The Effects of Border Enforcement on Migrants' Border Crossing Choices: Diversion, or Deterrence?[1]

Carmen Carrión-Flores
Department of Economics
The University of Arizona

Todd Sorensen[2]
Department of Economics
The University of Arizona

This Draft: March 2006
Preliminary Version; Please Do Note Cite Without Authors Permission

Abstract

Using an aggregate discrete choice model from the Industrial Organization literature, we estimate the impact of an increase in intensity of enforcement by the US Border Patrol on rates of migration from Mexico. As our measure of enforcement intensity varies across both time and space, we are able to decompose the effect of increased enforcement in one location into two parts: the *deterrence* of migration (migrants who are shifted back to their home village by an increase in border enforcement) and the *diversion* of migration (migrants who continue to migrate, but through a different crossing). While more intense border enforcement in a given sector does succeed in deterring migrants, it also diverts a significant number of migrants to other sectors. These findings suggest that estimates identifying the effect of border enforcement on migration into a particular geographic zone may overstate the impact of enforcement on migrations into the US.

Keywords: Illegal Migration, Border Enforcement, Aggregate Discrete Choice Models.

JEL Classification: F22, C35, C25.

---

[1] The authors wish to thank Gordon Hanson for providing us with enforcement data from Hanson and Spilimbergo (1999). Valueable comments have been provided by Manuela Angelucci, Greg Crawford, Price Fishback, Alfonso Flores-Lagunes, Kei Hirano, Ron Oaxaca, Alex Shcherbokov, and participants at presentations at the University of Arizona seminar series, the Western Economics Association 2005 Meetings and the Southern Economics Association.

[2] Corresponding Author:
Todd Sorensen,
Department of Economics
The University of Arizona
1130 E. Helen Street
Tucson, AZ 85721-0108
Phone: 520.621.6224; Fax 520.621.8450
E-mail: sorensen@u.arizona.edu

1 Introduction

There is increasing interest by academics and policy makers in Mexican migration to the United States. The migration flow from Mexico is relevant to the United States since Mexico has been the most important source country for United States' illegal immigration. The Pew Hispanic Center estimates that out of the 10.3 million of unauthorized immigrants in the United States, 5.9 million are unauthorized Mexican immigrants. Moreover, it is well established that the majority of new arrivals from Mexico enter the United States Illegally.

In recent years, the migration literature has increased its attention on issues surrounding undocumented migration from Mexico to the United States. A number of previous studies have focused on the effects of border enforcement on rates and duration of migration. It is well established in the literature that increased enforcement along the entire border decreases rates of migration, ceterius paribus. However, anecdotal evidence as well as one prior econometric study suggest that tightening enforcement along one portion of the border has the effect of shifting migrants into another sector. If this is true, estimates of the effect of enforcement that do not take into account this shifting of migrants may overstate the impact of increased enforcement.

In this paper, we identify the effect of enforcement on the rate of migration through 9 US Border Patrol sectors. Using an aggregate discrete choice model from the IO literature, we identify the effect of enforcement and distance travelled on the utility of a potential migrant. In order to take account of the endogensiety between migration rates and enforcement, we instrument for the level of enforcement in each sector with variables that should affect the ability of the sectors' political representatives to obtain additional funding for border enforcement.

After obtaining these parameter estimates, we compute marginal effects of enforcement on migration rates. We separately identify two effects of enforcement on migration rates: (1) the share of potential migrants that stay in Mexico as a result of increased enforcement (migration deterrence) and (2) the share of potential migrants who continue to travel to the US, but are diverted into a different border crossing sector (migration diversion). We find that, once we take account of group effects, approximately 53% of migrants shifted away from a given crossing sector by increased enforcement substitute to not migrating, while 47% choose to cross through other sectors into the US.

The rest of the paper proceeds as follows. Section 2 discusses the literature on border enforcement and the literature on location choice. Section 3 presents our econometric model, and Section 4 describes our data. Section 5 presents our results, Section 6 discusses future work, and Section 7 concludes.

2 Border Enforcement and Location Choice Literature

A number of recent papers have sought to identify effects of border enforcement rates on border crossings. Gordon Hanson (2005) provides a comprehensive review of the literature on illegal migration from Mexico to the United States. Hanson and Spilimbergo (1999) was one of the first papers to look at the estimates of the effects of border enforcement on apprehensions. As Hanson and Spilimbergo noted, enforcement is endogenous to apprehensions so they instrument using U.S. government expenditures on national defense and the timing of U.S. presidential, gubernatorial and senate elections. They found that there is a causal relationship between the number of apprehensions and enforcement, though OLS estimates are biased upwards as would be expected.

Hanson and Spilimbergo (2001) model the importance of political lobbying on border enforcement. They find that price shocks to sectors employing large numbers of undocumented workers are negatively correlated with the level of border enforcement, consistent with an effect of behind the scenes lobbying by employers in these industries. Hanson, Robertson and Spilimbergo (2002) study the impact of border enforcement on labor markets on the U.S. and Mexican border. They find that increased border enforcement has little impact on labor markets on either the United States or Mexican side of the border. This is consistent with two hypotheses: either enforcement has little impact on migration levels, or immigration has little impact on wages

Cornelius (2001) found that border enforcement has resulted in rechanneling flows of unauthorized migrants towards more hazardous areas and discouraging unauthorized migrants already in the United States from returning to their places of origin. However, Cornelius did not find evidence that the strategy is deterring or preventing significant numbers of new illegal entries, particularly given the absence of a serious effort to curtail employment of unauthorized migrants through work-site enforcement.

Orrenius and Zavodny (2003) examined whether mass legalization programs reduce future undocumented immigration. They found that right after the 1986 Immigration Reform and Control Act, the apprehensions of persons attempting to cross the U.S.-Mexico border illegally declined but the army program did not change long-term patterns of undocumented immigration from Mexico.

