A Comparison of Three Models of Internal Migration in Mexico

Estela Rivero-Fuentes

Introduction

An accurate forecast of interstate migration and an understanding of the factors that motivate individuals to move from one place to another are very important for policy makers. Future health, education and employment needs of the region depend on the number of persons who move in or out of a state. At the same time, any attempt to predict what will be the impact of particular economic and social policies on future migration flows demands an understanding of the forces that drive migration.

There is a vast theoretical and empirical literature that helps to understand the causes of internal migration in developing countries. However, one limitation of the majority of existing empirical studies is that they focus only on one migratory flow. Any individual who migrates can decide between many different destinations. Furthermore, people from different states vary in their probability of moving out and on their destination choices. Concentrating exclusively on the migration between one state and one origin can give a very limited picture of the factors that affect migratory decisions, and is of little use when one wants to forecast what are the states that will lose more population, and what are the states that will receive more migrants.

Studies that look simultaneously at the out-migration from more than one origin, and that consider competing destinations are becoming more common in the Demographic and Sociological literature. Nevertheless, these studies vary significantly in their methodology. A review of the empirical literature of migration to competing destinations in developed and developing countries during the last 15 years reveals that authors use indistinctly conditional logit models, nested logit models or hierarchical logit models when analyzing this phenomenon. These three models make different assumptions about the way that individuals decide whether to move or not and where to move to, and may result in different predictions of the level of out-migration and immigration. However, no study to date has contrasted the results of these three models for one single country.

In this paper I compare the results of applying the three statistical models more used in the literature of competing destinations to the case of interstate migration in Mexico during 1975-1980, 1985-1990 and 1995-2000. I compare the results of the three models in terms of their conclusions about the factors affecting migration, of their fit to all state-to-state migratory flows, of their accuracy for predicting the overall probability of out-migration from each state, and of their accuracy for predicting the number of migrants to each state.

My results show a great variation in the results and conclusions obtained from the three models. Nested logit models and hierarchical logit models, which treat the probability of out-migration and choice of destination as two different processes, consistently yield the most accurate predictions and their conclusions are consistent with existing theories of migration. On the contrary, conditional logit models assuming that the decision of out-migration and the choice of destination are simultaneous and similar yield the largest errors and lead, in some cases, to conclusions that are inconsistent with the prediction of migration theories.

Migration to Competing Destinations and Analytical Models of Migration

Theories of internal migration recognize the multiplicity of choices that individuals face when making a migratory decision. The most accepted perspective argues that individuals decide where to live based on a comparison of the expected long-term costs and benefits of living in different places, including the place they currently reside in {Sjaastad, 1962 #364}{Lucas, 1997 #292}. Thinking about inter-state migration, the destinations that individuals consider as residential alternatives depend on the information they have about each state. This, in turn, depends on the infrastructure and trade ties of their state of origin with the different regions in the country {Rivero-Fuentes, 2003 #366} and on the migratory behavior of other individuals before them {Rivero-Fuentes, 2003 #366} {Davis, 2002 #430}. The costs and benefits that individuals would obtain in each destination depend on its income and unemployment level {Harris, 1970 #185}, on the distance they would have to travel {Harris, 1970 #185} {Greenwood, 1985 #81}, on how many people they know there {De Jong, 1999 #399; García-España, 1994 #40}, and on how familiar they are with the destination {Hugo, 1991 #158}.

However, the literature is not clear about how exactly the migratory decision is made. Do individuals decide to leave their place of origin first, and then decide where to move to? Or do they decide that they want to migrate only after they have compared the conditions in all the destinations they are aware of and know where they will be moving to? And if so, do they value the conditions in their place of origin more than they value the conditions in other destinations, or would they be willing to move to any place that has relatively better living conditions? Different ways to think about these questions lead to different ways to model statistically the migration process.

A review of empirical studies of internal migration in developed and developing countries during the last 15 years shows the common use of three different statistical models to represent the migration process. These models are conditional logit models, hierarchical logit models, and nested logit models. The assumptions about the way that individuals decide whether to move or not and where to move to differ between these three models. In addition, these models may result in different predictions of the level of out-migration and immigration and in contradictory conclusions about the factors affecting migration. Different authors use different models, but none of them discusses the relative advantages and disadvantages of their choice compared to others, nor do they talk about how their results would be different if they had used other model.

In this paper I attempt to fill this gap in the literature by applying these three models of migration to the case of interstate migration in Mexico during 1975-1980, 1985-1990 and 1995-2000 and comparing them in terms of their conclusions about the factors affecting internal migration, of the predicted probability of leaving each state, and of the number of migrants predicted to go to each state.

