

Inter-State Migration in Mexico, 1995-2000: A Flexible Poisson
Gravity-Modeling Approach

Non Peer-reviewed author version

PEETERS, Ludo (2008) Inter-State Migration in Mexico, 1995-2000: A Flexible
Poisson Gravity-Modeling Approach.

Handle: <http://hdl.handle.net/1942/7973>

Inter-State Migration in Mexico, 1995-2000: A Flexible Poisson Gravity-Modeling Approach

Ludo Peeters

Abstract

In this paper, I present an empirical study of aggregate inter-state migration in Mexico in the period 1995-2000. The study contributes to the empirical literature on migration in a number of ways. Firstly, I show that the gravity model in levels (Poisson model) provides a more realistic description of the migration process than a model in logs (log-linear model). Secondly, I develop a more flexible model by relaxing the standard assumption of spatial invariance (constancy) of the distance-decay parameter. This added flexibility should allow one to account for the effects of 'idiosyncratic' ties between origin and destination states (imperfect information). Thirdly, the model attempts to bring together elements from both the neo-classical (economic opportunities) and the cumulative-causation (migrant networks) perspectives on migration. Finally, the model is estimated by using the 'non-classical' generalized maximum-entropy method to cope with over-parameterization in a cross-sectional setting. An interesting finding is that inter-state migration in Mexico is generally inelastic with respect to geographical distance measure after correcting for the effects of some structural (historical) components of migration, such as migrant stocks.

JEL: R23; C14

Key words: Internal migration; Poisson gravity model; Spatial structure; Imperfect information; Migrant stocks; Entropy estimation

1. Introduction

Internal migration is one of the most important factors determining the spatial distribution of the population in many Latin American and other developing countries. Despite its importance, our understanding of internal migration in these countries is still rather limited. This paper attempts to contribute to the literature on internal migration, taking the case of inter-state migration in Mexico in the period 1995-2000.¹

Most prior studies of aggregate internal and international migration consider only (observable) ‘push’ and ‘pull’ characteristics in origin and destinations, and assume that the effects of (unobservable) migration costs, measured by geographic distance, are constant across all origin-destination pairs. Only recently researchers have started realizing that the presence (or absence) of economic opportunities in origin and destination states (neo-classical theory) is not enough to explain the directionality of migration flows (i.e., what causes migration from certain states/countries to be directed to some particular states/countries, and not to others?).

In a recent (unpublished) paper, [Rivero-Fuentes \(2005\)](#) highlighted the important role of ‘idiosyncratic’ (historical, socio-cultural, economic, administrative, etc.) ties between origin and destination states in explaining inter-state migration in Mexico, emphasizing the bearing of the cumulative-causation perspective for studying migration patterns. Her argument is that such ties are related to migrants ‘awareness space’; that is, to their knowledge about potential destinations. In fact, the neo-classical theory of migration assumes that individuals have complete information about all potential alternative destinations they can migrate to, and that prospective migrants know the (inherently unobservable) ‘distance’ that separates them from all possible destinations. However, individuals might have imperfect information about conditions in all possible destinations, and they might consider only *some* states as alternative destinations. Thus, when making their migratory decisions, individuals choose among those places they have some information about, even though other possible destinations might actually offer better economic opportunities (giving rise to seemingly ‘anomalous’ behavior from a strictly neo-classical view).

This idea is not new, though, but has generally been ignored in the empirical migration literature. In an early paper, [Mueser \(1989\)](#) already hinted to the importance of such ties, or relative attachments, between regions (which may be fairly stable over a considerable period of time), and he showed that the impact of geographical distance on migration flows varies across

¹ Earlier studies of aggregate inter-state migration in Mexico are rather scarce. I found only a few studies in the late 1970s and early 1980s [[Greenwood and Ladman \(1978\)](#); [King \(1978\)](#); [Cole and Sanders \(1983\)](#)] and some more recent ones [[Aroca and Maloney \(2005\)](#); [Lara and Soloaga \(2005a, 2005b, unpublished\)](#)].

origin-destination pairs. Thus, not ‘controlling’ for such spatial variance in distance-decay effects may mask meaningful heterogeneity in the migration data, and may be an important source of estimation biases (misleading inferences).²

In this paper, I examine different specifications of a gravity model of aggregate inter-state migration in Mexico, and discuss their empirical implications. In doing so, I (hope to) contribute to the migration literature in the following ways. Firstly, I show that the gravity model in levels (Poisson specification) provides a more realistic description of the migration process than a model in logs (log-linear specification) [Flowerdew and Aitkin (1982); Shen (1999); see also Santos Silva and Tenreyro (2006)]. Secondly, I develop a more general and flexible Poisson model that departs from the standard assumption of spatial invariance of distance decay, by allowing for origin-destination-specific distance-decay elasticities [Gordon (1985); Mueser (1989)]. Thirdly, I look at the migration process not only from a neo-classical perspective (economic ‘push’ and ‘pull’ factors in origin and destination states), but also from a cumulative-causation perspective (e.g., migrant stocks) [Rivero-Fuentes (2005)]. Finally, the flexible Poisson model is estimated by using the ‘non-classical’ (semi-parametric) method of generalized maximum entropy [Golan et al. (1996)], which allows me to deal with ‘over-parameterized’ models in a cross-sectional setting.

The remainder of this paper is organized as follows. In Section 2, I provide a summary of some major characteristics of internal migration in Mexico. In Section 3, I give a brief overview of the explanatory variables included in the analysis, and the corresponding hypotheses I want to test empirically. In Section 4, I present three alternative gravity-model specifications, along with the methods used to estimate the models. In Section 5, I report the empirical findings and the results of some tests of model adequacy. In Section 6, I provide a summary and formulate some conclusions.

2. Inter-state migration in Mexico, 1995-2000

The migration data used in this paper refer to migrations among the 32 states in Mexico (for a map of the Mexican states, see Appendix 2) in the period 1995-2000. The data are obtained from the 2000 Population Census and the 1995 Population Count. Although the migration of Mexicans to the U.S. is a particularly important issue, this phenomenon as such will be excluded from the analysis in this paper.

² Also in the context of international trade some recent papers have focused on the variability of distance-decay parameters in gravity models [e.g., Fratianni and Kang (2006); Henderson and Millimet (2008)].

Migration is measured as the *gross* aggregate migration flow from origin state i to destination state j ($MIGR_{ij}$), simply defined as the number of people migrating from state i to state j . There are a total of 992 potential migration flows, with no zero-cells in the flow matrix (i.e., the actual number of migration flows is equal to the potential number). Over the study period, some 3.5 million Mexicans changed their state of residence, which is about 4.0% of the total population.

For illustrative purposes, Figure 1 shows the net in-migration *rates* ($NIMR_j$). In relative terms, the three most important net ‘gainers’ are Quintana Roo (+12.5%), Baja California (+7.8%), and Baja California Sur (+6.5%), while the three most important net ‘losers’ are Distrito Federal (-4.8%), Veracruz (-3.3%), and Guerrero (-3.0%).

Insert Figure 1 about here

In contrast with the past, Distrito Federal no longer takes a prominent position as an (net) ‘attraction pool’. Over the period 1995-2000, about 780,000 people moved out from Distrito Federal, mostly into its ‘neighboring’ states, of which 57.5% to México, and another 14.1% to Puebla, Veracruz, Hidalgo, and Morelos. Conversely, about 376,000 people still moved into Distrito Federal, mostly also from its ‘neighboring’ states, of which 46.2% are coming from México, and another 30.1% from Veracruz, Puebla, Oaxaca, Hidalgo, and Guerrero. In other words, the most important migration flows in Mexico are related to (two-way) ‘short-distance’ moves over a small number of states in the ‘center’ of the country (out-flows from and in-flows into Distrito Federal represent 20.5% of the total number of out-movements and 10.5% of the total number of in-movements in the country, respectively).

Insert Figure 2 about here

The size distribution of the migration flows is heavily (positively) skewed, as shown in Figure 2. The smallest migration flow (15) goes from Aguascalientes to Campeche, while the largest migration flow (448,548) goes from Distrito Federal to the neighboring state of México. The heavily skewed form of the size distribution poses some challenging estimation problems, as I will discuss shortly.

3. Explanatory variables

The gravity model of internal migration is calibrated with both origin- and destination-state-specific attributes and other aspects related to their relative locations as explanatory variables. The forces affecting the gross migration flow from origin state i to destination state j are broken down into ‘pull’ and ‘push’ forces, mostly derived from the neo-classical theory of migration. These forces are further sub-divided [see, for example, King (1978)].

Pull or attraction forces may be either retention or attraction forces, which are usually understood as follows: If retention forces are operating in a given region i , then an increase in the value of any of the independent variables associated with these forces is seen as increasing the attractiveness or benefits associated with this region, thus decreasing migration flows M_{ij} , *ceteris paribus*. On the other hand, if attraction forces are operating in a given region j , much the same can be said, only in this instance migration flows M_{ij} tend to increase, *ceteris paribus*. On the other hand, push or emission forces may be either expulsion or repulsion forces. If expulsion forces are operating in a given region i , an increase in the value of any of the independent variables associated with these forces decreases the desirability of residence in this region, thus inducing a rise in migration flows M_{ij} , *ceteris paribus*. On the other hand, if repulsion forces are operating in a given region j , then migration flows M_{ij} tend to fall, *ceteris paribus*.

I briefly discuss the measurement and theoretical role of the explanatory variables included in the gravity models. All variables are measured in the base year 1995, thus at the *beginning* of the study period, to avoid any problems of endogeneity (simultaneity). It should be mentioned, though, that by no means I intend to develop a ‘complete’ or ‘comprehensive model’ of inter-state migration in Mexico, even if this were possible. I only seek to explore whether or not some widely-accepted notions about migration are relevant to understanding the Mexican case.³

Distance ($DIST_{ij}$) – The distance between each origin-destination pair of states is measured as the kilometer distance between the capitals of the states involved. Distance enters in the gravity model as a measure of the (unobservable) costs of moving, including the direct economic costs,

³ The particular choice of explanatory variables was also partly motivated by limited data availability. For example, the data set I disposed of does not include information about costs of living and nominal wages across Mexican states. Neither was I able to incorporate trade/FDI and *maquila* value-added variables in the proposed model, which, hence, precludes a comparison of my results with those found by Aroca and Maloney (2005). Nevertheless, the model includes at least a basic set of variables [see also Davies et al. (2001)], which should allow me to present to potentials of the modeling and estimation approaches adopted in this paper.

the indirect psychic costs, and the costs incurred for obtaining information about the contemplated destination state [Greenwood (1997)]. Distance is viewed as a ‘resistance’ or ‘decay’ factor, where it is hypothesized that the greater the distance involved, the higher the costs of moving, and the more prospective migrants will be deterred from moving. As a result, the migration flow from i to j is lower if the distance between i to j is larger, *ceteris paribus*. Several authors [e.g., Yano et al. (2003)], however, have pointed out that distance deterrence is more likely related to *information* decay, given that the out-of-pocket money costs of migrating are very small in relation to the expected gains to the migrant over the lifetime of the move. This means that the distance variable is not necessarily a measure of ‘distance’ *per se*. For example, two contemplated destination states at equal distance from the origin state do not necessarily imply equal decay from the prospective migrant’s point of view, even after holding all other things constant. Therefore, the assumed constancy of the distance-decay parameter in migration models may be overly restrictive.