Bean, et al (1994) studies the effect of “Operation Hold the Line” in the Border Patrol’s El Paso sector. This operation marked a sharp change in the enforcement strategy, shifting the focus of enforcement from internal checkpoints to line watch. They consider this change to be an exogenous shock to the level of border enforcement. In contrast to other findings, the increase in enforcement intensity reduces the number of apprehensions. They find evidence that half of the decrease in the flow of migrants, as measured by the level of apprehensions, was offset by an increase in flows to other border sectors.

Carrión-Flores (2005) finds that increased border enforcement may increase the duration of migrants engaging in repeat migrations. Angelucci (2005) develops a model of return migration, and estimates the impacts of increased border enforcement on both the probability of a migrant undertaking a trip from Mexico to the United States, and of migrants already living in the United States returning to Mexico. She finds that while increased border enforcement discourages migrants from crossing into the United States, it may discourage the return to Mexico of migrants already in the United States.

Previous research on the determinants of the location choice of migrants within the US is also relevant for our study. If the economic outcomes of migrants depend upon their location choice within the US, and enforcement shifts migrants between enforcement sectors, border policy may have important economic consequences not only for potential migrants but also for migrants who continue to migrate into the US.

The literature on the location choice of migrants has focused on the importance of clustering and networks. Using a conditional logit analysis of the internal location choices of international migrants, Bartel (1989) finds that international migrants are more likely to choose locations with large numbers of other migrants from their home regions. From 1990 census data, Chiswick and Miller (2004) find continued agglomeration of international migrants by language. Jaeger (2004) studies the location choice of newly arrived international migrants. He finds that migrants are more likely to locate in areas with disproportionately high foreign born populations. These results can be explained by the literature on network externalities. Using data from the MMP, Munshi (2003) finds that the probability of a migrant being employed increases with the number of members of the home community in the destination.

Durand et al. (2000) found that early in the century, Mexicans went primarily to Texas, but after 1910, California emerged as a growing pole of attraction. California continued to gain at the expense of Texas through the 1920s and 1930s, but it did not surpass Texas until the Bracero Program of 1942-1964. Following the demise of this program, California came to dominate all other destinations; but since 1990, Mexican immigration has shifted away from it toward new states that never before have received significant numbers of Mexicans.

Kanbur and Rapoport (2005) explored an alternative route, based on a theoretical and empirical proposition of the migration literature, namely, that migration is a selective process. Focusing on skilled migration, they demonstrated the different forces in play that make selective migration a force for both divergence and convergence, and characterizes where each set of forces dominates. They also analyzed the consequences for convergence of combining migration selectivity and agglomeration effects arising from migrant networks.

Bauer T., et al. (2005), also using MMP data, find the existence of a network of individuals from the migrant's home community in a potential destination significantly increases the probability of choosing that destination. Moreover, they show that migrants choose smaller networks as their English language proficiency improves. Garcia (2005) found three distinct yet disconnected subnetworks in an immigrant network in a rural town in northeastern Oklahoma, a traditional subnetwork, a church subnetwork, and a contract subnetwork. Migrants tend to migrate to those areas where social networks exist.

3 Econometric Model and Identification

3.1 The Border Crossing Choice of a Representative Migrant

Consistent with Sjaastad (1962), we view migration as an investment, with a potential migrant maximizing return on the investment. In our model, a potential migrant chooses between ten possible choices. The outside option is to remain in the origin village; the utility of this option is normalized to 0. The other nine options in the potential migrant's choice set represent migrating to an unspecified location in the United States, and crossing through one of the nine border crossing sectors. The utility of the i^{th} potential migrant, originating in village l, to migrating through the j^{th} border crossing sector in year t can be expressed as

$$U_{ij} = \pi(l, j, t) - C(\ln(D_{jl}), Enf_{jt}, UR_{jt}, l, j, t) \text{ for } j = 1, 2, ..9 \quad (1)$$

$$U_{ij} = 0 \text{ for } j = 0 \quad (2)$$

where π represents the gross benefits to migrating, and is a function of the origin of the migrant, the border crossing sector, and the year in which the migration takes place. C represents the costs to migrating, which will also be a function of the sector, year, and home village. Also in C is D_{lj}, the log of the distance between the origin village and the crossing sector, and Enf_{jt}, a measure of border enforcement (we specify this as the log of annual line watch hours per mile) in the j^{th} sector during the year of the migration and t represents the year of the migration. The variable UR_{jt} is a measure of unemployment in destinations associated with each border crossing. The choice variables of the migrant are then j; whether or not to migrate, and conditional on migrating, the choice of the sector where he crosses. We assume that the migrant will choose the utility maximizing j.

3.2 Functional Form Specifications

In order to estimate this model in any parametric way, we must first specify a functional form of the utility function. The key parameter of interest in our utility function is the marginal disutility to the potential migrant of an additional unit of border enforcement. Additionally, we estimate the marginal disutility to the potential migrant of distance traveled to the border. Unobservable characteristics of the origin, crossing sector, and year of migration may also affect the utility of the trip. These effects are captured in an error term,

ξ_{jtl}. Additionally, we include ε_{ijlt}, an idiosyncratic error term representing the unobserved utility to of each choice to the individual model:

$$U_{ij} = \alpha Enf_{jt} + \beta_1 \ln(D_{jl}) + \beta_2 UR_{jt} + \xi_{jtl} + \varepsilon_{ijlt} \text{ for } j = 1, 2, ..., 9 \quad (3)$$

$$U_{ij} = 0 \text{ for } j = 0 \quad (4)$$

3.3 An Aggregate Discrete Choice Model Approach

3.3.1 Conditional Logit Specification

While we ultimately use an aggregate discrete model to estimate our model, first consider how our parameters would be identified using a discrete choice model with individual data. Let Y_{ijtl} represent a dummy variable, which is equal to 1 if individual i from village l chooses to migrate through sector j in year t, and zero otherwise. McFadden (1973) shows that if the ε terms are distributed extreme value type 1, then the probability of the representative individual choosing the k^{th} choice can be represented as

$$\Pr(Y_{iktl} = 1) = \frac{\exp(\delta_{ktl})}{1 + \sum_{j=1}^{9} \exp(\delta_{jtl})} \text{ for } k = 1, 2, .., 9 \quad (5)$$

$$\Pr(Y_{i0tl} = 1) = \frac{1}{1 + \sum_{j=1}^{9} \exp(\delta_{jtl})} \text{ for } k = 0 \quad (6)$$

where $\delta_{jtl} = \alpha Enf_{jt} + \beta \ln(D_{jl}) + \beta_2 UR_{jt} + \xi_{jtl}$.

