J Popul Econ (2002) 15:357–380

Journal of
Population
Economics

Learning about migration decisions from the migrants: Using complementary datasets to model intra-regional migrations in Spain

Olympia Bover[1], **Manuel Arellano**[2]

[1] Banco de España, Servicio de Estudios, Alcalá 50, 28014 Madrid, Spain
(Fax: +34-91-338-56 78; e-mail: bover@bde.es)
[2] CEMFI, Casado del Alisal 5, 28014 Madrid, Spain
(Fax: +34-91-429-10 56; e-mail: arellano@cemfi.es)

Received: 16 May 1999/Accepted: 26 September 2000

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.

JEL classification: C25, J6, J61

Key words: Intra-regional migration, endogenous sampling, employment in services

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 Figs. 1 and 2). In contrast, until the early 1980's it evolved around a more or less constant level.

All correspondence to Olympia Bover. We are grateful to Samuel Bentolila, Angel Estrada, Enrique Sentana, two referees, the editor, and seminar audiences at University College London, Bank of Spain, Universidad Pablo de Olavide, and Universitat Autònoma de Barcelona for their comments, to Pilar Velilla for her excellent research assistance, and to the Instituto Nacional de Estadística for providing the data. *Responsible editor:* Christoph M. Schmidt.

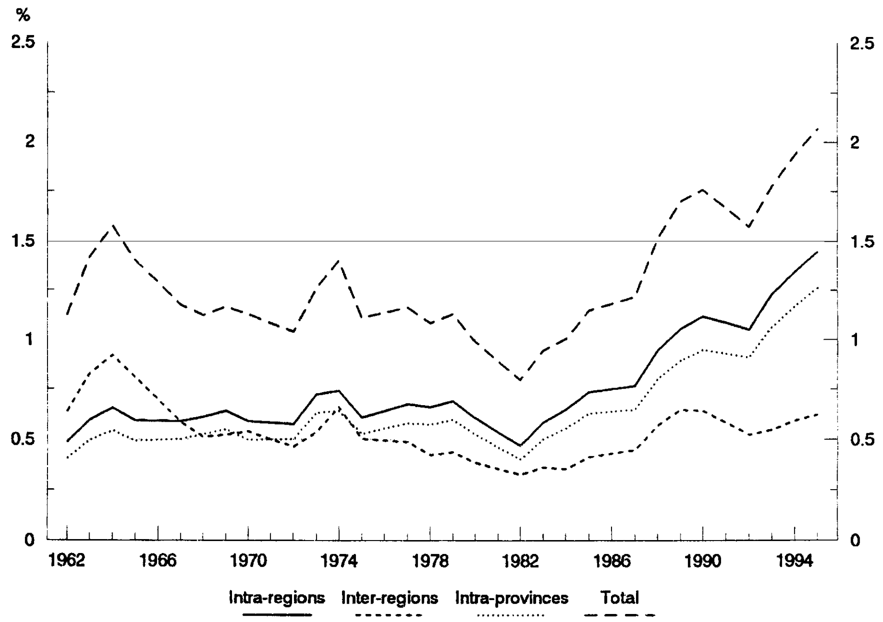


Fig. 1. Intra-regional migrations (in per capita terms) 1962–1995

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]

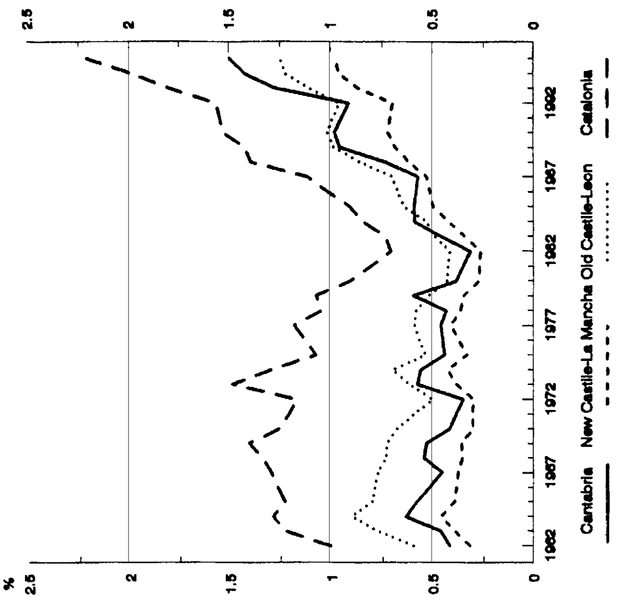
Employment in services climbed from 42% of total employment in 1977 to 61% in 1995. While from 1964 to 1978 the service share of employment grew at an annual rate of 0.79%, from 1980 to 1993 the annual rate was 1.12%, the highest among OECD countries together with Portugal. Furthermore, this increase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

