View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Repositorio Institucional de la Biblioteca del Banco de España

LEARNING ABOUT MIGRATION DECISIONS FROM THE MIGRATS

AN EXERCISE IN ENDOGENOUS SAMPLING AND COMPLEMENTARY DATASETS

Olympia Bover and Manuel Arellano

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9908

LEARNING ABOUT MIGRATION DECISIONS FROM THE MIGRANTS

AN EXERCISE IN ENDOGENOUS SAMPLING AND COMPLEMENTARY DATASETS (*)

Olympia Bover and Manuel Arellano

(*) We are grateful to Samuel Bentolila, Ángel Estrada, Enrique Sentana, and seminar audiences at University College London, Bank of Spain, and Universidad Pablo de Olavide for their comments, to Pilar Velilla for her excellent research assistance, and to the Instituto Nacional de Estadística for providing the data.

Banco de España - Servicio de Estudios
Documento de Trabajo nº 9908

In publishing this series the Banco de España seeks to disseminate studies of interest that will help acquaint readers better with the Spanish economy.

The analyses, opinions and findings of these papers represent the views of their authors; they are not necessarily those of the Banco de España.

The Banco de España disseminates its principal reports and economic indicators via INTERNET and INFOVÍA.PLUS at: http://www.bde.es

ISSN: 0213-2710
ISBN: 84-7793-661-7
Depósito legal: M. 15565-1999
Imprenta del Banco de España

Abstract

We investigate the determinants of the remarkable increase in intra-regional migrations since the 1980's in Spain, using a large administrative micro dataset on migrants. Conditional migration probabilities are identified by comparing the migrants' joint distribution of characteristics to the corresponding distribution from the Spanish Labour Force Survey. The proportion of employment in the service industry, unemployment, house prices and education, all have an important positive effect on the individual probabilities of intra-regional migration.

1 Introduction

Intra-regional migration increased spectacularly in all Spanish regions since 1982 and it was in 1995 at an all time high, with per capita intra-regional migration being three times higher than in 1982 (see Figures 1 and 2). In contrast, until the early 1980's it evolved around a more or less constant level.

This fact was first noted by Olano (1990) and more recently by Bover and Velilla (1997), but has not received much attention so far. It is, nevertheless, interesting to study what are the forces behind this steady and unprecedented increase in short distance moves, specially considering that nowadays in Spain high regional unemployment or own unemployment no longer trigger substantial inter-regional migrations from people in poor regions towards better off ones (cf. Antolin and Bover, 1997). Bover and Velilla (1997) conjectured that the increase in intra-regional migration might respond to the change in the pattern of employment opportunities, presumably prompting moves of mainly skilled workers towards larger towns where the new, mostly service sector, jobs were.[1]

Employment in services climbed from 42 percent of total employment in 1977 to 61 percent in 1995. While from 1964 to 1978 the service share of employment grew at an annual rate of 0.79 percent, from 1980 to 1993 the annual rate was 1.12 percent, the highest among OECD countries together with Portugal. Furthermore, this increase in the share of services has taken place in all regions. Breaking down services into its main groups, we see that

[1] Ródenas (1994) found that the employment share of services in the origin and destination regions is a significant determinant of inter-regional migrations.

the increase has been mainly due to increases in "services provided to firms", "public administration", "trade and repairs", and "education and research", which generally are activities that tend to concentrate in larger towns.

Bover and Velilla reported results from pooled time-series regional data which provided some evidence in support of their view. The aggregate regional data, however, compound effects of this type with moves away from the high housing costs associated with large towns, which will also increase intra-regional migration. These effects produce migrations in opposite directions, and their magnitude is likely to differ across demographic groups. As a result, the true extent of the effects may be difficult to identify from the aggregates. With aggregate data it may not be possible to pin down the potential role of individual characteristics like education or age, and their interactions with aggregate variables. In particular, the increase in the education level in Spain during the 1980's has been noteworthy.[2]

In this paper we resort to individual data in order to obtain more precise measures of the factors behind the changes in the cross-sectional probabilities of intra-regional migration over time and size of town of residence. The focus of this paper is the study of the determinants of short distance migrations. This notion can be made operational in several ways, and any of them involves a certain degree of arbitrariness. Here we have chosen within region migrations as a measure of short distance migrations, which facilitates a straightforward matching with regional-level economic variables.

Despite its increase, the absolute number of intra-regional migrants in

[2]In 1998 31 percent of the population had 11 or more years of education, as compared to 12 percent in 1980.

a given year is nevertheless a very small percentage of the total population (1.4 percent in 1995). So there will not be many of them in a typical representative sample. In Spain the quarterly Labour Force Survey includes once a year some questions about migration, but in spite of its large sample size, it is not enough to conduct a conditional analysis of migration by origin and destination. In contrast, the Census of Residential Variations provides exhaustive information on the migrants' moves and on some of their characteristics. Thus, our empirical strategy is to identify conditional migration probabilities from a comparison of the distribution of characteristics of the migrants (in a sample from the Census of Residential Variations) with the distribution of characteristics of the entire population (migrants and nonmigrants from the Labour Force Survey), using Bayes theorem. Estimation is, therefore, based on a choice-based sample. See Manski and Lerman (1977), and Amemiya (1985, pp. 319-338) for a survey of choice-based sampling in discrete choice models and further references. Identification of our model can also be regarded as arising from the use of complementary datasets or complementary population characteristics (see, for example, Angrist and Krueger, 1992, Arellano and Meghir, 1992, and Imbens and Lancaster, 1994, on the use of complementary datasets in different contexts).

The paper is organized as follows. We begin by explaining in Section 2 the econometric methods and the models used in the empirical analysis. From the comparison between the conditional distribution of characteristics given migration and the marginal distribution of characteristics, only odd ratios of migration are nonparametrically identified. Given the odd ratios, the conditional migration probabilities can be determined given the knowledge , of the unconditional migration probabilities. We consider multinomial models of migration by considering migration to small, medium, and large towns as separate alternatives.We discuss two asymptotically equivalent estimation methods. First, a minimum chi-square method for multinomial logit which can be implemented as a nonlinear weighted least-squares estimator. Second, a maximum likelihood estimator in which the intercepts are determined through implicit nonlinear constraints. In Section 3 we describe the data, which consists of a random sample from the Spanish Census of Residential Variations for the years 1988-1992 (excluding 1991), and aggregate statistics from the Labour Force Surveys for the same years. In Section 4 we present the empirical results from the various models and report estimated migration probabilities. Finally, Section 5 contains the conclusions.

2 Econometric Methods

Identifying Migration Probabilities from the Migrants We begin by presenting the basic ideas in the simpler context of binary choice (although we shall not report results for binary migration probabilities in the paper), and subsequently we extend them to multinomial choice.

Let the probability of migration for an individual with characteristics x be $\Pr(y = 1|x)$, and let $f(x)$ and $f(x|y = 1)$ be the marginal and the conditional probability distributions of x given migration, respectively. We then have

$$\Pr(y = 1|x) = \frac{f(x|y = 1)\Pr(y = 1)}{f(x)}. \tag{1}$$

Thus, the migration probabilities can be determined from equation (1) given knowledge of $f(x|y = 1)$, $f(x)$ and $p = \Pr(y = 1)$. Note that if p were unknown only relative probabilities or odd ratios could be identified.

Let us now consider a standard binary choice model for the probability of migration:

$$\Pr(y = 1|x) = G(\alpha + z'\beta) \tag{2}$$

where G is some known cdf (eg. logistic) and z is a vector of explanatory variables which contains some of the x's or functions of them, so that $z = z(x)$.

We are interested in estimating α and β from a sample of migrants and the knowledge of $f(x)$ and p. The set of explanatory variables in our empirical analysis consists of discrete individual characteristics and continuous aggregate regional-level variables. Since the latter can be regarded as linear combinations of region-specific time dummies, our dataset is one with many observations per cell (i.e. x will include a full set of region-specific time dummies, and aggregate variables will be elements of z). Thus, in our case $f(x)$ is a multinomial distribution with known probabilities. The information on these probabilities comes from Labour Force Survey (LFS) aggregates to which population elevation factors have been applied. The information on p comes from the census statistics on population and residential variations. Alternatively, we could assume that $f(x)$ and/or p are observed with sampling error. In such case, the estimators discussed below would be reinterpreted as being conditional on estimated quantities, and there would be an additional source of uncertainty in them; but since the LFS sample size is large, the standard errors that we report are calculated assuming that $f(x)$ and p are known. Estimation when $f(x)$ is estimated as opposed to known with certainty is discussed in Appendix 3.

Binary Minimum Distance Estimation If the vector of variables x can take q different values $\{\xi_1, ..., \xi_q\}$, an unrestricted estimate of $\Pr(y = 1|x = \xi_\ell)$ from a random sample of n migrants with observations $\{x_i\}$ $(i = 1, ..., n)$ is given by

$$\widehat{\Pr}(y = 1|x = \xi_\ell) = \frac{\widehat{\Pr}(x = \xi_\ell|y = 1)p}{\Pr(x = \xi_\ell)} \quad (\ell = 1, ..., q) \tag{3}$$

where

$$\widehat{\Pr}(x = \xi_\ell|y = 1) = \frac{1}{n}\sum_{i=1}^{n} 1(x_i = \xi_\ell) = \widehat{\phi}_\ell, \text{ say.} \tag{4}$$

The sample frequencies $\widehat{\phi}_\ell$ are consistent and asymptotically normal estimates of the corresponding probabilities $\phi_\ell = \Pr(x = \xi_\ell|y = 1)$. Specifically, letting $\phi = (\phi_1, ..., \phi_{q-1})'$ and $\widehat{\phi} = (\widehat{\phi}_1, ..., \widehat{\phi}_{q-1})'$, by the central limit theorem we have

$$\sqrt{n}(\widehat{\phi} - \phi) \xrightarrow{d} N(0, \Omega) \tag{5}$$

where $\Omega = \Lambda - \phi\phi'$ and $\Lambda = diag\{\phi_1, ..., \phi_{q-1}\}$.

