Embarrassing variety of choice: Modeling Mexican return migration decisions[1]

Claudia Masferrer, McGill University[2]

Abstract

This work investigates the utility of applying various models for return migration decisions. I study alternative models that fall in a broad class of polytomous choice models widely used when outcomes consist of categories of choices. The three conceptually plausible models considered are the classical multinomial logit, the nested logit and the sequential logit models. The statistical concepts underlying these models are described and investigated, with a special focus on the assumption of independent irrelevant alternatives, and then an application to data of Mexican return U.S. migration is provided. I use the complete set of individual records of the 2005 Mexican Population Count. I find that these models are dependent on how researchers think of these decision processes and that, for this specific type of application, the sequential logit model offers more flexibility in terms of defining the decision structure and in terms of comparison and interpretation.

1. Introduction

This work uses the complete set of individual and household records of the 2005 Mexican Population Count, to investigate the utility of applying various models for return migration decisions. Knowing who is returning to Mexico after a migratory experience in the United States and where they are coming back is important in order to characterize the different profiles of Mexican returnees and try to provide information useful for policy and demographic inquiry, for example, to design social programs for the reincorporation of returnees.

The lack of appropriate data limits the knowledge of both the out and return migration decision processes. Census data does not capture enough information to model, for example, what made migrants decide to return to a border city, a tourist destination or their hometown. Plus, different selection processes are interplaying: both for out and return migration (Borjas & Bratsberg, 1996; Cohen & Haberfeld, 2001; Lam, 1986, 1994). Different people may consider different set of options of places for return. Just like social networks, *push* and *pull* factors, and individual motivations and expectations can affect out migration decisions, they can affect return migration decisions (Cassarino, 2004; de Jong & Gardner, 1981; Fawcett, 1985). For simplicity and computational limitations, we will focus on return at the state level. We identify four types of states in Mexico that seem to have different patterns of return migration: traditional migration sending states, states in the Northern border with the United States, states that have more returnees than out migrants, and the rest.

Nested and sequential models have been used to model migration decision processes (Christiadi & Cushing, 2007; Pellegrini & Fotheringham, 2002). In this work, I examine possible ways of studying the Mexican return migration decision process using models that fall in a broad class of polytomous choice models widely used when outcomes consist of categories of choices: multinomial, nested and sequential logit models. Then, I compare the results of these models and discuss the different implications of these differences. Thus, this paper will try to answer the following research questions: Which of the

[1] ***DRAFT: Please DO NOT CITE without authorization***

[2] Sociology Department, McGill University. claudia.masferrer@mail.mcgill.ca

three discrete-choice models is more appropriate for modelling Mexican return migration decisions? What different results do we obtain if we apply different statistical models and how do the results depend on the choice of the model?

French people have the expression *l'embarras du choix*[3] about the difficulty in choosing. In our case, we are dealing with the difficulty of modeling the choice of where to return in Mexico after having being in the United States. We know the decisions that have been made, and from there, our statistical modeling will try to provide tools to answer the overarching question of "who decided to go to what kind of place?"

The outline of this paper is as follows. Section 2 provides an overview of Mexican return migration. It is followed by the presentation of the data and measures (section 3). In section 4 the methodology on which the paper is based is presented. I include first the motivation and discussion of the statistical methods used and then the empirical models. The results are contained in section 5: following the descriptive statistics there is a subsection with the results of the three statistical models. Section 6 comprises a discussion of these results and the paper concludes in section 7.

2. Mexican Return Migration: an Overview

Although return migration has been a constant feature of Mexico-US migration patterns, its significance and characteristics has changed sharply with time. For most of the twentieth century, most migrants to the US from Mexico were temporary, moving for seasonal work in the United States and returning to villages and small towns where the rest of their family had remained (Durand, Massey, & Zenteno, 2001). Over time, rural origin migration was increasingly complemented by urban origin migration (Roberts, Frank, & Lozano-Asencio, 1999), and just like the origins of Mexican migrants diversified away from traditional sending areas (Escobar, 2008; Quinn, 2006; Tuirán, Fuentes, & Ávila, 2002), such as the Center-West of Mexico, new destinations appeared (Donato, Tolbert, Nucci, & Kawano, 2008; Leach & Bean, 2008; Roberts & Hamilton, 2007), such as the North-East and North-West of the US.

In the latter half of the twentieth century, migration to the US from Mexico has taken on an increasingly permanent character as migrants obtained year-round jobs, often in the cities, and were joined by other family members in face of the increasing difficulties of household subsistence in Mexico's rural and urban economies (Roberts, et al., 1999). Return migration has continued, but at substantially lower levels. For example, return decreased from 290,944 people in 1995 and 267,150 people in 2000[4] to 242,533 returnees in 2005. The regions that attracted more returnees in 2005 were the central region of Mexico, followed by the northern region.

The migration flow that originated in the beginning of the 20th century in rural communities of Western Mexico constitutes now a well established flow such that we can identify states as traditional sending regions. The states that are considered under this category are Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas (Tuirán, et al., 2002). The studies emanating from the Mexican Migration Project have shown that long established migration flows

[3] *L'embarras du choix* is translated in English as "embarrassing variety of choice."

[4] Publications from INEGI report a total number of 267,150 returnees for 2000 and give the totals by state and size of locality. However, the weighted data from the 10 percent sample of the 2000 Population Census provides an estimation of 260,650 returnees because it does not include all the localities.

generate a cumulative causation in which earlier migrants facilitate the migration of subsequent ones, making it easier for women or those with less skills or material resources to migrate (Massey, 1990, 1999) although these mechanisms differ in urban areas (Fussell & Massey, 2004). Long-established flows are also likely to mean that a larger proportion of the community's migrants have documents to enter the US, gained through the provisions of IRCA and family reunification.

These factors are likely to have contradictory consequences for return migration since the presence of well-established migrant communities in the US may encourage migrants to stay, but strong transnational networks and legal documentation may support circular migration. The border region of Mexico and the United States will deserve special attention in this analysis. This geographical region is also a region of economic development in Mexico with job opportunities in the industrial sector, mainly in the *maquiladora* industry. Elizabeth Fussell (2004a, 2004b) finds that Tijuana plays a two-fold role in the migration flow: a destination for internal migrants (and we would say, for returnees) and a home base for migrants that make repeated trips to the United States. In her work, she analyses the role of this region, focusing in Tijuana, complementing the established rural and newer urban flows.

Once the migrant decides to leave the United States (voluntarily or forced, after a deportation for example) he faces the question: "But where in Mexico?" Return does not have to occur to the same locality of origin. The experience he has gained during the migratory experience, the new traditions and values that have been adopted, the perceptions of Mexico (both the country and his community) and of the United States have made this individual very different than the one he was before (Berry, 2001). In addition, family ties, pressures, expectations and perceptions may be also an element in the decisions he has to make (Fawcett, 1985); besides, taking into account the economic, political, and legal conditions of both his stay in the United States and his return to Mexico. Going back to the same community of origin may not be an option for the same reasons that the migrant had for leaving in the first place. The options that the migrant has while being in the United States are: stay in that particular place, move to another place in the United States, go back to the Mexican community where he was living before, go to Mexico to a different community, or go to somewhere else in the world (see Fig. 1).

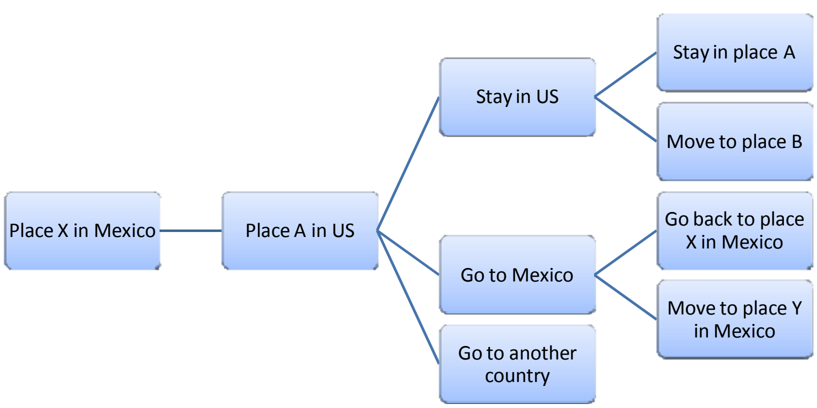

Figure 1. Options of a Mexican migrant in the United States

Recent work has shown evidence for an increase of people that do not go back to the community they originally left from (Masferrer & Roberts, 2009). The localities that are more attractive for return are urban areas with the following characteristics: they are border cities, metropolitan areas, localities in traditional migration states or new tourist areas. Building on the findings of previous work by the author and colleague (Masferrer, 2009; Masferrer & Roberts, 2009), I defined the four types of states already mentioned. Besides the traditional migration sending states and the border region, in this work I classify the rest of the states according to their rate of return migration in order to indirectly estimate the characteristics of those that may not be returning to the place they originally left from. In this sense, a rate of return greater than one tells us that the state is attracting more people than it sent out to the U.S.

Only Quintana Roo, Campeche, Baja California Sur and Baja California have rates greater than 1. However, the states of Yucatán, Tabasco, Chiapas, Sonora, Tamaulipas, Querétaro, Colima, Aguascalientes, Jalisco and Nayarit have a rate of return above the national average (equal to 0.2 for 2005). The states with a rate of return greater than one are not states considered traditional sending migrants states and they had in 2000 very low, low and moderate migration intensities (Tuirán, et al., 2002). The states with a rate of return above the national average are distributed in all the migration intensity degrees. The state with the highest rate of return is Quintana Roo and in 2000 this state had very low migration intensity. This implies that Quintana Roo is an attractive place for returnees. Note that the states that have a rate of return greater than one are also states that are attractive for internal migration. Data from the 2005 Population Count also show that the state with the highest percentage of its population of internal migrants is Quintana Roo (about 11.5%), followed by Baja California Sur (9%) and Baja California (7%).

The analysis of the circulatory and return migration patterns for different states show different patterns for different regions (Masferrer & Roberts, 2009). People are more likely to be circular migrants in traditional states and less likely to be absent migrants while the reverse is true for other regions. In contrast with the traditional states, in all other regions except the North (for 1995-2000) there is a greater preponderance of absent migrants than circulatory migrants. Northern states in Mexico have a higher propensity to have documented migrants, for example people living in Ciudad Juárez can easily go back and forth to the El Paso and people in Tijuana may even commute daily to work in San Diego.

The above led to the following classification of states in four types.

- *Traditional migration sending states*: Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas.

- *Border states*: those that share a physical border with the United States, i.e. Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas.

- *States that attract more returnees than they send (i.e. with rate of return above one)*: The states with a rate of return above one (in 2005) are Quintana Roo (5.6), Campeche (1.7) and Baja California Sur (1.1). Note that the rate of return of Baja California was 1.2 but it is considered as a border state and is included in the previous category.