Unemployment ($UNEMPL_i$; $UNEMPL_j$) – The unemployment rate in a state is measured as the population of a state, 12 years old and over, who are unemployed prior to the 1995 census, divided by the economically active population of the state. Due to data limitations (1995 values are not contained in the data set), I use the average of the 1990 and 2000 values for this variable. The hypothesis is that a lower value for this variable means a higher probability of finding a job (i.e., better economic opportunities), so that the coefficients of the destination $UNEMPL_j$ and origin $UNEMPL_i$ are expected to be negative and positive, respectively. However, the origin $UNEMPL_i$ may also measure the probability of not being able to finance migration: unemployed persons may be very desirous of moving to a state where more jobs are available, but given that these persons are unemployed their abilities to finance the move might be impaired. Consequently, as the value of this variable rises, forces of retention may also be at work in the origin state, so that the sign of the coefficient associated with $UNEMPL_i$ is indeterminate *a priori*.

Income per capita (GDP_PC_i ; GDP_PC_j) – The per-capita income in a state is measured as gross domestic product (GDP) per capita. It is used as a proxy for the expected potential economic gains or losses from migration (for the average real income/wage stream), and for the ability to finance migration. The hypothesis is a rise in the value of GDP_PC_i generates retention and expulsion forces, on the one hand operating to decrease the probability of out-migration and on the other to increase it. Thus, there is no *a priori* sign expectation. However, the variable GDP_PC_i does not affect the ability to finance migration directly; it is a proxy for the expected

real income stream in i . But if a person borrows money in i to migrate to j , then in this sense GDP_PC_i affects the ability to finance migration [see also King (1978)]. A rise in GDP_PC_j , on the other hand, increases the probability of out-migration from i to j since both the probable return to labor and the probable ability to repay a loan improve, generating attraction forces. One can thus expect a positive sign to be attached to the coefficient of this variable.⁴

Manufacturing ($MANU_i$; $MANU_j$) – Employment in the manufacturing sector is a ‘natural’ indicator of a state’s level of industrialization, and is measured as the share of the manufacturing sector in a state’s total employment. The hypothesis is that more non-agricultural employment opportunities may attract migrants in search for formal employment or escaping from ‘tight’ economic conditions in rural areas (decline in rural job opportunities) [see, for example, Araujo (2004)]. Therefore, a positive (negative) sign is expected for the coefficient of both $MANU_j$ and $MANU_i$.

Population-centrality index ($POPCENT_i$; $POPCENT_j$) – The population-centrality index of a given state i is a measure of population concentration in a *cluster* of states, composed of the state i itself and its (primarily) neighboring states, and is defined as follows:

$$POPCENT_i = \sum_{k=1}^n \frac{POP_k}{DIST_{ik}}, \text{ for origin state } i$$

$$POPCENT_j = \sum_{k=1}^n \frac{POP_k}{DIST_{jk}}, \text{ for destination state } j$$

where POP_k is the population in all other origin states k and destination states k , respectively.⁵ This index can be considered as a ‘catch all’ for numerous factors. Obviously, it reflects the spatial distribution of population (urbanization). By implication, this index also (at least partly) reflects the spatial distribution of the (quality and quantity of the) supply of service-based private

⁴ Unfortunately, due to the lack of price-level data at the level of the Mexican states, it is impossible to know to what extent variations in nominal wages reflect variations in real wages.

⁵ Note that the sum on the right-hand side of the centrality includes also the states $k = j$ or i . The intra-state distance is calculated as $DIST_{kk} = ((\pi - 1)/\pi) \sqrt{s_k/\pi}$, where s_k is the surface area (in square kilometres) of state k [see also, for example, Rietveld and Bruinsma (1998); Sá et al. (2004)]. This formula assumes ‘circular’ or ‘disk-shaped’ states, along with a homogeneous distribution of population within each state. In fact, the formula calculates the average distance between two arbitrary points within a given state.

and public *amenities* (consumptive and productive functions, facilities, etc.), as these are found precisely in those places where population is most concentrated [see, for example, [Greenwood \(1997\)](#)].⁶ Larger (central) states offer more ‘connection points’ than smaller (remote) states, and thus might either attract more migrants and/or retain local residents or, conversely, expel more local residents and/or repel migrants. In the first case, there would be a tendency of ‘urbanization’, due to the fact that migrants are searching for the many benefits associated with (high-amenity) urban areas; in the second case, there would be a tendency of ‘counter-urbanization’ due to, for example, congestion effects. It can thus be hypothesized that both retention and expulsion forces may be at work in origin states, and both attraction and expulsion forces may operate in destination states, so that it is difficult to predict the signs of the coefficients.

Migrant stock ($MSTOCK_{ij}$) – The migrant stock is used as a proxy for the ‘intensity’ of migrant networks, and is measured as a simple count of the number of people in the destination state j who had migrated earlier from the home state i , divided by the population of the home state i [[Lucas \(1997\)](#)]. The importance of migrant networks has received relatively little attention in empirical studies of internal migration. Migrant networks are defined as social ties that bind former and future migrants; the more developed these networks, the higher the propensity toward migration. Such a variable is included in many other papers on migration, which may act as an ‘accessibility’ or ‘inverse-resistance’ factor. This variable can be regarded as a proxy for some of the direct and opportunity costs of migration that are implied by the prospect of living in a new state and searching for employment [[Greenwood \(1969\)](#)]. These may be lowered significantly if there are friends and relatives (social and family connections) from one’s state of origin that can provide material support, free or low-cost housing, and aid (information) in obtaining employment at the potential destination. It also may be a proxy for the psychic costs of moving and adjustment; these may be lower if friends and relatives are present [[King \(1978\)](#)]. Thus, one can expect a positive sign to be attached to the coefficient of this variable: larger migrant stocks make additional moves more likely. Also, it can be expected that adding the migrant-stock

⁶ Some authors [e.g., [Aroca and Maloney \(2005\)](#)] would rather use a composite measure based on a host of ‘amenity variables’, such as the percentage of the population living in urban areas, mortality rates, health infrastructure, education, and infrastructure, etc.

variable to the gravity model (to capture the spatial-structure effect) will have a strong moderating effect on the distance-deterrence parameter.⁷

Border states ($BORDER_i$; $BORDER_j$) – The migration to the U.S. may interact with the distribution of internal migration in Mexico, particularly with the internal migration to the Mexican border states. Specifically, migrants without any access to a network of relatives already living in the U.S. and moving to these border states may possibly be looking for jobs in *maquiladoras* or work as *jornaleros*, in preparation for out-migration to the U.S. (when they have accumulated enough money and information).⁸ Thus, the border-state effects may at least partly capture such interactions that may lead to this kind of internal transmigration through the northern border states into the U.S. (i.e., the ‘gateway’ function of the northern border states). Although Mexico has witnessed a growing de-centralization of *maquiladora* jobs in the 1980s and 1990s (dispersion from the Mexican border states into central Mexico), these jobs are still heavily concentrated along the U.S.-Mexican border. Using data from INEGI, Jones (2001) reported that, in 1998, the Mexican Border States (Baja California, Coahuila, Chihuahua, Nuevo León, Sonora, and Tamaulipas) were still hosting about 84% of the *maquiladora* jobs in Mexico (of which 64% in border *municipios*).⁹ Thus, the hypothesis is that these borders states attract migrants from other parts of Mexico (i.e., positive sign for the coefficient of the $BORDER_j$ dummy), and that at the same time retention forces related to internal migration are at work (i.e., negative sign for the coefficient of the $BORDER_i$ dummy).

⁷ It should be noted that the migrant-stock variable is (as the labeling of this variable evidently suggests) a *stock* variable, which, therefore, is not to be viewed as a proxy for lagged migration [see, for example, Greenwood (1997)], which is a *flow* variable.

⁸ Besides, the scarcely available evidence seems to indicate that having already made an internal migration increases the probability of emigration to the U.S. [see OECD (2004)]. An in-depth analysis of this issue is severely hampered (or even completely ruled out) by the fact no (or, at best, only fragmentary) official data are available on origin-specific out-migration flows to the U.S.

⁹ In this paper, Baja California Sur is also considered as a ‘Border State’, following Rogers et al. (2006), who sub-divided the country in four regions (Border, North-Central, Central, and South). These four regions were chosen to reflect different economic and historical zones. The Border region contains most of the nation’s formal-sector employment. The North Central region has an economy focused on manufacturing and export agriculture. The Central region, home to Mexico City, and formerly the most dynamic area in Mexico, is still today the core of finance and politics. The South region, the nation’s poorest, has an economy based on tourism and petroleum.

Table 1 lists the various independent variables included in the models, along with summary statistics and the expected direction (sign) of their marginal effects. A map of the spatial distribution of the independent variables is given in Figure 3.

Insert Table 1 about here

Insert Figure 3 about here

4. Alternative specifications of the gravity model

In its simplest form, the gravity model of migration is

$$M_{ij} = \beta_0 X_i^{\beta_1} X_j^{\beta_2} D_{ij}^{\beta_3} u_{ij} \quad (1)$$

where M_{ij} is the gross migration flow from state i to state j (with $M_{ij} \geq 0$), which is assumed to be proportional to push and pull factors in the two states, denoted by X_i and X_j , respectively, and inversely proportional to the inter-state distance D_{ij} . In Equation (1) it is commonly assumed that u_{ij} is a log-normal error, with $E(u_{ij} | X_i, X_j, D_{ij}) = 1$ and a constant variance, thus independent of the regressors (homoskedasticity), which leads to $E(M_{ij} | X_i, X_j, D_{ij}) = \beta_0 X_i^{\beta_1} X_j^{\beta_2} D_{ij}^{\beta_3}$.