This conditional logit model can then be estimated by maximum likelihood, where the log likelihood function is

$$\ln L(\alpha, \beta | Y, Enf, \ln D) = \sum_{i=1}^{N} \sum_{j=0}^{J} Y_{ij} * \Pr(Y_{ijtl} = 1) \quad (7)$$

A major issue concerning the estimation of this conditional logit is potential endogeneity. If, as is the case here, we believe that enforcement is correlated with the potential migrants unobserved taste for a specific location, then we must instrument for enforcement. A benefit of the aggregate discrete choice model we use is that it is invertible to a linear model; therefore the Two-Stage Least Squares (2SLS) estimation can be used when using instrumental variables to control for endogeneity.

Using individual level choice data, estimation of the conditional logit model would also prove to be computationally burdensome: we observe data on 93 villages over 22 years with the pool of potential migrants in each year averaging in the hundreds. Estimation of this model involves 10 computations per iteration for each observation in the model (the computation of the utility of each element in the choice set for all individuals). Supposing that the average potential migrant pool is 500, 1 million individual decisions would be observed, requiring 10 million computations for each iteration of the likelihood function.

The aggregate discrete choice model is based on Berry (1994). Over the last ten years, it has become common in the empirical industrial organization literature. The motivation for this model was the need to identify the marginal utilities of product characteristics when the market share of each product is observed but the decisions of individual households are not. In our model, destinations are treated like products, and villages as markets. Berry proposes that we interpret shares as expected values of probabilities: the share of individuals from that village who choose to cross through sector 3 in 1994 should, on average, equal the probability that a representative migrant from a given village chooses to cross through sector 3 in 1994 should on average represent. We can then express the probability of observing a migrant from village l in year t choosing to cross in sector k as:

$$\Pr(Y_{ikl} = 1) = s_{ktl} \text{ for } j = 1, 2, ..., 9 \quad (8)$$
$$\Pr(Y_{i0tl} = 1) = s_{0tl} \quad (9)$$

Normalizing the share choosing each border crossing through each sector by the share who stay at home,

$$\frac{s_{ktl}}{s_{otl}} = \exp(\delta_{ktl}) \quad (10)$$

Taking the log of the odds ratio yields a linear econometric equation, a linear equation with stochastic term ξ_{ktl}; this model may now be estimated by OLS.

$$\ln(\frac{s_{ktl}}{s_{otl}}) = \delta_{ktl} = \alpha Enf_{kt} + \beta \ln(D_{kl}) + \xi_{ktl} \quad (11)$$

3.4 Nested Logit Specification

One source of potential specification error is the assumption that the ε terms are distributed *i.i.d.* If there exist correlations in the unobserved utility across choices, the above inversion will be incorrect. Without introducing a correlation between the utility received from choosing among a number of close substitutes, we also risk calculating implausible substitution patterns. To solve this problem, Berry (1994) also inverts a nested logit based on the model in Cardell (1997). Writing to random utility term as the sum of an individual/group specific term, ζ_{ig}, and the ε term, the utility function becomes:

$$U_{ij} = \gamma + \alpha Enf_{jt} + \beta \ln(D_{jl}) + \xi_{jtl} + \zeta_{ig} + (1-\sigma)\varepsilon_{ijlt} \text{ for } j = 1, 2, ...(9) \quad (12)$$
$$U_{ij} = 0 \text{ for } j = 0 \quad (13)$$

The probability of a representative individual choosing a given border crossing sector k can now be expressed as the product of the probability of migrating and the probability of choosing sector k conditional upon migrating:

$$\Pr(Y_{iktl} = 1) = \Pr(Y_{iktl} = 1|Y_{i0tl} \neq 1) * \Pr(|Y_{i0tl} \neq 1) \quad (14)$$

$$\Pr(Y_{iktl} = 1) = \frac{\exp(\delta_{ktl})}{\sum_{j=1}^{9} \exp(\delta_{jtl})} * \frac{[\sum_{j=1}^{9} \exp(\delta_{jtl})]^{1-\sigma}}{[\sum_{j=1}^{9} \exp(\delta_{jtl})]^{1-\sigma} + 1} \text{ for } k = 1, 2, ..., 9 \quad (15)$$

Note that when is equal to zero, the denominator of the first term cancels with the numerator of the second term, and the brackets in the denominator can be dropped, leaving the probability of making the choice identical to the probability in the conditional logit model.

The probability of choosing option 0, i.e. not migrating and staying in origin, is

$$\Pr(Y_{i0tl} = 1) = \Pr(Y_{i0tl} = 1) = \frac{1}{[\sum_{j=1}^{9} \exp(\delta_{jtl})]^{1-\sigma} + 1} \quad (16)$$

The observed shares are now

$$\Pr(Y_{ikl} = 1) = s_{ktl|migrate*} s_{migrate,tl} \text{ for } j = 1, 2, ..., 9 \quad (17)$$

$$\Pr(Y_{i0tl} = 1) = s_{0tl} \quad (18)$$

Again taking the log odds ratio, we find

$$\frac{\ln(s_j)}{\ln(s_0)} = \alpha Enf_j + \sigma * (\ln_{j|migrate}) + \xi_j \quad (19)$$

$$\ln(\frac{s_{ktl}}{s_{otl}}) = \alpha Enf_{kt} + \beta \ln(D_{kl}) + \sigma \ln(s_{k|migrate}) + \xi_{ktl} \quad (20)$$

where s_{k_g} measures the share of the choice within an arbitrary group. Here, we group our choices into two groups: staying in Mexico, and migrating to the US. Clearly there will be a correlation between the utility of migrating to the US through different border crossings; whether or not the migrant arrives in Los Angeles by crossing through San Diego or Yuma, they will receive higher labor market income than they would had they stayed at home in Mexico.