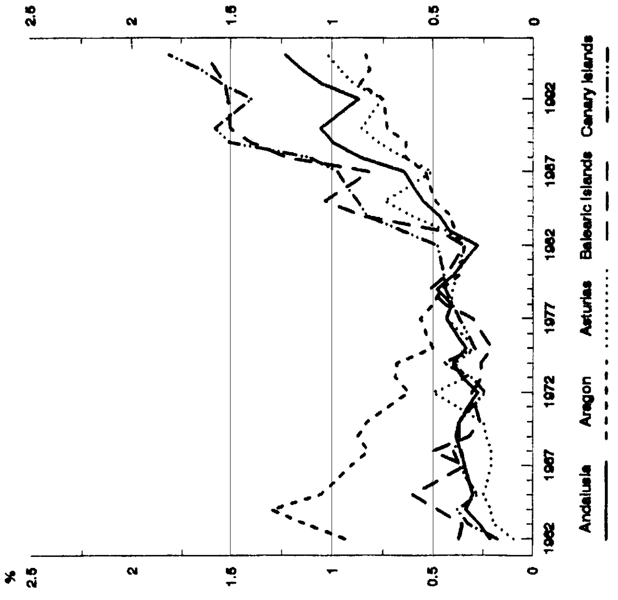
Bover and Velilla reported results from pooled time-series regional data which provided some evidence in support of their view. The aggregate regional data, however, compound effects of this type with moves away from the high housing costs associated with large towns, which will also increase intra-regional 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 a given year is nevertheless a very small percentage of the total population (1.4% 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 Sect. 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, 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 a minimum distance method for multinomial logit which can be implemented as a nonlinear weighted least-squares estimator, and it is asymptotically equivalent to maximum likelihood. In Sect. 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 Sect. 4 we present the empirical results from the various models and report estimated migration probabilities. Finally, Sect. 5 contains the conclusions.


%
2.5
2
1.5
1
0.5
0
1962
1967
1972
1977
1982
1987
1992
Cantabria
New Castile-La Mancha
Old Castile-Leon
Catalonia


%
2.5
2
1.5
1
0.5
0
1962
1967
1972
1977
1982
1987
1992
Andalusia
Aragon
Asturias
Balearic Islands
Canary Islands