The model specifies that

$$\phi_\ell = \phi_\ell(\alpha, \beta) = \frac{\pi_\ell}{p} G\left(\alpha + z(\xi_\ell)'\beta\right) \tag{6}$$

where $\pi_\ell = \Pr(x = \xi_\ell)$. Then, the optimal minimum distance estimator of α and β minimizes

$$s(\alpha, \beta) = [\widehat{\phi} - \phi(\alpha, \beta)]'\widehat{\Omega}^{-1}[\widehat{\phi} - \phi(\alpha, \beta)] \tag{7}$$

where $\widehat{\Omega}$ is the sample counterpart of Ω. Moreover, $\widehat{\Omega}^{-1} = \widehat{\Lambda}^{-1} - \iota\iota'/\widehat{\phi}_q$, where ι denotes a $(q-1) \times 1$ vector of ones. Upon substitution, the minimum distance estimation criterion can be written as[3]

$$s(\alpha,\beta) = \sum_{\ell=1}^{q} \frac{1}{\widehat{\phi}_\ell} \left(\widehat{\phi}_\ell - \frac{\pi_\ell}{p} G\left(\alpha + z(\xi_\ell)'\beta\right) \right)^2 . \tag{8}$$

From the theory of minimum distance estimation we know that the minimizer of $s(\alpha,\beta)$ is asymptotically equivalent to maximum likelihood. The asymptotic covariance matrix of both minimum distance and maximum likelihood estimates can be shown to be given by

$$V = \left[\sum_{\ell=1}^{q} \frac{1}{\phi_\ell} \left(\frac{\pi_\ell}{p} G'(\alpha + z(\xi_\ell)'\beta) \right)^2 z^*(\xi_\ell) z^*(\xi_\ell)' \right]^{-1} \tag{9}$$

where $z^*(\xi_\ell) = (1, z(\xi_\ell)')'$, and G' denotes the first derivative of G.

Maximum Likelihood Estimation The log likelihood of the sample of n independent observations of migrants is given by

$$L = \sum_{i=1}^{n} \ln f(x_i | y_i = 1) = n \sum_{\ell=1}^{q} \widehat{\phi}_\ell \ln \phi_\ell \tag{10}$$

[3]The form of Ω^{-1} results from the matrix inversion lemma:

$$(\Lambda - \phi\phi')^{-1} = \Lambda^{-1} + \frac{\Lambda^{-1}\phi\phi'\Lambda^{-1}}{(1 - \phi'\Lambda^{-1}\phi)}$$

and the fact that $\Lambda^{-1}\phi = \iota$ and $\phi_q = 1 - \iota'\phi$. As for the minimum distance criterion, we have

$$\begin{aligned} s(\alpha,\beta) &= [\widehat{\phi} - \phi(\alpha,\beta)]'\widehat{\Lambda}^{-1}[\widehat{\phi} - \phi(\alpha,\beta)] + \frac{1}{\widehat{\phi}_q}[\widehat{\phi} - \phi(\alpha,\beta)]'\iota\iota'[\widehat{\phi} - \phi(\alpha,\beta)] \\ &= \sum_{\ell=1}^{q-1} \frac{1}{\widehat{\phi}_\ell} \left(\widehat{\phi}_\ell - \phi_\ell(\alpha,\beta) \right)^2 + \frac{1}{\widehat{\phi}_q}[\iota'\widehat{\phi} - \iota'\phi_\ell(\alpha,\beta)]^2. \end{aligned}$$

Expression (8) follows from the fact that

$$[\iota'\widehat{\phi} - \iota'\phi_\ell(\alpha,\beta)]^2 = \left((1 - \widehat{\phi}_q) - [1 - \phi_q(\alpha,\beta)] \right)^2 = [\widehat{\phi}_q - \phi_q(\alpha,\beta)]^2.$$

where the ϕ_ℓ are as specified in (6), and $\sum_{\ell=1}^{q}\phi_\ell = 1$, or equivalently

$$\sum_{\ell=1}^{q} \pi_\ell G(\alpha + z(\xi_\ell)'\beta) = p. \tag{11}$$

Substituting (6) and (11) in (10) we obtain

$$L(\alpha,\beta) = n\sum_{\ell=1}^{q} \widehat{\phi}_\ell \ln G(\alpha + z(\xi_\ell)'\beta) - n\ln\left(\sum_{\ell=1}^{q} \pi_\ell G(\alpha + z(\xi_\ell)'\beta)\right). \tag{12}$$

Maximum likelihood estimates of α and β are obtained by maximizing $L(\alpha,\beta)$ subject to (11). To implement this method, we solve numerically (11) for the intercept as a function of the slope coefficients $\alpha = \alpha(\beta)$, say. Then, we first obtain estimates of the slope parameters as $\widehat{\beta} = \arg\max L(\alpha(\beta),\beta)$, from which the estimated intercept can be calculated as $\widehat{\alpha} = \alpha(\widehat{\beta})$. The estimated covariance matrix and the standard errors for $\widehat{\beta}$ are obtained from the hessian matrix of $L(\alpha(\beta),\beta)$. Given this, the standard error for $\widehat{\alpha}$ is calculated using the delta method. In both instances, numerical derivatives of $\alpha(\beta)$ are employed.[4] Indeed, one advantage of ML estimation over MD is that it enforces the restriction $\sum_{\ell=1}^{q}\phi_\ell = 1$ whereas MD does not.

When there are continuous explanatory variables, the nature of the estimation problem changes. This situation does not arise in our empirical analysis, because our continuous variables only vary by region and time, and

[4] Alternatively, substituting $\phi_q = 1 - \sum_{\ell=1}^{q-1}\phi_\ell$ in (10) we obtain

$$\begin{aligned} L &= n\sum_{\ell=1}^{q-1}\widehat{\phi}_\ell \ln\phi_\ell + n\widehat{\phi}_q \ln\left(1 - \sum_{\ell=1}^{q-1}\phi_\ell\right) \\ &\propto n\sum_{\ell=1}^{q-1}\widehat{\phi}_\ell \ln G(\alpha + z(\xi_\ell)'\beta) + n\widehat{\phi}_q \ln\left(1 - \frac{1}{p}\sum_{\ell=1}^{q-1}\pi_\ell G(\alpha + z(\xi_\ell)'\beta)\right). \end{aligned}$$

The problem with this way of enforcing the restriction (11) is that the expression whose log is taken in the last term could be negative for some values of α and β.

so they are regarded as functions of dummy variables. In Appendix 3, however, we present some discussion on the problem of estimation in the presence of continuous characteristics.

Multinomial Models In this paper we actually consider a multinomial choice among four different alternatives: (1) migration to a small town, (2) migration to a medium size town, (3) migration to a large town, and (4) no migration. Since we only observe migrants, all the individuals in our sample fall in one of the first three categories. This is represented by the indicator variable y, which in the multinomial case we redefine as taking on values in the set $\{1, 2, 3\}$ for each of the migration classes. In the event of no migration we assign the value $y = 0$.

The probability of migration to destination j can be determined from $f(x|y=j)$, $p_j = \Pr(y=j)$ and $f(x)$:

$$\Pr(y=j|x) = \frac{f(x|y=j)\Pr(y=j)}{f(x)} \quad (j=1,2,3). \tag{13}$$

We model these probabilities using a multinomial logit specification of the form

$$\Pr(y=j|x) = G_j(z;\alpha,\beta) \equiv \frac{e^{\alpha_j + z'\beta_j}}{1 + e^{\alpha_1+z'\beta_1} + e^{\alpha_2+z'\beta_2} + e^{\alpha_3+z'\beta_3}} \quad (j=1,2,3). \tag{14}$$

where $\alpha = (\alpha_1, \alpha_2, \alpha_3)'$, and $\beta = (\beta_1', \beta_2', \beta_3')'$.

The purpose of our analysis is to study how migration probabilities vary with the characteristics of the individuals and of their region of residence. Note that in our specification, the log-odd ratios between two alternatives contain a set of unrestricted coefficients. This is so because our explanatory variables vary with individuals but not with alternatives.[5]

Multinomial Minimum Distance Estimation An unrestricted estimate of $\Pr(y=j|x=\xi_\ell)$ is given by

$$\widehat{\Pr}(y=j|x=\xi_\ell)=\frac{\widehat{\phi}_{j\ell}p_j}{\pi_\ell}\ (j=1,2,3) \tag{15}$$

where $\pi_\ell=\Pr(x=\xi_\ell)$ and

$$\widehat{\phi}_{j\ell}=\widehat{\Pr}(x=\xi_\ell|y=j)=\frac{1}{n_j}\sum_{i=1}^{n}1(y_i=j)1(x_i=\xi_\ell), \tag{16}$$

and n_j is the number of observations with $y=j$.

As before, the sample frequencies $\widehat{\phi}_{j\ell}$ are consistent and asymptotically normal estimates of the corresponding probabilities $\phi_{j\ell}=\Pr(x=\xi_\ell|y=j)$. Letting $\phi_j=(\phi_{j1},...,\phi_{j(q-1)})'$ and $\widehat{\phi}_j=(\widehat{\phi}_{j1},...,\widehat{\phi}_{j(q-1)})'$ for $j=1,2,3$, by the central limit theorem we have

$$\sqrt{n}\begin{pmatrix}\widehat{\phi}_1-\phi_1\\ \widehat{\phi}_2-\phi_2\\ \widehat{\phi}_3-\phi_3\end{pmatrix}\xrightarrow{d}N\left[0,\begin{pmatrix}\frac{1}{r_1}\Omega_1 & 0 & 0\\ 0 & \frac{1}{r_2}\Omega_2 & 0\\ 0 & 0 & \frac{1}{r_3}\Omega_3\end{pmatrix}\right] \tag{17}$$

where $\Omega_j=\Lambda_j-\phi_j\phi_j'$, $\Lambda_j=diag\{\phi_{j1},...,\phi_{j(q-1)}\}$, and $r_j=p\lim_{n\to\infty}(n_j/n)$. Sample frequencies for different values of j are independent because they are based on different subsamples.

[5] If observations on destination-specific variables were available, their shadow value to the migrants could be measured (cf. McFadden, 1981). In such type of model, however, multinomial logit would lack a realistic pattern of similarity across alternatives, since for example we might expect that migrations to small or medium towns on the one hand, and migrations to medium or large towns on the other, could be perceived as alternatives with a relatively high similarity. Moreover, a two-stage decision process by which individuals first decide whether to migrate or not and if so to where, would not be very meaningful here, so that the pattern of similarity would lack a tree structure.