Hierarchical Logit Models [1]

This is the simplest of the three models considered. An example of this model can be found in Rogers, Willekens and Raymer's study of the changes in interstate migration in the U.S. during the last three decades {Rogers, 2001 #347}.

This model assumes that individuals decide first whether to migrate or not, and later they decide where to move to. The probability of out-migration is modeled with a logistic regression where the only two alternatives are staying in or moving out of state *i*, and the decision depends solely on characteristics of the state of residence like employment. Among those individuals who decide to migrate, the choice of destination is later modeled with a conditional logit model, where the choice of destination depends on a comparison of the conditions in all potential destinations, not including the state of origin. The unconditional probability of an individual

[1] I am following (Heiss 2002) and (Hunt 1999) terminology. Note that these are not "multilevel" models.

migrating from state *i* to state *j* is given by the probability of the individual migrating times the probability of the individual going to state *j*, given that he/she decided to migrate.

Expressed mathematically, the probability of moving out of state *i* is given by:

$$P(M_i) = \frac{\exp(\alpha + \beta X_i)}{1 + \exp(\alpha + \beta X_i)} \qquad \text{eq (1)}$$

where *Xi* are characteristics of the place of origin; and α and β are parameters that represent the baseline odds of out-migration and the effect of the covariates, respectively. The conditional probability of moving to state *j* given that the individual decides to migrate out of state *i* is given by:

$$P(mij / j \neq i) = \frac{\exp(\gamma_j + \lambda * X_j + \delta * X_{ij})}{\sum_{k=1}^{J} \exp(\gamma_k + \lambda * X_k + \delta * X_{ik})} \qquad \text{eq (2)}$$

where *Xj* are characteristics of state *j* like unemployment rate; *Xij* are characteristics of state *j* that vary according to the state of origin *i*, like distance; and γ, λ and δ are parameters. Finally, the unconditional or overall probability of migrating from state *i* to state *j* is given by:

$$P(Mij) = \frac{\exp(\alpha + \beta X_i)}{1 + \exp(\alpha + \beta X_i)} * \frac{\exp(\gamma_j + \lambda * X_j + \delta * X_{ij})}{\sum_{k=1}^{J} \exp(\gamma_k + \lambda * X_k + \delta * X_{ik})} \qquad \text{eq (3)}$$

Conditional Logit Models

One of the most common forms to model migration is through a conditional logit (see for example {Davies, 2001 #427; Grusky, 1998 #397; Lin, 1998 #396}). This model assumes that individuals decide simultaneously whether to migrate or not, and where to migrate to. When deciding, individuals compare the conditions in all their possible destinations (including their current residence) and choose the place that has more favorable conditions, given their state of residence. Since individuals might have a preference to stay where they are or there might be unobserved costs to migration, the option of "not migrating" is distinguished from the options that involve a change of residence with a “nonmigration” categorical variable that takes the value of one when the state of origin and the state of destination are the same (see (Davies, Greenwood et al. 2001)). However, individuals value the conditions in their place of origin just as much as they value the conditions elsewhere. For example, the effect on an increase in the income of the state of origin *i* on the odds of staying in *i* versus going to *k* will be the same than the effect an increase in the income of state *j* on the odds of going to *j* rather than to *k*.

The probability of moving from *i* to *j* in this model (or of staying in *i*, when *j=i*) is given by:

$$P(M_{ij}) = \frac{\exp(\alpha_j + \beta * X_j + \gamma * X_{ij} + \delta * NM_{ij})}{\sum_{k=1}^{J} \exp(\alpha_k + \beta * X_k + \gamma * X_{ik} + \delta * NM_{ik})} \qquad \text{eq(4)}$$

NMij is the nonmigration dummy and the coefficient of the nonmigration dummy (δ) represents the log-odds ratio of staying in the state of origin *versus* moving to a state with exactly the same characteristics as the origin. This coefficient has also been interpreted as the holding power of the state of origin (Lin and Xie 1998), and as an indicator of the costs of migration (Davies, Greenwood et al. 2001).

