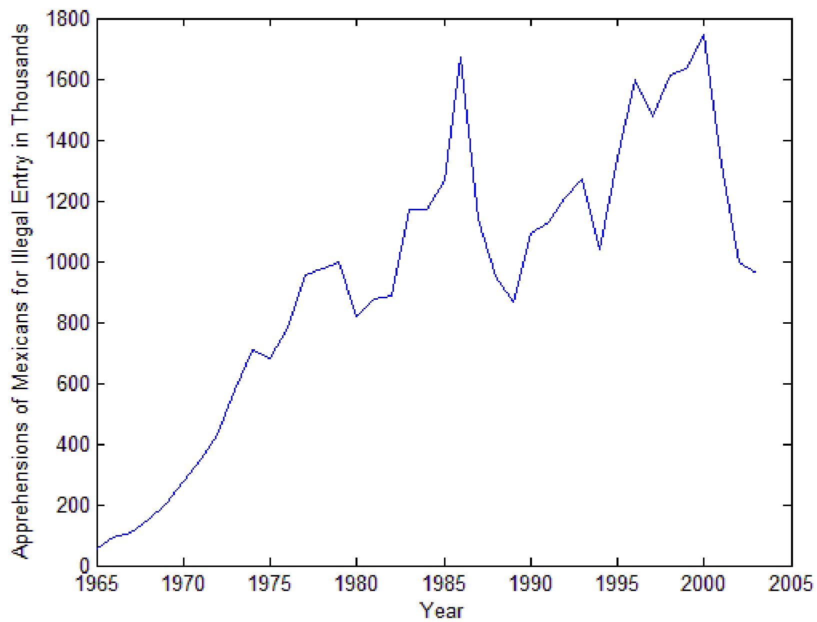

Figure 6: The Likelihood of Arrest while Attempting to Cross the Border Illegally