The model specifies that

$$\phi_{j\ell} = \frac{\pi_\ell}{p_j} G_j\left(z(\xi_\ell); \alpha, \beta\right). \tag{18}$$

Then the optimal minimum distance estimates of α and β minimize

$$\begin{aligned} s(\alpha,\beta) &= \frac{n_1}{n}[\widehat{\phi}_1 - \phi_1(\alpha,\beta)]'\widehat{\Omega}_1^{-1}[\widehat{\phi}_1 - \phi_1(\alpha,\beta)] \\ &+\frac{n_2}{n}[\widehat{\phi}_2 - \phi_2(\alpha,\beta)]'\widehat{\Omega}_2^{-1}[\widehat{\phi}_2 - \phi_2(\alpha,\beta)] \\ &+\frac{n_3}{n}[\widehat{\phi}_3 - \phi_3(\alpha,\beta)]'\widehat{\Omega}_3^{-1}[\widehat{\phi}_3 - \phi_3(\alpha,\beta)] \end{aligned} \tag{19}$$

where $\widehat{\Omega}_j$ is the sample counterpart of Ω_j. Moreover, as in the binary case $s(\alpha, \beta)$ can be written as

$$\begin{aligned} s(\alpha,\beta) &= \frac{n_1}{n}\sum_{\ell=1}^{q}\frac{1}{\widehat{\phi}_{1\ell}}\left(\widehat{\phi}_{1\ell} - \frac{\pi_\ell}{p_1}G_1\left(z(\xi_\ell);\alpha,\beta\right)\right)^2 \\ &+\frac{n_2}{n}\sum_{\ell=1}^{q}\frac{1}{\widehat{\phi}_{2\ell}}\left(\widehat{\phi}_{2\ell} - \frac{\pi_\ell}{p_2}G_2\left(z(\xi_\ell);\alpha,\beta\right)\right)^2 \\ &+\frac{n_3}{n}\sum_{\ell=1}^{q}\frac{1}{\widehat{\phi}_{3\ell}}\left(\widehat{\phi}_{3\ell} - \frac{\pi_\ell}{p_3}G_3\left(z(\xi_\ell);\alpha,\beta\right)\right)^2. \end{aligned} \tag{20}$$

The resulting estimates will be asymptotically equivalent to maximum likelihood.

Multinomial Maximum Likelihood Estimation Using similar arguments as in the binary case, the log-likelihood for the sample of n migrants is given by

$$L(\alpha,\beta) = \sum_{j=1}^{3}\left\{n_j\sum_{\ell=1}^{q}\widehat{\phi}_{j\ell}\ln G_j(z(\xi_\ell);\alpha,\beta) - n_j\ln\left(\sum_{\ell=1}^{q}\pi_\ell G_j(z(\xi_\ell);\alpha,\beta)\right)\right\}. \tag{21}$$

Maximum likelihood estimates are obtained maximizing $L(\alpha, \beta)$ with respect to α and β subject to the constraints

$$\sum_{\ell=1}^{q} \pi_\ell G_j(z(\xi_\ell); \alpha, \beta) = p_j \ (j = 1, 2, 3). \tag{22}$$

For multinomial logit, maximum likelihood estimation is implemented as in the binary case. That is, we solve numerically the three nonlinear equations (22) for the intercepts as functions of the slope coefficients (using a Gauss-Newton iteration), and maximize the log-likelihood as a function of the slope coefficients alone. Afterwards the estimated intercepts are obtained from the estimated slopes using the implicit functions.

3 The data

To study internal migrations in Spain there are two main data sources, aside from very low frequency Census data that take place every ten years. The first one is the annual Residential Variations Data (RVD) ("Estadística de Variaciones Residenciales"), which has traditionally recorded new arrivals (and departures) at the municipality level. This is the only source on migration flows inside Spain beginning in the 1960's, and has therefore been the main source for work on aggregate data. Its drawback for micro studies, as we shall detail below, is that it has scarce information on the characteristics of the migrants. The second source is the Migration Survey (MS), included in the second quarters of the Labour Force Survey (LFS), which takes as migrants those individuals whose municipality of reference is different from the one in the previous year. However, the small proportion of migrants in the population results in a very small sample of migrants in the LFS. Moreover, as reported by Ródenas and Martí (1997) the design of the MS may produce a severe underestimation of migration probabilities. For example, the MS does not show the (substantial) increase observed with the RVD in intra-regional migrations since the 1980's. Individual MS data are available since 1987 (2nd quarter), but they do not contain information on the size of town of origin, only on the province of origin, which is another limitation for our purposes.[6]

In this paper we use the RVD from 1988, when computerized individual records started to be available. The characteristics for internal migrants available in the RVD are: sex, province (or country) of birth, age, education, province of origin and destination, and size of towns of origin and destination.[7] Inspection of the data revealed lack of compatibility in the education variable from 1993 (possibly due to changes in the educational categories used), and as a consequence our sample period ends in 1992. Furthermore, we do not use 1991 observations because in this year the municipal census was renewed and as a result migrations dropped artificially. The reason is that during the months the renovation takes place, migrants are considered as new records to the census as opposed to immigrants. Therefore, the years of data we use are 1988, 1989, 1990, and 1992.

Given the lack of household characteristics, specially relevant for women, we restrict our attention to men, aged between 20 and 64, that have moved within region (with all the characteristics of interest available). The resulting

[6]From 1980 to 1986 the MS was also conducted as part of the LFS. However, the data for this period are not comparable with the data from 1987 because, among other things, the old MS took place every quarter instead of every second quarter.

[7]The Spanish provinces are an administrative division of the regions. There are 17 regions and 52 provinces.

dataset of intra-regional migrants varies between 120,000 and 145,000 individuals per year. From there we draw a 10 percent random sample, leading to a sample size of 52,135 intra-regional male migrants. Details on the characteristics of this sample and on the exact categories of the variables can be found in the Data Appendix. Unfortunately, we only have three categories for the education variable due to the new coding in the RVD after 1990, which aggregates together all individuals with eleven or more years of education.

The focus of this paper is the study of short distance moves. There are several potential definitions of short distance moves, for example, within province moves, within regions, within regions with the addition of moves to adjacent provinces or regions, etc.We eventually decided to use within region migrations as our measure of short distance moves because the aggregate economic variables are mainly available at the regional level, coupled with the fact that over 85 percent of within region moves are within provinces, and that moves to adjacent regions account for only about a quarter of inter-regional migrations (which in turn are less than half the volume of intra-regional migrations).

The source for the distribution of characteristics of the total population (migrants and non-migrants) consists of aggregate LFS probabilities. The LFS is conducted every quarter on all members of around 60,000 households. From there the Statistical Office (INE), after applying the corresponding population weights, provides the aggregate figures for the relevant population according to a set of characteristics; in our case, prime-age males by year, region, size of town of residence, age, and education.

We should point out that our LFS population includes inter-regional migrants, in addition to non-migrants and intra-regional migrants. Ideally one would prefer to exclude them to enable a cleaner comparison between intra-regional migrants and non-migrants. However, inter-regional migrants can only be observed in the LFS waves corresponding to the second quarters, when the MS takes place. Given that the between-region migrants in the MS are less than 0.3 percent of the male population, we preferred to keep them rather than reduce by four the size of the dataset on which our population probabilities are based. We also considered the possibility of including inter-regional moves in the analysis as additional alternatives, but this would involve modelling inter-regional migration, which would change the focus of the paper.

Turning to aggregate and regional economic variables, we consider the effects on intra-regional migration of unemployment, house prices, and the employment share of services. We use time series of regional unemployment, and the regional share of employment in services. As a variable for real house prices, we use nominal regional data deflated by the nationwide CPI. The reason for this choice is that regional CPI's (which are all set to 100 in the base year) cannot be used to take into account differences in cost of living across regions (on this point see, for example, Deaton, 1998). Differences across regions in our house price variable will therefore reflect not only house price differences but also differences in living costs. All regional economic variables are dated at $t-1$.

4 Empirical Results

Nonlinear minimum distance estimates of the parameters in the multinomial logit model are presented in Table 1. As initial values we used consistent but asymptotically inefficient linear MD estimates (linear MD estimation is discussed in Appendix 3). The calculation of maximum likelihood estimates turned out to require much more computing time than minimum distance. This was due to having to solve numerically for the intercepts the system of nonlinear constraints (22) at each iteration. Since the two methods are asymptotically equivalent and they provided very similar results in the cases where both were calculated, we only report the MD results.

Separate estimates for each of the three town of origin sizes (small, medium, large) are provided for a three equation system, which consists of the log odd ratios for each of the three town of destination sizes relative to the probability of non-migration. Aside from parameter estimates, to have a clearer picture of the magnitude of the effects, we provide in Table A2.1 in Appendix 2 an extensive calculation of the probabilities predicted by the estimated equations reported in Table 1. In Table 2 we present an illustrative selection of these probabilities.

The effect of age goes in the expected direction. In general, the younger the person the more mobile he is. For example, at sample means of the economic variables, a person aged 20 to 29 has between 15 and 20 percent higher probability of doing a short distance move than a person aged 30 to 44. As for the effect of education, the more educated the more they are likely to move (except to small towns, particularly if they live in large ones).

Overall, at sample means of the economic variables, people with 11+ years of education are 40 to 50 percent more likely to move within their region than people with 8 years of education. It is interesting to note that at average economic conditions, the probability of migrating is higher for people living in small towns than for people living in medium or large cities.[8] More educated people tend to move to larger towns: those with 11+ years of education move mostly towards medium size towns (if aged 20 to 29) or large towns (if aged 30 to 64), while those with 8 years of education tend to move to small towns (if aged 20 to 29) or medium ones (if aged 30 to 64). Note that the moves from small (or medium) to small towns may be reflecting moves towards the outskirts of large towns.

We now turn to consider the effects of the region's economic conditions.[9] The results show that high regional unemployment rates encourage people to move from small or medium towns to small or medium ones, but discourage moves from small or medium towns to large ones. These effects are stronger for people with little education. Specifically, the probability of moving to medium size towns increases by around 85, 60, and 40 percent, respectively, according to level of education when the unemployment rate is set at its sample period peak.

[8] At the same time, given the large fraction of the population living in large cities, in the sample of migrants we may well observe that the proportion of migrants coming from large cities is higher than the proportion of migrants leaving small towns.