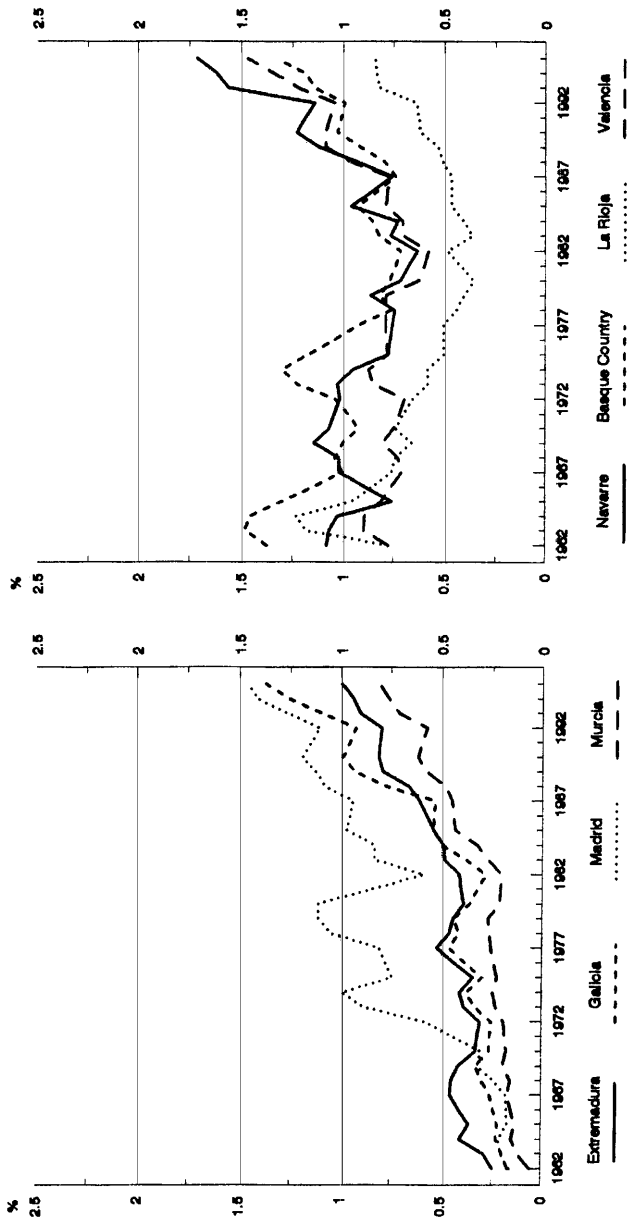


Fig. 2. Intra-regional migrations (in per capita terms) 1962–1995

2. Econometric methods

The purpose of our analysis is to study how migration probabilities vary with the characteristics of the individuals and of their region of residence. We 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. This is represented by the indicator variable y, which we define 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 $y = 0$. Thus our focus is in modelling

$$\Pr(y = j \mid x)$$

for $j = 1, 2, 3$, and estimating the parameters of the chosen models. However, since we only observe migrants, all the individuals in our sample fall in one of the first three categories. Therefore, our strategy is to identify migration probabilities using Bayes theorem and complementary datasets as we explain next.

2.1. Identifying migration probabilities from the migrants

In standard discrete choice analysis with many observations per cell, the probability of a particular type of migration given certain characteristics can be simply estimated as the relative frequency of migrants of the required type among those with the characteristics. Here we lack a suitable sample containing migrants and non-migrants to be able to perform this calculation, but we have two complementary datasets from which the migration probabilities can be estimated in an indirect way. The first one is a database that contains only migrants with information on their characteristics. The other is a representative sample of the total population (migrants and non-migrants) that contains similar characteristics of its members as the first one, but insufficient information on migration.

With these two datasets we cannot calculate estimates of the probabilities of migration as relative frequencies of migrants. Nevertheless, we can use Bayes theorem to estimate them from the ratio of relative frecuencies of the characteristics for those in the first (migrants) and second (migrants and non-migrants) samples. The required estimate is precisely this ratio times the overall migration probability, which is known to us from census information.

Once we have estimates of the probabilities of migration, we can fit models to them by minimum distance methods. That is, by considering a non-linear GLS regression of the estimated probabilities on the functional form of characteristics specified by the model. The form of the optimal MD estimator in our case, however, differs from the standard one. The reason is that relative frequencies for different cells are statistically independent whereas migration probabilities estimated from complementary datasets are not. Below we derive the form of the appropriate MD estimator for our complementary datasets.

Thus the method works in two steps. Firstly one obtains unrestricted estimates of migration probabilities and secondly a model is fitted to them by MD. An alternative approach is to estimate directly the model from the data by maximum likelihood. This is in fact the standard practice in discrete choice analysis, although MD and ML methods are asymptotically equivalent. With our type of complementary datasets ML estimation is also possible, but it is computationally more burdensome than minimum distance. We describe the appropriate form of the likelihood and ML estimation in Appendix 2.

We argue that the complementary datasets approach could be useful for migration analysis more generally, where we often have administrative data sources on migrants only, complemented with general population or labour surveys.

Specifically, let the probability of migration to destination j for an individual with characteristics x be $\Pr(y = j \mid x)$, and let $f(x)$ and $f(x \mid y = j)$ be the marginal and the conditional probability distributions of x given migration to j, respectively. We then have

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

Thus, the migration probabilities can be determined from Eq. (1) given knowledge of $f(x \mid y = j)$, $f(x)$ and $p_j = \Pr(y = j)$.