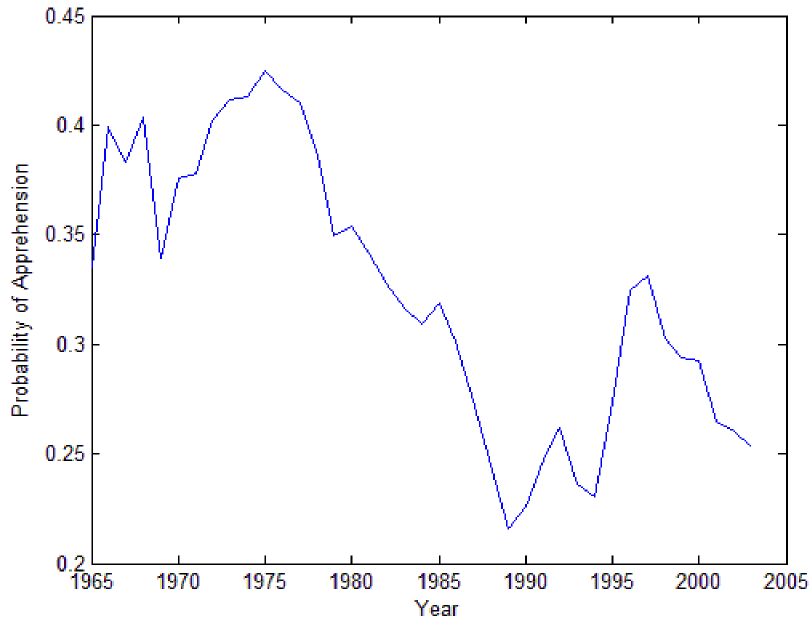

Figure 7: Number of Available Visas by Year

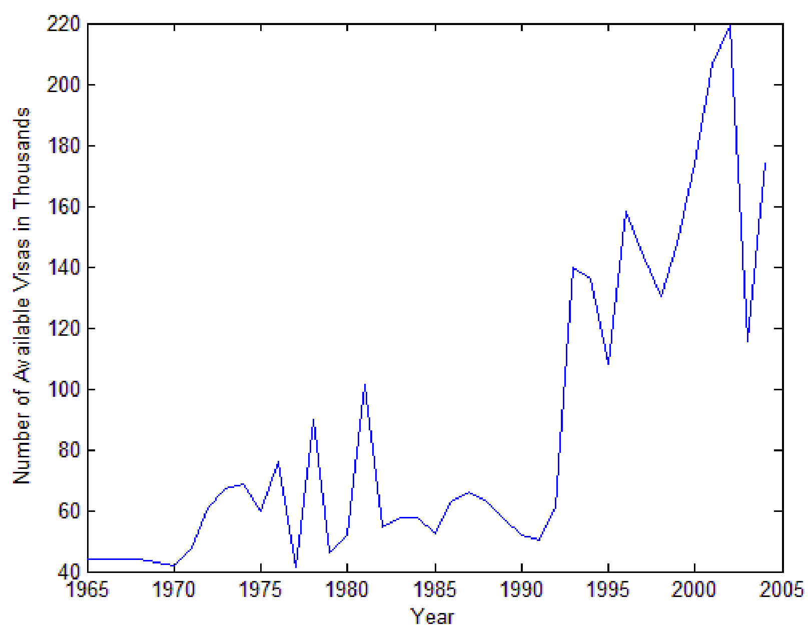

Figure 8: Model Fit - Percentage of U.S. Migrants

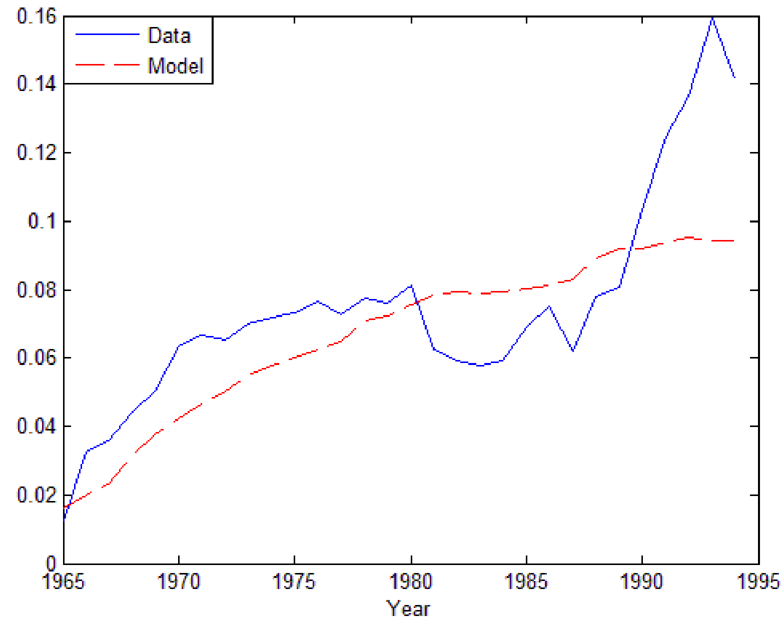

Figure 9: Model Fit - Percentage of Undocumented Migrants in the U.S.