One issue with our current results is the treatment of s_{k_g} as exogenous. Clearly, unobserved characteristics of the choice which affect the share of individuals choosing it should also affect the share of individuals within its group who choose the choice. We believe that this will create an upward bias in σ, which will affect our estimates of the marginal effects. Future work will focus on developing valid instruments for the within-group share.

3.5 Instrumental Variables

3.5.1 Instrumenting for Enforcement

Because we expect there to be a simultaneous relationship between border enforcement and migration rates, we seek instruments to identify a 2SLS estimation of the main equation. Relevant and valid instruments will be correlated with the level of enforcement in a given Border Patrol sector, but uncorrelated with the unobservables affecting the share of migrants from a village who cross through that sector. For a set of potential instruments, we turn to the political process determining the level of funding in each sector. Hanson and Spilimbergo (2001) provide evidence of the importance of local politicians on the level of enforcement in the sector. For example, they cite a case of farmers in Texas complaining to their Congressional representatives about a crack down on migrant labor in the 1950's.

Variables such as the party of the Congressional representatives of border sector constituents may be correlated with the level of funding for the sector. While relevant, these instruments will likely be invalid: the election of these representatives is likely a function of unobservable characteristics of the district. A constituency which votes for a representative who is "tough on illegal migration" is likely composed of individuals who have a distaste for migration; which may directly affect the utility of migration. For example, a district populated by individuals who elect a representative who promises to increase funding to the Border Patrol may be more likely to be populated by individuals who will discriminate against migrants in the housing market than a district that does not elect such a representative.

As argued above, variables measuring a politician's desire to appropriate funds to the local Border Patrol sector are likely to be invalid. But suppose that all elected officials have some demand for increased funding in their home sector; if not to crack down on illegal immigration, simply to increase the amount of federal funds and jobs provided to constituents in their district. Then variables which measure a politicians *ability* to increase funding to the Border Patrol in their sector may be valid. We create two such "political supply" instruments to identify our model.

Our first instrument for border enforcement is the size of the Congressional delegation representing the Border Patrol sector. Specifically, we count how many Congressional districts share the sector's border with Mexico. For example, suppose that one Congressional district in Arizona includes Yuma and Pima county, while a second encompasses Santa Cruz and Cochise counties. As shown in Table I, the Yuma county Border Patrol sector includes only Yuma county, while the Tucson sector includes Pima, Santa Cruz and Cochise counties. Thus the Tucson sector has two representatives who have an interest in securing funding for the sector, while the Yuma sector has only one. As the size of the sector's lobby in Congress grows, so should its budget and accordingly the level of enforcement. We refer to this variable as NUMREPS

Our second instrument for enforcement is the strength of the state Congressional delegation representing a Border Patrol sector. Here, we count how many members from the state's delegation are on the House Appropriations committee. We then match these state delegations to Border Patrol sectors. Note that the El Paso sector includes counties in both Texas and New Mexico, while the Marfa sector includes only counties in Texas. This variable is named STATEAPP. Thus, the delegation which has an interest to procure funds for the El Paso sector is larger than the delegation representing Marfa. Finally, we use the variable SENDIST as a simple measure of the seniority of the delegation representing the sector. This variable is defined as the minimum of within party seniority ranking among all members of the house representing the district. The most senior person in each party has a seniority ranking of 1, thus we would expect enforcement to be negatively correlated with SENDIST. These variables were computed using data from the Congressional Quarterly Almanac.

3.5.2 Instrumenting for Within Group Shares

One complication of using the inverted nested logit model is that the $\ln(s_{k|migrate})$ term is likely to be endogenous. If we believe that unobservables which affect utility of migrating through a given border sector also affect the utility of migrating through that sector, conditional on migrating, the variable will be endogenous. Suggested instruments for this variable from Berry (1994) are: 1) the number of choices in the group, *and* 2) characteristics of other choices in the group. The motivation for the first variable in the IO literature is that price margins will be a function of the level of competition in the market. Here, we can imagine that subdividing a choice set into smaller and smaller choices should decrease the probability of an individual choosing any one of the given choices. However, each migrant in our dataset faces a choice of 9 Border Patrol sectors through which to cross, thus we have no variation in this potential instrument.

The second set of instruments capture the effect of "competing product" characteristics on market share. These variables will be relevant as by changing the utility of alternative choices, they will directly affect the share of individuals choosing a given choice. Recall that for an instrument to be valid it must be uncorrelated with the unobservable characteristics of the product. If we believe that these characteristics are jointly determined, these potential instruments would likely be invalid. For example, as the level of border enforcement in all sectors is greatly influenced by policy at a national level, enforcement in El Paso is likely to be an invalid instrument for enforcement in San Diego. However, it is reasonable to believe that distance is a predetermined variable, as the location of origin communities are fixed and the border patrol sectors do not change their geographic boundaries throughout the period of our study.

We try a number of instrumental variables for the "in group share", $\ln(s_{k|migrate})$, however strong instruments remain illusive. This is a caveat when interpreting our estimation of the σ parameter. As unobservables which increase the conditional (upon migrating) share of migrating through a given crossing are also likely to positively affect the unconditional share of individuals choosing the crossings, we would suspect that our estimates of σwould be biased upwards.

3.6 Marginal Effects

Now that we have obtained consistent estimates of the effect of enforcement on the marginal utility of a potential migrant, we can address our main research question: does an increase in the level of enforcement along one sector of the border divert migrants into other crossing sectors, or does it lead them to stay at home rather than migrating? For notational convenience, we drop origin and time subscripts in this section.

3.7 Conditional Logit

In the conditional logit model, there are two types of marginal effects: own effects and cross effects. An own effect is the change in the probability (or share) of choice k when a characteristic of k is changed. The marginal effect of an increase in enforcement on the share of potential migrants choosing a crossing sector is represented as

$$\frac{\partial s_k}{\partial Enf_k} = (1 - s_k) * s_k * \alpha \tag{21}$$

This marginal effect can be interpreted as the total effect of an increase in enforcement on the number of migrants crossing through the sector. As discussed earlier, we can decompose this effect into two parts: the share of migrants who are deterred from migrating, and the share of migrants who are diverted into another sector, but who continue to migrate.