- *The rest (i.e. states that are not traditional migration states, are not border states and have a rate of return below one)*: Chiapas, Distrito Federal, Guerrero, Hidalgo, Estado de México, Morelos, Oaxaca, Puebla, Querétaro, Sinaloa, Tabasco, Tlaxcala, Veracruz and Yucatán.

Note that the states of the last group show a big heterogeneity in terms of migration intensity index, marginality index and poverty (Anzaldo & Prado, 2007; Consejo Nacional de Evaluación de la Política de Desarrollo Social, 2007). They are also spread in different regions of the country and are likely to be in different stages of a migration tradition or pattern. For specific purposes of this work, and as a result of the previous theoretical and description of the general trend, this classification of states in four different types was defined. However, other classifications will be considered in future work by the author as well as the sensibility of the results by different classifications.

3. Data and Measures

This paper uses the ten percent sample of the 2000 Population Census and the complete set of individual and household records of the 2005 Mexican Population Count (Conteo). The only question regarding migration that is available in the 2005 Population Count is "In which state of Mexico or in which country were you living five years ago?" ("Hace 5 años, en octubre de 2000, ¿en qué estado de la República o en qué país vivía?"). Therefore, we define a returnee as somebody that was living in the United States five years before the census and is living in Mexico at the moment of the census. And consequently, we define return migration for 2005 as the population who 5 years ago or more lived in the United States in October of 2000 but is living in Mexico in October 2005. Return migration for 2000 is defined analogously.

For analytical purposes, we will only include in the analysis non-institutionalized individuals that were in 2000 in the United States, i.e. 238,331 returnees. Thus, this excludes 6,095 returnees that are homeless or that are living in collective dwellings like hotels, hospitals, orphanages, care homes for the elderly, religious institutions, jails, army facilities, refugee camps, etc. In order to apply the nested logit model, a 5% random sample by state of the total data set of returnees had to be constructed because of computational limitations. For the purposes of this work, the sample was generated randomly by state to assure that the distribution of returnees by state was maintained. We used an equal frequency weight equal to 20 for each observation in the sample of 11,917 observations. The weighted sample had in total 238,340 returnees (9 more due to rounding effects in the weighted sample).

The rate of return is defined as those who were in Mexico in 2005, having been in the U.S. in 2000 over those who went to the United States from 1995 to 2000 and did not come back by 2000 (Masferrer & Roberts, 2009). The concept of rate of return, as defined here, relates the total population that was living in the U.S. in 2000 and is living in Mexico in 2005, and the population that left for the U.S. before 1995 and was living in Mexico in 1995. Note that the populations reported as living in the US in 1995 in the 2000 Census and those living in the US in 2000 in the Population Count of 2005 can include people who migrated to the US many years before either 1995 or 2000, but who have only recently returned to Mexico. Thus, places with a large and long-standing stock of migrants in the US have a potentially much bigger base for generating return migrants. The return migration rate cannot be calculated at the municipality or locality level due to restrictions of data availability in the 2000 census.

The lack of appropriate data limits the knowledge of both the out and return migration decision processes. For modeling purposes and due to computational limitations, we will focus on return at the state level. This means that what will be assumed indirectly is that a migrant while being in the United States has the following options: stay in that particular place, move to a different place in the U.S., move to the state in Mexico he left from originally (possibly to a different municipality or even a different locality), move to a different state in Mexico or move somewhere else in the world. Recall that there is no way to know where he left from originally.

The Marginality Index (*Índice de marginación*) produced by CONAPO is generated at the locality level using the technique of Principal Components and summarizes educational characteristics of the population (population that does not know how to write and read and population with incomplete basic education), as well as dwelling characteristics (access to drainage, electricity and water; crowdedness, material of the floor and existence of refrigerator). The Degree of Marginality is the categorical version of the Marginality Index (Anzaldo & Prado, 2007) and it takes the following values: very low (1), low (2), moderate (3), high (4) and very high (5). Before an official measure of poverty existed, that was accepted by researchers and policy makers, the index of marginality was the indicator of social exclusion most often used in Mexico (Cortés, 2002; Hernández & Székely, 2005).

The unordered variable *type* is used to classify individuals that are living in different types of Mexican states in 2005 but where living five years ago in the United States, i.e. in 2000. The variable *type* consists of four categories that we have labeled *traditional* (the reference category), *border*, *above one* and *below one*.[5] These labels are merely for classification in a mutually exclusive way and they do not assume any order at all between categories. Table 1 shows the distribution of returnees by state and type of state.

Table 1
Number of cases (returnees) in each type of state

Type of state							
Traditional		Border		Above one		Below one	
Aguascalientes	4,922	Baja California	19,434	Baja California Sur	1,272	Chiapas	1,636
Colima	3,306	Coahuila	3,762	Campeche	623	Distrito Federal	8,260
Durango	6,288	Chihuahua	12,950	Quintana Roo	2,064	Guerrero	5,276
Guanajuato	14,829	Nuevo León	6,046			Hidalgo	5,818
Jalisco	29,429	Sonora	6,290			Estado de México	12,355
Michoacán	21,351	Tamaulipas	7,220			Morelos	3,883
Nayarit	6,361					Oaxaca	9,633
San Luis Potosí	8,363					Puebla	6,640
Zacatecas	9,737					Querétaro	3,599
						Sinaloa	5,165
						Tabasco	599
						Tlaxcala	938
						Veracruz	8,834
						Yucatán	1,448
Total	104,586		55,702		3,959		74,084
% total returnees	43.9%		23.4%		1.7%		31.1%

Source: 2005 Population Count

Table 1. Number of returnees in each type of state, 2005

[5] Note, however, that the category called *above* or *above one* actually refers to states with rate of return above one that are not traditional sending states and that are not border states. In the same way, the category called *below* or *below one* actually refers to states with rate of return below one that are not traditional sending states and that are not border states.

The independent variables will focus on individual characteristics, family relationship and type of household and migration experience in the household. For gender, the variable *male* is an indicator for males. For age, we use a set of indicator variables of age categories: 5 to 19 years old, 20 to 34 years old (reference category), 35 to 49 years old, 50 to 64 years old, and 65 and older. For level of education we also use a set of indicator variables for the following levels: no formal education, primary school, secondary school, high school, professional education and above (reference category). For the relationship with the head of household we use the following indicator variables: head of household (not unipersonal household), spouse, and son or daughter. The type of household is characterized by three indicator variables: nuclear household, extended or mixed family household, and unipersonal household. Return migration experience in the household will be characterized by two indicators: one to indicate that the individual is living in a household (non unipersonal) where all the members are returnees and another to indicate that the returnee is the only returnee in the household. Finally, at the state level, we will incorporate the dimension of social exclusion in the state (using the degree of marginality as a proxy). Thus, we assume that economic opportunities (indirectly captured by the degree of marginality) of the state affect the decision to return.

It is important to note some limitations and advantages of the data, the whole set of records of the 2005 Population Count. One of its limitations is associated to the problem of measurement and the inherent underestimation produced by counting only households present in Mexico. The data from the Population Count of 2005 does not have the date of departure and arrival, neither the place of origin or last emigration so we cannot determine if they returned to the original place they departed from. In absence of longitudinal data, the information cannot be completely related at the individual or household level to the 2000 Mexican Census. Neither do we have the causes of the emigration or remigration, nor the time of stay to assist in evaluating causality. Another analytical consideration that is worth noting is that since the data does not include place of birth, some individuals considered as returnees may be American expatriates. However, having the data at the individual and household level presents advantages, mainly because we are not dealing with a sample, but with the universe of returnees.

4. Methods

4.a. Motivation and discussion

The most widely used model for a categorical outcome is the logit model, implemented when the independent variable is dichotomous. Unordered qualitative variables appear in different contexts and the models used for these are generally called discrete-choice models. An advantage of discrete choice analysis other than the standard multinomial logit models is the flexibility to include variables of the choices and then acknowledge that the attributes of the choices matter, as well as individual characteristics. The classic examples include the decision of alternative ways of transportation (train, bus or car), union membership, Dutch elections, or labor-force participation (Mare, 1981; McFadden, 1973; van den Berg & Groot, 1992). The alternatives in unordered variables represent different categories that allow for a classification without assuming that there is any nominal value related to them.

When we apply decision making models and define the choice set (in this case, the choice set is the set defined by the types of states we have presented) we will be choosing an arbitrary classification that we hope to be reasonable enough. Ben-Akiva and Lerman (1985, p. 101) write about the choice set generation process saying that "we will assume that each individual's choice set can be specified by the analyst using some reasonable, deterministic rules. [...] This imputation of the choice-set by the analyst is in effect a potentially crude model of a complex interaction between an individual decision maker and his or her environment."

The discussion of whether certain decisions are sequential or simultaneous is not new. For example, van den Berg and Groot (1992) show that the model specification, as well as the estimations and conclusions differ if they assume that the decisions to join a union and the decision to which one are sequential or simultaneous. In this sense, a clear illustration of the importance of wondering whether a decision is sequential or simultaneous is that provided by van Ophem and Schram (1997) as an example. Imagine someone has to decide whether to eat meat or not and whether to spend more or less than $50 monthly in meat. In this case, the choices are: spend more than $50 monthly, spend less, or do not eat meat at all. The sample could include vegetarians for which the question of spending more or less on meat makes no sense.

Van Ophem and Schram (1997) show how a nested model has the sequential and multinomial logit as special cases and how to test the validity of these specifications. They show that to analyze how people decide to join a labor union it makes sense to model first the decision to become a member of a union independent of the choices of unions available and then model which union given the first "step" in the sequence. Extending this work, Nagakura and Kobayashi (2008) show that the sequential logit model, which can be characterized as a sequence of independent multinomial logit models, is a limiting case of the nested logit model.

Note that in sequential models, there exists somehow the notion of time such that some decisions are taken first, followed by others; i.e. educational transitions (Mare, 1981; Buis, 2009), worker's labor force transitions (Jiménez-Martin et al, 2006) or residential location change (Ben-Akiva and de Palma, 1986). Then, for example, in the example of meat expenditure, someone before deciding on the expenditure has to decide to eat meat or to be a vegetarian on the first place. Although he uses a different method, in the case of return migration, an example of a conception of a sequential model is that applied by Kit-Chun Lam studying the impacts of imperfect information and schooling on rates of return (Lam, 1986).

About the use of polythomous logit models in the migration literature, Christiadi and Cushing (2007) note that conditional logit models have been used in many recent migration studies to model migration choice as a multinomial discrete choice because they can be applied at the individual level, and thus can better represent migration as an individual's utility maximization decision. The first applications of conditional logit models in migration literature were used for internal migration (Davies, Greenwood, & Li, 2001; Mueller, 1985).

4.b. The models

Here, I study alternative models that fall in a broad class of polytomous choice models widely used when outcomes consist of categories of choices. The three conceptually plausible models considered are the classical multinomial logit, the nested logit and the sequential logit models. Logit models will be considered within the discrete choice models framework where the set of alternatives, or choice set, will obey the following characteristics: (i) the alternatives are mutually exclusive, (ii) alternatives are exhaustive, and (iii) the number of alternatives is finite (Ben-Akiva & Lerman, 1985; Train, 2003). Also, our discrete choice models will derive from a random utility model (RUM) framework in which decision makers are assumed to maximize a utility.