In the migration literature, there is a long-standing tradition of log-linearizing Equation (1) and estimating the parameters of interest by OLS, using the equation

$$m_{ij} = \ln \beta_0 + \beta_1 x_i + \beta_2 x_j + \beta_3 d_{ij} + \ln u_{ij} \quad (2)$$

where the lower-case variables are the log forms of the corresponding upper-case variables in Equation (1), so that Equation (2) is a constant-elasticity model. The validity of this procedure depends critically on the assumption that the error terms u_{ij} , and, therefore, $\ln u_{ij}$, are statistically independent of the regressors.

However, in a recent paper [Santos Silva and Tenreyo \(2006\)](#) have convincingly shown that the standard practice of log-linearizing and estimating the parameters of gravity models by OLS is inappropriate for a number of reasons, and that estimation in *levels* is recommended. While

their paper is dealing with trade flows, their argumentation may equally well be applicable to migration flows. They provided evidence that the usual assumption of independence between determinants of trade flows and the error term in the log-linear gravity model (i.e., the required ‘orthogonality’ condition for a consistent estimator) may not hold due to the fact that unobservable determinants of trade flows in the gravity model expressed in levels are *heteroskedastic* (see the positive skewness of the size distribution of migration flows, as shown in Figure 2).

Given these considerations, I write Equation (1) as

$$M_{ij} = \exp(\ln \beta_0 + \beta_1 x_i + \beta_2 x_j + \beta_3 d_{ij}) + \varepsilon_{ij} = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta}) + \varepsilon_{ij} \quad (3)$$

where $E(M_{ij} | \mathbf{z}_{ij}) = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})$. Hence, the error term u_{ij} in Equation (1) can be re-written as $u_{ij} = 1 + \varepsilon_{ij} / \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})$, where the variance of u_{ij} is $V(u_{ij}) = V(\varepsilon_{ij}) \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^{-2}$. From this, it follows that the variance of u_{ij} is constant *only* if ε_{ij} can be written as $\varepsilon_{ij} = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta}) v_{ij}$, where v_{ij} is a random variable that is statistically independent of the explanatory variables. In other words, the assumption of constant variance (homoskedasticity) in Equation (1) is valid *only* if $V(\varepsilon_{ij}) \equiv V(M_{ij} | \mathbf{z}_{ij}) \propto E(M_{ij} | \mathbf{z}_{ij})^2$, which means that the conditional variance of M_{ij} (and, thus, ε_{ij}) is proportional to $\exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^2$. Since there is no *a priori* reason to assume that this very specific condition holds, they advocated estimating the gravity model in levels by using a Poisson pseudo-maximum likelihood estimator, while maintaining that, in practice, it is sufficient to assume that the variance is proportional to the (conditional) mean, $V(\varepsilon_{ij}) \propto E(M_{ij} | \mathbf{z}_{ij})$, in order to obtain consistent estimates.

I now move on to discuss three gravity-model specifications, which will be empirically estimated and tested in this paper.

4.1 Model 1: Log-linear model (logs)

The first specification is the usual log-linear model, which will be estimated by using Ordinary Least Squares (log-linear OLS):

$$m_{ij} = \alpha + \mathbf{x}'_i \boldsymbol{\beta}_1 + \mathbf{x}'_j \boldsymbol{\beta}_2 + \gamma s_{ij} + \delta d_{ij} + \theta_{i \in B} + \phi_{j \in B} + \ln u_{ij} \quad (4)$$

where m_{ij} is the (log) migration flow from origin state i to destination state j , \mathbf{x}_i and \mathbf{x}_j are vectors containing the selected (log) attributes of origin states i and destination states j (unemployment rates, incomes per capita, shares of manufacturing sector, and centrality indexes), respectively, d_{ij} is the (log) distance between origin state i to destination state j , s_{ij} is the (log) of migrant stocks, and $\theta_{i \in B}$, $\phi_{j \in B}$ are the coefficients of border-state dummies, as region of origin and region of destination, respectively.¹⁰

4.2 Model 2: Basic Poisson model (levels)

The second specification is a basic Poisson model:

$$M_{ij} = \exp(\alpha + \mathbf{x}'_i \boldsymbol{\beta}_1 + \mathbf{x}'_j \boldsymbol{\beta}_2 + \gamma s_{ij} + \delta d_{ij} + \theta_{i \in B} + \phi_{j \in B}) + \varepsilon_{ij} \quad (5)$$

where the variables are defined similarly as in Model 1.

This model will be estimated by using Poisson pseudo-maximum likelihood (PML). Specifically, under the assumption that $V(\varepsilon_{ij}) \propto E(M_{ij} | \mathbf{z}_{ij})$, the parameters of the model can be estimated by solving the following set of first-order conditions:

$$\sum_{i=1}^n \sum_{j=1}^n [M_{ij} - \exp(\mathbf{z}'_{ij} \tilde{\boldsymbol{\beta}})] \mathbf{z}_{ij} = 0 \quad (6)$$

Solving this set of first-order conditions implies that all observations are given the same weight to all observations, rather than emphasizing those for which $\exp(\mathbf{z}'_{ij} \tilde{\boldsymbol{\beta}})$ is large (as would be the case with NLS).

Even though the Poisson PML estimator does not take full account of the heteroskedasticity in the data (i.e., the data do not have to be Poisson at all), the Poisson PML estimator is consistent

¹⁰ The deterring effect of distance may decline at greater distances, because the marginal cost of moving one kilometer farther is lower at greater distances [e.g., Davies et al. (2001, p. 339)]. However, including a non-linear (quadratic) distance term in the model did not yield sensible results.

[see [Gourieroux et al. \(1984\)](#)]. Correct statistical inference, however, has to be based on an Eicker-White robust covariance estimator.¹¹

4.3 Model 3: Flexible Poisson model (levels)

The third specification is a more general and considerably more *flexible* version of the Poisson gravity model:

$$M_{ij} = \exp(\alpha + \mathbf{x}'_i \boldsymbol{\beta}_1 + \mathbf{x}'_j \boldsymbol{\beta}_2 + \gamma s_{ij} + \delta_{ij} d_{ij} + \theta_{i \in B} + \phi_{j \in B} + v_i + \omega_j + \varepsilon_{ij}) \quad (7)$$

where δ_{ij} are origin-destination-specific distance parameters (to capture the unobserved spatial structure of migration), v_i and ω_j are the coefficients of state-specific dummies (to capture unobserved state fixed effects on migration), respectively, and ε_{ij} is the ‘genuine’ error term.

Equation (7) can then be re-formulated as

$$M_{ij} = \exp(\alpha + \mathbf{x}'_i \boldsymbol{\beta}_1 + \mathbf{x}'_j \boldsymbol{\beta}_2 + \gamma s_{ij} + \delta_{ij} d_{ij} + \theta_{i \in B} + \phi_{j \in B}) \eta_{ij} \quad (8)$$

where $\eta_{ij} = \exp(v_i + \omega_j + \varepsilon_{ij})$ is a ‘composite’ error, which has mean 1 and variance $V(\eta_{ij}) = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^{-1}$. All three components of this error are supposed to *jointly* capture the ‘noise’ (i.e., unobserved heterogeneity) present in the system.

This Poisson model will be estimated by using the ‘non-classical’ method of generalized maximum entropy (GME). Whilst sidestepping the technical details of entropy econometrics, the GME formulation of the Poisson gravity model is outlined in Appendix 1 [for more details, see [Golan et al. \(1996\)](#); for another application in the context of migration, see [Peeters \(2008\)](#)]. Our principal motivation for applying GME is that this method allows me to deal with ‘over-parameterized’ models (note that the number of parameters is equal to $32 \times 2 + 992 = 1,056$, whereas the number of observations is 992). Estimating such a model by using ‘classical’ estimation methods is thus ruled out, whilst GME is known to be ‘immune’ to this dimensionality problem.

¹¹ Besides being consistent in the presence of heteroskedasticity, the Poisson PML method also provides a natural way to deal with (possible) zero values of the dependent variable.

State-fixed effects (v_i, ω_j) – The model in Equation (8) includes state-fixed effects to account for unobserved economic and non-economic factors that may play an important role in the migration decision [see, for example, also Davies et al. (2001)]. These factors may lead to expulsion or retention forces in origin states and attraction or repulsion forces in destination states, *regardless* of the potential destination or origin states, respectively. This approach is basically motivated by the fact that in any empirical model it is simply impossible to account for all the heterogeneity in the population by including a small number of explanatory variables. Note also the difference between the (common) *border-state* (*state-group*) dummies and the (specific) *individual-state* dummies; so, there is no problem of parameter redundancy.

Spatial variation in distance-decay effects (δ_{ij}) – The model in Equation (8) also includes pair-wise flow-specific distance-deterrence coefficients. As already discussed in the introductory section, I depart from the standard assumption of *spatial invariance* in distance-deterrence functions. These origin-destination-specific coefficients account for *spatial heterogeneity* in the (unobserved) migration-cost effects or the (relatively stable) idiosyncratic ‘migration attachments’ between origin and destination states, in the parlance of Mueser (1989). I further impose the restrictions that $\delta_{ij} \leq 0$ (non-positivity) and $\delta_{ij} \neq \delta_{ji}$ (asymmetry). Thus, the model allows me to estimate a *full matrix* of (992) distance-deterrence elasticities, dispensing with the need to impose any particular (arbitrary) functional form.¹² Furthermore, the *column* totals of the matrix of distance-deterrence elasticities, $\sum_i \delta_{ij} = \delta_{\bullet j}$, enable me to distinguish between destination states with a rather local or regional ‘recruitment’ area (i.e., $\delta_{\bullet j}$ is large in absolute value) and those with a wider or nationwide scope – perhaps due to some unique attractive characteristics (i.e., $\delta_{\bullet j}$ is small in absolute value), *ceteris paribus*. Conversely, the *row* totals, $\sum_j \delta_{ij} = \delta_{i \bullet}$, enable us to distinguish between origin states with a rather local or regional ‘sprawling’ area (i.e., $\delta_{i \bullet}$ is large in absolute value) and those with a wider or nation-wide orientation (i.e., $\delta_{i \bullet}$ is small in absolute value), *ceteris paribus*.

¹² In the migration literature very few migration studies allow for varying distance-decay parameters. While they are more common in competitive-destinations models, those studies do so only for either origin-specific or destination-specific coefficients, not for pair-wise origin-destination-specific coefficients..

5. Empirical results

The estimated coefficients from the three model specifications are presented in Table 2.¹³ I begin with discussing the results obtained from the basic Models 1 (logs) and 2 (levels), before moving to the results obtained from the flexible Model 3.

5.1 Basic Models 1 and 2

A first observation is that the estimates obtained from the basic Models 1 (log-linear) and 2 (Poisson) are strikingly different in magnitude and, in some instances, also in sign. As a result, the two models show a different picture of how state-specific push and pull factors are shaping inter-state migration flows, which clearly demonstrates the impact of model (mis)specification.