Cross effects represent the change in the share choosing choice j when a characteristic of k changes. The cross effects allows us to estimate the effect of an increase in enforcement in one border sector on the share of migrants substituting to all other choice, including the outside option of staying home:

$$\frac{\partial s_j}{\partial Enf_k} = -s_j * s_k * \alpha \tag{22}$$

3.7.1 Nested Logit

Berry derives the marginal own effects in his paper. The effect of an increase in enforcement in crossing k on the number of migrants crossing at k is now

$$\frac{\partial s_k}{\partial Enf_k} = \frac{1}{1 - \sigma} s_k [1 - s_{k|g} - (1 - \sigma) s_k] \alpha \tag{23}$$

note that this term will equal the own marginal effects from the conditional logit when the σ term is 0.

As we explain below, the only cross effect of interest to us is the effect of enforcement in sector k on the share of individuals who choose to remain in Mexico. In 1, we show that this marginal effect is identical to the same figure in the conditional logit. This result does not hold for other cross effects, only for the share substituting to the "outside" option, when the characteristic of another choice changes.

$$\frac{\partial s_0}{\partial Enf_k} = -s_0 * s_k * \alpha \tag{24}$$

3.7.2 Decomposing the Marginal Effects: Deterrence and Diversion

Because the migrants' choice set represents mutually exclusive options, the number of migrants substituting away from an option will equal the sum across choice of migrants substituting to an alternative choice:

$$-\frac{\partial s_k}{\partial Enf_k} = \sum_{j\neq k} \frac{\partial s_j}{\partial Enf_k} \tag{25}$$

$$-\frac{\partial s_k}{\partial Enf_k} = \frac{\partial s_0}{\partial Enf_k} + \sum_{j\neq k,0} \frac{\partial s_j}{\partial Enf_k} \tag{26}$$

$$\text{Total Effect} = \text{Deterrence+Diversion} \tag{27}$$

The deterrence effect, as a share of the total effect, can then be expressed as

$$PCT_{Deter} = \frac{\frac{\partial s_0}{\partial Enf_k}}{\frac{\partial s_0}{\partial Enf_k} + \sum_{j\neq k,0} \frac{\partial s_j}{\partial Enf_k}} = \frac{\frac{\partial s_0}{\partial Enf_k}}{-\frac{\partial s_k}{\partial Enf_k}} \tag{28}$$

Substituting in 24 and 23,

$$PCT_{Deter} = \frac{-s_0 * s_k * \alpha}{-\frac{1}{1-\sigma} s_k[1 - \sigma s_{k|migrate} - (1-\sigma)s_k]\alpha} \tag{29}$$

$$= \frac{(1-\sigma) * s_0}{[1 - \sigma s_{k|migrate} - (1-\sigma)s_k]} \tag{30}$$

Note that when $\sigma = 1$, this term is zero: if "within-group substitution" is perfect, increased enforcement in one border sector will only have the effect of pushing migrants to other crossings and no migrants will stay in Mexico as a result of the policy change. On the other hand, when $\sigma = 0$ we observe that

$$PCT_{Deter} = \frac{s_0}{1 - s_k}$$

which is consistent with the conditional logit model.[3] The percent of the total

effect which can be attributed to deterrence is decreasing in σ. This relationship between percent deterrence and the σ term is illustrated in Figure 1, where for an s_0 value of .9 and an s_j value of .05, we trace the graph of sigma on percent deterrence.

This relationship is similar to the "market-stealing" vs. "market-expansion" analysis found in Berry and Waldfogel (2001). This analysis is of the effect of changes in radio broadcasting variety. An important question in that market is whether this increased product choice resulted in creating more listeners, or simply *diverting* (stealing) listeners from a pre-existing station to a new station.

4 Data

4.1 Distance

The data on the choices of migrants comes from the Mexican Migration Project, a survey constructed jointly by researchers in Mexico and the United States. Separate waves of the survey have taken place annually, starting in 1982. Due to confidentiality concerns, the MMP does not disclose the home village of survey respondents, but the state of each village is given. Distance from the home village to the border crossing is proxied by distance from the capital of the home state to the border. The MMP gives data on the Municipio of the migrants crossing. As our choice set for migrants is one of nine border crossing sectors, rather than a continuum of crossings, we compute the distance from the state capital to the primary crossing in each sector.

4.2 Migrations

In order to construct our dependent variable, we must determine how many individuals migrated from each village in each year, and where they crossed. One of the data sets created by the MMP includes a comprehensive migration history. We are then able to count the number of migrations taking place from the village through each of the Border Patrol sectors in years prior to the survey. Our sub-sample of migrants includes all males between the ages of 18 and 60.

4.3 Potential Migrants

To construct the log-share ratios, we also must know what the pool of potential migrants was for a village in a given year. An MMP data set gives

[3] As the total effect of an increase in enfrocment in k on the share of migrants substituting away from k is $-\frac{\partial s_k}{\partial Enf_k} = -(1-s_k)*s_k*\alpha$ and the share who are detered into staying in Mexico is $\frac{\partial s_0}{\partial Enf_k} = -s_0*s_k*\alpha$; the ratio of these two terms gives us $\frac{s_0}{1-s_k}$

individual level data for each household in the sample. Individual members of the household are included in the data set whether or not they were present at the time of the interview, or absent during a migration spell. We are able to use this data set to construct a count of the number of individuals in each village during the year of the survey. A separate MMP data set provides information on the population of the larger community from which the sample was drawn. This statistic is available only for Mexican census years, which are not necessarily the same years as MMP survey years.

To estimate the size of the pool of potential migrants in each year, we first interpolate the community population between census years. We then find the ratio of the pool of potential migrants (males between 18 and 60) to the community in the survey year. By assuming that this ratio remains constant over time[4], we can estimate the pool of potential migrants in all years.