A decision maker *n* has J, {1, 2, …, J} alternatives from where to choose one, and from any alternative j the individual obtains utility $U_{nj,}\ j = 1, \cdots, J$. Then, the decision maker chooses the alternative such that $U_{ni} > U_{nj}, \forall j \neq i$. Note that the researcher cannot observe the utility for each individual and therefore we can think of this as latent for each decision maker. What the researcher may know are some characteristics of the decision maker, S_n, and some characteristics of the alternatives, $x_{nj}, \forall j$. Therefore, we can specify a relationship between the observed factors and the utility of the decision maker with a function $V_{nj} = V(x_{nj}, s_n)$. And since there are unobserved factors, we can decompose the utility function with systematic and stochastic components V_{nj} and ε_{nj} such that $\forall j$, $U_{nj} = V_{nj} + \varepsilon_{nj} \Rightarrow \varepsilon_{nj} = U_{nj} - V_{nj}$ where ε_{nj} are random with joint density function $f(\varepsilon_n)$ where $\varepsilon_n = \{\varepsilon_{n1}, \ldots \varepsilon_{nJ}\}$.

Therefore, we can get different discrete choice models if we specify different density functions to the stochastic term ε_n. Logit and nested logit models have a close form and they are obtained when the unobserved part of the utility is assumed to be distributed iid extreme value or with a type of generalized extreme value. If we assume that $f(\cdot)$ is a multivariate normal, then we get the Probit model and the probability below does not have a close form.

In a multinomial logit model the utility function for individual *n* choosing alternative *j*, with systematic component V_{nj} is $U_{nj} = z_{nj}\gamma_j + \varepsilon_{nj}$. Note that the parameters γ_j relate to individual-specific characteristics and the effect of the independent variable will vary across all the choices. Also, z_n have nothing to do with the alternatives that are available. Then, the probability that an individual *n* chooses alternative *i* is

$$P_{ni} = \frac{e^{V_{ni}}}{\sum_j e^{V_{nj}}} = \frac{e^{z_{ni}\gamma_i}}{\sum_j e^{z_{nj}\gamma_j}}.$$

To solve an identification problem (similar to the need of defining *n-1* dummy variables from an *n* category variable) we need to consider a reference (or baseline) category against which the results are compared.

Independence of Irrelevant Alternatives Assumption

One of the main limitations of decision making models is the independence of irrelevant alternatives assumption that implies proportional substitution across alternatives. For any alternatives *i* and *k*, the ratio of logit probabilities is:

$$\frac{P_{ni}}{P_{nk}} = \frac{\dfrac{e^{V_{ni}}}{\sum_j e^{V_{nj}}}}{\dfrac{e^{V_{nk}}}{\sum_k e^{V_{nk}}}} = \frac{e^{V_{ni}}}{e^{V_{nk}}} = e^{V_{ni}-V_{nk}} \; .$$

This means that the ratio of two logit probabilities for choosing two alternatives only depends on these two alternatives *i* and *k*, and not on the rest of the possibilities in the choice set; i.e. the ratio is independent of other alternatives than *i* and *k* and this is what is known as independence of irrelevant alternatives (IIA).

The IIIA assumption becomes problematic when two or more alternatives are substitutes for each other and if the assumption does not hold, then the estimated coefficients will become inconsistent. In addition, if we omit a variable in the model that is common to two alternatives, the omitted variable information will be captured by the error term and then correlating these errors. Therefore, the errors will not be independent and the IIA assumption will be violated. But in this case, the problem comes from the omitted variable, and the results will be specific for this model. It is a mistake to generalize that this would apply in a different context, say with a different set of data but the same model specification.

Hausman and McFadden proposed in 1984 a test (known as Hausman test) to decide whether the IIA assumption holds or not. In this paper, we will use this test, however there are other tests available: McFadden, Train and Tye's test and Small and Hsiao test and these are compared in detail in Cheng and Long (2007). We will see how the use of estimation results without checking if this assumption holds could be problematic.

In our problem of return migration, the IIA assumption may hold if some people may consider only one or two possible destinations independent on any of the rest. The assumption may be realistic for people who transfer their job and have a fixed destination assigned or for retirees that have a small number of possible destinations where they would like to retire (Christiadi & Cushing, 2007). Focusing in our example, maybe this could be true for American expatriates that move to retire in Mexico in very specific places like Chapala, Jalisco or San Miguel de Allende, Guanajuato. But the assumption may not be true for the rest of the possible returnees that can choose between a bigger set of choices. Then, models that assume IIA may be too restrictive. The larger the number of alternatives the easier it is to violate the IIA assumption because the larger the number of alternatives, the higher the likelihood of finding at least one restricted model that is significantly different from the unrestricted model which includes all the alternatives (Christiadi & Cushing, 2007).

Models that do not assume IIA

Probit and mixed models are often used in econometrics; mainly because they do not require the IIA assumption to hold and because it provides flexibility in the modeling. However, their main drawback is that even if they are theoretically attractive, they are a computationally burdensome (Davidson & MacKinnon, 2004). Probit models can have any substitution pattern among the alternatives and this pattern will depend on the covariance matrix that is specified. I will only concentrate in the nested logit and sequential logit models.

The nested logit model is a special case (maybe the most used) of the Generalized Extreme Values (GEV) Models which are a generalization of the univariate extreme value distribution that results in the logit model. The choice set of the decision maker with alternatives can be grouped in nests or subsets such that we have the following substitution patterns:

i. IIA assumption holds within each nest: for any two alternatives in the same nest, the ratio of probabilities is independent of the attributes or the existence of all other alternatives in the nest.

ii. IIA assumption may not hold for alternatives in different nests: for any two alternatives in different nests, the ratio of probabilities may depend on the attributes of other alternatives in the two nests considered.

Let the set of *J* alternatives be partitioned into *K* mutually exclusive subsets $B_1,...,B_k$ that will be from now on called nests such that $\{1,2,\ldots,J\}=\bigcup_{i=1}^{k} B_i$. Let $U_{nj}=V_{nj}+\varepsilon_{nj}$ be the utility that individual *n* obtains from alternative *j* in B_k. The nested logit model is a special case of the GEV obtained by assuming that the cumulative distribution of the vector of the unobserved component of the utility ε_n is

$$\exp\left(-\sum_{k=1}^{K}\left(\sum_{j\in B_k} e^{-\varepsilon_{nj}/\lambda_k}\right)^{\lambda_k}\right).$$

For the logit model, each ε_{nj} is iid, i.e. independent with a univariate extreme value distribution. However, in this case the marginal distribution of each ε_{nj} is univariate extreme value, although they do not have to be independent: actually they are correlated within nests and are independent between nests. The parameter λ_k is a measure of the degree of independence in the unobserved utility among the alternatives in the nest B_k such that a high value of λ_k implies greater independence and less correlation. If $\lambda_k = 1$ there is complete independence within B_k, then the GEV becomes the product of independent univariate extreme value distributions and the nested logit model reduces itself to the multinomial logit model presented before. It is generally argued that λ_k should be between 0 and 1 in order to be consistent with the utility-maximizing behavior. If $\lambda_k > 1$, then the model is consistent only for some range of independent variables and if $\lambda_k < 0$ then it is inconsistent because it implies that increasing the utility of an alternative decreases the probability of being chosen. Recall also that λ_k is fixed for each nest such that the decision makers have the same correlations among unobserved factors. But this may not be the case.[6] It is possible to do the hypothesis testing of whether the nested specification is correct, or whether it reduces to a multinomial logit.

In the nested logit model, the choice probability of alternative *i* in nest B_k is

$$P_{ni} = \frac{e^{V_{ni}/\lambda_k}\left(\sum_{j\in B_k} e^{V_{ni}/\lambda_k}\right)^{\lambda_k - 1}}{\sum_{l=1}^{K}\left(\sum_{j\in B_l} e^{V_{ni}/\lambda_l}\right)^{\lambda_l}}$$

The IIA holds within nests but not across nests because it can be shown that the ratio of probabilities $\frac{P_{ni}}{P_{nm}} = \frac{e^{V_{ni}/\lambda_k}}{e^{V_{nm}/\lambda_l}}$ is independent on all the other alternatives within the nest. But for two alternatives in different nests, the terms in parenthesis do not cancel and then the ratio depends on the attributes in the nests that contain *i* and *m*.

It is possible to discompose the utility of alternative *j* in nest B_k in two parts: one that is constant for all alternatives within a nest (W) and another that varies over alternatives within a nest (Y). This allows us to conceive nested logit models as having two levels: the first level where alternative *k* is chosen and the second one where alternative *i* is chosen. For a thorough discussion of the nested logit, its representations and the derivations of its probability see (Train, 2003).

Kenneth Train (2003) suggests the visualization of the substitution patterns of a nested logit model with a tree diagram where each branch denotes a subset of alternatives and every leaf in every branch denotes an alternative. That is, there is proportional substitution across leaves (or alternatives) within a branch (or nest), but not across branches (or nests). Christiadi and Cushing (2007) highlight the importance of developing an acceptable, coherent, nesting pattern and that while nesting may seem to model decision making in a sequential way, it is not generally intended to represent sequential decisions, but to categorize. They also point out that the literature on discrete choice has still to develop a well defined methodology to determine which of the nesting patterns best represents reality.

The aim of a sequential logit model[7], as explained by Maarten Buis (2007), the creator of the *seqlogit* package for Stata, is to study a process that can be described as a series of choices between a small numbers of options that eventually lead to an end result. The package allows studying the effect of explanatory variables and the final outcome where the effect of the explanatory variables is a weighted sum of the log odds of passing the necessary transitions. The general idea is that the process defines a series of transitions where choices are made at each stage. Also, one has to be at risk of passing the transition in order to make the next transition; i.e. one has had to pass through all lower transitions.

[6] To allow the parameter to vary across decision makers, it is possible to define a parametric distribution on λ_k that is a function of observed characteristics of the decision makers. However, we will not do this in this work.

[7] Also called a sequential response model or model for nested dichotomies.

Therefore, the approach is to estimate a separate standard logit model at each transition where the transitions are assumed to be completely independent. That is, each choice is modeled separately using a logit model on the sub-sample that is at risk.

In order to be in stage two you have had to pass stage one, in order to pass to stage three the previous stages have to be passed through, and so on, and each transition can be considered as a binary choice. Therefore, the probability that $y_i \geq j, j > 1$ given that the previous stages have passed is:

$$P_{ij}^{+} = P(y_i \geq j \mid y_i \geq j-1) = \frac{\sum_{l=j}^{J} P_{il}}{\sum_{h=j-1}^{J} P_{ih}}, \; j > 1$$

Thus, this model does not assume the IIA in the sense that was presented before since the logit models to be estimated depend on the structure of transitions defined. For an illustration of the application of a sequential model that consists of two transitions and three alternatives studying educational attainment, see (Powers & Xie, 2000). For an analysis of the effects of unobserved heterogeneity in sequential logit models, see (Buis, 2009).