Conditional logit models can also be interpreted in terms of a utility maximizing decision process[2] (Moss 1979; McFadden). One can think that individuals living in state *i* can derive a utility *Uij* from any state *j* (including *i*) where they consider living. This utility is a linear function on the characteristics *Xi* and *Xij*, and on the case of no migration, in an additional premium for not-moving *(δ*NMii)*. According to this model, individuals decide to live in the

[2] This is not the case for hierarchical logit models (Moss 1979; Hunt 2000; Heiss 2002).

state where they obtain the maximum expected utility, given by

$$E(Uij) = \alpha_j + \beta * X_j + \gamma * X_{ij} + \delta * NM_{ij}.$$

Nested Logit Models

These models have been increasingly used in the literature of internal migration during the last six years, and good examples of their application are Liang and White's study of the impact of market transition in interprovincial migration in China in 1983-1988 {Liang, 1997 #400}, and Frey et al. study of the factors affecting population redistribution in the U.S. {Frey, 1996 #411}.

The modeling procedure of nested logit models is similar to hierarchical logit models in the sense that they divide the migration process in two stages. However, contrary to hierarchical logit models, the overall probability of out-migration in nested logit models does not depend exclusively on the characteristics of the place of origin but also on the conditions in all other states in the country. This is because the probability of out-migration is conceptualized as a logistic process that includes a summary measure of the conditions in all the other states in the country. This summary measure is known as the "inclusive value" (IV) and represents how attractive is the option of migration for each state.

In addition, nested logit models differ from hierarchical logit models because these models are flexible enough to allow the assumption that individuals value the conditions in their place of origin differently than the conditions in other states[3]; and also because they allow the possibility that individuals do not consider the same characteristics of origin and destination when making their migratory decisions[4].

---

[3] The β coefficients are different for sending and receiving states.

[4] The variables included to describe the characteristics in the state of origin need not be the same then the variables included to describe the characteristics of the potential destinations.

As in hierarchical logit models, the unconditional probability of migrating to any particular state *j* is the product of the overall probability of leaving the state of origin times the probability of going to *j* given that a migration occurs. When a nested logit model is fitted, one needs to calculate first the the probability of moving from origin *i* to destination *j*, given that individuals are moving out of *i*. This is given, as in hierarchical logit models by:

$$P(mij\,/\,j \neq i) = \frac{\exp(\gamma_j + \lambda * X_j + \delta * X_{ij})}{\sum_{k=1}^{J} \exp(\gamma_k + \lambda * X_k + \delta * X_{ik})} \qquad \text{eq (2)}$$

Secondly, one calculates the inclusive value for each state of origin, which is given by:

$$IV_i = \sum_{\substack{k=1 \\ k \neq i}}^{J} \exp(\gamma_k + \lambda * X_k + \delta * X_{ik}) \qquad \text{eq (5)}$$

And thirdly, one calculates the overall probability of out-migrating from state *i*, as given by:

$$P(M_i) = \frac{\exp(\alpha + \beta * X_i + \varphi * IV_i)}{1 + \exp(\alpha + \beta * X_i + \varphi * IV_i)} \qquad \text{eq(6)}$$

Nested logit models can also be interpreted in terms of an utility maximizing decision process (Hunt 2000; Heiss 2002). One can think that the utility that individuals living in state *i* would get from moving to state *j* would be given by the function:

$$E(Uij) = \gamma_j + \lambda * (X_j) + \delta * (X_{ij}), \text{ for } i \neq j \qquad \text{eq(7)}$$

And that the utility that individuals would get from staying in their place of origin would be given by:

$$E(U_{ii}) = \alpha + \beta * X_i \qquad \text{eq(8)}$$

Individuals decide whether to leave the state they live in or not by comparing their expected utility there *(E(Uii))* with the maximum utility they could get in any other state (max(*E(Uij)*). If they decide to migrate, they will move to the place that maximizes their utility.

Data and Variables

The dependent variable in the statistical analyses in this paper is the probability of individuals 20 years and older out-migrating from their state of residence, and moving from state *i* to state *j* in Mexico during 1975-1980, 1985-1990 and 1995-2000. The number of migrants in each period comes from the Mexican Censuses of Population of 1980, 1990 and 2000 {Instituto Nacional de Estadística Geografía e Informática, 1984 #377; Instituto Nacional de Estadística Geografía e Informática, 1992 #378; Instituto Nacional de Estadística Geografía e Informática, 2001 #379}.

Interstate migration in Mexico is a good example to compare the fit of these three models because the rates of out-migration vary notably between the 32 states in the country and because the states migrants go to differ significantly from one state to another. In addition, the level of interstate migration and the variation in migrants' destinations changed from 1975 to 2000. Table 1 presents the mean probability of out-migration and the mean probability of migrating to a specific destination for the three periods of observation. The diversity in migration patterns in these three periods serves to illustrate how well each model does in different migration regimes.