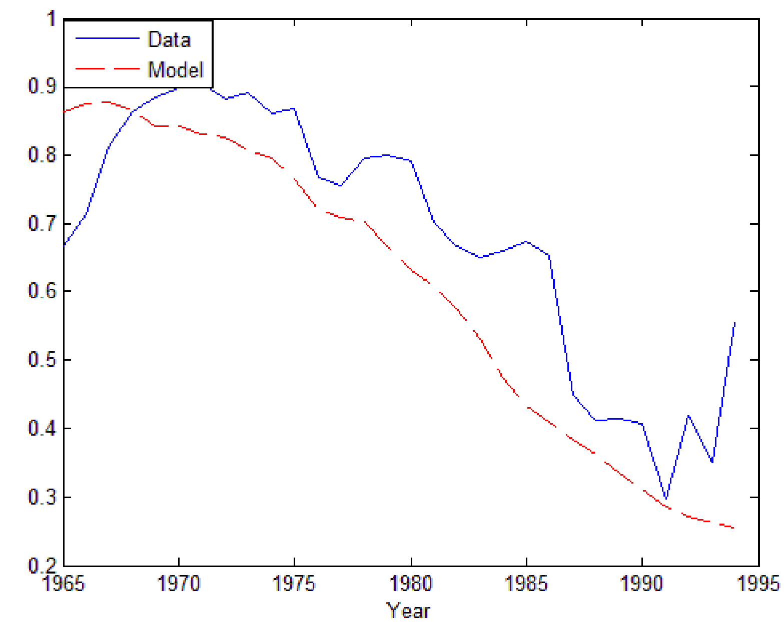

Figure 10: Policy Experiment (1) - Increase in Number of Available Visas by 100%

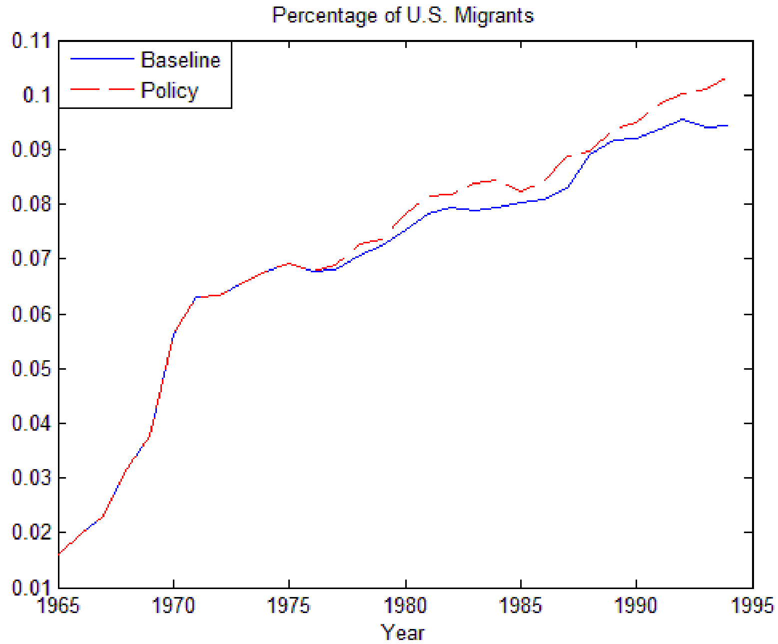

Figure 11: Policy Experiment (1) - Increase in Number of Available Visas by 100%

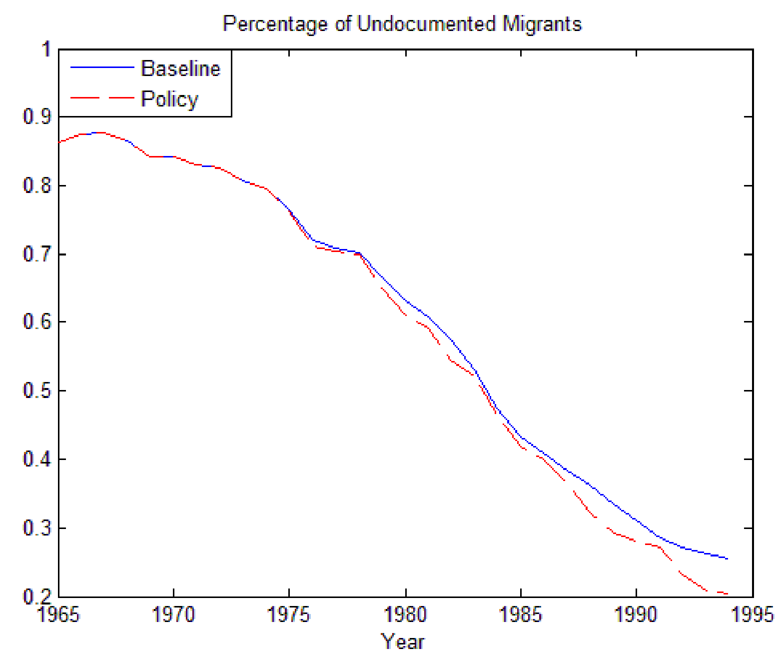

Figure 12: Policy Experiment (2) - Permanent Increase in the Number of Border Patrol Officers by 100%

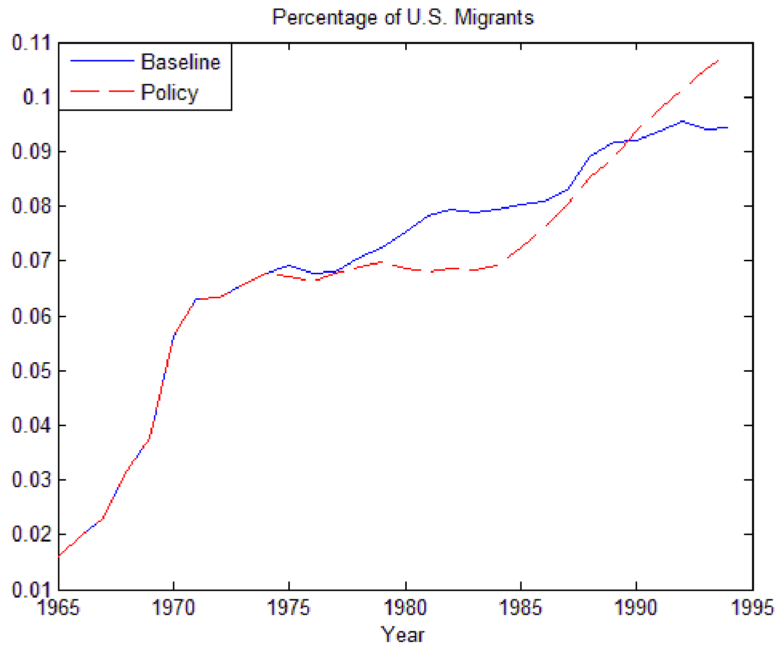

Figure 13: Policy Experiment (2) - Permanent Increase in the Number of Border Patrol Officers by 100%