A nested model has the sequential and multinomial logit as special cases and the validity of these specifications can be tested (van Ophem & Schram, 1997). In their work, van Ophen and Schram are interested in testing whether individuals act in a sequential way restricting the choice set without taking the characteristics of the other options into account. In their definition, a sequential structure is a process where the utility of options in later stages do not influence the choice from a certain set. It is argued that if $\lambda_k \notin [0,1]$ then there is no economic interpretation of the model although it can take any real value. However, in the sequential logit model "the economic restriction that $\lambda \in [0,1]$ does not yield a corresponding statistical restriction. In the nested logit model, on the other hand, […] the estimated correlation can only lie in the interval [0,1] [and] in this case it is a statistical restriction" (van Ophem & Schram, 1997). The test of hypotheses that $\lambda = 0$ or $\lambda = 1$ can be calculated but the t-ratios are not plagued.

The difference between a sequential and a simultaneous process is the utility attributed to different alternatives in each stage. Not only should the utilities attributed to each alternative be considered, but also the underlying structure. Nagakura and Kobayashi (2009) explain that the fact that the values of the utilities of the second stage of the sequential logit model do not influence the decisions at the first stage can be interpreted in two ways. The first one says that if the between-group differences in the utilities are much larger than the within-group differences, then the choice process can be regarded as having two stages that can be treated almost independently. Their second interpretation is that an individual does not know the values of the utilities of the second-stage alternatives while being at the first stage because the information could be too costly or it could be impossible to have them beforehand. They also point out later, in their application, that different tests may yield different results such that it is usually hard to know (or inconclusive) as to which model (the nested logit model or multinomial logit model) is more suitable for the data.

For this application, we are considering only people that were living in Mexico in 2005 and were living in the United States in 2000. We know where they are in 2005 so we will assume that at some point in time, for some reasons, they decided to live in that state so that they were captured in the Population Count. Therefore, we will model the decision of choosing a state where to live in Mexico for those that were in 2000 in the United States.

For the multinomial logit model, we will consider the structure shown in figure 2. This structure assumes that returnees can choose between four major types of states. Recall that the multinomial logit model assumes the independence of irrelevant alternatives, so in this case for example this assumption would imply that the probability of choosing to live in a border state is independent to that of choosing a traditional migration sending state.

For the nested logit model, the structure is as shown in figure 3. Note that in this model, we assume that the IIA does not hold between types of states, but it does hold within each nest, or type of state. In other words, for example, the ratio of probabilities for choosing Guanajuato and Michoacán (both traditional sending states) is independent on the attributes or existence of all the other alternatives within the nest of traditional sending states. However, for Guanajuato and Baja California (two alternatives in different nests) the ratio of probabilities can depend on the attributes of other alternatives in the traditional sending states' and U.S. Border States nests.

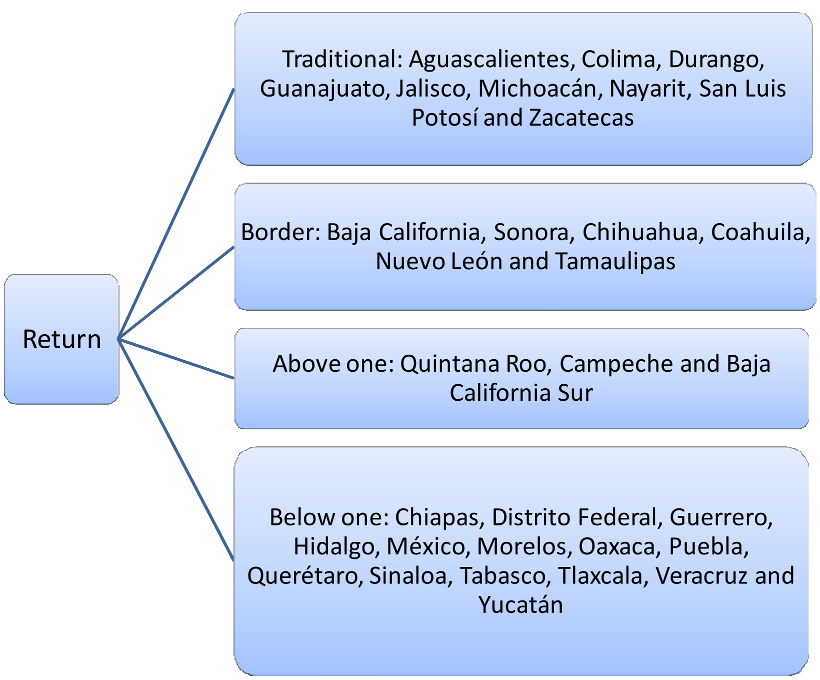

Figure 2. Decision process structure for multinomial logit model

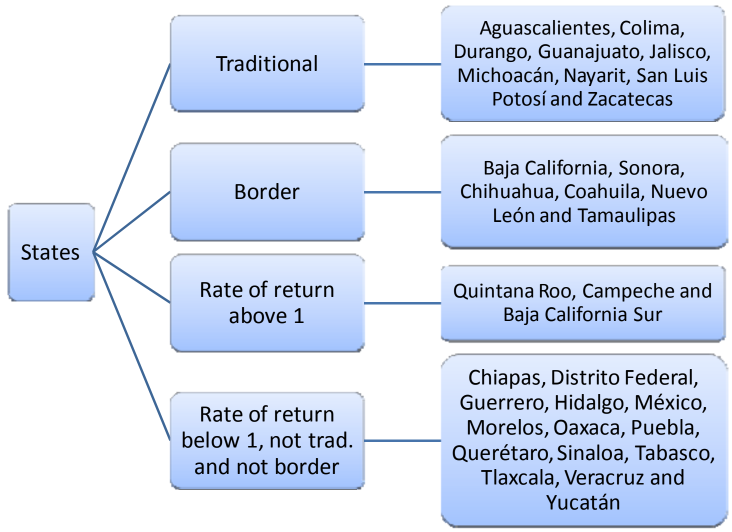

Figure 3. Decision process structure for nested logit model

Finally, for the sequential logit model we will try the following four structures that are defined from the four types of states that were defined previously (see figure 4). The selection of the sequential logit model that is more appropriate will depend on the definition of the structure of the decision process.

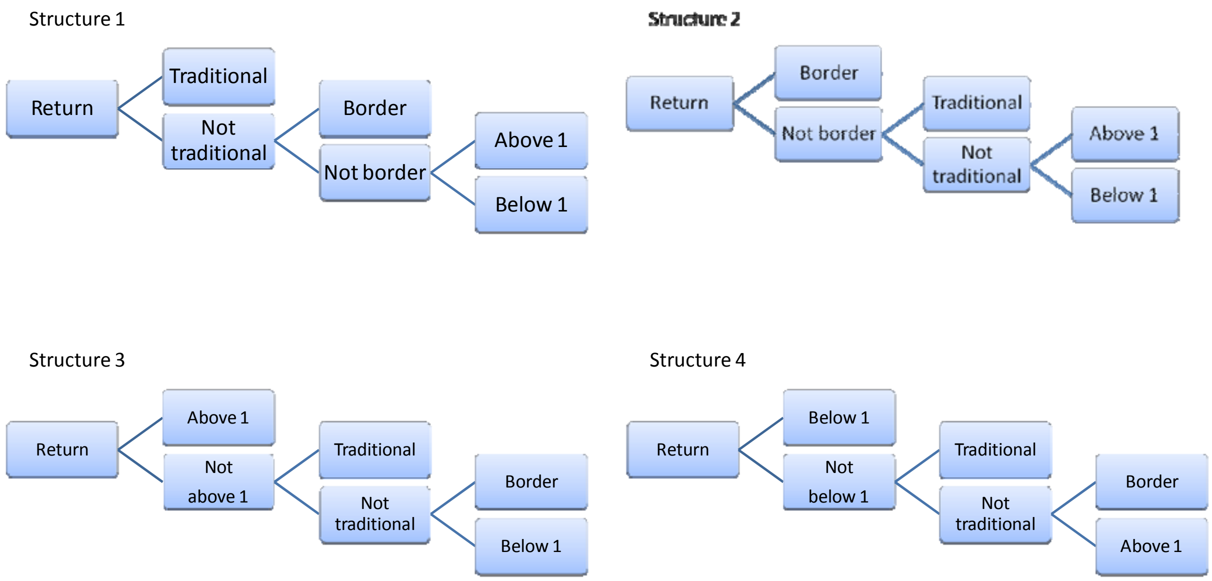

Figure 4. Four decision process structures for sequential logit models.

5. Results

5.a. Different Returnees in Four Different Types of States: a Description

Now, let us examine some descriptive statistics of the characteristics of the returnees by type of state (tables 2, 3, 4 and 5). Table 2 shows the distribution of returnees for selected characteristics of the returnees and the localities where they are living. From this table we can notice some differences. For example, almost half of the returnees are living in traditional migration sending states. The highest percentage of male returnees by type of state is living what we have called traditional states, which is coherent with the idea of subsistent circular returnees. Also the highest percentage of returnees living in rural areas is in localities in traditional sending states. The returnees living in states in the U.S. border are much more likely to be living in urban areas and metropolitan areas than those in other types of states. This was expected from the concentration of returnees in places like Tijuana, Juárez and other border cities.

Table 3 shows the age and educational characteristics of returnees by type of state. Note that there are different distributions for male and women for age. Table 4 shows the family relationship of returnees with the head of their households as well as the type of household where they are living; this is shown by gender and type of state. As expected, men are more likely to be head of households followed by sons while women are more likely to be daughters and then head of households. Also, note that the distribution of the type of households for different types of states seems fairly similar, except for the extended family households and for unipersonal households.

Table 5 shows that the types of states we have defined differ in the presence of return migration at the household level in various ways. The percentage of returnees living in households where they are the only returnee in traditional sending states and in those with rate of return below one is much higher than in border states or in those with rate of return below one. This may point out to a very different pattern of migration. On the other side, of those returnees living in households where all its members are returnees they are mostly in states that have a rate above one; i.e. these states seem to attract whole households of returnees. Also, for nuclear households with all its members being returnees, the highest percentage by type of state is living in states that attract more people than they sent out. Unipersonal households tend to be concentrated in border states or in those with a rate of return above one.