Moreover, looking at the models where the migrant-stock variable has been omitted (columns 1 and 3 of Table 2), it can be seen that distance-deterrence elasticity is negative and significantly different from zero (-1.132 and -0.985, respectively). On the other hand, including the *MSTOCK* variable produces two different pictures for the two model specifications (columns 2 and 4 of Table 2). For the log-linear model (Model 1), the distance elasticity is *positive* (0.136) and significant (at the 10% level), whereas for the basic Poisson model (Model 2), the distance elasticity remains negative (-0.316), but considerably smaller in magnitude and not significantly different from zero. Obviously, a positive sign for the estimated distance elasticity is not in conformity with theoretical expectations, questioning the appropriateness of the log-linear specification.¹⁴ The results clearly show that the migrant-stock variable has an important ‘inverse-resistance’ effect on internal migration flows.

Insert Table 2 about here

Also, all other coefficients do not seem to be robust, both in sign and magnitude, across the two model specifications. Only looking at the models including the *MSTOCK* variable, it can be seen that the signs are fairly robust, however, except for the distance-deterrence coefficient.

¹³ For the estimation of the Models 1 and 2, I used STATA, while for the estimation of Model 3, I used GAMS.

¹⁴ A positive sign of the distance-deterrence parameter of the log-linear model was also found by [Lara and Soloaga \(2005a, 2005b, unpublished\)](#).

The coefficients of the Models 1 and 2 (see columns 2 and 4 of Table 2) indicate that unemployment has an expulsion (repulsion) effect in the origin (destination) state, and that income per capita has a retention (attractive) effect in the origin (destination) state.¹⁵ In other words, unemployment (job opportunities) and income per capita (wage potentials) are found to be significant in explaining the migration pattern in Mexico. These results are consistent with prior expectations from the neo-classical theory. Also, a higher share of the manufacturing sector in the origin (destination) state exerts a retention (attraction) effect. This result is also consistent with the coefficients of the border-state dummies (representing *maquiladora* effects), suggesting that retention effects and attraction effects are simultaneously at work in the Border States, ‘reinforcing’ each other (the retention effect is not significantly different from zero, though). Finally, the population-centrality index produces two ‘opposing’ effects. The positive value of the centrality index in the origin state has an expulsion effect, which might indicate that congestion forces are at work (counter-urbanization), whereas a positive value in the destination state indicates an attraction effect (urbanization). In Model 1, the agglomeration effect is stronger (larger in magnitude) than the dispersion effect, whereas in Model 2 the reverse situation seems to occur. The latter, however, is more in line with recently developments observed in the Mexican case.

The goodness-of-fit of the models is measured by the pseudo- R^2 , which is defined as the square of the correlation between the actual and the fitted migration flows. Somewhat surprisingly, the pseudo- R^2 of Model 2 without migrant-stock variable is *higher* than for the model including the migrant-stock variable. More important, however, is that the coefficient of the *MSTOCK* variable is positive and highly significant (as could be expected), highlighting the important role of migrant stocks in shaping aggregate migration patterns in Mexico. Also, the Poisson specification persistently predicts internal migration flows better than its log-linearized counterpart.

To check the adequacy of the specification of the Models 1 and 2 (including the migrant-stock variable), I perform a Ramsey heteroskedasticity-robust RESET-test. The p -values are reported at the bottom of Table 2. Clearly, both models pass the RESET-test, so there is no evidence of misspecification of either of the two gravity-model specifications. Although the RESET-test is not really conclusive, a slight ‘superiority’ of the Poisson regression over the log-linear regression can be observed. In addition, I test the importance of heteroskedasticity by using the ‘two-degrees-of-freedom’ White’s test for heteroskedasticity [Wooldridge (2002, p.

¹⁵ This finding contrast with the common observation, as mentioned by Aroca and Maloney (2005), of the fact that most previous studies on migration failed to obtain such a result.

127)]. The p -values are reported at the bottom of Table 2. The null-hypothesis of homoskedastic error terms is unequivocally rejected in both cases. Finally, I test whether the particular patterns of heteroskedasticity assumed by Models 1 and 2, respectively, are appropriate. The test results are presented in Table 3. The adequacy of the log-linear model (Model 1) is checked by using the Park test, while the adequacy of the Poisson model (Model 2) is checked by using a Weighted Least Squares (WLS) test. Both tests are extensively described in Santos Silva and Tenreyro (2006). The test results are not really conclusive, but, again, a slight ‘preference’ for the Poisson model over the log-linear model emerges.

Insert Table 3 about here

Given the various test results and (above all) given the incorrect positive sign of the distance-decay elasticity for the log-linear specification (which is also statistically significant), I take the Poisson model as the preferred specification of the gravity model (including the migrant-stock variable).

5.2 Flexible Poisson Model 3

Next, I turn to the results obtained from Model 3 (estimated by using GME), where I imposed the theoretical restriction that all $\hat{\delta}_{ij} \leq 0$ (i.e., that all distance-deterrence elasticities are forced to be non-positive). The adequacy of the Poisson GME model has also been checked by using the WLS test (Table 3), and the test results indicate that Model 3 is (slightly) superior to the other model specifications. A remarkable finding, though, is that the improvement in fit (pseudo- R^2) offered by allowing for varying distance parameters is relatively modest.

The average distance-decay elasticity is -0.205, which is rather low; that is, internal migration in Mexico is quite inelastic with respect to distance: as the distance between origin and destination increases by 1%, migration drops on average by only 0.2%, *ceteris paribus*. However, the elasticities exhibit a wide variation, with a standard deviation of 0.123 (with a maximum, in absolute value, of -0.669). Interestingly, a plot of the elasticities against distance in Figure 4 suggests that the distance decay decreases (logarithmically) with distance, which possibly reflects decreasing marginal costs. The number of zero-valued distance elasticities equals 55 (roughly 5.5%) of the total number (992) of calculated elasticities. Given this relatively small number of ‘corner-point’ solutions, I am quite confident that the negativity of the average distance-

deterrence elasticity is not merely an ‘artifact’ of the imposed non-positivity constraints. Also, about 38% of these zero-values are related to México as destination state of migration, and about 31% of them are associated with migration to contiguous states (and about 62% to either of both). Finally, the probability of finding zero-values falls sharply with increasing distance.

Insert Figure 4 about here

The distance-decay coefficients also reveal the existence of asymmetry in the distance-deterrence effects. Just to give a few examples: the distance-decay elasticity of migration from Distrito Federal to México, Morelos, and Chihuahua, are zero, -0.006, and -0.129, respectively, whereas in the opposite direction, the coefficients are -0.429, -0.532, and -0.236, respectively. So, apparently, people moving out from Distrito Federal to other states are clearly less deterred by distance, *ceteris paribus*, than people from the same states moving into Distrito Federal.

Based on the results for the origin-destination-specific coefficients, I can further calculate the *average* distance-deterrence elasticities for each individual state, both as sending (row-totals) and as receiving (column totals) location. The spatial distribution of the average elasticities is shown in Figure 5, where the black and dark-grey shaded states have a ‘local’ and ‘narrow regional’ orientation (out-migration) or recruitment (in-migration) scope, *ceteris paribus*, respectively, and the white and light-grey shaded states have a more ‘national’ and ‘wide regional’ scope, *ceteris paribus*, respectively. Panel A of Figure 5 shows that distance decay is more elastic for out-migration from most of the states in the center of the country, and hence, the migrants from these states have a more regional ‘orientation’ than those from the remote states in the Border region and in parts of the North-Central and South regions. Panel B of Figure 5 presents to some degree a ‘mirror image’ of Panel A: most of the states in the Central and North-Central regions have a wider migrant ‘recruitment’ area than the more remote states in the Northern and Southern regions.

Insert Figure 5 about here

In addition, it is instructive to look at the geographical pattern of the distance-deterrence elasticities both within and between the regions in Mexico previously identified (see panel E in Figure 3). As shown in Figure 6, a tendency emerges for *higher* (absolute) values, on average, to be found for the pair-wise distance-deterrence elasticities in the Central and North-Central regions, between states at relative short geographical distance and, thus, with considerably *more*

interaction potentials. There are two possible explanations I can think of. A first possible explanation is that for highly ‘substitutable’ locations in the Central and North-Central regions (i.e., states that are *perceived* as ‘similar’ to some extent, in terms economic opportunities), prospective migrants may, with a higher probability, choose to migrate to the closest among the ‘competitive’ destinations. Obviously, if such is the case, one would expect higher ‘resistance’ of migrants to geographical distance, *ceteris paribus*. A second possible explanation is that migration networks may play only a minor role in the more urbanized areas of the Central and North-Central regions [see [Fussell and Massey \(2004\)](#)]. While such explanations may seem reasonable, they cannot account for the observed asymmetries found within these two regions. I conjecture that these asymmetries may be due to differences in the (unobserved) composition (e.g., in terms of education, work status, income, etc.) the migration flows, depending of their directionality; that is, the composition of the migration flow from i to j can be different from the flow from j to i , due to differences in the migrants’ preferences.

Insert Figure 6 about here

Several estimates from Model 3 are also different from those that resulted from Model 2. In particular, the push and push effects of income per capit (GDP_PC variable) turn out to be noticeably stronger. This result, however, contradicts the prior expectation that the role of economic factors should diminish after controlling for non-economic, cumulative-causation forces. Also, the opposing forces of urbanization and counter-urbanization (POP_CENT variable) seem to be equally strong. It is to be expected that the composition of the two migration streams (in terms of, say, incomes, skills, age, etc., of the migrants) will be different, however. Furthermore, the retention forces operating in the northern border states seem to be rather weak.

Finally, Figure 7 provides a map of the spatial distribution of the unobserved origin- and destination-specific (state fixed) effects. The average values are zero, by construction, and their distributions are noticeably positively skewed. Veracruz has the maximum value for $\hat{\nu}_i$ of 0.055 (strong expulsion forces), whereas Tlaxcala has the minimum value of -0.018 (strong retention forces). The maximum value for $\hat{\omega}_j$ of 0.136 (a clear ‘outlier’!) is associated with México (extremely strong attraction forces), whereas the minimum value of -0.018 corresponds to Distrito Federal and also Colima (strong repulsion forces). Overall, there seem to emerge two ‘clusters’ of states, one in the northern Border region (Baja California, Sonora, Chihuahua, Coahuila, Nuevo

Leon, and also Sinaloa) and another in the South-Central region (Guerrero, Oaxaca, Chiapas, and also Veracruz and Distrito Federal), showing relatively strong expulsion forces.