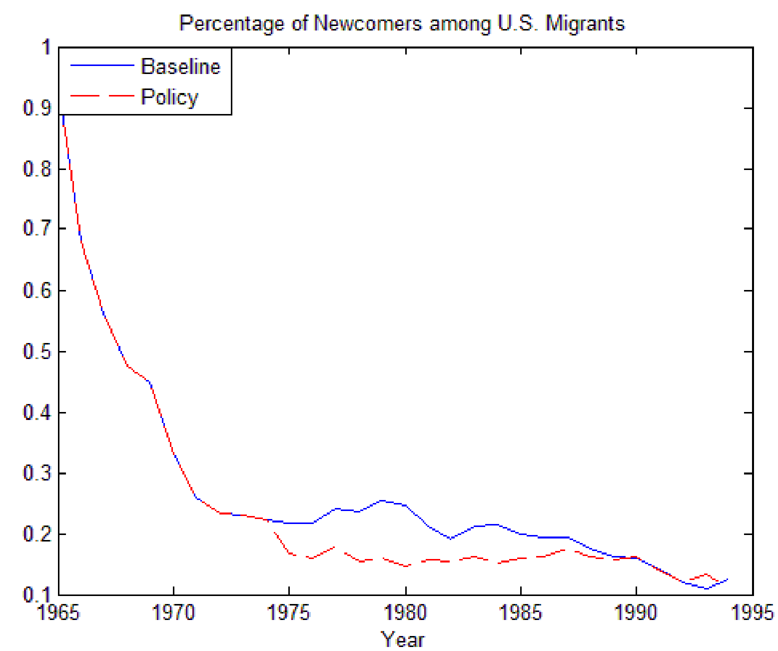

Figure 14: Policy Experiment (2) - Permanent Increase in the Number of Border Patrol Officers by 100%

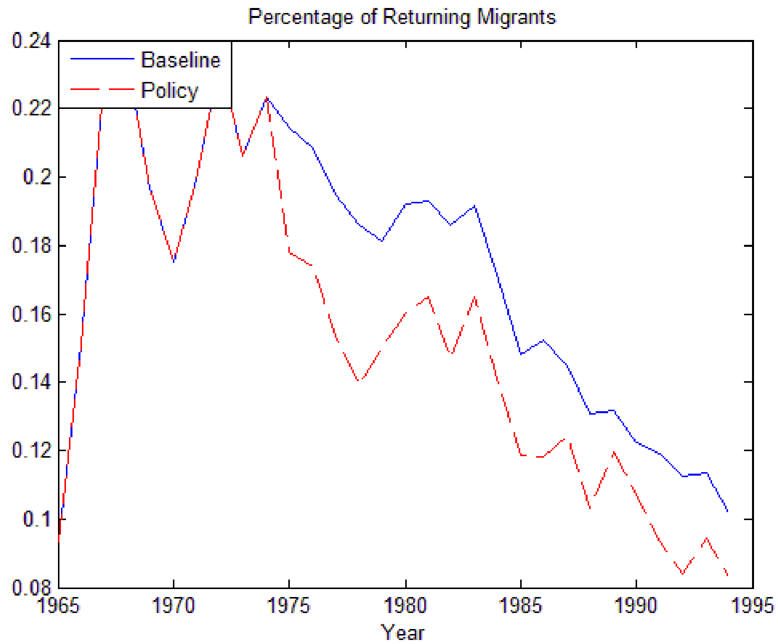

Figure 15: Policy Experiment (2) - Permanent Increase in the Number of Border Patrol Officers by 100%