[9] Elasticities with respect to migration probabilities can be constructed for the continuous economic variables. Let $\widehat{p}_{kj}$ be the predicted probability of moving from k to j, z the economic variable of interest, and $\widehat{\beta}_{kj}$ its associated estimated coefficient in the odd ratio for destination j from k. The estimated elasticity at z will be given by:

$$\varepsilon_{zkj} = \widehat{\beta}_{kj} z (1 - \widehat{p}_{kj}).$$

The effect of the proportion of regional employment in the service sector is clear cut, increasing significantly the probability of moving to large towns, where most of the new service sector jobs are, from towns of any size, and diminishing the probabilities of moves from small or medium to small or medium towns. It is a sizeable effect that more than doubles the probability of going to large towns from towns of any size, when the share of employment in services is changed from the average to the maximum value observed in the sample period. At its maximum, it brings the probability of moving to a large town in a given year, for a man aged 30 to 44 with 11+ years of education, to 4.43 percent. The positive effects on the probability of a short distance move of the share of employment in services and the unemployment rate show how people move in response to economic incentives, and in particular to employment prospects.[10]

High house prices are also associated with larger migration probabilities, but in a different direction, making people leave large cities towards smaller towns, where house prices are usually lower.[11] The predicted probabilities indicate that the probability of migrating from a large town to a small or medium one approximately trebles when house prices are at their peak; for example, taking it to 3.89 percent, for an individual aged 20 to 29 with 11+ years of education. In general, older people tend to move more than younger people when house prices are high, presumably because a higher

[10] Regretedly there is no information in the RVD on whether individuals are unemployed or not.

[11] Increasing house prices have been associated with increasing house price differentials between small and large towns (which would be a better variable if available), according to house price data by size of town of residence for the period 1987-1995 published by the Ministry of Public Works and Transport at the national level.

fraction of them own a house or command higher income. The estimated effects of house prices also tend to show that high house prices decrease the probability of moving to large towns (although the estimated coefficients are not significant), and increase the probability of moving from small or medium to small or medium towns. Again, these moves to small towns may indicate moves to small towns in the outskirts of large cities when house prices are high. We suspect, nevertheless, that the estimated house price effects may be somewhat upward biased, because they may be picking up the effect of an omitted activity or real per capita income variable.[12] Unfortunately, given the inability to have level measures of such variables that are comparable in real terms across regions, it is difficult to pin down the extent of the bias.

5 Conclusions

We estimated a multinomial model of the probability of intra-regional migration by town size of origin and destination. The model is identified from a comparison of the distribution of characteristics in a sample of migrants with the corresponding distribution of a representative sample of the population of migrants and nonmigrants. Our explanatory variables are either discrete individual indicators or continuous aggregate regional variables. Since the latter can be regarded as linear combinations of region-specific time dummies, our dataset is one with "many observations per cell". We discussed two asymptotically equivalent nonlinear estimation methods based, respectively, on minimum chi-square and maximum likelihood principles. We only

[12]Results in Bover (1993) indicate that the increase in real per capita income has been the major source of increase in house prices during the second half of the eighties in Spain.

reported results for the former, since it was computationally simpler, and the two produced very similar results in the instances where we calculated both of them.

We found that house prices have a positive effect on intra-regional migration, making people leave large cities towards small and medium ones, where housing costs are lower. The share of employment in the services industry is also found to have a positive effect on short distance moves, inducing moves towards large cities where most of the employment opportunities in the service sector are. Unemployment induces also movements, mainly among the people with low education, towards medium size towns. Finally, an increasing educational level is found to lead to increasing mobility.

Some of these moves, prompted by high house prices, from large cities to smaller towns do not necessarily involve a change of job. However, the estimated responses to unemployment and, mainly, to the share of employment in services indicate that (in contrast to the extended view of low mobility) many Spaniards move in response to economic activity, in particular in search of better employment prospects. These moves are not necessarily between regions as they used to be, since employment opportunities in the service, non-manual sector have increased substantially within all regions, but mainly in large cities.

FIGURE 1: INTRA-REGIONAL MIGRATIONS
(In per capita terms)
1962-1995

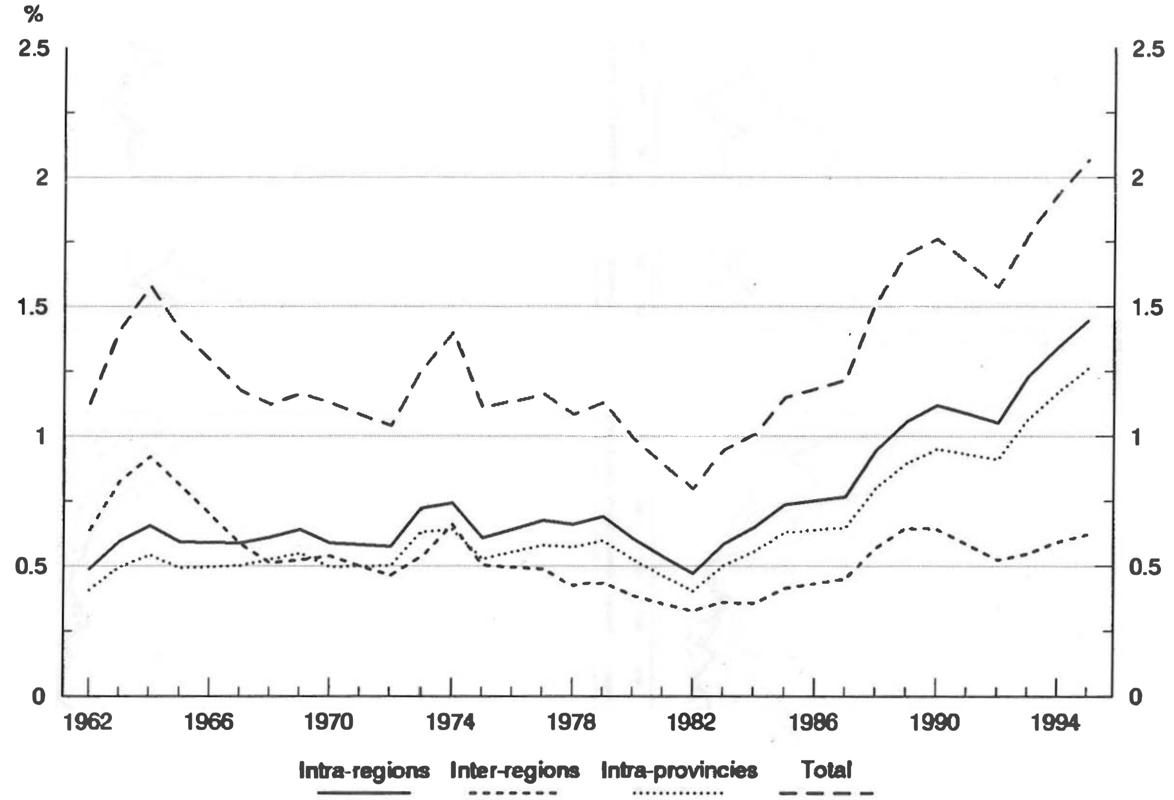

FIGURE 2: INTRA-REGIONAL MIGRATIONS
(in per capita terms)
1962-1995

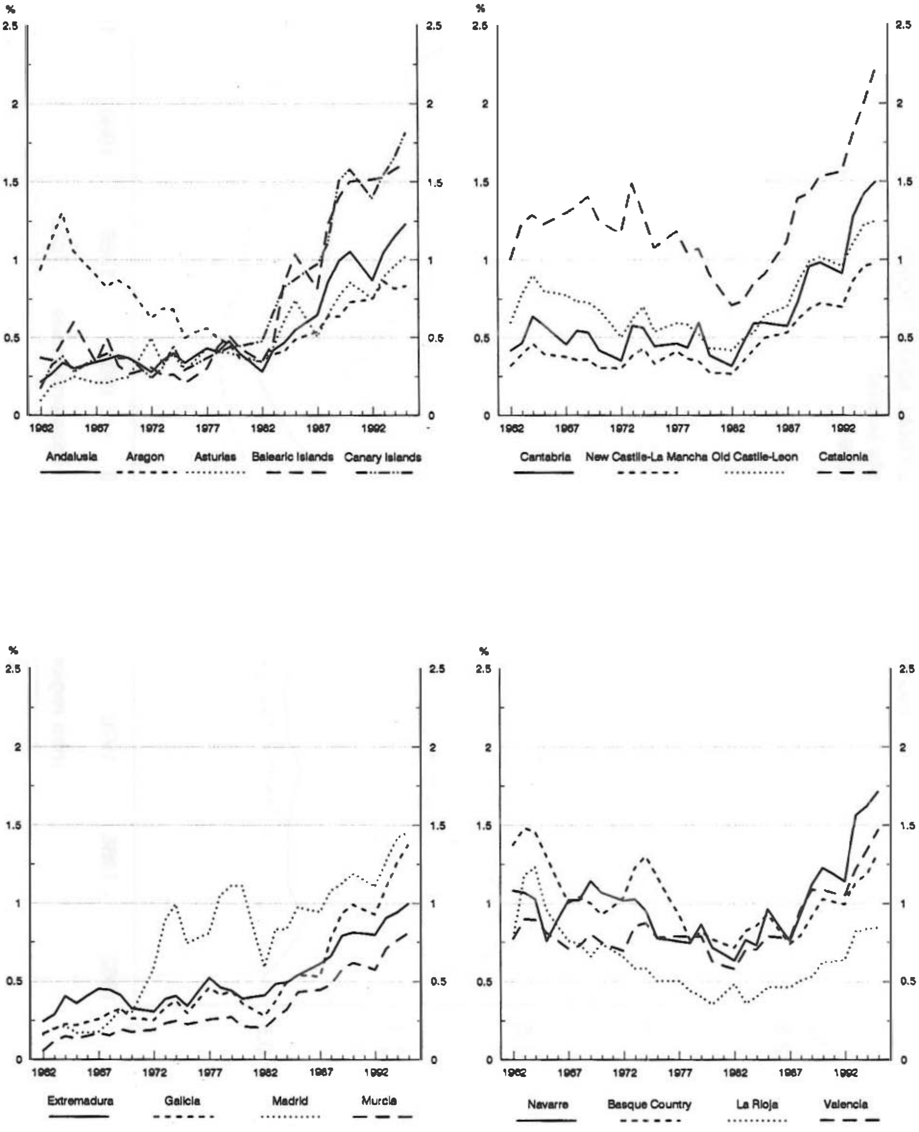

Table 1
Minimum Distance Estimates for the Probability
of Intra-Regional Migration,
by Size of Town of Origin and Destination