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

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

where $\alpha = (\alpha_1, \alpha_2, \alpha_3)'$, $\beta = (\beta_1', \beta_2', \beta_3')'$, and z is a vector of explanatory variables which contains some of the x's or functions of them, so that $z = z(x)$.

With such model we can analyze how migration probabilities vary with individual and regional characteristics. 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.[3]

We are interested in estimating α and β from the sample of migrants and the knowledge of $f(x)$ and the p_j. 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 the p_j comes from LFS population statistics and the census of residential variations.[4] Alternatively, we could assume that $f(x)$ and/or p_j 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_j are known. Estimation when $f(x)$ is estimated as opposed to known with certainty is discussed in Appendix 2.

2.2. *Minimum distance estimation*

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

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

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

$$\hat{\phi}_{j\ell} = \widehat{\Pr}(x = \xi_\ell \mid y = j) = \frac{1}{n_j} \sum_{i=1}^{n} \mathbf{1}(y_i = j) \mathbf{1}(x_i = \xi_\ell), \tag{4}$$

n_j is the number of observations with $y = j$, and $\mathbf{1}(.)$ is an indicator function.

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

$$\sqrt{n} \begin{pmatrix} \hat{\phi}_1 - \phi_1 \\ \hat{\phi}_2 - \phi_2 \\ \hat{\phi}_3 - \phi_3 \end{pmatrix} \xrightarrow{d} N \left[0, \begin{pmatrix} \frac{1}{r_1} \Omega_1 & 0 & 0 \\ 0 & \frac{1}{r_2} \Omega_2 & 0 \\ 0 & 0 & \frac{1}{r_3} \Omega_3 \end{pmatrix} \right] \tag{5}$$

where $\Omega_j = \Lambda_j - \phi_j \phi_j'$, $\Lambda_j = \operatorname{diag}\{\phi_{j1}, \ldots, \phi_{j(q-1)}\}$, 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.

The model specifies that

$$\phi_{j\ell} = \frac{\pi_\ell}{p_j} G_j(z(\xi_\ell); \alpha, \beta). \tag{6}$$

Then the optimal minimum distance estimates of α and β minimize

$$s(\alpha,\beta) = \sum_{j=1}^{3} \frac{n_j}{n} [\hat{\phi}_j - \phi_j(\alpha,\beta)]' \hat{\Omega}_j^{-1} [\hat{\phi}_j - \phi_j(\alpha,\beta)] \tag{7}$$

where $\hat{\Omega}_j$ is the sample counterpart of Ω_j. Moreover, $\hat{\Omega}_j^{-1} = \hat{\Lambda}_j^{-1} - \iota\iota'/\hat{\phi}_{jq}$ where ι denotes a $(q-1) \times 1$ vector of ones. Upon substitution, the minimum distance estimation criterion can be written as[5]

$$s(\alpha,\beta) = \sum_{j=1}^{3} \frac{n_j}{n} \sum_{\ell=1}^{q} \frac{1}{\hat{\phi}_{j\ell}} \left(\hat{\phi}_{j\ell} - \frac{\pi_\ell}{p_j} G_j(z(\xi_\ell);\alpha,\beta) \right)^2 \tag{8}$$

From the theory of minimum distance estimation we know that the minimizer of $s(\alpha,\beta)$ is asymptotically equivalent to maximum likelihood (See Ferguson 1958).

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 dataset of intra-regional migrants varies between 120,000 and 145,000 individuals per year. From there we draw a 10% 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% 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% 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.

A summary of the datasets, and the exact information used from each of them in order to obtain the various probabilities involved in identifying conditional migration probabilities, can be found in Table A1.3.