---Table 1 about here --

The covariates used in the analyses were chosen according to the neoclassical, social capital, and historical-structural theories of migration and to the empirical literature on the determinants of internal migration in developing countries. One of the most common assumptions in the literature of internal migration is that migrants go from areas with low wages and high unemployment to areas with high wages and low unemployment {Todaro, 1980 #483}. In this case I use the unemployment rate, the proportion of the labor force that is self-employed, and the proportion of the labor force that are unpaid family workers to measure the employment conditions in each state. I use the proportion of the labor force that earns more than twice the minimum wage to measure the wage conditions in each state. All these indicators come from Population Censuses and are lagged ten years to reflect the conditions in the state before the period of observation.

I also include GDP growth to control for other non-observed economic conditions. The models for migration in 1975-1980 include GDP growth in each state between 1970 and 1975, while the models for migration in 1985-1990 include GDP growth between 1980 and 1985, and the models for migration in 1995-2000 include GDP between 1990 and 1993.

The probability of migration decreases with the distance that separates the state of origin from the state of destination {Lucas, 1997 #292} {Greenwood, 1985 #81}. Distance in this case is measured with two variables. The first variable is the shortest distance in kilometers through a modern, paved, road between the capital cities of the two states. The second is a categorical variable that takes the value of 1 when the two states are contiguous and the value of 0 otherwise. Both of these indicators come from Road Atlas for 1975, 1985 and 1995. {Arbiganst, 1975 #436; McNally, 1984 #435; Guia Roji, 1995 #380}. The distance of any state to itself is considered to be cero, and its indicator of contiguity is considered to be 1.

Independently of the labor market conditions and the distance separating them from different destinations, individuals are more likely to migrate to places they know more about. Some factors that have been found to increase the knowledge of a particular destination are the presence of individuals who have migrated there before and infrastructure, trade and administrative links between the state of origin and the destination.

An extensive literature demonstrates that migration is cumulative and that the probability of migrating from one place to another increases with the number of individuals who have migrated before {Massey, 1994 #106; Massey, 1990 #86}. However, in the case of internal migration, the cumulative causation of migration is not linear (Rivero-Fuentes, 2003). To approach the effect of past migrants in the probability of migration, I include in the analyses the proportion of the population born in state *i* who was living in state *j* in 1970, 1980 and 1990. For the option of staying in the state of destination, the proportion of past migrants is equal to the proportion of individuals born in the state that had not migrated elsewhere.

To capture the non-linearity of cumulative causation, I divide this indicator in four variables. The first variable is the percentage of the population born in *i* and living in *j*, when the percentage is lower than 0.5%. This variable takes the value of 0 for all the cases when more than 0.5% of the population born in *i* lives in *j*. The second variable is the percentage of the population born in *i* and living in *j* when the percentage is between 0.5% and 1%, and 0 otherwise. The third variable is the percentage of the population born in *i* and living in *j* when the percentage is between 1% and 5%, and 0 otherwise. Finally, the fourth variable is the percentage of the population born in *i* and living in *j* when the percentage is larger than 5%, and 0 otherwise. The data for these indicators comes from Population Censuses.

I include two indicators of infrastructure, trade and administrative links between the state of origin and the state of destination. The first indicator refers to ties during the XIX and early XX century because the trade and administrative linkages between states in this era determined the future shape of internal markets and the construction of highways and railroads {Coatsworth, 1984 #333; Moreno Toscano, 1972 #306}. I control for the importance of these ties with a categorical variable that takes the value of 1 if any two states were connected by a main road or railroad anytime before 1910; or if they belonged to the same economic and productive region during the XIX and early XX century[5]. States are considered to have a historical tie with themselves, and this indicator takes the value of 1. The data for this indicator comes from Mexican archived historical maps {Ortíz Hernán, 1994 #338}{Florescano, 1983 #295}{Coatsworth, 1984 #333}; and from a review of Economic History Literature in Mexico {Coello Salazar, 1965 #301}{Duhau, 1988 #302}.

The second indicator refers to current trade ties. This indicator approaches the level of merchandise exchange between two states by measuring the intensity of cargo traffic between them. The indicator of recent trade between two states *i* and *j* for the observations in 1975-1980 takes the value of 1 if the total number of cargo vehicles whose trip started in *i* and ended in *j* (or *vice versa*) between 1962 and 1968 represent more than 4%[6] of all cargo vehicles that started or ended their trip in *i*. For the observations in 1975-1980 and 1985-1990 the indicator is cumulative. The indicator of recent trade between *i* and *j* for 1985-1990 takes the value of 1 if the total number of cargo vehicles whose trip started in *i* and ended in *j* (or *vice versa*) between 1977 and 1985 represent more than 4%[7] of all cargo vehicles that started or ended their trip in *i*, or if the indicator of trade for the period 1975-1980 is 1. The indicator of recent trade between *i* and *j* for 1995-2000 takes the value of 1 if the total number of cargo vehicles whose trip started