Insert Figure 7 about here

6. Summary and conclusions

In this paper, I employ a gravity-modeling approach to study aggregate inter-state migration in Mexico in the period 1995-2000. The purpose of this study is threefold: (1) to show that the gravity model in levels (Model 2) provides a more realistic description of the migration process than a model in logs (Model 1); (2) to develop a more general and flexible Poisson model (Model 3) that departs from the standard assumption of spatial invariance of distance decay; (3) to show that the migration process should not only be explained from a neo-classical perspective (e.g., socio-economic ‘push’ and ‘pull’ factors), but also from a cumulative-causation perspective (e.g., migrant stocks and idiosyncratic ties between locations). Also, because the flexible Poisson model (Model 3) is over-parameterized, I used the ‘non-classical’ generalized maximum entropy estimation method.

Many prior studies of aggregate internal migration are overly restrictive in the way that they specify spatial relations among locations, and some studies also found troublesome results. The flexible Poisson gravity-model framework adopted in this paper, however, yields the expected signs for all the estimated coefficients. The major findings of the present study can be summarized as follows. One important finding is that the estimated impact of distance (unobserved migration costs) depends crucially on the choice of model specification. In the case of a constant distance-decay parameter (Models 1 and 2), this impact appeared to be overstated, when migrant stocks were excluded from the model, or it turned out to be either insignificant or to have an incorrect sign, when migrant stocks were included in the model. In the case of varying distance-decay parameters (Model 3), the estimated distance-decay coefficients range from zero to -0.67, with a mean value of -0.20.

Thus, based on the results of Model 3 (which is the preferred model specification in the present study) internal migration in Mexico turns out to be quite inelastic with respect to geographical distance. On the other hand, social networks (measured as migrant stocks) are an important facilitating (inverse-resistance) factor of migration, reducing considerably the distance-deterrence-effect. Thus, the flexible Poisson gravity model presented in this paper to some extent

accounts (implicitly) for structural or historical components of inter-state migration and for the heterogeneous preferences among migrants.

Other important findings of this study are: (1) unemployment works as an expulsion (repulsion) force in origin (destination) states, whereas income per capita and the share of manufacturing act as retention (attraction) forces in origin (destination) states – which is consistent with prior predictions from neo-classical theory; (2) the population-centrality index (partially reflecting the supply of amenities in different locations) gives rise to opposing effects in origin and destination states: in origin states, it acts as an expulsion factor (congestion), whereas in destinations it acts as an attraction factor (demand for public and private services); (3) the northern border states appear to be economically attractive locations for people to reside and work (at least temporarily).

In conclusion, I would like to underline the exploratory nature of the present study, which can be improved in several ways. Evidently, a more substantive analysis is needed to explain the observed spatial variation in the distance-decay effects, focusing on functional and economic differences between states as well as the differences in the composition (e.g., in terms of skill, age, gender, income, etc., composition) of the migration flows that are associated with the spatial variation in distance-decay parameters. Also, the study could be improved by controlling for cross-sectional (spatial) correlation in the error terms. Finally, the study could be expanded by investigating the (large-sample) statistical properties of the GME estimator.

Despite these obvious ‘limitations’ and the focus of this paper on the gravity equation for modeling internal migration in Mexico, I am confident that the approach adopted in this paper holds great promise and can fruitfully be extended to a broader range of economic applications (including the study of international trade).

References

- Araujo, C. (2004). Can non-agricultural employment reduce rural poverty? Evidence from Mexico. *Cuadernos de Economía*, 41, 383-399.
- Aroca, P. and Maloney, W.F. (2005). Migration, trade, and foreign direct investment in Mexico. *The World Bank Economic Review*, 19, 449-472.
- Cole, W.E. and Sanders, R.D. (1983). Interstate migration in Mexico: variations on the Todaro theme. *Journal of Development Economics*, 341-354.
- Curran, S.R. and Rivero-Fuentes, E. (2003). Engendering migrant networks: the case of Mexican migration. *Demography*, 40, 289-307.
- Davies, P.S., Greenwood, M.J., and Li, H. (2001). A conditional logit model approach to U.S. state-to-state migration. *Journal of Regional Science*, 41, 337-360.
- Davis, B., Stecklov, G., and Winters, P. (2002). Domestic and international migration from rural Mexico: disaggregating the effects of network structure and composition. *Population Studies*, 38(2), 1-26.
- del Rey Poveda, A. (2007). Determinants and consequences of internal and international migration: the case of rural populations in the south of Veracruz, Mexico. *Demographic Research*, 16, 287-313.
- Flowerdew, R. and Aitkin, M. (1982). A method of fitting the gravity model based on the Poisson distribution. *Journal of Regional Science*, 22, 191-202.
- Fotheringham, A.S. and O'Kelly, M.E. (1989). *Spatial Interaction Models: Formulations and Applications*. Kluwer Academic Publishers, Dordrecht.
- Fratianni, M. and Kang, H. (2006). Heterogeneous distance elasticities in trade gravity models. *Economics Letters*, 90, 68-71.
- Fussell, E. and Massey, D.S. (2004). The limits to cumulative causation: international migration from Mexican urban areas. *Demography*, 41, 151-171.
- Golan, A., Judge, G., and Miller, D. (1996). *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York: John Wiley & Sons.
- Gordon, I.R. (1985). Economic explanations of spatial variation in distance deterrence. *Environment and Planning A*, 17, 59-72.
- Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood methods: applications to Poisson models. *Econometrica*, 52, 701-720.
- Greenwood, M.J. (1969). An analysis of the determinants of geographical labor mobility in the United States, *The Review of Economics and Statistics*, 51, 189-194.

- Greenwood, M.J. (1997). Internal migration in developed countries. In: M. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*, Vol. 1B, North-Holland: Elsevier, pp. 647-719.
- Greenwood, M.J. and Hunt, G.L. (2003). The early history of migration research. *International Regional Science Review*, 26, 3-37.
- Greenwood, M.J. and Ladman, J.R. (1978). An economic analysis of migration in Mexico. *Annals of Regional Science*, 12, 16-31.
- Henderson, D.J. and Millimet, D.L. (2007). Is gravity linear? Forthcoming in *The Review of Economics and Statistics*.
- Jones, R.C. (2001). Maquiladoras and U.S.-bound migration in Central Mexico. *Growth and Change*, 32, 193-216.
- King, J. (1978). Interstate migration in Mexico. *Economic Development and Cultural Change*, 27, 83-101.
- Lara, G. and Soloaga, I. (2005a). What drives internal migration in Mexico? A gravity model approach. Manuscript, Universidad de las Americas, Puebla.
- Lara, G. and Soloaga, I. (2005b). Determinants of migration in Mexico: a spatial econometrics approach. Manuscript, Universidad de las Americas, Puebla, Mexico.
- Lee, S.W. and Zhee, W. (2001). Independent and linked migration: individual return of employment opportunity and household return of poverty status to African American interstate migration. *Annals of Regional Science*, 35, 605-635.
- Lozano Ascencio, F. (2001). Nuevos orígenes de la migración mexicana a los Estados Unidos: migrantes urbanos versus migrantes rurales. *Revista Electrónica de Geografía y Ciencias Sociales*, No. 94(14), August [www.ub.es/geocrit/sn-94-14.htm].
- Lucas, R.E.B. (1997). Internal migration in developing countries. In: M.R. Rosenzweig and O. Stark (eds.), *Handbook of Population and Family Economics*, Volume 1B. North-Holland: Elsevier, pp. 721-798.
- Maier, G. (1985). Cumulative causation and selectivity in labour market oriented migration caused by imperfect information. *Regional Studies*, 19, 231-241.
- Massey, D.S., Arango, J., Hugo, G., Kouaouci, A., Pellegrino, A., and Taylor, J.E. (1994). An evaluation of international migration theory: the North American case. *Population and Development Review*, 20, 699-751.
- Mueser, P.R. (1989). The spatial structure of migration: an analysis of flows between the states in the USA over three decades. *Regional Studies*, 23, 185-200.
- Mueser, P.R. and Graves, P.E. (1995). Examining the role of economic opportunity and amenities in explaining population redistribution. *Journal of Urban Economics*, 37, 176-200.

- OECD (2004). *Economic Surveys 2002-2003: Mexico*, Volume 2003 – Supplement No. 1 – January 2004. Organisation for Economic Co-operation and Development, Paris.
- Peeters, L. (2008). Selective in-migration and convergence and divergence of incomes per capita across Belgian municipalities. Forthcoming in *Regional Studies*.
- Rietveld, P. and Bruinsma, F. (1998). *Is Transport Infrastructure Effective? Transport Infrastructure and Accessibility: Impacts on the Space Economy*. Springer-Verlag, Heidelberg.
- Rivero-Fuentes, E. (2005). *Beyond Income Differentials: Explaining Migrant's Destinations in Mexico*, Paper presented at the 25th International Population Conference, Tours, France, July 18-23, 2005 [<http://iussp2005.princeton.edu/download.aspx?submissionId=51591>].
- Rogers, A., Liu, J., Partida, V., and Tilton, M. (2006). Inferring directional flows from net migration data: Mexico. Paper presented at the Meeting of the Western Regional Science Association, Santa Fe, New Mexico (February 22-25, 2006).
- Rogers, A., Willekens, F., Little, J., and Raymer, J. (2002). Describing migration spatial structure. *Papers in Regional Science*, 81, 29-48.
- Sá, C., Florax, R.J.G.M., and Rietveld, P. (2004) Determinants of the regional demand for higher education in The Netherlands: a gravity model approach. *Regional Studies*, 38, 375-392.
- Santos Silva, J.M.C. and Tenreyro, S. (2006). The log of gravity, *The Review of Economics and Statistics*, 88, 641-658.
- Shen, J. (1999). Modelling regional migration in China: estimation and decomposition, *Environment and Planning A*, 31, 1226-1238.
- Wooldridge, J.M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA/London, England: The MIT Press.
- Yano, K., Nakaya, T., Fotheringham, A.S., Openshaw, S., and Ishikawa, Y. (2003). A comparison of migration behaviour in Japan and Britain using spatial interaction models. *International Journal of Population Geography*, 9, 419-431.