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 (see Bover and Arellano 1999 for a discussion of linear MD estimation). The calculation of maximum likelihood estimates (as discussed in Appendix 2) 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 (A.2) 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. We expect that the effects of the economic variables and individual characteristics may be different depending on the size of town of origin. We, therefore, allow for different coefficients for the three town sizes of origin. Aside from parameter estimates, to have a clearer picture of the magnitude of the effects, we present in Table 2 an illustrative selection of the probabilities predicted by the estimated equations reported in Table 1.[8]

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% 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% 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.[9] 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.[10] The results show that high regional unemployment rates encourage people to

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)
≥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. ($t-1$) * ≥ 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

Origin	Destination 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

Origin	Destination 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

Origin	Destination 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

Origin	Destination 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

(e)
$\geq$11 years of education, age 20–29, economic variables at sample means

Origin	Destination 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

Table 2 (cont.)

(f)
≥11 years of education, age 30–44, economic variables at sample means

Origin	Destination 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)
≥11 years of education, age 20–29, % employment in services at maximum

Origin	Destination 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)
≥11 years of education, age 30–44, % employment in services at maximum

Origin	Destination Small	Medium	Large	Total
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

(i)
8 years of education, age 20–29, house prices at maximum

Origin	Destination 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

Origin	Destination 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

Table 2 (cont.)

(k)
8 years of education, age 20–29, unemployment rate at maximum

Origin	Destination 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

Origin	Destination 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

(m)
≥11 years of education, age 20–29, house prices at maximum

Origin	Destination 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

Origin	Destination 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

Origin	Destination 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

Table 2 (cont.)

(p)
≥11 years of education, age 30–44, unemployment rate at maximum

Origin	Destination 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

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%, respectively, according to level of education when the unemployment rate is set at its sample period peak.

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%. 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.[11]

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.[12] 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%, 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 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.[13] 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

To investigate the determinants of the increase in intra-regional migrations since the 1980's in Spain, 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 estimation methods for complementary datasets based, respectively, on minimum distance and maximum likelihood principles. We only 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.

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		LFS
	Absolute frequency	Relative frequency	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 (cont.)

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

Table A1.3. Datasets and information used for the various probabilities

Probabilities	Data Sources
Migrants only (RVD)	
• Conditional distribution of characteristics x given migration choice j $f(x \mid y = j)$	• Random sample from the annual Census of Residential Variations (RVD). • Characteristics: age, education, region & year by town size of origin and destination.[a]
Migrants and non-migrants (LFS)	
• Marginal distribution of characteristics x $f(x)$	• Annual Labour Force Survey (LFS) • Aggregate information on characteristics by town size of origin.
RVD and LFS	
• Unconditional migration probability to j p_j	• Numerator: Total no. of migrants to destination j from RVD by town size of origin. • Denominator: Population by town size of origin from LFS aggregates.

[a] Since we have 17 regions, 4 years, 3 education categories, and 3 age groups, x takes on 612 different values for each of the 3 town sizes of origin.

Appendix 2

A2.1. Maximum likelihood estimation

The log likelihood for the sample of n independent observations of migrants is given by

$$L = \sum_{j=1}^{3} \sum_{i=1}^{n_j} \ln f(x_i \mid y_i = j) = \sum_{j=1}^{3} n_j \sum_{\ell=1}^{q} \hat{\phi}_{j\ell} \ln \phi_{j\ell} \tag{A.1}$$

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

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

Substituting (6) and (A.2) in (A.1) we obtain

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

ML estimates of α and β are obtained by maximizing $L(\alpha, \beta)$ subject to (A.2). To implement this method, we solve numerically the three nonlinear equations (A.2) for the intercepts as functions of the slope coefficients $\alpha = \alpha(\beta)$, say (using a Gauss-Newton iteration). Then, we first obtain estimates of the slope parameters as $\hat{\beta} = \arg\max L(\alpha(\beta), \beta)$, from which the estimated intercepts can be calculated as $\hat{\alpha} = \alpha(\hat{\beta})$. The estimated covariance matrix and the standard errors for $\hat{\beta}$ are obtained from the hessian matrix of $L(\alpha(\beta), \beta)$. Given this, the standard error for $\hat{\alpha}$ is calculated using the delta method. In both instances, numerical derivatives of $\alpha(\beta)$ are employed.[14] Indeed, one advantage of ML estimation over MD is that it enforces the restrictions $\sum_{\ell=1}^{q} \phi_{j\ell} = 1$ 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 so they are regarded as functions of dummy variables. Some discussion on the problem of estimation in the presence of continuous characteristics is contained in Bover and Arellano (1999).