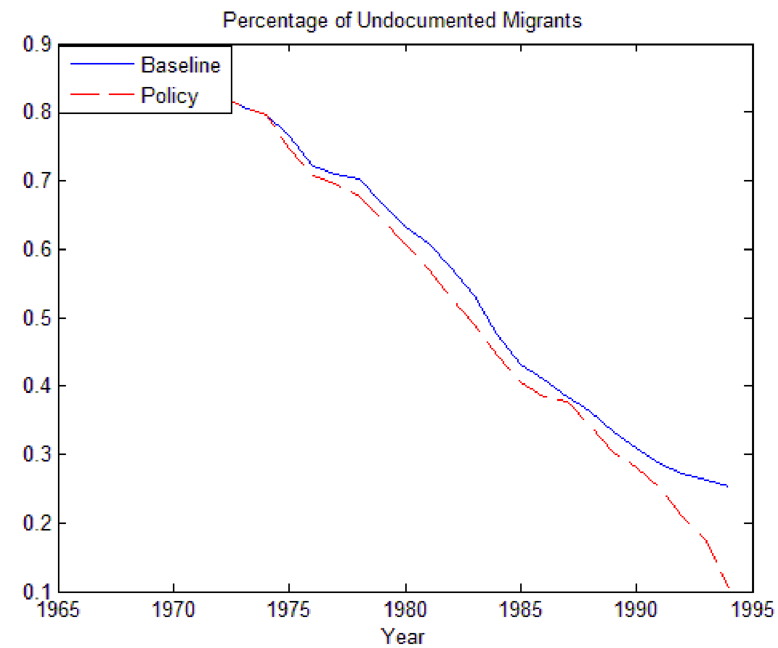

Figure 16: Policy Experiment (3) - Permanent Reduction in U.S. Wages of Undocumented Workers

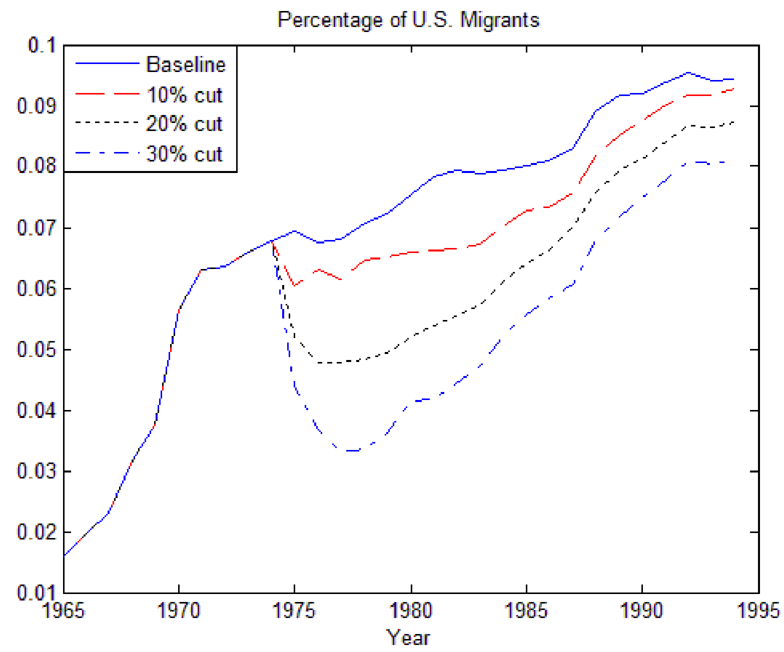

Figure 17: Policy Experiment (3) - Permanent Reduction in U.S. Wages of Undocumented Workers

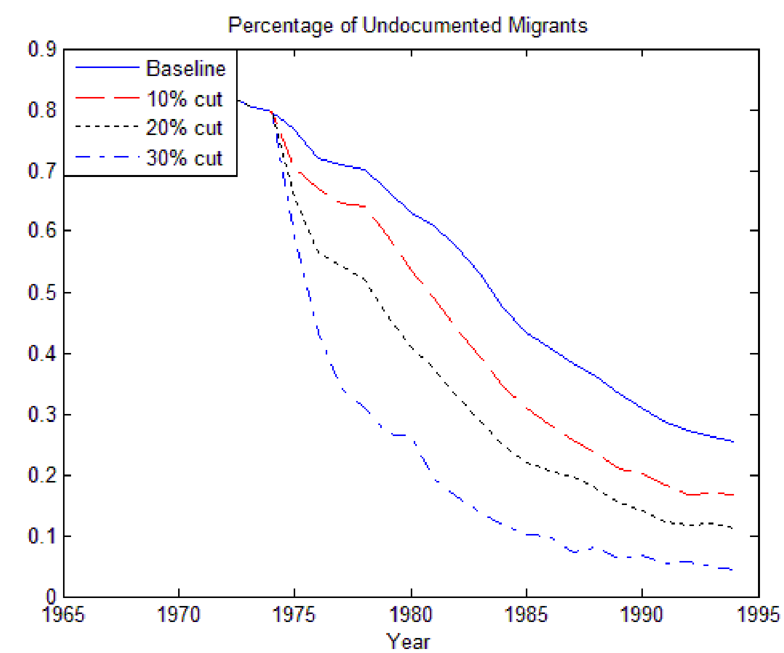

Figure 18: Out-of-Sample Fit - Percentage of U.S. Migrants