[5] I use Moreno Toscano's definition of economic and productive regions during the XIX century {Moreno Toscano, 1998 #305} and Duhau and Coello Salazar's definition of economic and productive regions during the profiriato {Coello Salazar, 1965 #301} {Duhau, 1988 #302}.
[6] 5% was chosen as the cutting point to denote an exchange between *i* and *j* as important, because it is above 3%, which would be the percentage expected if all exchanges were random.
[7] 5% was chosen as the cutting point to denote an exchange between *i* and *j* as important, because it is above 3%, which would be the percentage expected if all exchanges were random.

in i and ended in j (or *vice versa*) between 1985 and 1994 represent more than 4%[8] of all cargo vehicles that started or ended their trip in i, or if the indicator of trade for the period 1985-1990 is 1. The data for this indicator comes from the Studies of Origin-Destination of Cargo Traffic of the Mexican Ministry of Transport {Secretaria de Comunicaciones y Transportes, 2002 #337}.

Finally, population size is included as a control for geographical size. More populated states might be larger and more able to capture migrants than less populated states.

Table 2 presents the mean and standard deviation of these characteristics for all states in Mexico during each of the three periods of observation.

-- Table 2 about here --

Statistical Models Used in this Paper

In the hierarchical logit models fitted in this paper, the probability of out-migrating from any state ($P(M_i)$) is determined by its unemployment rate, the proportion of the labor force that is self-employed (SE_i) or a family worker (FW_i), the proportion of the labor force earning less than twice the minimum wage (W_i), GDP growth (GDP_i), and its population size (P_i). Among those who migrate, the probability of going to a specific state ($P(m_{ij})$) is determined by the population size in the state of destination (P_j), the unemployment rate in j (UR_j), the proportion of the labor force in j that is self-employed (SE_j) or a family worker (FW_j), wages in j (W_j), GDP growth in j (GDP_j), the population size in j (P_j), distance and contiguity between i and j (D_{ij} and C_{ij}, respecitvely), the proportion of individuals born in i who live in j ($PM0_{ij}$, $PM5_{ij}$, $PM10_{ij}$, and $PM50_{ij}$), the presence of historical ties between i and j (H_{ij}) and by recent trade between i and j

[8] 5% was chosen as the cutting point to denote an exchange between i and j as important, because it is above 3%, which would be the percentage expected if all exchanges were random.

(T_{ij}). In consequence, in the hierarchical logit model the predicted overall probability of moving from i to j ($P(M_{ij})$) is calculated as follows:

$$P(M_i) = \frac{\exp(\alpha + \beta_1 UR_i + \beta_2 SE_i + \beta_3 FW_i + \beta_4 W_i + \beta_5 GDP_i + \beta_6 P_i)}{1 + \exp(\alpha + \beta_1 UR_i + +\beta_2 SE_i + \beta_3 FW_i + \beta_4 W_i + \beta_5 GDP_i + \beta_6 P_i)},$$

$$P(m_{ij} / j \neq i) = \frac{\exp(\beta_7 UR_j + \beta_8 SE_j + \beta_9 FW_j + \beta_{10} W_j + \beta_{11} GDP_j + \beta_{12} P_j + \beta_{13} D_{ij} + \beta_{14} C_{ij} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ij} + \beta_{17} PM10_{ij} + \beta_{18} PM50_{ij} + \beta_{19} H_{ij} +}{\sum_{\substack{k=1 \\ k \neq i}}^{J} \exp(\beta_7 UR_k + \beta_8 SE_k + \beta_9 FW_k + \beta_{10} W_k + \beta_{11} GDP_k + \beta_{12} P_k + \beta_{13} D_{ik} + \beta_{14} C_{ik} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ik} + \beta_{17} PM10_{ik} + \beta_{18} PM50_{ik} + \beta_{19}}$$

and $P(M_{ij}) = P(M_i) * P(m_{ij} / j \neq i)$

Similarly, the conditional logit model the overall probability of moving from i to j is given by:

$$P(Mij) = \frac{\exp(\beta_7 UR_j + \beta_8 SE_j + \beta_9 FW_j + \beta_{10} W_j + \beta_{11} GDP_j + \beta_{12} P_j + \beta_{13} D_{ij} + \beta_{14} C_{ij} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ij} + \beta_{17} PM10_{ij} + \beta_{18} PM50_{ij} + \beta_{19} H_{ij} + \beta}{\sum_{\substack{k=1 \\ k \neq i}}^{J} \exp(\beta_7 UR_k + \beta_8 SE_k + \beta_9 FW_k + \beta_{10} W_k + \beta_{11} GDP_k + \beta_{12} P_k + \beta_{13} D_{ik} + \beta_{14} C_{ik} + \beta_{15} PM0_{ik} + \beta_{16} PM5_{ik} + \beta_{17} PM10_{ik} + \beta_{18} PM50_{ik} + \beta_{19} H_i}$$

where NM_{ij} is an indicator of non-migration for the cases when $i=j$.