Table 2

Distribution of returnees for selected characteristics by type of state, 2005

Charactersitic		Traditional	Border	Rate of return above one	Rate of return below one	Total
Total	N	104,586	55,702	3,959	74,084	238,331
	%	43.9	23.4	1.7	31.1	100
Gender						
Female	N	33,137	22,197	1,573	25,366	82,273
	%	31.68	39.85	39.73	34.24	34.52
Male	N	71,449	33,505	2,386	48,718	156,058
	%	68.3	60.2	60.3	65.8	65.5
Size of the locality						
Urban (Population >= 15,000)	N	46,564	47,488	2,988	35,142	132,182
	%	44.5	85.3	75.5	47.4	55.5
Rural (Population <15,000)	N	58,022	8,214	971	38,942	106,149
	%	55.5	14.8	24.5	52.6	44.5
Living in a Metropolitan Area						
	N	33,057	33,209	1,189	31,319	98,774
	%	31.6	59.6	30.0	42.3	41.4
Degree of marginality of the locality						
Very high	N	397	32	14	1,617	2,060
	%	0.4	0.1	0.4	2.2	0.9
High	N	10,443	520	180	16,816	27,959
	%	10.1	0.9	4.7	22.8	11.8
Moderate	N	13,915	737	252	10,753	25,657
	%	13.4	1.3	6.5	14.6	10.9
Low	N	30,556	2,310	579	13,641	47,086
	%	29.5	4.2	15.0	18.5	19.9
Very low	N	48,232	51,628	2,843	30,894	133,597
	%	46.6	93.5	73.5	41.9	56.5

Notes: the traditional sending states include Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas; U.S. border states include Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas; those with rate of return above one include Quintana Roo, Campeche and Baja California Sur; and those with rate of return below one include the rest: Chiapas, Distrito Federal, Guerrero, Hidalgo, Estado de México, Morelos, Oaxaca, Puebla, Querétaro, Sinaloa, Tabasco, Tlaxcala, Veracruz and Yucatán.
Source: 2005 Mexican Population Count at the individual level, and degree of marginality at the locality level (CONAPO).

Table 2. Distribution of returnees for selected characteristics by type of state, 2005

Table 3
Age and educational characteristics of returnees by gender and by type of state, 2005

Charactersitic		Male: Traditional migration state	Male: Border state	Male: Rate of return above one	Male: Rate of return below one	Male: Total	Female: Traditional migration state	Female: Border state	Female: Rate of return above one	Female: Rate of return below one	Female: Total
Total	N	71,449	33,505	2,386	48,718	156,058	33,137	22,197	1,573	25,366	82,273
	%	45.8	21.5	1.5	31.2	100	40.3	27.0	1.9	30.8	100
Age											
Mean		32.81	32.53	33.84	31.94	32.49	29.52	29.32	33.25	29.20	29.44
Age distribution											
Age: 5 to 19	N	10,879	7,020	413	7,171	25,483	10,614	6,917	355	7,193	25,079
	%	15.2	21.0	17.3	14.7	16.3	32.1	31.2	22.6	28.4	30.5
Age: 20 to 34	N	32,046	12,089	928	22,517	67,580	11,657	7,860	545	10,119	30,181
	%	44.9	36.1	38.9	46.3	43.3	35.2	35.5	34.7	39.9	36.7
Age: 35 to 49	N	19,388	9,657	640	14,428	44,113	5,972	4,527	387	5,208	16,094
	%	27	29	27	30	28	18.0	20.4	24.6	20.6	19.6
Age: 50 to 64	N	6,368	3,492	302	3,631	13,793	3,299	2,074	209	1,944	7,526
	%	8.9	10.4	12.7	7.5	8.8	10.0	9.4	13.3	7.7	9.2
Age: 65+	N	2,734	1,217	101	936	4,988	1,573	793	75	876	3,317
	%	3.8	3.6	4.2	1.9	3.2	4.8	3.6	4.8	3.5	4.0
Education											
Mean years of schooling		6.80	8.59	9.99	7.83	7.55	7.09	8.40	10.33	7.95	7.76
Education distribution (15 years and older)											
No formal education	N	5,288	1,910	133	2,902	10,233	3,134	1,467	86	2,375	7,062
	%	7.5	5.8	5.7	6.0	6.6	9.6	6.8	5.6	9.6	8.8
Primary	N	33,564	9,723	509	18,561	62,357	13,335	6,316	312	8,336	28,299
	%	47.4	29.6	21.9	38.5	40.4	40.8	29.2	20.5	33.5	35.1
Secondary or equivalent	N	19,425	8,768	457	14,059	42,709	8,087	5,737	253	6,107	20,184
	%	27.4	26.7	19.7	29.2	27.7	24.8	26.5	16.6	24.6	25.0
Highschool or equivalent	N	8,414	7,893	548	7,529	24,384	5,272	5,181	384	4,491	15,328
	%	11.9	24.0	23.6	15.6	15.8	16.1	23.9	25.2	18.1	19.0
Profesional or more	N	4,159	4,562	678	5,180	14,579	2,845	2,943	490	3,554	9,832
	%	5.9	13.9	29.2	10.7	9.5	8.7	13.6	32.1	14.3	12.2

Notes: the traditional sending states include Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas; U.S. border states include Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas; those with rate of return above one include Quintana Roo, Campeche and Baja California Sur; and those with rate of return below one include the rest: Chiapas, Distrito Federal, Guerrero, Hidalgo, Estado de México, Morelos, Oaxaca, Puebla, Querétaro, Sinaloa, Tabasco, Tlaxcala, Veracruz and Yucatán.

Source: 2005 Mexican Population Count at the individual level

Table 3. Age and educational characteristics of returnees by gender and type of state, 2005

Table 4
Family relationship and type of households of returnees by gender and by type of state, 2005

Characteristic		**Male**					**Female**				
		Traditional	**Border**	**Rate of return above one**	**Rate of return below one**	**Total**	**Traditional**	**Border**	**Rate of return above one**	**Rate of return below one**	**Total**
Total											
	N	71,449	33,505	2,386	48,718	156,058	33,137	22,197	1,573	25,366	82,273
	%	45.8	21.5	1.5	31.2	100	40.3	27.0	1.9	30.8	100
Family relationship with the head of the household											
Head	N	43,311	19,301	1,482	28,810	92,904	5,320	3,300	302	4,201	13,123
	%	60.6	57.6	62.1	59.1	59.5	16.1	14.9	19.2	16.6	16.0
Spouse, husband or partner	N	1,675	1,264	104	1,171	4,214	12,087	9,009	680	8,852	30,628
	%	2.3	3.8	4.4	2.4	2.7	36.5	40.6	43.2	34.9	37.2
Son or daughter	N	20,640	8,641	517	14,169	43,967	11,457	6,997	395	8,325	27,174
	%	28.9	25.8	21.7	29.1	28.2	34.6	31.5	25.1	32.8	33.0
Domestic worker(s)	N	7	6	0	17	30	25	18	2	80	125
	%	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.3	0.2
Without family relationship	N	397	754	82	354	1,587	180	276	43	202	701
	%	0.6	2.3	3.4	0.7	1.0	0.5	1.2	2.7	0.8	0.9
Other relationship	N	5,328	3,464	195	4,129	13,116	3,984	2,536	146	3,634	10,300
	%	7.5	10.3	8.2	8.5	8.4	12.0	11.4	9.3	14.3	12.5
Guest	N	6	5	0	2	13	7	1	0	5	13
	%	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Not specified	N	85	70	6	66	227	77	60	5	67	209
	%	0.1	0.2	0.3	0.1	0.2	0.2	0.3	0.3	0.3	0.3
Type of household of returnees											
Nuclear family	N	49,895	20,488	1,435	31,759	103,577	22,293	14,847	1,048	15,651	53,839
	%	69.8	61.2	60.1	65.2	66.4	67.3	66.9	66.6	61.7	65.4
Extended family	N	16,371	7,667	465	13,510	38,013	8,866	5,689	333	8,166	23,054
	%	22.9	22.9	19.5	27.7	24.4	26.8	25.6	21.2	32.2	28.0
Mixed family	N	554	712	60	443	1,769	311	438	33	283	1,065
	%	0.8	2.1	2.5	0.9	1.1	0.9	2.0	2.1	1.1	1.3
Family, not specified	N	510	280	14	354	1,158	268	183	12	272	735
	%	0.7	0.8	0.6	0.7	0.7	0.8	0.8	0.8	1.1	0.9
Unipersonal, not fami	N	3,630	3,489	297	2,205	9,621	1,172	776	93	787	2,828
	%	5.1	10.4	12.5	4.5	6.2	3.5	3.5	5.9	3.1	3.4
Corresident, non family	N	441	804	108	404	1,757	190	240	50	183	663
	%	0.6	2.4	4.5	0.8	1.1	0.6	1.1	3.2	0.7	0.8
Not specified	N	48	65	7	43	163	37	24	4	24	89
	%	0.1	0.2	0.3	0.1	0.1	0.1	0.1	0.3	0.1	0.1

Notes: the traditional sending states include Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas; U.S. border states include Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas; those with rate of return above one include Quintana Roo, Campeche and Baja California Sur; and those with rate of return below one include the rest: Chiapas, Distrito Federal, Guerrero, Hidalgo, Estado de México, Morelos, Oaxaca, Puebla, Querétaro, Sinaloa, Tabasco, Tlaxcala, Veracruz and Yucatán.

Source: 2005 Mexican Population Count at the individual level

Table 4. Family relationship and type of households of returnees by gender and by type of state, 2005

Table 5
Number of returnees and selected type of household in households with return migration by type of state, 2005

Charactersitic		Traditional	Border	Rate of return above one	Rate of return below one	Total
Total	N	73,195	35,312	2,718	55,808	167,033
	% type	43.8	21.1	1.6	33.4	100
Number of returnees						
Unipersonal household	N	4,802	4,265	390	2,992	12,449
	% type	38.6	34.3	3.1	24.0	100
	% char.	6.6	12.1	14.4	5.4	7.5
Only one returnee	N	51,326	20,050	1,581	42,126	115,083
	% type	44.6	17.4	1.4	36.6	100
	% char.	70.1	56.8	58.2	75.5	68.9
Not all returnees	N	11,359	6,620	347	7,885	26,211
	% type	43.3	25.3	1.3	30.1	100
	% char.	15.5	18.8	12.8	14.1	15.7
All returnees	N	5,708	4,377	400	2,805	13,290
	% type	43.0	32.9	3.0	21.1	100
	% char.	7.8	12.4	14.7	5.0	8.0
For nuclear households						
Only one returnee	N	35,957	12,153	957	27,330	76,397
	%	74.2	60.2	62.3	78.9	72.9
Not all returnees	N	7,295	4,178	229	4,774	16,476
	%	15.0	20.7	14.9	13.8	15.7
All returnees	N	5,241	3,855	350	2,528	11,974
	%	10.8	19.1	22.8	7.3	11.4
For extended family households						
Only one returnee	N	14,189	6,445	446	13,587	34,667
	%	77.6	72.3	78.5	81.5	78.0
Not all returnees	N	3,771	2,110	96	2,885	8,862
	%	20.6	23.7	16.9	17.3	20
All returnees	N	321	356	26	203	906
	%	1.8	4.0	4.6	1.2	2.0

Notes: the traditional sending states include Aguascalientes, Colima, Durango, Guanajuato, Jalisco, Michoacán, Nayarit, San Luis Potosí and Zacatecas; U.S. border states include Baja California, Sonora, Chihuahua, Coahuila, Nuevo León and Tamaulipas; those with rate of return above one include Quintana Roo, Campeche and Baja California Sur; and those with rate of return below one include the rest: Chiapas, Distrito Federal, Guerrero, Hidalgo, Estado de México, Morelos, Oaxaca, Puebla, Querétaro, Sinaloa, Tabasco, Tlaxcala, Veracruz and Yucatán.