Variable	Moves to small towns			Moves to medium towns			Moves to large towns		
	From small	From medium	From large	From small	From medium	From large	From small	From medium	From large
Constant	-5.04	-5.42	-5.86	-5.49	-5.31	-7.19	-7.88	-7.92	-8.36
	(41.5)[1]	(39.8)	(33.6)	(33.2)	(33.6)	(36.8)	(40.5)	(43.1)	(39.1)
Aged 30 to 44	-0.67	-0.63	-0.44	-0.32	-0.34	-0.40	-0.08	-0.13	-0.36
	(6.4)	(5.3)	(3.2)	(2.8)	(3.2)	(3.0)	(0.6)	(1.0)	(2.0)
Aged 45 to 64	-1.99	-1.96	-1.59	-1.84	-1.91	-2.11	-1.12	-1.15	-1.67
	(13.0)	(11.4)	(8.6)	(12.1)	(13.5)	(11.6)	(6.3)	(6.9)	(7.8)
8 years of education	0.22	0.28	-0.14	0.45	0.47	0.38	0.40	0.43	0.02
	(1.6)	(1.8)	(0.9)	(3.0)	(3.3)	(2.7)	(1.8)	(2.2)	(0.1)
$\geq$ 11 years of education	0.12	0.11	-0.49	1.00	1.05	0.82	1.25	1.17	1.22
	(0.8)	(0.6)	(2.8)	(7.2)	(7.7)	(6.0)	(7.2)	(6.9)	(6.8)
% of employment in services $(t-1)$	-0.01	-0.01	0.00	-0.02	-0.03	0.01	0.03	0.03	0.03
	(3.5)	(3.4)	(0.0)	(4.5)	(6.8)	(1.4)	(9.0)	(9.6)	(9.0)
Unemployment rate $(t-1)$	0.03	0.03	0.00	0.04	0.05	0.04	-0.02	-0.02	0.00
	(6.0)	(5.5)	(0.1)	(8.6)	(10.4)	(10.0)	(2.7)	(2.8)	(0.6)
Unemp.$(t-1)$ * 8 years of education	-0.01	-0.01	-0.00	-0.01	-0.01	-0.01	0.01	0.01	0.02
	(2.1)	(2.1)	(0.2)	(1.1)	(1.3)	(2.3)	(0.8)	(0.9)	(1.8)
Unemp.$(t-1)$ * $\geq$ 11 years of education	-0.01	-0.01	0.01	-0.02	-0.02	-0.02	0.01	0.02	0.01
	(1.6)	(1.4)	(1.1)	(2.8)	(3.1)	(3.7)	(1.8)	(2.3)	(1.3)
House prices $(t-1)$	0.28	0.31	0.70	0.34	0.51	0.74	-0.05	-0.05	-0.02
	(4.0)	(3.8)	(10.1)	(4.1)	(6.2)	(10.2)	(0.6)	(0.5)	(0.3)
House prices $(t-1)$ * Aged 30 to 44	0.20	0.19	-0.06	0.17	0.18	0.10	0.15	0.18	0.25
	(2.8)	(2.4)	(0.6)	(2.1)	(2.5)	(1.1)	(1.6)	(2.0)	(2.0)
House prices $(t-1)$ * Aged 45 to 64	0.18	0.18	-0.17	0.40	0.44	0.43	0.21	0.23	0.49
	(1.5)	(1.4)	(1.2)	(3.6)	(4.5)	(3.3)	(1.8)	(2.1)	(3.2)

[1]t-ratios in parentheses.

Table 2
Some Predicted Probabilities of Intra-Regional Migration (%)

(a)
8 Years of Education, Age 20-29, Economic Variables at Sample Means

	Destination			
Origin	Small	Medium	Large	Total
Small	0.72	0.57	0.28	1.57
Medium	0.50	0.54	0.28	1.32
Large	0.59	0.62	0.22	1.43
Total	1.81	1.73	0.78	4.32

(b)
8 Years of Education, Age 30-44, Economic Variables at Sample Means

	Destination			
Origin	Small	Medium	Large	Total
Small	0.48	0.52	0.31	1.31
Medium	0.34	0.49	0.31	1.14
Large	0.35	0.47	0.21	1.03
Total	1.17	1.48	0.83	3.48

(c)
8 Years of Education, Age 20-29, % Employment in Services at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.55	0.37	0.57	1.49
Medium	0.37	0.27	0.57	1.21
Large	0.59	0.70	0.44	1.73
Total	1.51	1.34	1.58	4.43

(d)
8 Years of Education, Age 30-44, % Employment in Services at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.37	0.33	0.63	1.33
Medium	0.25	0.25	0.64	1.14
Large	0.35	0.53	0.43	1.31
Total	0.97	1.11	1.70	3.78

Table 2 (contd.)
Some Predicted Probabilities of Intra-Regional Migration (%)

(e)

$\geq$ 11 Years of Education, Age 20-29, Economic Variables at Sample Means

	Destination			
Origin	Small	Medium	Large	Total
Small	0.68	0.83	0.72	2.23
Medium	0.45	0.79	0.68	1.92
Large	0.50	0.83	0.62	1.95
Total	1.63	2.45	2.02	6.10

(f)

$\geq$ 11 Years of Education, Age 30-44, Economic Variables at Sample Means

	Destination			
Origin	Small	Medium	Large	Total
Small	0.45	0.75	0.81	2.01
Medium	0.31	0.71	0.76	1.78
Large	0.29	0.63	0.60	1.52
Total	1.05	2.09	2.17	5.31

(g)

$\geq$ 11 Years of Education, Age 20-29, % Employment in Services at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.52	0.53	1.47	2.52
Medium	0.33	0.40	1.39	2.12
Large	0.49	0.93	1.28	2.70
Total	1.34	1.86	4.14	7.34

(h)

$\geq$ 11 Years of Education, Age 30-44, % Employment in Services at Maximum

	Destination			Total
Origin	Small	Medium	Large	
Small	0.34	0.48	1.65	2.47
Medium	0.23	0.35	1.54	2.12
Large	0.29	0.71	1.24	2.24
Total	0.86	1.54	4.43	6.83

Table 2 (contd.)
Some Predicted Probabilities of Intra-Regional Migration (%)

(i)
8 Years of Education, Age 20-29, House Prices at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	1.09	0.95	0.26	2.30
Medium	0.80	1.17	0.26	2.23
Large	1.68	1.86	0.20	3.74
Total	3.57	3.98	0.72	8.27

(j)
8 Years of Education, Age 30-44, House Prices at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.99	1.10	0.36	2.45
Medium	0.73	1.38	0.38	2.49
Large	0.91	1.66	0.29	2.86
Total	2.63	4.14	1.03	7.80

(k)
8 Years of Education, Age 20-29, Unemployment Rate at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.86	0.91	0.25	2.02
Medium	0.61	0.98	0.25	1.84
Large	0.58	0.91	0.29	1.78
Total	2.05	2.80	0.79	5.64

(l)
8 Years of Education, Age 30-44, Unemployment Rate at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.57	0.83	0.28	1.68
Medium	0.42	0.88	0.28	1.58
Large	0.34	0.70	0.28	1.32
Total	1.33	2.41	0.84	4.58

Table 2 (contd.)
Some Predicted Probabilities of Intra-Regional Migration (%)

(m)
≥ 11 Years of Education, Age 20-29, House Prices at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	1.02	1.38	0.67	3.07
Medium	0.72	1.70	0.63	3.05
Large	1.40	2.49	0.58	4.47
Total	3.14	5.57	1.88	10.59

(n)
≥ 11 Years of Education, Age 30-44, House Prices at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.92	1.60	0.94	3.46
Medium	0.66	1.99	0.92	3.57
Large	0.76	2.21	0.84	3.81
Total	2.34	5.8	2.7	10.84

(o)
≥ 11 Years of Education, Age 20-29, Unemployment Rate at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.84	1.16	0.70	2.7
Medium	0.58	1.21	0.69	2.48
Large	0.56	1.09	0.74	2.39
Total	1.98	3.46	2.13	7.57

(p)
≥ 11 Years of Education, Age 30-44, Unemployment Rate at Maximum

	Destination			
Origin	Small	Medium	Large	Total
Small	0.56	1.05	0.79	2.4
Medium	0.40	1.09	0.77	2.26
Large	0.33	0.83	0.72	1.88
Total	1.29	2.97	2.28	6.54

Appendix 1: Database description

Individual variables

Source: "Estadística de Variaciones Residenciales", INE.

Size of town. Three groups:

- Small: less than 10,000 inhabitants.
- Medium: 10 to 100 thousand inhabitants.
- Large: more than 100,000 inhabitants.

Education. Three categories:

- Five or less years of education
- Eight years of education.
- Eleven or more years of education.

Age. Three groups:

- 20 to 29 years old.
- 30 to 44 years old.
- 45 to 64 years old.

Aggregate and regional variables

Share of employment in the service sector, by regions.
Source: "Encuesta de Población Activa", INE.

Unemployment rates, by regions.
Source: "Encuesta de Población Activa", INE.

House prices. Numerator: Average regional house price of new dwellings per square meter in capitals of provinces. (Source: Sociedad de Tasación.) Denominator: National CPI (base 1992), INE.

Table A1.1

Frequencies of the Variables in the 10% Random Sample from the Residential Variation Data (Size=52135) and Population Frequencies from the Labour Force Survey

Variable	RVD Absolute Frequency	RVD Relative Frequency	LFS Relative Frequency
Year			
1988	11474	22.01	24.64
1989	12940	24.82	24.97
1990	14034	26.92	25.05
1992	13687	26.25	25.34
Region			
Andalusia	8009	15.36	17.31
Aragon	1075	2.06	3.15
Asturias	1164	2.23	2.99
Balearic Islands	1287	2.47	1.70
Canary Islands	2896	5.55	3.83
Cantabria	640	1.23	1.37
New Castile-La Mancha	1466	2.81	4.29
Old Castile-Leon	3457	6.63	6.95
Catalonia	11769	22.57	15.78
Basque Country	2706	5.19	5.97
Extremadura	1051	2.02	2.84
Galicia	3132	6.01	7.19
Madrid	6720	12.89	12.44
Murcia	733	1.41	2.52
Navarre	764	1.47	1.37
La Rioja	187	0.36	0.66
Valencia	5079	9.74	9.63

Table A1.1 (contd.)