Table 1. Summary statistics for explanatory variables and expected direction of their effects

	Mean	Standard deviation	Minimum	Maximum	Expected sign of coefficient	Effect
$UNEMPL_i$	0.019	0.004	0.011	0.027	+	Expulsion
$UNEMPL_j$	0.019	0.004	0.011	0.027	-	Repulsion
GDP_{PC}_i	12.138	5.499	5.893	30.870	?	Expulsion/Retention?
GDP_{PC}_j	12.138	5.499	5.893	30.870	+	Attraction
$MANU_i$	0.229	0.084	0.053	0.375	-	Retention
$MANU_j$	0.229	0.084	0.053	0.375	+	Attraction
$POPCENT_i$	207.7	171.9	26.7	940.1	?	Expulsion/Retention?
$POPCENT_j$	207.7	171.9	26.7	940.1	?	Attraction/Repulsion?
$MSTOCK_{ij}$	0.005	0.016	0.000	0.326	+	Expulsion/Attraction
$BORDER_i$	0.219 ^a		0	1	-	Retention
$BORDER_j$	0.219 ^a		0	1	+	Attraction
$DIST_{ij}$	1370	1077	33	5961	-	Deterrence

^a The following seven states (21.9% of the total number of states in Mexico) are considered here as 'border states' [following Rogers et al. (2006)]: Baja California, Baja California Sur, Sonora, Chihuahua, Coahuila, Nuevo León, and Tamaulipas.

Table 2: Empirical results for three specifications of the gravity model of migration

	Model 1 Log-linear OLS		Model 2 Poisson PML		Model 3 Poisson GME
Constant	5.050 (2.095)**	3.538 (1.497)**	5.695 (3.033)*	5.227 (3.389)	2.822
$\ln UNEMPL_i$	-0.347 (0.239)	0.449 (0.177)**	-0.030 (0.393)	0.237 (0.392)	0.205
$\ln UNEMPL_j$	-0.920 (0.234)***	-1.475 (0.177)***	-0.539 (0.412)	-0.938 (0.456)**	-1.192
$\ln GDP_PC_i$	-0.239 (0.149)	-0.760 (0.100)***	0.397 (0.306)	-0.075 (0.497)	-0.793
$\ln GDP_PC_j$	0.218 (0.137)	0.213 (0.099)**	0.523 (0.222)**	0.349 (0.154)**	1.063
$\ln MANU_i$	-0.388 (0.135)***	-0.254 (0.093)***	-0.395 (0.219)*	-0.531 (0.182)***	-0.446
$\ln MANU_j$	-0.115 (0.135)	0.160 (0.094)*	-0.004 (0.224)	0.330 (0.209)	0.446
$\ln POPCENT_i$	0.436 (0.130)***	0.300 (0.089)***	0.563 (0.175)***	0.493 (0.160)***	0.499
$\ln POPCENT_j$	0.214 (0.122)*	0.664 (0.090)***	0.113 (0.169)	0.305 (0.162)*	0.500
$\ln MSTOCK_{ij}$		0.859 (0.025)***		0.544 (0.130)***	0.664
$BORDER_i$	1.067 (0.191)***	-0.208 (0.147)	-0.011 (0.398)	-0.356 (0.300)	-0.075
$BORDER_j$	1.399 (0.195)***	1.314 (0.142)***	0.850 (0.270)***	1.006 (0.239)***	0.807
$\ln DIST_{ij}$	-1.132 (0.071)***	0.136 (0.061)**	-0.985 (0.133)***	-0.316 (0.317)	
Mean					-0.205
Standard deviation					0.123
Minimum (several)					0.000
Maximum (Tlaxcala-Colima)					-0.669
Median					-0.198
$STATE_i$ (origin)					
Mean					0.000
Standard deviation					0.016
Minimum (Tlaxcala)					-0.018
Maximum (Veracruz)					0.055
Median					-0.005
$STATE_j$ (destination)					
Mean					0.000
Standard deviation					0.029
Minimum (Distrito Federal)					-0.018
Maximum (México)					0.136
Median					-0.009
Pseudo- R^2	0.073	0.172	0.343	0.275	0.332
RESET-test, p -values		0.232		0.749	
White's 2-df test, p -values		0.042		0.000	0.000

Note: Robust standard errors are given in parentheses. Statistical significance at the 10%, 5%, and 1% level is indicated with *, **, and ***, respectively. The number of observations is 992. The pseudo- R^2 is defined as the square of the correlation between actual and predicted migration flows. The pseudo- R^2 and White's test for Model 3 are based on observables (covariates) only. The RESET-test is applied only to Models 1 and 2.

Table 3. Results of tests for type of heteroskedasticity

Test	<i>t</i> -Statistic	<i>p</i> -Value
Park $H_0 : \lambda_1 = 2$	1.145	0.252
Log-linear OLS (Model 1) is valid $V[Y_i x] = E[Y_i x]^2$		
WLS $H_0 : \lambda_0(\lambda_1 - 1) = 0$	0.970	0.332
Poisson PML (Model 2) is valid $V[Y_i x] = E[Y_i x]$		
WLS $H_0 : \lambda_0(\lambda_1 - 1) = 0$	0.926	0.355
Poisson GME (Model 3) is valid $V[Y_i x] = E[Y_i x]$		

Note: The tests have been described extensively in Santos Silva and Tenreiro (2006, p. 646).

Figure 1. Spatial distribution of net in-migration rates in Mexico, 1995-2000

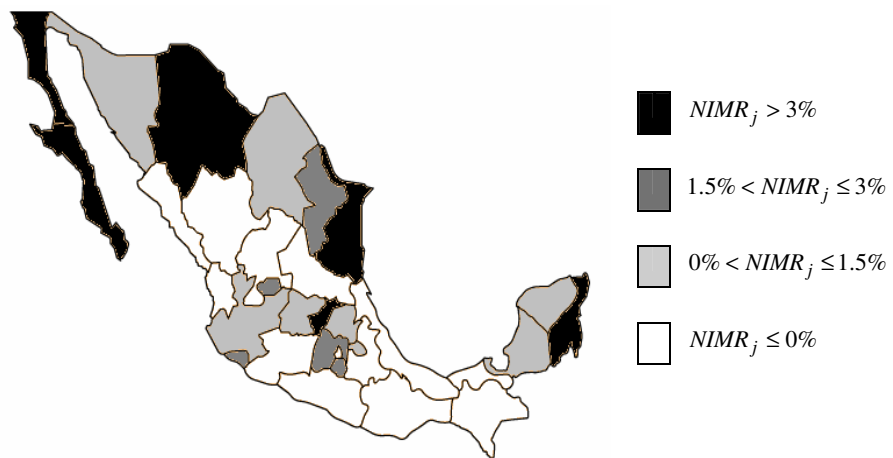


Figure 2. Size distribution of internal migration flows in Mexico, 1995-2000

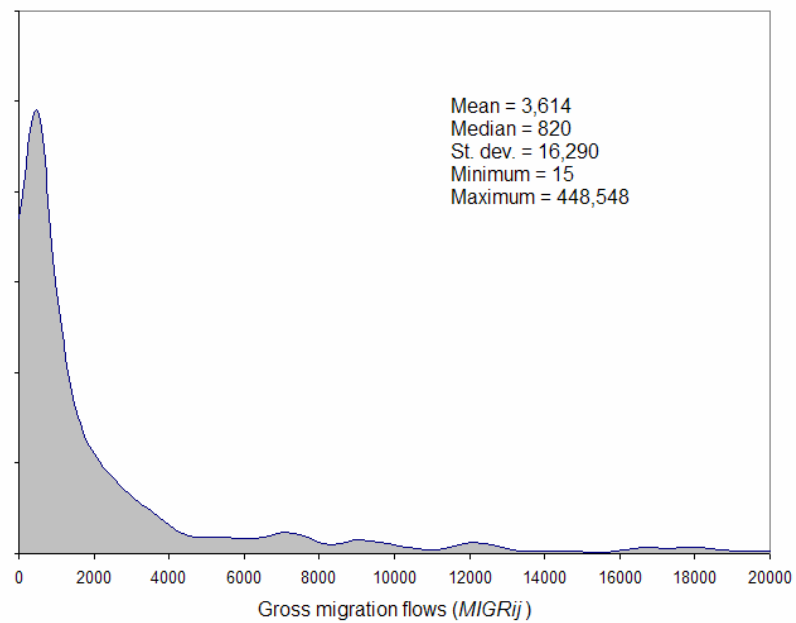
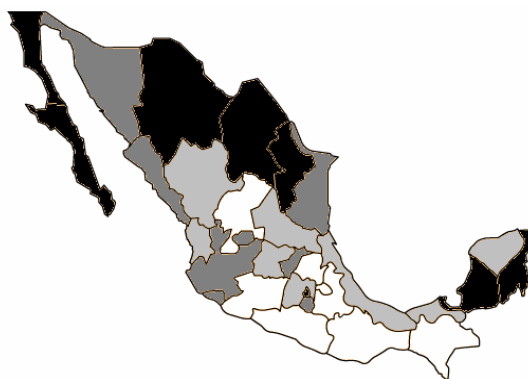


Figure 3. Spatial distribution of covariates, base year 1995
(quartiles)