For their preponderance, the selected types of households are nuclear, extended family and unipersonal.

Source: 2005 Population Count at the individual and household level

Table 5. Number of returnees and selected type of household in households with return migration by type of state, 2005

5.b. Results: Statistical models

The results of the multinomial model, both the full and intermediate models (with different set of independent variables), will not be shown here. However, the results of the Hausman test (see table 6) show evidence against the independence of irrelevant alternatives assumption and it is therefore likely that the estimated coefficients are inconsistent[8]. Table 6 shows the Hausman test for two models: one with the variable of degree of marginality of the state (model 2), and one without it (model 1). Note that marginality is not individual specific, but is an attribute of the alternative (the chosen state) and is therefore used here in a different way than in the nested and sequential logit models where the probabilities change with attributes of the alternatives.

Table 6

Hausman test for Independence of Irrelevant Alternatives (Multinomial Logit Model)

Null hypothesis (Ho): Odds(Outcome J) vs Odds(Outcome K) are independent of other alternatives

Omitted	chi2	df	P>chi2	Evidence
Model 1				
Traditional	-2.20E+03	34	1.000	**For Ho**
Border	59.459	34	0.004	**Against Ho**
Rate above 1	123.874	34	0.000	**Against Ho**
Rate below 1	1337.779	34	0.000	**Against Ho**
Model 2				
Traditional	3.20E+04	36	0.000	**Against Ho**
Border	790.823	36	0.000	**Against Ho**
Rate above 1	8464.234	36	0.000	**Against Ho**
Rate below 1	9923.741	36	0.000	**Against Ho**

Table 6. Hausman test for the indepdendence of irrelevant alternatives in two multinomial logit models

The evidence against the IIA assumption may reflect the fact that individuals do not consider different types of states independently and that actually two or more alternatives could be substitutes for each other. Relating this to the discussion presented before, this could be reflecting that individuals decide on where to return in some hierarchical way discarding some options and not deal with all of them simultaneously.

Next, in order to apply the nested logit model, a specific data structure is needed: it has to be in long format. That is, if there are in total J alternatives or end results, each individual has to be repeated J times in the data base with an indicator variable showing which of these alternatives was chosen. In our problem, we have 32 states so the 238,331 returnees had to be repeated 32 times. When the original database of individuals is expanded, the result is a database too big to work with[9].

Therefore, it was decided to work with a 5 percent sample[10] of the set of 238,331 individual records. For the purposes of this work, the sample was generated randomly by state to assure that the

[8] Unfortunately, in many studies where the multinomial logit model is used the IIA assumption is not generally tested, or the test is not made explicit.

[9] First I tried running the nested logit models using a UNIX server at the Population Research Center and I had to assign 3GB of memory to Stata. The model that only controlled for gender, age and education took about two weeks to run and the full model had to be interrupted at the second week. It has been pointed out in Stata discussion forums that the nested logit estimation can take a long time to converge.

[10] Although we will be working with a relatively large number of observations, having to use a sample instead of the whole set of records due to computational complexity is one of the major limitations of this procedure.

distribution of returnees by state was maintained. We used an equal frequency weight equal to 20 for each observation in the sample of 11,917 observations. The weighted sample had in total 238,340 returnees (9 more due to rounding the weighted sample).

Table 7 shows the results of the nested logit model that has the traditional sending states as the baseline. At the first level equation, we have that the degree of marginality (alternative-specific) of the state is statistically significant (although close to zero) and its estimated coefficient is negative. This means that states with higher degree of marginality are less likely to be chosen to return to.

If we focus on the comparison of individuals that chose to live in traditional sending states with those at border states (and states below one), we have that all the variables included are statistically significant (5% level); while for the states with rate above one there are some variables that are not. In the second part of table 7 we can see the results for the states that have more returnees than out migrants (rate of return above one) and those that have less (rate below one). We will not provide a full interpretation of all the estimated coefficients because this goes beyond the purpose of this work since this work intends to focus in the expository work of the methodological issue of modeling return migration decisions with different polytomous models.

For example, let us consider the indicator variable of a unipersonal household, i.e. individuals living alone. From the results of the nested logit models (and from the descriptive statistics shown before) we have that unipersonal households are more likely to be in border states and states with rate below one than in traditional sending states. On the other hand, returnees living alone are less likely to be in states with rate above one than in traditional sending state. For returnees living in non unipersonal households, it is less likely for them to choose traditional sending states than border states or states with rate of return above one while the opposite is true for states with rate below one.

We can reject the null hypothesis that all of the log-sum coefficients are one from the likelihood ratio test for the IIA assumption. This indicates that in fact, a nested logit model is more appropriate in this case than a standard model. This is coherent with our previous results of the Hausman test for IIA in the multinomial logit. However, note from the dissimilarity parameters that two (for traditional sending states and for states with rate of return above one) are statistically significant negative and the other two are close to zero (although they are statistically significant different to zero). As we mentioned in the previous section, this is inconsistent with the random utility maximization theory where these parameters, in order to make sense had to be in (0, 1].[11]

---

[11] One of the advantages of the nested logit model is that if there are reasons to believe some specific structure, constraints can be specified on the dissimilarity parameters. For example, if we think that the dissimilarity parameters should be all non-negative and we want to constrain to avoid the negative parameters for *traditional* and *above*, then we can run the model again specifying different constraints. Constraints are usually used when there are degenerate branches (with only one alternative) and thus the dissimilarity parameter should be equal to one. It is also possible to constrain two or more log-sum coefficients to be equal. However, due to numerical reasons, a constraint of a parameter equal to zero cannot be defined. Although van Ophem and Schram (1997) show that a nested logit model with log-sum coefficients equal to zero is theoretically a sequential logit model, they had to use simulation for this. With the *nlogit* command in Stata, there is no way of constraining log-sum coefficients equal to zero.
We tried running a nested logit model constraining the dissimilarity parameters for *traditional* and *above* equal to 0.001 (which we thought to be close to zero) and the estimation could not be carried out since it is too small. However, with the constraint of them being equal to 0.01, the procedure (results not shown here) estimated negative log-sum coefficients for the other two types (*border* and *below*) and several standard errors of *border* could not be calculated. Although there is evidence that suggests that this model is closer to a sequential logit than to a multinomial logit model, we did not calculate this explicitly.

Table 7
RUM-consistent nested logit estimations with traditional sending states as baseline

	Estimated coefficients					
Level 1						
Degree of Marginality (chosen state)	-0.003	**				
Level 2 (type)						
	Border		**Above**		**Below**	
Male	0.039	*	0.068	-	-0.114	***
Age: 5 to 19	0.682	***	0.707	***	0.145	***
Age: 35 to 49	0.208	***	0.081	-	0.060	***
Age: 50 to 64	0.210	***	-0.077		-0.246	***
Age: 65 and more	0.102	**	0.090		-0.454	***
No formal education	-1.034	***	-1.589	***	-0.606	***
Primary school	-1.062	***	-1.988	***	-0.794	***
Secondary school	-0.520	***	-1.998	***	-0.628	***
High school	0.141	***	-0.727	***	-0.415	***
Head of household (not unipersonal)	-0.111	***	-0.332	***	0.182	***
Spouse or husband of head	0.169	***	0.057		0.220	***
Son or daughter of head	-0.301	***	-0.611	***	0.111	***
Nuclear household	-0.327	***	-0.926	***	-0.184	***
Extended or mixed family household	-0.120	***	-1.192	***	0.090	*
Unipersonal household	0.177	***	-0.510	***	0.104	*
All the members of the household are returnees	0.264	***	0.505	***	-0.186	***
Only returnee in household	-0.274	***	0.107	*	0.307	***

Dissimilarity parameters				
Type of state	**Coef.**	**Std. Err.**	**[95% Conf.**	**Interval]**
Traditional_tau	-0.027	0.008	-0.043	-0.011
Border_tau	0.007	0.002	0.003	0.012
Above one_tau	-1.046	0.072	-1.187	-0.905
Below one_tau	0.030	0.009	0.013	0.048
LR test for IIA (tau =1)	chi2(4) = 5234.55		Prob > chi2 = 0.000	
Log likelihood	-786867			

Notes: - means p < 0.1, * means p < 0.05, ** means p < 0.01 and *** means p < 0.001.
The model corresponds to 7626880 observations and 238340 cases using the frequency weight equal to 20 for each case. The Wald statistic (chi2(52)) is 30211. There are 32 alternatives per case.
Source: ten percent sample of the 2005 Population Count generated by the author

Table 7. Nested logit estimations for return migration

Finally, we include next the results of the sequential logit models with the four structures of decision process that were presented previously (see tables 8, 9, 10 and 11). Note that they differ in how the transitions at each step are defined (see fig. 4).

Table 8
Estimated Odds Ratio for the Sequential Logit Model for Structure 1

Variables	***Transition 1: Traditional vs not traditional (border, above one, below one)***		***Transition 2: Border vs not border (above one or below one)***		***Transition 3: Above one vs Below one***	
Male	0.892	***	0.864	***	0.891	**
Age: 5 to 19	1.433	***	0.776	***	0.626	***
Age: 35 to 49	1.130	***	1.062	**	1.056	
Age: 50 to 64	0.992		0.983		0.698	***
Age: 65 and more	0.782	***	1.001		0.656	***
No formal education	0.435	***	0.525	***	2.215	***
Primary school	0.410	***	0.496	***	2.608	***
Secondary school	0.557	***	0.511	***	2.286	***
High school	0.789	***	0.467	***	1.221	***
Head of household (not unipersonal)	1.012		0.925	**	0.984	
Spouse or husband of head	1.181	***	0.792	***	0.753	**
Son or daughter of head	0.936	***	1.252	***	1.341	***
Nuclear household	0.705	***	1.313	***	2.149	***
Extended or mixed family household	0.909	**	1.229	***	2.161	***
Unipersonal household	1.000		1.073		0.994	
All the members of the household are returnees	1.001		0.898	***	0.484	***
Only returnee in household	1.170	***	1.657	***	1.096	*
Degree of Marginality	0.842	***	8.356	***	1.475	***

Notes: * means $p < 0.05$, ** means $p < 0.01$ and *** means $p < 0.001$.
Source: 2005 Population Count

Table 8. Estimated odds ratios for the sequential logit model for structure one.