After computing both the number of migrants for each village/year/crossing sector combination, we are able to compute the share of each choice for each year and village. In order to allow us to calculate the log odds ratio, we bound the share of migrants at an arbitrarily small number greater than zero for all observations.

4.4 Enforcement

Our border enforcement data was originally used in Hanson and Spilimebergo (1999). It provides rich information on both the number of apprehensions and line watch hours (the number of labor hours spent patrolling the border) in the nine Border Patrol sectors. Monthly data by sector is provided from 1977 to 2000. Table 2 provides enforcement hours per mile of border for 1980, 1990 and 2000. Note the variation both between and within sectors.

4.5 Unemployment

Our unemployment variable is constructed from BLS data from both the city and state level. For each of the 9 border crossing sectors, we tabulate the most popular inside the United States for migrants crossing through the sector over a number of years. We then create a weighted average of the unemployment rate for each of the three most popular destinations, where the weights are the proportion of migrants choosing the location over the entire period of our study. When city unemployment rates are unavailable, we use state unemployment rates as a proxy.

[4] Essentially, we assume that the sample of males age 16 to 60 grows at the same rate as the population over a period of 15 to 20 years. This assumption could be relaxed accounting for differential growth rates for our key demographic and the Mexican population as a whole during this period.

5 Results

To identify the marginal utility parameters of the model, we estimate various specifications of the following model. Summary statistics are reported in 3. To identify the parameters of our model, we first run an OLS regression of the log-share ratio from the conditional logit inversion on enforcement, distance and unemployment. Next, we add in fixed effects for the crossing sector, year, and origin. This is followed by an IV estimation of the conditional logit inversion with the fixed effects. Finally, we repeat these three estimations for the inversion of the nested logit model.

5.1 Parameter Estimates

Both distance and unemployment are negative and significant in all of our models. The positive and significant slope coefficient on enforcement in our simple model clearly points to the endogeneity of enforcement. After introducing fixed effects to the model, enforcement becomes insignificant. Finally, after instrumenting for enforcement, we obtain negative and significant results, as we would predict. We report the estimates from the first stage of our two instrumental variable estimations in the 4th and 8th columns of the results reported in Table IV. When regressing the "political supply" variables on the log of enforcement, each of our instruments is significant and of the expected sign. We can reject the null hypothesis of an F-test on all excluded instruments being insignificant at greater that the .001% level. We also report the p-value from the Hanson overidentification test; we are not able to reject the null that our instruments were properly omitted from the second stage.

The results from the first stage reveal that the addition of an extra member of congress representing an area which includes part of the Mexican border in a given sector will increase enforcement by 24%. If one member of the state congressional delegation were to join the House Appropriations Committee, we would expect enforcement to rise by 6%. When the rank of the most senior member of the sectors delegation increases by 100 places, enforcement increases by 3%. Hanson and Spilembergo (2001) show how political *demand* can change levels of enforcement; when output prices of goods requiring undocumented labor for production increase, enforcement decreases. Here we have found that political *supply* has significant impacts on enforcement: border sectors represented by a more powerful congressional delegation tend to have higher levels of enforcement.

Our estimates of σ in the regressions based on the inversion of the nested logit find σ to be significantly greater than 0, allowing us to reject the hypothesis that there is no correlation between the error terms in the group of choices which represent migrating to the US. Conditioning on fixed effects for origin, year, and crossing significantly increases the estimate of σ. After instrumenting for enforcement we find no additional change in σ which remains around .5.

5.2 Decomposed Marginal Effects

Table V presents estimates of average marginal effects, both across all observations and broken down by decade. Due to the non-linear nature of the model and these marginal effects, the average of marginal effects across all observations will not necessarily equal the marginal effect at the sample average values of X. We report only the average values of the marginal effects. While the substitution patterns do not change significantly over time, we find in all years that around 50% of the effect of additional enforcement can be accounted for by deference. This should serve as a caveat to policy makers who observe that crossings decrease when a particular sector of the border is tightened. The perceived effect on migration crossings in that sector may indeed be causal, but it may overstate the effect on crossings into the United States significantly.

6 Future Work

6.1 Using Individual Level Data to Identify σ

Without strong instruments for $\ln(s_{k|migrate})$, it may be necessary to estimate the model by maximum likelihood, using individual level data. Recall the likelihood of an individual choosing a given crossing in the nested model from equation14:

$$\Pr(Y_{iktl} = 1) = \frac{\exp(\delta_{ktl})}{\sum_{j=1}^{9} \exp(\delta_{jtl})} * \frac{[\sum_{j=1}^{9} \exp(\delta_{jtl})]^{1-\sigma}}{[\sum_{j=1}^{9} \exp(\delta_{jtl})]^{1-\sigma} + 1} fork = 1, 2, .., 9$$

Note that σ is no longer a coefficient on a variable, thus σ is not a coefficient on an endogenous variable requiring instruments. However, it is still necessary to instrument for the level of border enforcement, which is contained in the δ_{ktl} term. It will no longer be possible to use standard linear IV estimation. Future work will center on using a LIML model to estimate the nested logit model with individual level data, thus obtaining consistent estimates of σ.

6.2 Using Marginal Effects to Evaluate Effectiveness of Border Enforcement Policy

Having consistently identified parameters of a structural model, we are now in a position to conduct a wide array of policy experiments. Future work will evaluate counterfactual policies to answer important questions. For instance, what is the effect of increased enforcement on the welfare of a potential migrant? During the time period that our data spans, was the US optimally allocating hours of labor between sectors, or could a different allocation of resources between sectors reduced migration? What would be the effect of a 10% increase in enforcement intensity on the incidence of migration?

The Border Patrol uses more than just labor hours to patrol the border; inputs such as helicopters and fences also play a large part in border enforcement. Therefore hours of enforcement measured in our enforcement variable are not homogenous. Essentially, we are measuring the level of labor inputs in a production process to deter migration. The use of complementary capital inputs and different technologies will also have an impact on these deterrence rates. We find significant variation in the deterrence effect across time and space. In future work, we will be able to use this variation in deterrence to identify the effectiveness of other parts of the production function.