Frequencies of the Variables in the 10% Random Sample from the Residential Variation Data (Size=52135) and Population Frequencies from the Labour Force Survey

Variable	RVD		LFS
	Absolute Frequency	Relative Frequency	Relative Frequency
Size of town of origin and destination			
From small	15572	29.87	25.87
to small	6081		-
to medium	5693		-
to large	3798		-
From medium	16866	32.35	32.62
to small	5116		-
to medium	7091		-
to large	4659		-
From large	19697	37.78	41.51
to small	6476		-
to medium	8760		-
to large	4461		-
Age			
20 to 29 years old	22614	43.38	28.36
30 to 44 years old	20773	39.84	32.48
45 to 64 years old	8748	16.78	39.16
Education			
5 years or less	21055	40.39	55.98
8 years	11130	21.35	17.15
11 years or more	19950	38.27	26.87

Table A1.2

Summary Statistics for the Economic Variables
(1988, 1989, 1990, 1992)

Variable	Mean	Standard Deviation	Minimum	Maximum
Unemployment rate $(t-1)$	17.24	5.08	9.60	30.8
% of Employment in services $(t-1)$	52.76	7.94	37.95	73.21
House prices $(t-1)$	1.28	0.41	0.74	2.80

Appendix 2:
Detailed Predicted Probabilities of Intra-Regional Migration

Table A2.1
Predicted Probabilities of Moving to a Small Town (%)
(Within Region Migration)

			1. Economic variables at sample means	Economic variables at sample means[1] except for the following changes					
				2. Services min.=37.95	3. Services max.=73.21	4. Unemployment min.=9.6	5. Unemployment max.=30.8	6. House prices min.=0.74	7. House prices max.=2.8
From a small town	Aged 20 to 29	≤ 5 years of educ.	0.74	0.89	0.57	0.60	1.06	0.64	1.11
		8 years of educ.	0.72	0.87	0.55	0.65	0.86	0.62	1.09
		≥ 11 years of educ.	0.68	0.82	0.52	0.60	0.84	0.58	1.02
	Aged 30 to 44	≤ 5 years of educ.	0.49	0.59	0.38	0.40	0.71	0.38	1.01
		8 years of educ.	0.48	0.58	0.37	0.43	0.57	0.37	0.99
		≥ 11 years of educ.	0.45	0.54	0.34	0.40	0.56	0.35	0.92
	Aged 45 to 64	≤ 5 years of educ.	0.13	0.15	0.10	0.10	0.18	0.10	0.25
		8 years of educ.	0.12	0.15	0.10	0.11	0.15	0.10	0.25
		≥ 11 years of educ.	0.12	0.14	0.09	0.10	0.15	0.09	0.23
From a medium size town	Aged 20 to 29	≤ 5 years of educ.	0.49	0.61	0.36	0.39	0.73	0.42	0.78
		8 years of educ.	0.50	0.62	0.37	0.45	0.61	0.42	0.80
		≥ 11 years of educ.	0.45	0.57	0.33	0.39	0.58	0.38	0.72
	Aged 30 to 44	≤ 5 years of educ.	0.34	0.42	0.25	0.27	0.50	0.26	0.72
		8 years of educ.	0.34	0.43	0.25	0.31	0.42	0.26	0.73
		≥ 11 years of educ.	0.31	0.39	0.23	0.27	0.40	0.24	0.66
	Aged 45 to 64	≤ 5 years of educ.	0.09	0.11	0.06	0.07	0.13	0.07	0.19
		8 years of educ.	0.09	0.11	0.07	0.08	0.11	0.07	0.19
		≥ 11 years of educ.	0.08	0.10	0.06	0.07	0.11	0.06	0.17
From a large town	Aged 20 to 29	≤ 5 years of educ.	0.70	0.70	0.70	0.70	0.70	0.48	1.99
		8 years of educ.	0.59	0.59	0.59	0.60	0.58	0.41	1.68
		≥ 11 years of educ.	0.50	0.50	0.49	0.46	0.56	0.34	1.40
	Aged 30 to 44	≤ 5 years of educ.	0.42	0.42	0.42	0.42	0.42	0.30	1.08
		8 years of educ.	0.35	0.35	0.35	0.36	0.34	0.25	0.91
		≥ 11 years of educ.	0.29	0.29	0.29	0.28	0.33	0.21	0.76
	Aged 45 to 64	≤ 5 years of educ.	0.12	0.12	0.12	0.12	0.12	0.09	0.26
		8 years of educ.	0.10	0.10	0.10	0.10	0.10	0.07	0.22
		≥ 11 years of educ.	0.08	0.08	0.08	0.08	0.09	0.06	0.19

1. The means of the regional economic variables are: % of employment in services = 52.76, unemployment rate = 17.24, house prices =1.28.

Table A2.1 (contd.)
Predicted Probabilities of Moving to a Medium Size Town (%)
(Within Region Migration)

			1. Economic variables at sample means	Economic variables at sample means[1] except for the following changes					
				2. Services min.=37.95	3. Services max.=73.21	4. Unemployment min.=9.6	5. Unemployment max.=30.8	6. House prices min.=0.74	7. House prices max.=2.8
From a small town	Aged 20 to 29	≤ 5 years of educ.	0.42	0.58	0.27	0.30	0.74	0.35	0.70
		8 years of educ.	0.57	0.79	0.37	0.44	0.91	0.48	0.95
		≥ 11 years of educ.	0.83	1.14	0.53	0.69	1.16	0.69	1.38
	Aged 30 to 44	≤ 5 years of educ.	0.38	0.52	0.24	0.27	0.67	0.29	0.81
		8 years of educ.	0.52	0.71	0.33	0.40	0.83	0.39	1.10
		≥ 11 years of educ.	0.75	1.03	0.48	0.62	1.05	0.57	1.60
	Aged 45 to 64	≤ 5 years of educ.	0.11	0.15	0.07	0.08	0.20	0.07	0.34
		8 years of educ.	0.15	0.21	0.10	0.12	0.24	0.10	0.47
		≥ 11 years of educ.	0.22	0.31	0.14	0.18	0.31	0.15	0.68
From a medium size town	Aged 20 to 29	≤ 5 years of educ.	0.39	0.65	0.20	0.26	0.79	0.30	0.85
		8 years of educ.	0.54	0.89	0.27	0.39	0.98	0.41	1.17
		≥ 11 years of educ.	0.79	1.29	0.40	0.62	1.21	0.60	1.70
	Aged 30 to 44	≤ 5 years of educ.	0.35	0.58	0.18	0.24	0.71	0.24	1.00
		8 years of educ.	0.49	0.80	0.25	0.35	0.88	0.34	1.38
		≥ 11 years of educ.	0.71	1.16	0.35	0.55	1.09	0.49	1.99
	Aged 45 to 64	≤ 5 years of educ.	0.10	0.17	0.05	0.07	0.21	0.06	0.44
		8 years of educ.	0.14	0.24	0.07	0.10	0.26	0.09	0.61
		≥ 11 years of educ.	0.21	0.34	0.10	0.16	0.32	0.12	0.89
From a large town	Aged 20 to 29	≤ 5 years of educ.	0.55	0.50	0.62	0.39	0.98	0.37	1.64
		8 years of educ.	0.62	0.57	0.70	0.50	0.91	0.42	1.86
		≥ 11 years of educ.	0.83	0.76	0.93	0.71	1.09	0.56	2.49
	Aged 30 to 44	≤ 5 years of educ.	0.42	0.38	0.47	0.30	0.74	0.27	1.46
		8 years of educ.	0.47	0.43	0.53	0.38	0.70	0.30	1.66
		≥ 11 years of educ.	0.63	0.58	0.71	0.54	0.83	0.40	2.21
	Aged 45 to 64	≤ 5 years of educ.	0.12	0.11	0.13	0.08	0.21	0.06	0.69
		8 years of educ.	0.13	0.12	0.15	0.10	0.20	0.07	0.78
		≥ 11 years of educ.	0.18	0.16	0.20	0.15	0.23	0.09	1.05

1. The means of the regional economic variables are: % of employment in services =52.76, unemployment rate= 17.24, house prices =1.28

Table A2.1(contd.)
Predicted Probabilities of Moving to a Large Town (%)
(Within Region Migration)

			1. Economic variables at sample means	Economic variables at sample means[1] except for the following changes					
				2. Services min.=37.95	3. Services max.=73.21	4. Unemployment min.=9.6	5. Unemployment max.=30.8	6. House prices min.=0.74	7. House prices max.=2.8
From a small town	Aged 20 to 29	≤ 5 years of educ.	0.16	0.10	0.33	0.19	0.13	0.17	0.15
		8 years of educ.	0.28	0.17	0.57	0.30	0.25	0.29	0.26
		≥ 11 years of educ.	0.72	0.43	1.47	0.74	0.70	0.75	0.67
	Aged 30 to 44	≤ 5 years of educ.	0.18	0.11	0.37	0.21	0.14	0.17	0.21
		8 years of educ.	0.31	0.19	0.63	0.33	0.28	0.30	0.36
		≥ 11 years of educ.	0.81	0.49	1.65	0.83	0.79	0.77	0.94
	Aged 45 to 64	≤ 5 years of educ.	0.07	0.04	0.14	0.08	0.06	0.06	0.09
		8 years of educ.	0.12	0.07	0.25	0.13	0.11	0.11	0.15
		≥ 11 years of educ.	0.32	0.19	0.64	0.32	0.31	0.29	0.40
From a medium size town	Aged 20 to 29	≤ 5 years of educ.	0.16	0.09	0.32	0.18	0.12	0.16	0.15
		8 years of educ.	0.28	0.17	0.57	0.30	0.25	0.29	0.26
		≥ 11 years of educ.	0.68	0.41	1.39	0.68	0.69	0.70	0.63
	Aged 30 to 44	≤ 5 years of educ.	0.18	0.10	0.36	0.20	0.14	0.16	0.21
		8 years of educ.	0.31	0.19	0.64	0.33	0.28	0.29	0.38
		≥ 11 years of educ.	0.76	0.45	1.54	0.76	0.77	0.71	0.92
	Aged 45 to 64	≤ 5 years of educ.	0.07	0.04	0.14	0.08	0.05	0.06	0.09
		8 years of educ.	0.12	0.07	0.25	0.13	0.11	0.11	0.16
		≥ 11 years of educ.	0.29	0.17	0.60	0.29	0.30	0.27	0.39
From a large town	Aged 20 to 29	≤ 5 years of educ.	0.15	0.09	0.32	0.15	0.16	0.16	0.15
		8 years of educ.	0.22	0.13	0.44	0.18	0.29	0.22	0.20
		≥ 11 years of educ.	0.62	0.37	1.28	0.56	0.74	0.63	0.58
	Aged 30 to 44	≤ 5 years of educ.	0.15	0.09	0.31	0.15	0.16	0.13	0.21
		8 years of educ.	0.21	0.12	0.43	0.18	0.28	0.19	0.29
		≥ 11 years of educ.	0.60	0.36	1.24	0.55	0.72	0.54	0.84
	Aged 45 to 64	≤ 5 years of educ.	0.05	0.03	0.11	0.05	0.06	0.04	0.11
		8 years of educ.	0.08	0.04	0.16	0.06	0.10	0.06	0.16
		≥ 11 years of educ.	0.22	0.13	0.46	0.20	0.27	0.17	0.44