Table 9
Estimated Odds Ratio for Sequential Logit Model for Structure 2

Variables	***Transition 1: Border vs not border (traditional, above one, below one)***		***Transition 2: Traditional vs not traditional (above one or below one)***		***Transition 3: Above one vs Below one***	
Male	0.928	***	0.826	***	0.891	**
Age: 5 to 19	0.638	***	1.322	***	0.626	***
Age: 35 to 49	-		1.089	***	1.056	
Age: 50 to 64	-		0.923	***	0.698	***
Age: 65 and more	1.016		-		0.656	***
No formal education	0.903	**	0.329	***	2.215	***
Primary school	0.912	***	0.306	***	2.608	***
Secondary school	0.752	***	0.414	***	2.286	***
High school	0.571	***	0.554	***	1.221	***
Head of household (not unipersonal)	0.997		1.000		0.984	
Spouse or husband of head	-		1.107	***	0.753	**
Son or daughter of head	1.290	***	0.984		1.341	***
Nuclear household	-		-		2.149	***
Extended or mixed family household	1.178	**	0.994		2.161	***
Unipersonal household	1.059		0.993		0.994	
All the members of the household are returnees	0.914	***	0.893	***	0.484	***
Only returnee in household	1.381	***	1.383	***	1.096	*
Degree of Marginality	13.198	***	1.614	***	1.475	***

Notes: * means $p < 0.05$, ** means $p < 0.01$ and *** means $p < 0.001$.
The symbol - denotes variables that were removed due to collinearity issues
Source: 2005 Population Count

Table 9. Estimated odds ratios for the sequential logit model for structure two.

Table 10
Estimated Odds Ratio for Sequential Logit Model for Structure 3

Variables	*Transition 1: Above one vs not above one (traditional, border, below one)*		*Transition 2: Border vs not border (traditional or below one)*		*Transition 3: Traditional vs Below one*	
Male	1.005		0.891	***	0.843	***
Age: 5 to 19	0.611	***	1.418	***	0.744	***
Age: 35 to 49	0.979		-		1.077	**
Age: 50 to 64	0.739	***	0.985		0.903	**
Age: 65 and more	0.789	**	0.776	***	0.880	*
No formal education	3.928	***	0.455	***	0.566	***
Primary school	4.953	***	0.430	***	0.536	***
Secondary school	3.686	***	0.584	***	0.541	***
High school	1.830	***	0.813	***	0.468	***
Head of household (not unipersonal)	1.023		1.013		0.918	*
Spouse or husband of head	0.808	**	-		0.767	***
Son or daughter of head	1.273	***	0.942	***	1.278	***
Nuclear household	1.962	***	0.723	***	1.443	***
Extended or mixed family household	1.883	***	0.934	*	1.333	***
Unipersonal household	0.974		1.009		1.073	
All the members of the household are returnees	0.565	***	0.983		0.834	***
Only returnee in household	0.872	*	1.167	*	1.703	***
Degree of Marginality	0.981		0.841	**	8.935	***

Notes: * means $p < 0.05$, ** means $p < 0.01$ and *** means $p < 0.001$.
The symbol - denotes variables that were removed due to collinearity issues
Source: 2005 Population Count

Table 10. Estimated odds ratios for the sequential logit model for structure three.

Table 11
Estimated Odds Ratio for Sequential Logit Model for Structure 4

Variables	*Transition 1: Below one vs not below one (traditional, border, above one)*		*Transition 2: Traditional vs not traditional border or above one)*		*Transition 3: Border vs Above one*	
Male	1.223	***	0.976		1.052	
Age: 5 to 19	0.850	***	1.858	***	1.229	**
Age: 35 to 49	0.930	***	1.244	***	0.967	
Age: 50 to 64	1.120	***	1.241	***	1.543	***
Age: 65 and more	1.338	***	0.961		1.929	***
No formal education	2.260	***	0.391	***	0.350	***
Primary school	2.387	***	0.355	***	0.314	***
Secondary school	1.845	***	0.515	***	0.369	***
High school	1.631	***	0.892	***	0.555	***
Head of household (not unipersonal)	0.454	***	0.084	***	12.717	***
Spouse or husband of head	1.008		0.904	**	1.019	
Son or daughter of head	0.983		1.195	***	1.125	
Nuclear household	0.959	*	0.753	***	0.959	
Extended or mixed family household	0.999		0.564	***	0.635	***
Unipersonal household	0.858	***	0.739	***	0.646	***
All the members of the household are returnees	0.977		0.881		1.045	
Only returnee in household	1.220	***	1.146	***	1.642	***
Degree of Marginality	0.725	***	0.965		1.145	*

Notes: * means $p < 0.05$, ** means $p < 0.01$ and *** means $p < 0.001$.
Source: 2005 Population Count

Table 11. Estimated odds ratios for the sequential logit model for structure four.

Next, to compare the sequential logit models defined by the four possible decision structures, we have that the one with greatest log likelihood is the model with structure 4. Note that choosing this model against the others is also the result of applying different criteria using AIC or BIC (see table 12).

Therefore, for this data, it seems that the more appropriate sequential logit model is the last one, where the transitions are the following:

i. Decide to return to a state that has a rate of return below one (but which is not a traditional migration sending state or a border state) or decide to go to another state.

ii. Given that a migrant decided not to return to a state with rate below one, then decide whether to return to a traditional sending state or not.

iii. Given that a migrant decided not to return to a traditional migration sending state, then decide whether to return to a state in the U.S. border or to a state that has a rate of return above one.

Table 12
Statistics for Goodness of fit for the different structures estimated for the Sequential Logit Models

Sequential Logit Model	Obs.	Log likelihood of null model	Log likelihood of current model	df	AIC	BIC
Structure 1	238331	-269900.2	-215583.2	57	431280.5	431872.2
Structure 2	238331	-269900.2	-193805.1	57	387724.3	388316
Structure 3	238331	-269900.2	-215993.8	57	432101.5	432693.3
Structure 4	238331	-269900.2	**-186464.9**	57	**373043.9**	**373635.6**

Notes: structures of the sequential logit model are those shown in figure 4

Table 12. Goodness of fit for sequential logit models

Interpretation for the sequential logit model is done as you would in a standard logit model, but dependant on the outcome of each transition. For example, focusing only in the last sequential model (structure 4) that seems to be fitting this specific data better and in the indicator variable of unipersonal households, we have the following. In the first transition, returnees living alone (unipersonal households) are less likely to be go to states with rate of return below than to the rest (odds ratio is 0.858). In the second transition, given that a returnee did not go to a state with rate below one, they are less likely to go to traditional sending states (odds ratio is 0.739) than to the rest (border states or states with rate above one). Finally, on the last transition, unipersonal households are less likely to go to states in the northern border than to states with rate above one. What this is telling us is that returnees living alone are more likely to be living in those states pointed out as being attracting more returnees than the people that they have sent out, confirming what was shown in the descriptive analysis.

Due to the data structure of the nested logit model, it is impossible to compare the nested logit models and the sequential logit models through the Log likelihood function and the calculation of the AIC or BIC goodness of fit statistics. On the other hand, sequential logit models allow us to calculate fit statistics to compare the models and thus, the ability to test within the possible alternative sequential models. Note also that these models do not include ancillary parameters (like the dissimilarity parameters in the nested logit). In addition, the sequential logit models are tractable computationally[12] and do not present the complexity as do nested logit models.

[12] The sequential logit models (for the whole set of returnees) were estimated in a few minutes, and not hours like the nested logit model (for the 5% sample of returnees).

6. Discussion

We found evidence (both empirical -using the Hausman and likelihood ratio tests- and theoretical) that the first method, i.e. the multinomial logit model, is not appropriate because the independence of irrelevant alternatives assumption does not hold.

The nested logit model shows that for this specific setting, the estimated dissimilarity parameters go against the maximizing utility framework. We could think that there is a model misspecification in either, or both, of the levels of analysis (nests and alternatives) and/or that there is inconsistency in random utility maximization because migrants are not maximizers but satisficers while deciding where to return to. However, it could also be the case that a sequential decision process is more realistic than the simultaneous. Recall that it was pointed out earlier that in decision processes, the underlying structure should be taken into consideration.

For example, if there are three alternatives (A, B and C) and the decision maker can conceive the process in a simultaneous way (all three alternatives at the same time) or in a sequential way (1st: choose C or not C (A or B); if not C: choose A or B) In this way, in the utility maximization framework the fact that the highest utility is associated to an option A is necessary, but not sufficient, to choose A. It is possible that an individual does not arrive far enough in the structure (in a sequential model) to consider the alternative A because it may be possible that in the sequential structure he chose C. For this to happen, it is clear that the utility of C has to be lower than the utility of "NOT C". If the utility for C is higher than "NOT C" then the individual will chose C and not consider at all the alternatives A or B. Note that in this case, for option C, a highest utility attributed to it is a sufficient, but not necessary condition for it to be chosen.

From a migration perspective, it makes sense that for some individuals it could be the case that their decision can actually be reduced to going back to where they originally left from or not; and once they decided to go somewhere else, then they can choose from a choice set of different places. Plus, it is hard to believe that returnees have similar levels of information about all the states of Mexico. There is an equivalency with the meat expenditure problem. It could be possible that for someone, in a first stage, the choice is between only two alternatives: go back to place X or not to go to place X; just like for a vegetarian the meat expenditure problem can be thought first as "eat meat or not eat meat". For this people, the simultaneous model of considering all the options at the same time may be out of consideration.

About the inconsistency in random utility maximization, let us consider possible explanations for this. From the migration literature we know that returnees may have very diverse reasons for returning: from family ties, pressures and expectations, to kinship, and other economic, legal, social and political factors. Also, most of the deported Mexicans are sent to border cities and they then have to find their way in Mexico. Some may be staying in border cities while attempting to enter the United States, some may stay there because there are usually job opportunities in manufacturing in those cities, while others may go to the community they originally left from, or somewhere else. On the other hand, it could be possible that a father that left behind a wife and children in a rural and poor community may have different intentions, pressures and reasons to decide return back to his home. The "utilities" that individuals have to "maximize" may be completely different from one individual to another.

The nested logit model does not relax the IIA assumption completely since it assumes IIA within nests. Therefore, it could also be possible that the way states where classified in different types of states is not the most appropriate for the nested logit model that would best suit this data. The misspecification of the nested logit model could be derived from the violation of the IIA assumption within some nests. The assumption of IIA within nests assumes a specific substitution pattern for each type of state.

Although there is a way of testing which of the possible structures in the sequential logit model provides a better statistical fit, it is impossible to compare the sequential logit model and the nested logit we have estimated. In this sense, our results are inconclusive as to which is more appropriate for our problem. However, from the limitations and advantages of each model we can do some comments.

A limitation of the nested logit model is its computational complexity and the structure of the data that is required. In this sense, the sequential logit model is preferable since its routine is pretty fast and does not need the data in long form. Also, the sequential logit estimation routine allows for decompositions of the effects of the variables and of unobserved heterogeneity (Buis, 2009), which were not carried out at this point but remain as future work.