7 Conclusion

This paper studies the impact of border enforcement on illegal migration along the US-Mexico border. Using an aggregate discrete choice model, we are able to identify the parameters of a random utility model, where the agent chooses between crossing into the United States through one of 9 Border Patrol sectors and staying home. We tackle two important econometric questions in this analysis.

First, we must account for the endogenous nature of enforcement. This is done by developing instruments which will affect the level of enforcement in a sector, but will not otherwise be correlated with the utility that the potential migrant receives from crossing there. Specifically, we use "political supply" instruments which capture the ability of the congressional delegation representing the part of the border which overlaps with the Border Patrol sector. We find that the presence of a stronger delegation significantly increases the level of enforcement.

Second, we must account for the fact that the error term in the random utility model will likely not be independent across a given individual's choices; crossing the border means access to labor markets in the United States, regardless of which border crossing is chosen. To solve this problem we apply an aggregate discrete choice nested logit model. The σ in this model allows us to determine the nature of substitution patterns for migrants facing an increased level of enforcement at their preferred crossing: will they cross somewhere else, or will they stay home? While we suspect that σ may be biased upwards, overestimating the share of migrants who substitute to other crossings, our preliminary findings indicate that nearly half of migrants deterred away from one crossing by tighter enforcement will continue to cross into the United States, though through a different crossing. Future work will involve estimating an individual level choice model which should allow for a better estimate of the crucial σ parameter.

References

Angulucci, Manuela. “U.S. Border Enforcement and the Net Inflow of Mexican Illegal Migration.” UCL Working Paper, March 2004.

Berry, Steven T. "Estimating Discrete-Choice Models of Product Differentiation". The RAND Journal of Economics, 25(2), Summer 1994: 242-262.

Berry, Steven T.; Waldfogel, Joel. "Do Mergers Increase Product Variety? Evidence from Radio Broadcasting." The Quarterly Journal of Economics, November 2001: 1009-1025.

Bean, Frank D; Chanove, Roland; Cushing, Robert G; de la Garza, Rodolfo; Freeman, Gary P.; Haynes, Charles W.; Spencer, David. “Illegal Mexican Migration and the United States/Mexico Border: The Effects of Operation Hold the Line on El Paso/Juarez”. U.S. Commission on Immigration Reform research paper (July 1994).

Bartel, Ann P. “Where Do the New U.S. Immigrants Live?” Journal of Labor Economics: 7(4), Oct. 1989: pp.371-391.

Bauer T, GS Epstein and IN Gang (2005) Enclaves, language, and the location choice of migrants. Journal of Population Economics 18(4): 649-662

Blundell, Richard W.; Smith, Richard J. “Estimation in a Class of Simultaneous Equation Limited Dependent Variable Models”. The Review of Economics Studies: 56(1), Jan 1989: 37-57.

Cardell, Scott, “Variance Components Structures for Extreme-Value and Logistic Distributions with Application to Models of Heterogeneity”, Econometric Theory, 13(2), April 1997, pg. 185-213.

Carrión-Flores, Carmen E. “What Makes You Go Back Home? Determinants of the Duration of Migration of Mexican Immigrants in the U.S.” Mimeo, 2004.

Chiswick, Barry R; Miller, Paul W. “Where Immigrants Settle in the United States.” Journal of Comparative Policy Analysis: 6(2), 2004: 185-197.

Congressional Quarterly Almanac.: 1975-1997.

Congressional District Almanac: 1977, 1983, 1993.

Cornelius WA (2001) Death at the border: Efficacy and unintended consequences of US immigration control policy. Population and Development Review. 27(4):661-685.

Durand J, DS Massey and F Charvet (2000) The changing geography of Mexican immigration to the United States: 1910-1996. Social Science Quarterly 81(1):1-15

Garcia C (2005) Buscando Trabajo: Social networking among immigrants from Mexico to the United States. Hispanic Journal of Behavioral Sciences 27(1):3-22.

Hanson, Gordon. "Illegal Migration from Mexico to the United States." Working paper, Nov 2005.

Hanson, Gordon H.; Spilimbergo, Antonio. "Illegal Immigration, Border Enforcement, and Relative Wages: Evidence from Apprehensions at the U.S.-Mexico Border. The American Economic Review: 89(5), December 1999: 1337-1357.

______ "Political Economy, Sectoral Shocks, and Border Enforcement." The Canadian Journal of Economics: 34(3), August 2001: 612-638.

Hanson, Gordon H.; Robertson, Raymond; Spilimbergo, Antonio. "Does Border Enforcement Protect U.S. Workers from Illegal Immigration." Review of Economics & Statistics: 84(1), February 2002:

Jaeger, David A. "Local Labor Markets, Admission Categories, and Immigrant Location Choice." IZA working paper.

Kanbur R and H Rapoport (2005). Migration Selectivity and the evolution of spatial inequality. Journal of Economic Geography 5(1):43-57

McFadden, D. "Conditional Logit Analyses of Qualitative Choice Behavior". Frontiers of Econometrics, 1973.

McFadden, Daniel, "Econometric Analysis of Qualitative Choice Models." Handbook of Econometrics, VII: 1982.

Munshi, Kaivan. “Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market.” The Quarterly Journal of Economics: May 2003: 549-599.

Orrenius PM and M Zayodny (2003) Do amnesty programs reduce undocumented immigration? Evidence from IRCA. Demography 40(3):347-450

Sjastad, Larry A. “The Costs and Returns of Human Migration.” The Journal of Political Economy, 70(5), Part 2: Investment in Human Beings, October, 1962: 80-93.