1. The means of the regional economic variables are: % of employment in services =52.76, unemployment rate= 17.24, house prices =1.28

Appendix 3

A3.1 Linear Minimum Distance Estimation

Binary Models The MD estimators discussed in Section 2 require nonlinear optimization. Linear MD estimates of α and β can be obtained as follows. Let us first consider for the binary model (2) the transformed migration probabilities using the inverse function of G:

$$d_\ell \equiv G^{-1}[\Pr(y=1|x=\xi_\ell)] = \alpha + z(\xi_\ell)'\beta. \tag{A.1}$$

Also, lét us define their unrestricted estimates, which are given by

$$\widehat{d}_\ell = G^{-1}\left(\frac{p}{\pi_\ell}\widehat{\phi}_\ell\right). \tag{A.2}$$

Letting $d = (d_1, ..., d_{q-1})'$ and $\widehat{d} = (\widehat{d}_1, ..., \widehat{d}_{q-1})'$, by (5) and the delta method we have

$$\sqrt{n}(\widehat{d} - d) \xrightarrow{d} N(0, D\Omega D') \tag{A.3}$$

where $D = \partial d/\partial\phi' = diag\{\delta_1, ...\delta_{q-1}\}$ and $\delta_\ell = p/[\pi_\ell G'(p\phi_\ell/\pi_\ell)]$. Therefore, an alternative linear MD estimator minimizes

$$s_d(\alpha,\beta) = [\widehat{d} - d(\alpha,\beta)]'\left(\widehat{D}\widehat{\Omega}\widehat{D}'\right)^{-1}[\widehat{d} - d(\alpha,\beta)]. \tag{A.4}$$

In ordinary binary choice models, this is similar to the minimum chi-square method first proposed by Berkson (1944) for the logit model (cf. Amemiya, 1985, pp. 275-278.). There is a fundamental difference betwen the two cases, however, since in the case of exogenous sampling the unrestricted estimates of the transformed probabilities d_ℓ are mutually independent, because they are calculated from observations in different cells. In contrast, in our case the $\widehat{d}_\ell$ are functions of the $\widehat{\phi}_\ell$, which are correlated. Nevertheless, $s_d(\alpha,\beta)$ turns out to be a simple linear weighted least squares criterion of the form

$$s_d(\alpha,\beta) = \sum_{\ell=1}^{q-1} \frac{1}{\widehat{\delta}_\ell^2 \widehat{\phi}_\ell} \left(\widehat{d}_\ell - \alpha - z(\xi_\ell)'\beta\right)^2 + \frac{1}{\widehat{\phi}_q} \left(\sum_{\ell=1}^{q-1} \frac{1}{\widehat{\delta}_\ell} \left(\widehat{d}_\ell - \alpha - z(\xi_\ell)'\beta\right)\right)^2 . \tag{A.5}$$

Multinomial Models For the multinomial logit model (14) it is also possible to develop an asymptotically equivalent linear MD estimator along the lines of that considered for the binary model. Note that the log odd ratios are given by

$$d_{j\ell} \equiv \ln \left(\frac{\phi_{j\ell} p_j}{\pi_\ell - \phi_{1\ell} p_1 - \phi_{2\ell} p_2 - \phi_{3\ell} p_3} \right) = \alpha_j + z(\xi_\ell)'\beta_j \ (j = 1,2,3) \tag{A.6}$$

Unrestricted estimates $\widehat{d}_{j\ell}$ can be obtained replacing the $\phi_{j\ell}$ in (A.6) by their sample counterparts. However, since the $\widehat{d}_{j\ell}$ are functions of the cell sample frequencies for all destinations, they are not independent for different j as it happens with the $\widehat{\phi}_{j\ell}$. The implication is that the optimal linear MD criterion in the multinomial case cannot be written as a simple weighted least-squares function.

A3.2 Estimation with Continuous Explanatory Variables

The problem of estimation with continuous explanatory variables does not arise in our empirical analysis, because our continuous variables only vary by region and time, and so they are regarded as functions of dummy variables. It is still of some interest, however, to discuss how the methods used in the paper could be extended in the presence of continuous characteristics.

When there are continuous explanatory variables, one possibility is to assume a parametric density function for x, $f(x,\delta)$ say, with known (or previously estimated) parameter vector δ. The resulting log-likelihood will be of the form

$$L(\alpha,\beta) = \sum_{i=1}^{n} \ln G(\alpha + z_i'\beta) - n \ln \left(\int G(\alpha + \xi'\beta) f(\xi,\delta) d\xi \right). \qquad \text{(A.7)}$$

It contains a multiple integral which except in special cases will not have a closed form. One such special case is when $G(.)$ is the normal probability $\Phi(.)$ (the probit model), and x has a multivariate normal distribution $N(\mu,\Sigma)$, for in this situation it can be shown that

$$L(\alpha,\beta) = \sum_{i=1}^{n} \ln \Phi(\alpha + z_i'\beta) - n \ln \Phi \left(\frac{\alpha + \mu'\beta}{(1+\beta'\Sigma\beta)^{1/2}} \right), \qquad \text{(A.8)}$$

to be maximized subject to

$$\Phi \left(\frac{\alpha + \mu'\beta}{(1+\beta'\Sigma\beta)^{1/2}} \right) = p, \qquad \text{(A.9)}$$

or equivalently

$$\alpha = \alpha(\beta) \equiv -\mu'\beta + \Phi^{-1}(p)(1+\beta'\Sigma\beta)^{1/2}.$$

Although this result can be easily generalized to the case where the distribution of x is a mixture of multivariate normals, and normal mixtures can approximate a large class of distributions, it will often be impractical or too restrictive in applications.

Another possibility is to consider a nonparametric density function for x. In this situation the MD method can be extended by constructing the variable

$$\tilde{d}_i = G^{-1} \left(\frac{\tilde{f}(x_i|y_i=1)p}{\tilde{f}(x_i)} \right) \qquad \text{(A.10)}$$

where $\widetilde{f}(x_i|y_i = 1)$ and $\widetilde{f}(x_i)$ are nonparametric kernel estimators of the conditional and the unconditional densities of x, respectively. In the context of our paper, the former would be obtained from the sample of migrants and the latter from the labour force surveys. Note that $\widetilde{d}_i$ is a smoothed version of the variable $\widehat{d}_i$ introduced above for the discrete case. Linear least-squares estimates of α and β obtained by regressing $\widetilde{d}_i$ on x_i will be consistent and asymptotically normal under standard regularity conditions of the type discussed by Newey and McFadden (1995).

A3.3 MD when the Distribution of Characteristics is Estimated

Binary Models Suppose that π_ℓ is not known with certainty, but an unrestricted estimate $\widehat{\pi}_\ell$ (a sample frequency) is available from a complementary data set of size m, independent of the size-n sample of migrants. Let us first consider the form of the covariance matrix of the model's constraints evaluated at the true values of α and β:

$$\widehat{e}_\ell = \widehat{\phi}_\ell - \frac{\widehat{\pi}_\ell}{p} G\left(\alpha + z(\xi_\ell)'\beta\right) \quad (\ell = 1, ..., q). \tag{A.11}$$

Letting $G_\ell = G\left(\alpha + z(\xi_\ell)'\beta\right)$, we have

$$Var(\widehat{e}_\ell) = Var(\widehat{\phi}_\ell) + \frac{G_\ell^2}{p^2} Var(\widehat{\pi}_\ell) = \frac{1}{n}\phi_\ell(1-\phi_\ell) + \frac{1}{m}\frac{G_\ell^2}{p^2}\pi_\ell(1-\pi_\ell) \tag{A.12}$$

and

$$Cov(\widehat{e}_\ell, \widehat{e}_{\ell'}) = -\frac{1}{n}\phi_\ell\phi_{\ell'} - \frac{1}{m}\frac{G_\ell G_{\ell'}}{p^2}\pi_\ell\pi_{\ell'}. \tag{A.13}$$

Now let us define $\widehat{e} = (\widehat{e}_1, ..., \widehat{e}_{q-1})'$. Then by the central limit theorem and the delta method we have

$$\sqrt{n}\widehat{e} \xrightarrow{d} N\left(0, \Omega + sD_\pi(\Lambda_\pi - \pi\pi')D_\pi'\right) \tag{A.14}$$

where as before $\Omega = \Lambda - \phi\phi'$ and $\Lambda = diag\{\phi_1, ..., \phi_{q-1}\}$, and similarly $\pi = (\pi_1, ..., \pi_{q-1})'$ and $\Lambda_\pi = diag\{\pi_1, ..., \pi_{q-1}\}$. Moreover, $D_\pi = diag\{G_1/p, ..., G_{q-1}/p\} = diag\{\phi_1/\pi_1, ..., \phi_{q-1}/\pi_{q-1}\}$ and $s = p\lim(n/m)$.

In the analysis conducted in the paper we have assumed that $s = 0$. When $s \neq 0$, the additional term $D_\pi(\Lambda_\pi - \pi\pi')D_\pi'$ accounts for sampling error in the $\widehat{\pi}_\ell$. Also note that

$$D_\pi(\Lambda_\pi - \pi\pi')D_\pi' = \Lambda\Lambda_\pi^{-1}\Lambda - \phi\phi'. \tag{A.15}$$

When $s \neq 0$, the estimates discussed in the main text are not asymptotically efficient (since they do not use an optimal weight matrix) but remain consistent. Their asymptotic covariance matrix is given by

$$V_R = VM_RV \tag{A.16}$$

where V corresponds to the expression given in (9) and M_R is given by

$$M_R = \left(\frac{\partial\phi(\alpha,\beta)}{\partial(\alpha,\beta')'}\right)' \Omega^{-1}\left(\Omega + sD_\pi(\Lambda_\pi - \pi\pi')D_\pi'\right)\Omega^{-1}\left(\frac{\partial\phi(\alpha,\beta)}{\partial(\alpha,\beta')'}\right). \tag{A.17}$$

When $s = 0$ $M_R = V^{-1}$ and the formula in (9) is valid. However, when $s \neq 0$ the standard errors obtained under the assumption that $s = 0$ are inconsistent. Consistent standard errors can be calculated from the sample counterpart of V_R.

Asymptotically efficient estimates of α and β can be obtained as the minimizers of the following two-sample asymptotic least-squares criterion (see Gourieroux and Monfort, 1995, 9.1):

$$s_R(\alpha,\beta) = \widehat{e}'\left(\widehat{\Omega} + \frac{n}{m}\widehat{D}_\pi(\widehat{\Lambda}_\pi - \widehat{\pi}\widehat{\pi}')\widehat{D}_\pi'\right)^{-1}\widehat{e} \tag{A.18}$$

where the hats denote sample counterparts of the corresponding population characteristics. This criterion differs from that in (7) by the addition of the second term in the weight matrix. The difference with the estimates reported in the paper can be expected to be smaller the smaller is the value of n/m.