From the fit statistics we have that the sequential structure that fits better is the last one, but we could think of this as: "on average" the best sequence. However, it could be the case that the "best" decision process structure for each individual is different, just because people have different rationales. A sequence structure that is appropriate for someone may not make sense for someone else. It should be recognized that this analysis has been dependent on the four types of states that were defined. We had theoretical and empirical reasons for choosing this classification scheme, but nothing tells us we did the best we could. In this sense, it is important to recall what Ben-Akiva (1985) pointed out about the choice set generation process already discussed. The definition and classification of the types of states has been exogenously done by the researcher and it is not endogenous, or from the data.

As an extension, we could also try to define different sequences with an underlying latent class of sequences that hold from different people. For example, if we create a continuous variable that identifies or classifies people in different groups, then we could model a latent class of structure with this new classification, and then focus on a specific characteristic of returnees. This could be extended to be applied to study whether people go back to where they came from originally. For example, using information of who left using the 2000 Census we could try to summarize theory characteristics and then see what sequential structures are better for different migrants. An advantage here is that if we use sequential logit models it is possible to compare and chose a structure via fit statistics.

7. Conclusions

Finally, let us go back to our original questions of which of the three discrete-choice models is more appropriate for Mexican return migration decisions, what different results do we obtain if we apply different statistical models and how do the results depend on the choice of the model.

As discussed in the previous section, for this specific application, the sequential logit model offers more flexibility in terms of defining the decision structure and in terms of comparison and interpretation. The routines were carried out in a short time and it allows for decompositions of the effects of the variables and of unobserved heterogeneity. There is evidence showing that the multinomial logit model is not appropriate because the independence of irrelevant alternatives assumption does not hold. In the nested logit model the estimated dissimilarity parameters go against the maximizing utility framework and had the limitation of its computational complexity and the structure of the data that is required.

We find that these models are dependent on how researchers think of these decision processes and that, for this specific type of application, the sequential logit model offers more flexibility. As highlighted in the discussion, this implementation opened venues for future work for extensions of this model.

Finally, we could think of *l'embarras du choix* as the difficulty that migrants in the United States face while choosing where to return in Mexico. On the other hand, the expression could also reflect the difficulty that researchers face while choosing a statistical model that best suits the problem they want to solve. As a researcher studying return migration with this data, we cannot infer on the mechanisms put into place for those that returned about their decision of where to go. What researchers face is a set of different alternatives. At the same time, different researchers will conceive the problem differently and therefore, have different opinions about what structure and model to use. We all face somehow the embarrassment of choice.

Acknowledgements

I would like to thank the support of two institutions: the University of Texas at Austin and McGill University. I am very grateful to Daniel Powers and Bryan Roberts for supervising this work and to Thomas Pullum for the comments on a previous version of this paper. This work was carried out in the Population Research Center of the University of Texas at Austin while being a graduate student there. The access to the restricted data was possible due to an agreement between the University of Texas at Austin and the Mexican Ministry of Social Development (SEDESOL) to share data for research purposes. I would also like to thank Celine Le Bourdais and the Sociology Department of McGill University for the financial support for attending the 2010 Meeting of PAA.

References

Anzaldo, C., & Prado, M. (2007). *Índice de Marginación a Nivel Localidad 2005*. Mexico City: Consejo Nacional de Población.

Ben-Akiva, M., & Lerman, S. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge: MIT Press.

Berry, J. (2001). A Psychology of Immigration. *Journal of Social Issues, 57*(3), 615-631.

Borjas, G. J., & Bratsberg, B. (1996). Who Leaves? the Outmigration of the Foreign-Born. *The Review of Economics and Statistics, 78*(1), 165-176.

Buis, M. (2007). SEQLOGIT: Stata module to fit a sequential logit model Retrieved February 24th, 2009, from http://ideas.repec.org/c/boc/bocode/s456843.html

Buis, M. (2009). The consequences of unobserved heterogeneity in a sequential logit model. Retrieved February 24th, 2009, from http://home.fsw.vu.nl/m.buis/wp/unobserved_het.pdf

Cassarino, J.-P. (2004). Theorising Return Migration: The Conceptual Approach to Return Migrants Revisited. *International Journal on Multicultural Societies, 6*(2), 253-279.

Cheng, S., & Long, S. (2007). Testing for IIA in the Multinomial Logit Model. *Sociological Methods and Research 35*(4), 583-600.

Christiadi, & Cushing, B. (2007). *Conditional Logit, IIA, and Alternatives for Estimating models of Interstate Migration*. Paper presented at the 46th Annual Meeting of the SRSA. from http://www.rri.wvu.edu/pdffiles/wpcushing2007-4.pdf

Cohen, Y., & Haberfeld, Y. (2001). Self-Selection and Return Migration: Israeli-Born Jews Returning Home from the United States during the 1980s. *Population Studies, 55*(1), 79-91.

Consejo Nacional de Evaluación de la Política de Desarrollo Social (2007). *El CONEVAL Da a Conocer Mapas de Pobreza en México. (September 2007. Press Release and Data Bases)*. Mexico City: Consejo Nacional de Evaluación de la Política de Desarrollo Social.

Cortés, F. (2002). Consideraciones sobre la Marginalidad, Marginación, Pobreza y Desigualdad en la Distribución del Ingreso. *Papeles de Población, 8*(31), 9-24.

Davidson, R., & MacKinnon, J. (2004). *Econometric Theory and Methods*. New York: Oxford University Press.

Davies, P. S., Greenwood, M. J., & Li, H. (2001). A Conditional Logit Approach to U.S. State-to-State Migration. *Journal of Regional Science, 41*(2), 337-360.

de Jong, G. F., & Gardner, R. W. (1981). *Migration Decision Making. Multidisciplinary Approaches to Microlevel Studies in Developed and Developing Countries*. New York: Pergamon Press.

Donato, K. M., Tolbert, C., Nucci, A., & Kawano, Y. (2008). Changing Faces, Changing Places: The Emergence of New Nonmetropolitan Immigrant Gateways. In D. Massey (Ed.), *New Faces in New Places: the Changing Geography of American Immigration*. New York: Russell Sage Foundation.

Durand, J., Massey, D., & Zenteno, R. (2001). Mexican Immigration to the United States: Continuities and Changes. [Report]. *Latin American Research Review, 36*(1), 107-127.

Escobar, A. (2008). Pobreza y Migración Internacional en el México Rural: un Enfoque Antropológico. In A. Escobar (Ed.), *Pobreza y Migración Internacional* (pp. 9-38). Mexico City: Centro de Investigaciones y Estudios Superiores en Antropología Social.

Fawcett, J. (1985). Migration Psychology: New Behavioral Models. *Population and Environment, 8*(1-2), 5-14.

Fussell, E. (2004a). Sources of Mexico's Migration Stream: Rural, Urban and Border Migrants to the United States. *Social Forces, 82*(3), 937-967.

Fussell, E. (2004b). Tijuana's Place in the Mexican Migration Stream: Destination for Internal Migrants or Stepping Stone to the United States? In J. Durand & D. Massey (Eds.), *Crossing the Border: Research from the Mexican Migration Project*. New York: Russell Sage Foundation.

Fussell, E., & Massey, D. (2004). The Limits to Cumulative Causation: International Migration From Mexican Urban Areas. *Demography, 41*(1), 151-171.

Hernández, D., & Székely, M. (2005). Medición del Bienestar en México en los Inicios del Siglo XXI. In M. Székely (Ed.), *Números que Mueven al Mundo: La Medición de la Pobreza en México* (pp. 85-103). Mexico City: Miguel Ángel Porrúa.

Lam, K.-C. (1986). Imperfect Information, Specificity of Schooling and Rate of Return Migration. *Economics Letters, 21*, 283-289.

Lam, K.-C. (1994). Outmigration of Foreign-Born Members in Canada. *The Canadian Journal of Economics / Revue canadienne d'Economique, 27*(2), 352-370.

Leach, M. A., & Bean, F. D. (2008). The Structure and Dynamics of Mexican Migration to New Destinations in the United States. In D. Massey (Ed.), *New faces in new places: the changing geography of American immigration*. New York: Russell Sage Foundation.

Mare, R. d. (1981). Change and Stability in Educational Stratification. *American Sociological Review, 46*(1), 72-87.

Masferrer, C. (2009). *L'Embarras du Choix: Modeling Mexican Return Migration Decisions.* The University of Texas at Austin, Austin.

Masferrer, C., & Roberts, B. (2009). *Individual and Household Characteristics of Mexican Return Migration by Community of Return in 2005*. Paper presented at the 2009 Annual Meeting Population Association of America. from http://paa2009.princeton.edu/abstractViewer.aspx?submissionId=90657

Massey, D. (1990). Social Structure, Household Strategies, and the Cumulative Causation of Migration. *Population Index, 56*, 3-26.

Massey, D. (1999). Why Does Immigration Occur? A Theoretical Synthesis. In C. Hirschman, P. Kasinitz & J. DeWind (Eds.), *The Handbook of International Migration: The American Experience* (pp. 34-52). New York: Russell Sage Foundation.

McFadden, D. (1973). Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka (Ed.), *Frontiers of Econometrics*. New York: Academic Press.

Mueller, C. F. (1985). *The Economics of Labor Migration: A Behavioral Analysis*. New York: Academic Press.

Nagakura, D., & Kobayashi, M. (2009). Testing the Sequential Logit Model Against the Nested Logit Model. *The Japanese Economic Review, 60*(3), 345-361.

Pellegrini, P., & Fotheringham, S. (2002). Modelling Spatial Choice: A Review and Synthesis in a Migration Context. *Progress in Human Geography, 26*(4), 487-510.

Powers, D., & Xie, Y. (2000). *Statistical Methods for Categorical Data Analysis*. San Diego: Academic Press.

Quinn, M. A. (2006). Relative Depravation, Wage Differentials and Mexican Migration. *Review of Development Economics, 10*(1), 135-153.

Roberts, B., Frank, R., & Lozano-Asencio, F. (1999). Transnational Migrant Communities and Mexican Migration to the US. *Ethnic and Racial Studies, 22*(2), 238-266.

Roberts, B., & Hamilton, E. (2007). La Nueva Geografía de la Emigración: Zonas Emergentes de Atracción y Expulsión, Continuidad y Cambio. In M. Ariza & A. Portes (Eds.), *El País Transnacional: Migración Mexicana y Cambio Social a Través de la Frontera*. Mexico City: Universidad Nacional Autónoma de México.

Train, K. (2003). *Discrete Choice Methods with Simulation*. Cambridge: Cambridge University Press.

Tuirán, R., Fuentes, C., & Ávila, J. L. (2002). *Índice de Intensidad Migratoria México-Estados Unidos 2000*. Mexico City: Consejo Nacional de Población.

van den Berg, A., & Groot, W. (1992). Union Membership in the Netherlands: A Cross-Sectional Analysis. *Empirical Econometrics, 17*, 537-564.

van Ophem, H., & Schram, A. (1997). Sequential and Multinomial Logit: A Nested Model. *Empirical Econometrics, 22,* 131-152.