Figure 1: Within Group Substitution and Deterrence Share

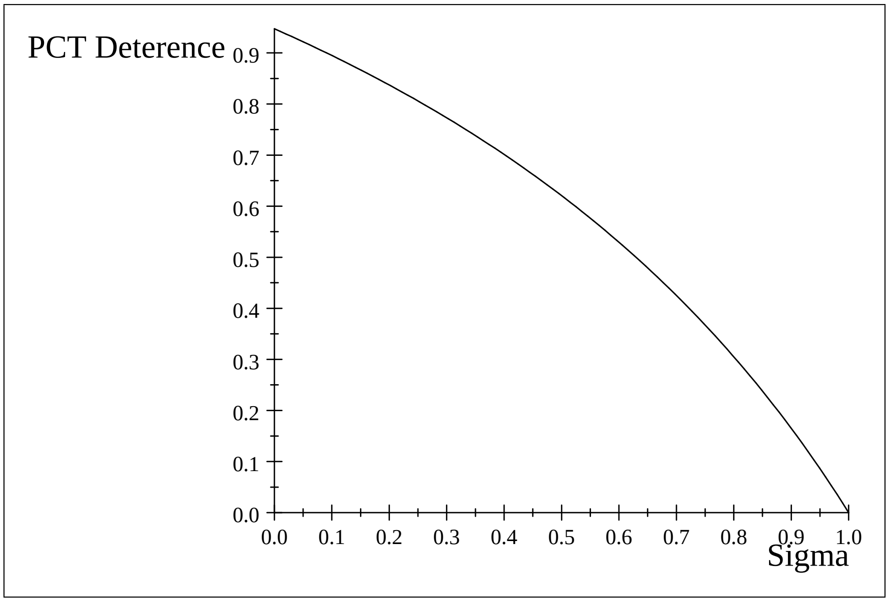

Table I: Border Patrol Zones

Sector	**Counties**	**Principal Crossing**	**Miles of Border**
San Diego	San Diego (CA)	Tijuana-San Diego	66
El Centro	Imperial (CA)	Mexicali-Calixico	75
Yuma	Yuma (AZ)	SL Rio Colorado-Yuma	118
Tucson	Pima, Santa Cruz, Cochise (AZ)	Nogales-Nogales	261
El Paso	Hidalgo, Luna, Dona Ana (NM), El Paso, Hudspeth (TX)	Ciudad Juarez-El Paso	289
Marfa	Presidio, Brewster (TX)	Ojinaga-Presidio	420
Del Rio	Terrell, Val Verde (TX)	Del Rio-Piedras Negras	205
Laredo	Kinney, Maverick, Webb, Zapata (TX)	Nuevo Laredo-Laredo	171
McAllen	Starr, Hidalgo, Cameron (TX)	McAllen-Reynosa	284

Table II: Border Enforcement Hours Per Mile

Sector	1980s	1990s	2000
San Diego	7236	7769	30764
El Centro	1422	2849	9210
Yuma	1955	1880	2680
Tucson	550	994	7950
El Paso	928	1456	2849
Marfa	195	137	248
Del Rio	1134	1372	3217
Laredo	573	1566	4848
McAllen	574	954	5183

Table III: Summary Stats

Variable	Mean	Std. Dev.
Share	0.002	0.01
Share Staying Home	0.978	0.031
Enforcement (Annual Hours/Mile)	2327.851	3794.108
Mexican Distance	1040.9	454.041
Destination Unemployment	7.016	1.664
Sector Reps	1.513	0.594
State Reps on App. Comm.	2.99	1.702
Senority of Most Senior Rep	53.135	44.188
N	17631	

Table IV: Results of Estimation on Log-Share Ratio

	Conditional Logit				Nested Logit			
Variable	OLS	OLS	IV		OLS	OLS	IV FE	
Enforcement	1.09***	-.08	-.93***		.70***	-.04	-.46**	
UR	-.04*	-.14***	-.21***	.-.09***	-.03*	-.07***	-.11***	-.09***
Distance	-.19***	-.45***	-.44***	.003	-.06*	-.22**	-.22***	.004
Share Within					.35***	.50***	.52***	.002
STATEAPP				.06***				.06***
NUMREPS				.24***				.24***
SENIORITY				-.0003***				-.0003*
Crossing FEs		X	X	X		X	X	X
Year FEs		X	X	X		X	X	X
Village FEs		X	X	X		X	X	X
R2	.083	.260			.2877	.56		
F Test on Z				1174***				1174***
Partial R2				.168				.168
Hanson P-val				.23				.54
N	17631	17631	17631	17631	17631	17631	17631	17631

Table V: Average Marginal Effects of Enforcement on Deterrence and Diversion (N-LOGIT)

	Total Effect	Deterrence	Diversion	%Deterrence
All	-.0014	-.0010	-.0004	53%
1970s	-.0015	-.0011	-.0004	53%
1980s	-.0015	-.0011	-.0004	53%
1990s	-.0012	-.0009	-.0003	53%

Appendix 1: Share Substituting to Outside Good in Nested Logit

We are interested in the effect of an increase in X_k on the share of consumers in the market who choose the outside option. My outside option is staying in Mexico, while my 9 inside goods represent migrating to the US through 1 of the 9 border crossing sectors. Then:

$$\begin{aligned} D_{US} &= \sum_{j=1}^{9} \exp(\frac{\delta_j}{1-\sigma}), \delta_j = X_j\beta + \xi_j \\ D_{MEX} &= \exp(\frac{\delta_0}{1-\sigma}), \delta_j = 0 \end{aligned} \tag{31}$$

Then the share choosing the outside option is

$$s_o = s_{0|g_o} * s_{g_0} = 1 * \frac{1}{1 + D_{US}^{1-\sigma}} = [1 + D_{US}^{1-\sigma}]^{-1}$$

Taking the partial with respect to X_k

$$\begin{aligned} \frac{\partial s_o}{\partial X_k} &= -[1 + D_{US}^{1-\sigma}]^{-2} * (1-\sigma) * D_{US}^{-\sigma} * \exp(\frac{\delta_k}{1-\sigma}) * \frac{\beta}{1-\sigma} \\ &= -\beta * \frac{1}{1 + D_{US}^{1-\sigma}} * \frac{\exp(\frac{\delta_k}{1-\sigma})}{D_{US}^{\sigma}(1 + D_{US}^{1-\sigma})} = -\beta * s_0 * s_k \end{aligned}$$