Multinomial Models Let us define

$$\widehat{e}_{j\ell} = \widehat{\phi}_{j\ell} - \frac{\widehat{\pi}_\ell}{p_j} G_j\left(z(\xi_\ell);\alpha,\beta\right) \;\; (j=1,2,3;\; \ell = 1,...,q) \tag{A.19}$$

and $\widehat{e}_j = (\widehat{e}_{j1},...,\widehat{e}_{j(q-1)})'$ for $j = 1,2,3$. The main difference with the case when the probabilities π_ℓ are known is that now the vectors $\widehat{e}_1$, $\widehat{e}_2$ and $\widehat{e}_3$ are not independent since they all depend on the same $\widehat{\pi}_\ell$ which are stochastic in this case. If we write $\widehat{e}_j = \widehat{\phi}_j - D_{\pi j}\widehat{\pi}$ where

$$D_{\pi j} = diag\{G_{j1}/p_j,...,G_{j(q-1)}/p_j\} = diag\{\phi_{j1}/\pi_1,...,\phi_{j(q-1)}/\pi_{q-1}\} = \Lambda_j\Lambda_\pi^{-1},$$

we thus have

$$Var(\widehat{e}_j) = \frac{1}{n_j}\Omega_j + \frac{1}{m}D_{\pi j}(\Lambda_\pi - \pi\pi')D_{\pi j}' \;\; (j=1,2,3) \tag{A.20}$$

and

$$Cov(\widehat{e}_j,\widehat{e}_k) = \frac{1}{m}D_{\pi j}(\Lambda_\pi - \pi\pi')D_{\pi k}' \;\; (j,k=1,2,3). \tag{A.21}$$

Therefore, letting $\mathfrak{D}_\pi = (D_{\pi 1}', D_{\pi 2}', D_{\pi 3}')'$, by the central limit theorem and the delta method we have

$$\sqrt{n}\begin{pmatrix}\widehat{e}_1\\ \widehat{e}_2\\ \widehat{e}_3\end{pmatrix} \xrightarrow{d} N\left[0, \begin{pmatrix}\frac{1}{r_1}\Omega_1 & 0 & 0\\ 0 & \frac{1}{r_2}\Omega_2 & 0\\ 0 & 0 & \frac{1}{r_3}\Omega_3\end{pmatrix} + s\mathfrak{D}_\pi(\Lambda_\pi - \pi\pi')\mathfrak{D}_\pi'\right] \tag{A.22}$$

Then the optimal two-sample asymptotic least-squares estimates will minimize the following criterion

$$\begin{pmatrix} \widehat{e}_1 \\ \widehat{e}_2 \\ \widehat{e}_3 \end{pmatrix}' \left[\begin{pmatrix} \frac{n}{n_1}\widehat{\Omega}_1 & 0 & 0 \\ 0 & \frac{n}{n_2}\widehat{\Omega}_2 & 0 \\ 0 & 0 & \frac{n}{n_3}\widehat{\Omega}_3 \end{pmatrix} + \frac{n}{m}\widehat{\mathfrak{D}}_\pi(\widehat{\Lambda}_\pi - \widehat{\pi}\widehat{\pi}')\widehat{\mathfrak{D}}_\pi' \right]^{-1} \begin{pmatrix} \widehat{e}_1 \\ \widehat{e}_2 \\ \widehat{e}_3 \end{pmatrix} \tag{A.23}$$

This criterion differs from that in (19) by the addition of the second term in the weight matrix.

Finally, as in the binary case, we can also obtain robust standard errors for the estimates discussed in the paper which remain consistent when $s \neq 0$.

References

Amemiya, T. (1985): *Advanced Econometrics*, Basil Blackwell, Oxford.

Angrist, J. and A. Krueger (1992): "The Effect of Age at School Entry on Educational Attainment: An Application of Instrumental Variables with Moments from Two Samples", *Journal of the American Statistical Association*, 87, 328-336.

Antolin, P. and O. Bover (1997): "Regional Migration in Spain: The Effect of Personal Characteristics and of Unemployment, Wage and House Price Differentials Using Pooled Cross-Sections", *Oxford Bulletin of Economics and Statistics*, 59, 215-235.

Arellano, M. and C. Meghir (1992): "Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets", *Review of Economic Studies*, 59, 537-559.

Berkson, J. (1944): "Application of the Logistic Function to Bio-Assay", *Journal of the American Statistical Association*, 39, 357-365.

Bover, O. (1993): "An Empirical Model of House Prices in Spain (1976-1991)", *Investigaciones Económicas*, 17, 65-86.

Bover, O. and P. Velilla (1997): "Migrations in Spain: Historical Background and Current Trends", paper presented at the CEPR Conference on European Migration, Munich, November.

Deaton, A. (1998): "Getting Prices Right: What Should Be Done?", *Journal of Economic Perspectives*, 12, 37-46.

Gourieroux, C. and A. Monfort (1995): "*Statistics and Econometric Models*", Vols. 1 and 2, Cambridge University Press.

Imbens, G.W. and T. Lancaster (1994): "Combining Micro and Macro Data in Microeconometric Models", *Review of Economic Studies*, 61, 655-680.

McFadden, D. (1981): "Econometric Models of Probabilistic Choice", in Manski, C.F. and D. McFadden (eds.): *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, Chapter 5, 198-272.

Manski, C.F. and S. R. Lerman (1977): "The Estimation of Choice Probabilities from Choice Based Samples", *Econometrica*, 45, 1977-1988.

Newey, W. and D. McFadden (1994): "Large Sample Estimation and Hypothesis Testing", in R.F. Engle and D.L. McFadden (eds.): *Handbook of Econometrics*, vol. 4, Chapter 36.

Olano, A. (1990): "Las Migraciones Interiores en Fase de Dispersión", *Revista de Economía y Sociología del Trabajo*, 8-9, 86-98.

Ródenas, C. (1994): "Migraciones Interregionales en España 1960-1989", *Revista de Economía Aplicada*, 2, 5-36.

Ródenas, C. and M. Martí (1997): "¿Son Bajos los Flujos Migratorios en España?", *Revista de Economía Aplicada*, 5, 155-171.

WORKING PAPERS (1)

9802 **Pilar Álvarez Canal:** Evolución de la banca extranjera en el período 1992-1996.

9803 **Ángel Estrada y Alberto Urtasun:** Cuantificación de expectativas a partir de las encuestas de opinión.

9804 **Soyoung Kim:** Monetary Policy Rules and Business Cycles.

9805 **Víctor Gómez and Agustín Maravall:** Guide for using the programs TRAMO and SEATS.

9806 **Javier Andrés, Ignacio Hernando and J. David López-Salido:** Disinflation, output and unemployment: the case of Spain.

9807 **Olympia Bover, Pilar García-Perea and Pedro Portugal:** A comparative study of the Portuguese and Spanish labour markets.

9808 **Víctor Gómez and Agustín Maravall:** Automatic modeling methods for univariate series.

9809 **Víctor Gómez and Agustín Maravall:** Seasonal adjustment and signal extraction in economic time series.

9810 **Pablo Hernández de Cos e Ignacio Hernando:** El crédito comercial en las empresas manufactureras españolas.

9811 **Soyoung Kim:** Identifying European Monetary Policy Interactions: French and Spanish System with German Variables.

9812 **Juan Ayuso, Roberto Blanco y Alicia Sanchís:** Una clasificación por riesgo de los fondos de inversión españoles.

9813 **José Viñals:** The retreat of inflation and the making of monetary policy: where do we stand?

9814 **Juan Ayuso, Graciela L. Kaminsky and David López-Salido:** A switching-regime model for the Spanish inflation: 1962-1997.

9815 **Roberto Blanco:** Transmisión de información y volatilidad entre el mercado de futuros sobre el índice Ibex 35 y el mercado al contado.

9816 **M.ª Cruz Manzano and Isabel Sánchez:** Indicators of short-term interest rate expectations. The information contained in the options market. (The Spanish original of this publication has the same number.)

9817 **Alberto Cabrero, José Luis Escrivá, Emilio Muñoz and Juan Peñalosa:** The controllability of a monetary aggregate in EMU.

9818 **José M. González Mínguez y Javier Santillán Fraile:** El papel del euro en el Sistema Monetario Internacional.

9819 **Eva Ortega:** The Spanish business cycle and its relationship to Europe.

9820 **Eva Ortega:** Comparing Evaluation Methodologies for Stochastic Dynamic General Equilibrium Models.

9821 **Eva Ortega:** Assessing the fit of simulated multivariate dynamic models.

9822 **Coral García y Esther Gordo:** Funciones trimestrales de exportación e importación para la economía española.

9823 **Enrique Alberola-Ila and Timo Tyrväinen:** Is there scope for inflation differentials in EMU? An empirical evaluation of the Balassa-Samuelson model in EMU countries.

9824 **Concha Artola e Isabel Argimón:** Titularidad y eficiencia relativa en las manufacturas españolas.

9825 **Javier Andrés, Ignacio Hernando and J. David López-Salido:** The long-run effect of permanent disinflations.

9901 **José Ramón Martínez Resano:** Instrumentos derivados de los tipos *Overnight: call money swaps* y futuros sobre fondos federales.

9902 **J. Andrés, J. D. López-Salido and J. Vallés:** The liquidity effect in a small open economy model.

9903 **Olympia Bover and Ramón Gómez:** Another look at unemployment duration: long-term unemployment and exit to a permanent job. (The Spanish original of this publication has the same number.)

9904 **Ignacio Hernando y Josep A. Tribó:** Relación entre contratos laborales y financieros: Un estudio teórico para el caso español.

9905 **Cristina Mazón and Soledad Núñez:** On the optimality of treasury bond auctions: the Spanish case.

9906 **Nadine Watson:** Bank Lending Channel Evidence at the Firm Level.

9907 **José Viñals:** El marco general de la política monetaria única: racionalidad, consecuencias y cuestiones pendientes.

9908 **Olympia Bover and Manuel Arellano:** Learning about migration decisions from the migrants: an exercise in endogenous samplig and complementary datasets.

(1) Previously published Working Papers are listed in the Banco de España publications catalogue.

Queries should be addressed to: Banco de España
Sección de Publicaciones. Negociado de Distribución y Gestión
Telephone: 91 338 5180
Alcalá, 50. 28014 Madrid