In the nested logit model the probability of migrating from i to j among those individuals who outmigrate ($P(m_{ij}$), the probability of outmigration ($P(M_i)$), and the overall probability of migrating from i to j ($P(Mij)$) are given by:

$$P(m_{ij} / i \neq j) = \frac{\exp(\beta_7 UR_j + \beta_8 SE_j + \beta_9 FW_j + \beta_{10} W_j + \beta_{11} GDP_j + \beta_{12} P_j + \beta_{13} D_{ij} + \beta_{14} C_{ij} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ij} + \beta_{17} PM10_{ij} + \beta_{18} PM50_{ij} + \beta_{19} H}{\sum_{\substack{k=1 \\ k \neq i}}^{J} \exp(\beta_7 UR_k + \beta_8 SE_k + \beta_9 FW_k + \beta_{10} W_k + \beta_{11} GDP_k + \beta_{12} P_k + \beta_{13} D_{ik} + \beta_{14} C_{ik} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ij} + \beta_{17} PM10_{ik} + \beta_{18} PM50_{ik} + \beta}$$

$$P(Mi) = \frac{\exp(\alpha + \beta_1 UR_i + \beta_2 SE_i + \beta_3 FW_i + \beta_4 W_i + \beta_5 GDP_i + \beta_6 P_i + \beta_{22} IV_i)}{1 + \exp(\alpha + \beta_1 UR_i + \beta_2 SE_i + \beta_3 FW_i + \beta_4 W_i + \beta_5 GDP_i + \beta_6 P + \beta_{22} IV)},$$

where $IV_i = \sum_{\substack{k=1 \\ k \neq i}}^{J} \exp(\beta_7 UR_k + \beta_8 SE_k + \beta_9 FW_k + \beta_{10} W_k + \beta_{11} GDP_k + \beta_{12} P_k + \beta_{13} D_{ik} + \beta_{14} C_{ik} + \beta_{15} PM0_{ij} + \beta_{16} PM5_{ij} + \beta_{17} PM10_{ik} + \beta_{18} PM50_{ik} + \beta_{19} H$

and $P(M_{ij}) = P(M_i) * P(m_{ij} / i \neq j)$

Results

The results of the three statistical models are compared in terms of their coefficients and their implications for the determinants for migration, their goodness of fit to the 1024 M_{ij} flows[9], their accuracy in predicting the overall probability of out-migration from each state ($P(M_i)$), and their accuracy in predicting the total number of individuals migrating to each state ($M_{.i}$).

Comparison of results in terms of β coefficients and goodness of fit

When working with grouped data coming from large samples, the deviance statistic can be an inadequate measure to compare goodness of fit across models because it is usually very large and responsive to trivial improvements in the model. An alternative measure of goodness of fit commonly used is the Index of Dissimilarity (Δ) between observed and predicted frequencies[10]. This is a descriptive measure can be interpreted as the smallest fraction of the population that would need to be reclassified in order to make the model fit adequately {Powers, 2002 #504, p. 106}{Firth, 1999 #506}. The Index of Dissimilarity varies between 0 and 1. Models that fit perfectly have a Δ of cero, and models that predict all the frequencies erroneously have a Δ of one.

Table 3 presents the β coefficients of the three models, the deviance[11] and the Index of Dissimilarity for the three periods of observation. The β coefficients presented refer to the variables standardized around their mean.

[9] This includes the migrants from the 32 states in Mexico to each other state, as well as the number of individuals who did not out-migrate.