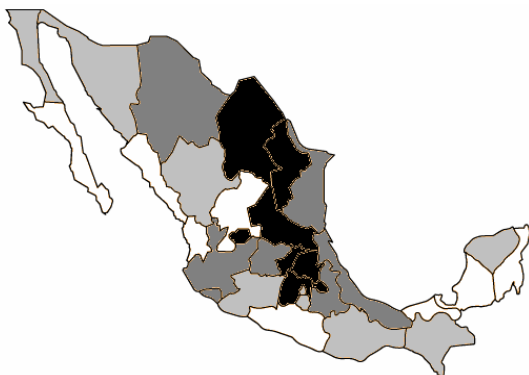
A: Unemployment rate



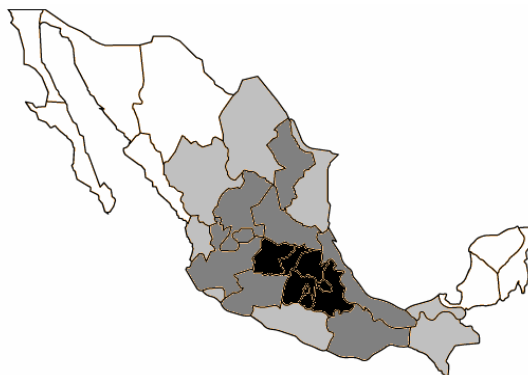
B: Income (GDP) per capita



C: Share of manufacturing



D: Population-centrality index



E: Four regions of Mexico

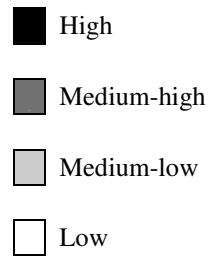


Figure 4. Distance-deterrence elasticities as a function of (log) distance

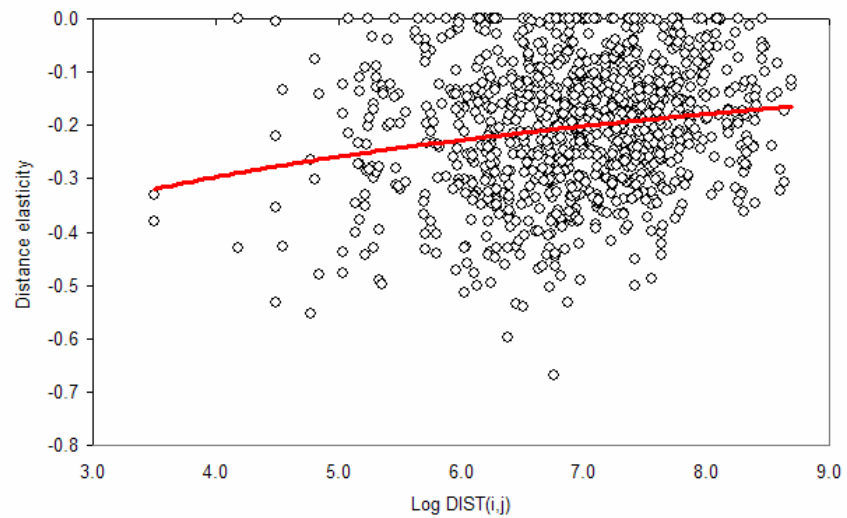
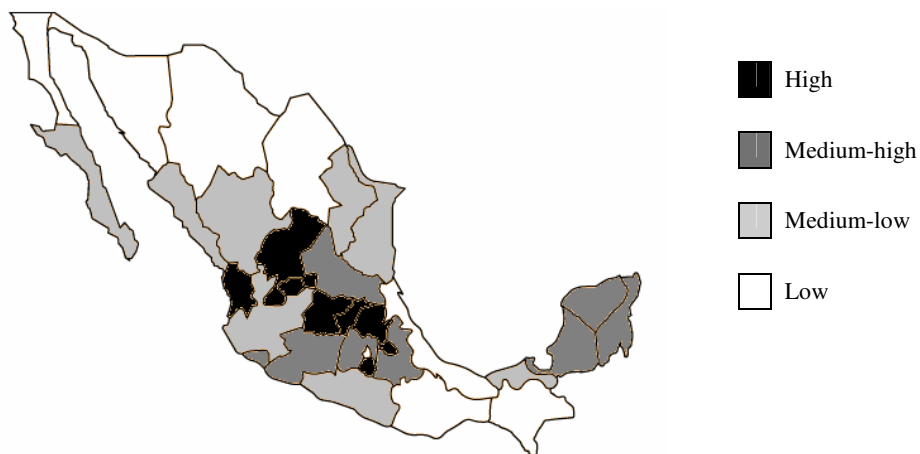


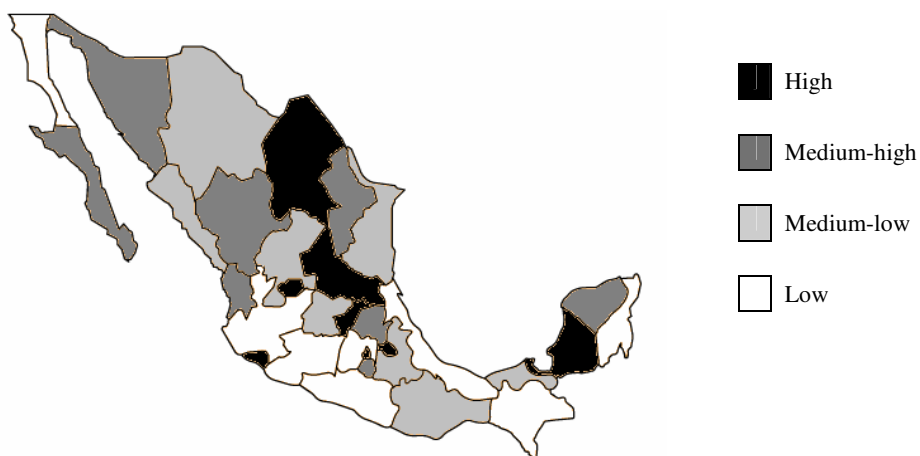
Figure 5. Spatial distribution of average origin- and destination-specific distance-deterrence elasticities (quartiles)

A: Average origin-specific distance-deterrence elasticities ($\hat{\delta}_{i\bullet}$)



High = distance *strongly* deters migrants from moving out to distant states, *ceteris paribus*;
 Low = distance *weakly* deters migrants from moving out to distant states, *ceteris paribus*.

B: Average destination-specific distance-deterrence elasticities ($\hat{\delta}_{\bullet j}$)



High = distance *strongly* deters migrants from distant states from moving in, *ceteris paribus*;
 Low = distance *weakly* deters migrants from distant states from moving in, *ceteris paribus*.

Figure 6. Geographical pattern of the variation in distance-decay of spatial interactions, averages across four Mexican regions (quartiles)

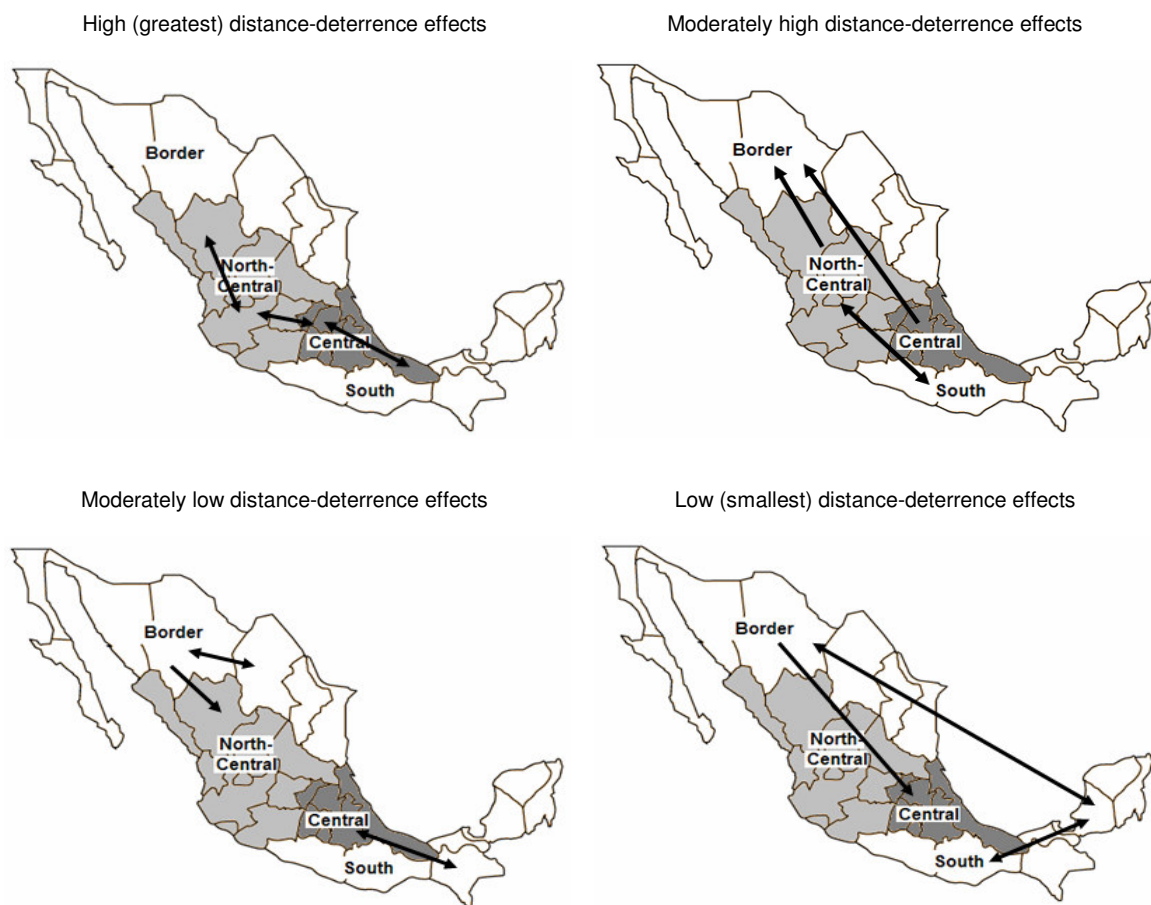
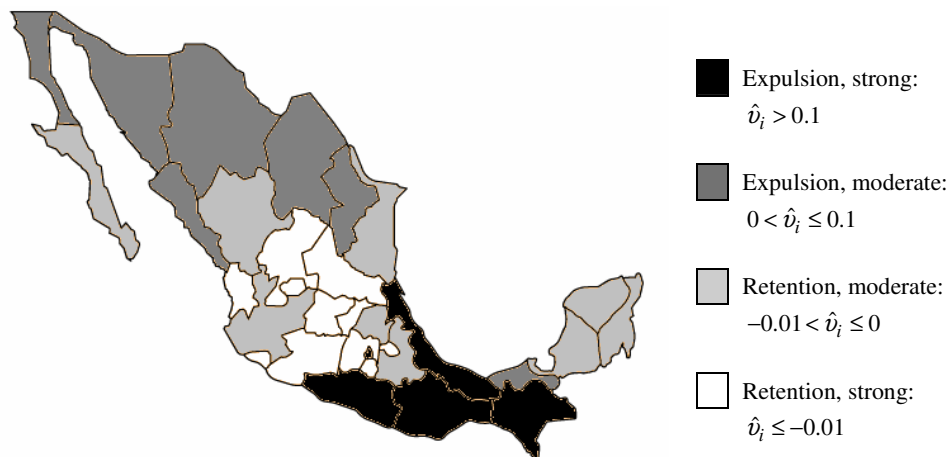
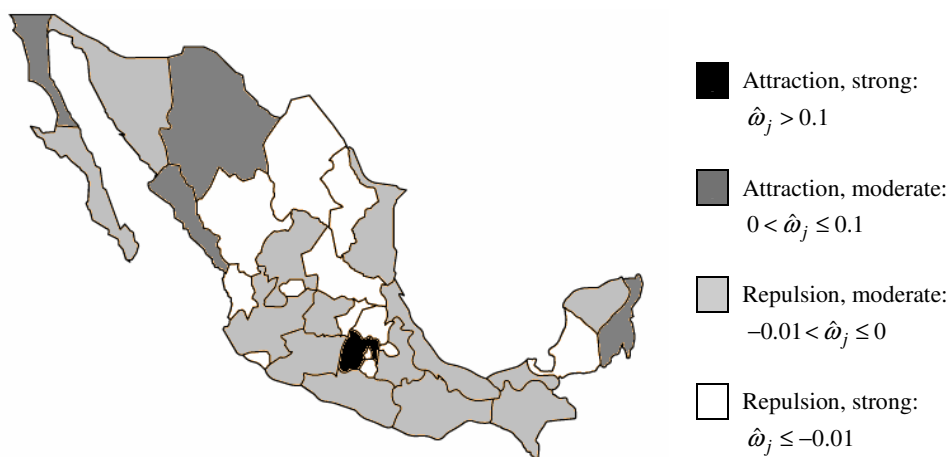


Figure 7: Spatial distribution of origin- and destination-specific effects

A: Origin-specific push (expulsion) or pull (retention) forces ($\hat{\nu}_i$)



B: Destination-state-specific pull (attraction) or push (repulsion) forces ($\hat{\omega}_j$)



Appendix 1: GME formulation of the flexible Poisson gravity model

In its simplest form, the flexible Poisson gravity model (Model 3) is

$$\begin{aligned}
 M_{ij} &= \exp(\alpha + \beta_1 x_i + \beta_2 x_j + \delta_{ij} d_{ij} + v_i + \omega_j + \varepsilon_{ij}) \\
 &= \exp(\alpha + \beta_1 x_i + \beta_2 x_j + \delta_{ij} d_{ij}) \times \exp(v_i + \omega_j + \varepsilon_{ij}) \\
 &= \exp(\alpha + \beta_1 x_i + \beta_2 x_j + \delta_{ij} d_{ij}) \eta_{ij}
 \end{aligned} \tag{A.1}$$

where $\eta_{ij} = \exp(v_i + \omega_j + \varepsilon_{ij})$.

To implement the GME method, this model in Equation (A.1) needs to be re-parameterized and converted into a constrained optimization problem, where the objective function consists of the joint entropy in equation (A.2) below. This objective function is to be maximized, subject to the appropriate data-consistency and normalization constraints, and relevant equality and inequality restrictions on the coefficients.

Re-parameterization

The coefficients $\{\alpha, \beta_1, \beta_2, \delta_{ij}\}$ and the unobserved-heterogeneity terms $\{v_i, \omega_j, \varepsilon_{ij}\}$ are defined as linear combinations of a set of unknown probability vectors $\mathbf{p}_\alpha = (p_{\alpha,1}, \dots, p_{\alpha,M})'$, $\mathbf{p}_{\beta_1} = (p_{\beta_1,1}, \dots, p_{\beta_1,M})'$, $\mathbf{p}_{\beta_2} = (p_{\beta_2,1}, \dots, p_{\beta_2,M})'$, $\mathbf{p}_{\delta_{ij}} = (p_{\delta_{ij},1}, \dots, p_{\delta_{ij},M})'$, $\mathbf{p}_{v_i} = (p_{v_i,1}, \dots, p_{v_i,M})'$, $\mathbf{p}_{\omega_j} = (p_{\omega_j,1}, \dots, p_{\omega_j,M})'$, $\mathbf{p}_{\varepsilon_{ij}} = (p_{\varepsilon_{ij},1}, \dots, p_{\varepsilon_{ij},M})'$, of dimension $M \geq 2$, and the corresponding (common) support vector $\mathbf{s} = (s_1, \dots, s_M)'$.

In addition, I define a multiplication factor for the ‘composite’ error term η_{ij} (which will be explained shortly), where $\mathbf{p}_{\sigma_{ij}} = (p_{\sigma_{ij},1}, \dots, p_{\sigma_{ij},G})'$, of dimension $G \geq 2$, and the (common) support vector $\mathbf{s}_\sigma = (s_{\sigma,1}, \dots, s_{\sigma,G})'$.

The coefficients, unobserved-heterogeneity terms, and multiplication factor, are then re-parameterized as $\alpha = \mathbf{p}'_\alpha \mathbf{s}$, $\beta_1 = \mathbf{p}'_{\beta_1} \mathbf{s}$, and so on, and $\sigma_{ij} = \mathbf{p}'_{\sigma_{ij}} \mathbf{s}_\sigma$.

Optimization problem

The GME formulation of the model is

$$\text{Max}_{\mathbf{p}} H(\mathbf{p}) = -\mathbf{p}'_{\alpha} \ln(\mathbf{p}_{\alpha}) - \mathbf{p}'_{\beta_1} \ln(\mathbf{p}_{\beta_1}) - \mathbf{p}'_{\beta_2} \ln(\mathbf{p}_{\beta_2}) - \sum_{i=1}^n \sum_{j=1}^n \mathbf{p}'_{\delta_{ij}} \ln(\mathbf{p}_{\delta_{ij}}) \quad (\text{A.2})$$

$$- \sum_{i=1}^n \mathbf{p}'_{v_i} \ln(\mathbf{p}_{v_i}) - \sum_{j=1}^n \mathbf{p}'_{\omega_j} \ln(\mathbf{p}_{\omega_j}) - \sum_{i=1}^n \sum_{j=1}^n \mathbf{p}'_{\varepsilon_{ij}} \ln(\mathbf{p}_{\varepsilon_{ij}}) - \sum_{i=1}^n \sum_{j=1}^n \mathbf{p}'_{\sigma_{ij}} \ln(\mathbf{p}_{\sigma_{ij}})$$

subject to

$$M_{ij} = \exp\left[(\mathbf{p}'_{\alpha} \mathbf{s}) + (\mathbf{p}'_{\beta_1} \mathbf{s})x_i + (\mathbf{p}'_{\beta_2} \mathbf{s})x_j + (\mathbf{p}'_{\delta_{ij}} \mathbf{s})d_{ij}\right] \eta_{ij} \quad (\text{A.3})$$

$$\eta_{ij} \geq \exp\left[(\mathbf{p}'_{v_i} \mathbf{s}) + (\mathbf{p}'_{\omega_j} \mathbf{s}) + (\mathbf{p}'_{\varepsilon_{ij}} \mathbf{s})\right] \quad (\text{A.4})$$

$$= (\mathbf{p}'_{\sigma_{ij}} \mathbf{s}_{\sigma}) / \exp\left[(\mathbf{p}'_{\alpha} \mathbf{s}) + (\mathbf{p}'_{\beta_1} \mathbf{s})x_i + (\mathbf{p}'_{\beta_2} \mathbf{s})x_j + (\mathbf{p}'_{\delta_{ij}} \mathbf{s})d_{ij}\right]^{1/2}$$

and

$$\sum_{i=1}^n \mathbf{p}'_{v_i} \mathbf{s} = 0; \quad \sum_{j=1}^n \mathbf{p}'_{\omega_j} \mathbf{s} = 0 \quad (\text{A.5})$$

$$\mathbf{p}'_{\delta_{ij}} \mathbf{s} \leq 0 \quad (\text{A.6})$$

$$\sum_{m=1}^M p_{\alpha,m} = 1; \quad \sum_{m=1}^M p_{\beta_1,m} = 1; \quad \sum_{m=1}^M p_{\beta_2,m} = 1; \quad \sum_{m=1}^M p_{\delta_{ij},m} = 1 \quad (\text{A.7})$$

$$\sum_{m=1}^M p_{v_i,m} = 1; \quad \sum_{m=1}^M p_{\omega_j,m} = 1; \quad \sum_{m=1}^M p_{\varepsilon_{ij},m} = 1; \quad \sum_{g=1}^G p_{\sigma_{ij},g} = 1$$

Equation (A.2) denotes the entropy objective, which is subject to the data-consistency constraints in equation (A.3). Equation (A.4) imposes the Poisson variance restriction that $V(\eta_{ij}) = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^{-1}$ or $SE(\eta_{ij}) = \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^{-1/2}$, where the inequality allows for possible ‘over-

dispersion'. The constraints in Equation (A.5) preserve a mean unobserved state-specific effect equal to zero, while the constraints in Equation (A.6) impose non-positivity on the distance-decay elasticities. Finally, the constraints in Equation (A.7) ensure that all probabilities add up to one (normalization).

After solving the entropy optimization problem in (A.2) through (A.7), the parameter estimates and the error terms can be recovered as $\hat{\alpha} = \hat{\mathbf{p}}'_\alpha \mathbf{s}_\alpha$, $\hat{\beta}_1 = \hat{\mathbf{p}}'_{\beta_1} \mathbf{s}_\alpha$, etc., and $\hat{\sigma}_{ij} = \hat{\mathbf{p}}'_{\sigma_{ij}} \mathbf{s}_\sigma$, where $\hat{\mathbf{p}}_\alpha$, $\hat{\mathbf{p}}_{\beta_1}$, etc., and $\hat{\mathbf{p}}_{\sigma_{ij}}$ are the estimated probabilities.

Support ranges

For estimation purposes, a common support vector for the coefficients is used. Since I have little prior knowledge about the 'true' value of the coefficients, the support vector is set as $\mathbf{s} = (-50, 50)'$ for the constant term, and $\mathbf{s} = (-10, 10)'$ for all the other coefficients ($M = 2$), which represent ranges wide enough (i.e., at least ten times the ranges of the magnitudes of the Poisson PML estimates) to include all possible outcomes.

On the other hand, the 'composite error' η_{ij} is defined in terms of the Poisson variance assumption and the widely-accepted 'three-sigma rule' [see Golan et al. (1996)]. Specifically, the Poisson $\text{SE}(\eta_{ij})$ is multiplied by a factor $\sigma_{ij} \in [-3, 3]$, which in turn is re-parameterized as $\mathbf{p}'_{\sigma_{ij}} \mathbf{s}_\sigma$, where $\mathbf{s}_\sigma = (-3, 0, 3)'$ ($G = 3$). Thus, I can write $\eta_{ij} \geq \sigma_{ij} \text{SE}(\eta_{ij}) = (\mathbf{p}'_{\sigma_{ij}} \mathbf{s}_\sigma) \exp(\mathbf{z}'_{ij} \boldsymbol{\beta})^{-1/2}$, which is Equation (A.4).

The estimates may be sensitive to the support values chosen. However, given the large number of observations ($n = 992$), the impact of the prior information is expected to be minimal.

Appendix 2: States of Mexico



<i>Code</i>	<i>State</i>	<i>Code</i>	<i>State</i>
01	Aguascalientes	17	Morelos
02	Baja California Norte	18	Nayarit
03	Baja California Sur	19	Nuevo León
04	Campeche	20	Oaxaca
05	Chiapas	21	Puebla
06	Chihuahua	22	Querétaro
07	Coahuila	23	Quintana Roo
08	Colima	24	San Luis Potosí
09	Distrito Federal	25	Sinaloa
10	Durango	26	Sonora
11	Guanajuato	27	Tabasco
12	Guerrero	28	Tamaulipas
13	Hidalgo	29	Tlaxcala
14	Jalisco	30	Veracruz
15	México	31	Yucatán
16	Michoacán	32	Zacatecas