[10] For a contingency table with K cells, $\Delta = \frac{\sum_{i=1}^{K} \left| Y_i - \hat{Y}_i \right|}{2N}$

[11] The hierarchical model is nested in the nested logit model because the two models are equal when β_{22} -the coefficient for the inclusive value (IV)- is zero. Consequently, the deviance of the two models can be formally compared to determine which of the two models fits the data better, and whether IV contributes or not to explaining migration. The conditional logit model is not nested in either of the two other models, and its deviance cannot be formally compared with the deviance of the nested or hierarchical models to test which variables should be included

-- Table 3 about here --

Considering the degrees of freedom left, the deviance of the three models is very large and indicative of the need to consider other explicative variables. This result, however, is common in the literature of internal migration where empirical studies generally report deviances in the millions because of over-dispersion and the extremely large samples needed to capture the phenomenon ({Congdon, 1993 #419; Congdon, 1992 #426}).

The index of dissimilarity, nevertheless, indicates that the three models do a good job at predicting the distribution of migrants and non-migrants across destinations. The three models fit the migration during 1985-1990 and 1995-2000 better than they fit migration during 1985-1990. During this period (1975-1980) the index of dissimilarity of the three models indicates that only around 3% of the population is predicted to be in an incorrect destination[12].

Of the three models explored, the conditional logit model is the least appropriate to explain the determinants and directionality of internal migration in Mexico. This model shows the worse fit to the data during the three periods of observation, as indicated both by the deviance and the index of dissimilarity. The nested logit model is the model that fits the data better during the three periods. But there is not much difference between the fit of the hierarchical and the nested logit model. The deviance of the conditional logit model is 12% larger than the deviance of the nested logit model for migration during 1975-1990, 65% larger for migration during 1985-1990, and 45% larger for migration during 1995-2000. In comparison, the deviance of the hierarchical logit model is less than 1% larger than the deviance of the nested logit model for migration during 1975-1980, and 1985-1990, and 4% larger for migration during 1995-2000.

---

in the model. However, an informal comparison of the magnitude of the deviance of the conditional model with the other two deviance indicates whether or not the conditional model fits the data as well as the other models.

[12] Similar studies of internal migration in other countries report indices of dissimilarity of above 10% (see for example {Lin, 1998 #396}).

The β coefficients of all the models are significant at the 1% for the three periods. And, with the exception of the coefficients for unemployment and self-employment of the nested logit model during 1995-2000, the signs of the coefficients of nested and hierarchical logit models coincide with the predictions of migration theories.

The coefficients of the conditional logit model during 1975-1980 also behave as predicted by migration theories. However, during 1985-1990 and 1995-2000 the signs of many coefficients of the conditional logit are contrary to theory predictions and would lead to erroneous conclusions. For example, the results of hierarchical and nested logit models confirm that the number of migrants moving to any state is negatively associated to its level of unemployment and distance ($\beta_8<0$ and $\beta_{13}<0$ for these models during the three periods), and positively associated to the number of past migrants in the state (β_{15} and $\beta_{16}>0$). However, the results of the conditional logit model in 1985-1990 would lead to the conclusion that the level of migration is positively associated to distance (β_{13}=0.068), and negatively associated to past migration *(*β_{15}=-0.092 and β_{16}=-0.047).

The only difference between hierarchical and conditional logit models is that the nested logit model includes an inclusive value and the hierarchical logit model does not. The coefficients for the inclusive value *(*β_{22}) are significant with a p>0.01 in the three periods, and imply that the population considers the conditions in other states before deciding to move out of their state of residence.

Comparison of results in terms of predicted probability of out-migration

Figure 1 compares how much the three models err in predicting the overall probability of out-migration. The error is expressed as a percentage of the observed probability. The errors of nested and hierarchical models are indistinguishable, except for two or three points in each period where the hierarchical model yields slightly larger errors.

-- Figure 1 about here --

Figure 1 evidences the difficulty of studying the determinants of out-migration at the aggregate level. The three models fail at predicting the probability of out-migration for most states, and the error in the prediction is commonly a misestimation of more than 50%.

Still, the hierarchical and nested logit models are more adequate than the hierarchical logit model for studying out-migration, especially in instances when the probability of out-migration is very low in all the states.

In 1975-1980, the period with the highest out-migration, the conditional logit model is not much worse than hierarchical and nested logit models, and it even predicts the highest and lowest probabilities of out-migration more accurately than nested and hierarchical logit models. However, in 1985-1990 and 1995-2000, when the probability of out-migration was lower and less diverse, the conditional logit model is less efficient than hierarchical and nested logit models. In two instances during these periods the conditonal logit model overestimates out-migration by more than 200%.

Comparison of results in terms of predicted number of immigrants to each state

Figure 2 shows the errors in the predicted number of immigrants for each of the three models, in each of the three periods of observation. The errors are expressed as a percentage of the observed number of immigrants to each state and graphed against the size of the flow.[13]

-- Figure 2 about here --

[13] In the graph I use a logarithmic scale in the x-axis to diminish the concentration of points in the very small flows.

The three models of migration explored in this paper fail to predict the number of migrants that each state receives. In general, the error is larger for smaller flows, if one considers that the three models overestimate many of the smaller flows by more than 50%.

The results of the hierarchical and the nested logit models are almost identical. And, as is the case when predicting the probability of out-migration, the conditional logit model is the model that gives the worst results. The errors of this model are commonly larger than the errors of nested and hierarchical models, and its predictions worsen as the overall level of migration decreases in the country. The difference between the conditional model and the nested and hierarchical models was smaller in 1975-1980 than in 1985-1990 and 1995-2000.

Conclusions

Being able to predict how many persons will change their state of residence and where will they go is a very important task for policy makers in many developing countries. In Latin-American countries, China, and South-East Asian countries, for example, discrepancies in population growth from one region to another are mainly explained by internal migration {Lucas, 1997 #292} {Poston, 1998 #494}.

The task of predicting the number of out-migrants and immigrants in any given state is particularly complicated because individuals from different states have different propensities to move out, and they also tend to move to different destinations. It seems that in the last 15years researchers and policy-makers have become more aware of the importance of considering this complexity when studying migration. A number of articles in Demographic and Sociological journals show that researchers are starting to distinguish between different destinations when studying the determinants of migration ({Davies, 2001 #427; De Jong, 1999 #399; Greenwood, 2003 #405; Grusky, 1998 #397} {Frey, 1996 #411}) and when predicting migration (Rogers). All of these studies are at the aggregate level, but their methodological approach is not uniform. Authors use indiscriminately one of three different statistical models: conditional logit models, hierarchical logit models and nested logit models. However, no study before has compared the results of these three approaches.

In this paper I apply each of these three models to state-to-state migration in Mexico during 1975-1980, 1985-1990, and 1995-2000, and compare their results in terms of what they conclude about the determinants of migration, of how well they predict destination-specific flows, the probability of out-migration for each state, and the number of immigrants to be received by each state. My analyses are, like those of other studies concerned with modeling destination-specific flows, at the aggregate level. My choice of explanatory variables was driven by economic and political economy theories of migration, and includes income and employment conditions in the states of origin and of destination, and distance, past migration, and past and current trade between origin and destination. The results of this paper are important to those interested in modeling and predicting migration flows, to those interested in understanding how migration decisions are made, and to those interested in understanding the determinants of migration.

The three models of migration I used are inadequate for predicting the probability of out-migration and the number of migrants to be expected by any state. For many states the error in the predictions of both out-migrants and immigrants surpasses 50%. Furthermore, the lack of fit of the models increases as the level of migration in the country decreases. The three models are relatively better at predicting migration in 1975-1980, the period with the highest probabilities of migration, than in 1985-1990 and 1995-2000.

Of the three models considered, the nested logit model is the one that has a better statistical fit, while the conditional logit model is the one with the worst fit. Furthermore, the conditional logit model can lead to erroneous conclusions about the factors affecting migration. The magnitude and direction of the effects predicted by this model differs in several instances from both the theory and the conclusions of the nested and the hierarchical model. For example, the conditional logit model would lead to the conclusion that past migration between two states decreases rather than increases the probability of migration.

The difference between nested and hierarchical logit models has more implications for theory and the understanding of how migration decisions are made than for the prediction of migration flows. The nested logit model shows that the probability of out-migration depends on the characteristics of all the potential destinations as well as on the characteristics of the place of origin. However, the hierarchical logit model predicts almost the same number of out-migrants and immigrants than the hierarchical logit model, even when it models the probability of out-migration based solely on the conditions in the state of origin. In addition, the two models lead to the same conclusions about the factors affecting the level of migration.

One possible explanation for the consistent lack of fit of the three models explored in this paper is the aggregate nature of the analysis. The literature on both internal and international migration has shown that the decision to migrate depends largely on individual and family characteristics like sex, family migrant networks, life stage, and education {Portes, 1999 #166}{Massey, 1999 #229}{Lucas, 1997 #292}.

Few studies to date have included individual and community or state level characteristics when studying migration to alternative destinations. The only cases that I know of are the articles by {Davis, 2002 #430} and Curran and Rivero (2003), and these are limited to only two migration flows (internal migration in Mexico and migration from Mexico to the U.S.). These studies show that individual and family characteristics affect the probability of migration differently, depending on the destination. It is pertinent to think that the same happens when individuals migrate within their country, and that we can greatly improve both our understanding and our predictions of interstate migration if we consider individual characteristics in future migration models.