A2.2. MD when the distribution of characteristics is estimated

Suppose that π_ℓ is not known with certainty, but an unrestricted estimate $\hat{\pi}_\ell$ (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. Define:

$$\hat{e}_{j\ell} = \hat{\phi}_{j\ell} - \frac{\hat{\pi}_\ell}{p_j} G_j(z(\xi_\ell); \alpha, \beta) \quad (j = 1, 2, 3; \ell = 1, \dots, q) \tag{A.4}$$

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

As before $\Omega_j = \Lambda_j - \phi_j \phi_j'$ and $\Lambda_j = \mathrm{diag}\{\phi_{j1}, \dots, \phi_{j(q-1)}\}$. Similarly $\pi = (\pi_1, \dots, \pi_{q-1})'$, $\Lambda_\pi = \mathrm{diag}\{\pi_1, \dots, \pi_{q-1}\}$ and $G_{j\ell} = G_j(z(\xi_\ell); \alpha, \beta)$. If we write $\hat{e}_j = \hat{\phi}_j - D_{\pi j} \hat{\pi}$ where

$$\begin{aligned} D_{\pi j} &= \mathrm{diag}\{G_{j1}/p_j, \dots, G_{j(q-1)}/p_j\} \\ &= \mathrm{diag}\{\phi_{j1}/\pi_1, \dots, \phi_{j(q-1)}/\pi_{q-1}\} = \Lambda_j \Lambda_\pi^{-1}, \end{aligned}$$

we have

$$\mathrm{Var}(\hat{e}_j) = \frac{1}{n_j} \Omega_j + \frac{1}{m} D_{\pi j} (\Lambda_\pi - \pi\pi') D_{\pi j}' \quad (j = 1, 2, 3) \tag{A.5}$$

and

$$\mathrm{Cov}(\hat{e}_j, \hat{e}_k) = \frac{1}{m} D_{\pi j} (\Lambda_\pi - \pi\pi') D_{\pi k}' \quad (j, k = 1, 2, 3). \tag{A.6}$$

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

$$\sqrt{n}\begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix} \xrightarrow{d} N\left[0, \begin{pmatrix} \frac{1}{r_1}\Omega_1 & 0 & 0 \\ 0 & \frac{1}{r_2}\Omega_2 & 0 \\ 0 & 0 & \frac{1}{r_3}\Omega_3 \end{pmatrix} + s\mathfrak{D}_\pi(\Lambda_\pi - \pi\pi')\mathfrak{D}'_\pi\right] \tag{A.7}$$

where $s = p\lim(n/m)$. In the analysis conducted in the paper we have assumed that $s = 0$. When $s \neq 0$, the additional term $\mathfrak{D}_\pi(\Lambda_\pi - \pi\pi')\mathfrak{D}'_\pi$ accounts for sampling error in the $\hat{\pi}_\ell$.

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

$$V_R = VM_R V \tag{A.8}$$

where

$$V = (\Delta'\bar{\Omega}^{-1}\Delta)^{-1}$$

$$M_R = \Delta'\bar{\Omega}^{-1}(\bar{\Omega} + s\mathfrak{D}_\pi(\Lambda_\pi - \pi\pi')\mathfrak{D}'_\pi)\bar{\Omega}^{-1}\Delta,$$

and we are using the notation $\Delta = (\Delta'_1, \Delta'_2, \Delta'_3)'$, $\Delta_j = \partial\phi_j(\alpha, \beta)/\partial(\alpha, \beta')$ and $\bar{\Omega} = \text{diag}\{r_j^{-1}\Omega_j\}$.

When $s = 0$ $M_R = V^{-1}$ so that the asymptotic covariance matrix is given by V. 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):

$$s_R(\alpha, \beta) = \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix}' \left[\begin{pmatrix} \frac{n}{n_1}\hat{\Omega}_1 & 0 & 0 \\ 0 & \frac{n}{n_2}\hat{\Omega}_2 & 0 \\ 0 & 0 & \frac{n}{n_3}\hat{\Omega}_3 \end{pmatrix} + \frac{n}{m}\hat{\mathfrak{D}}_\pi(\hat{\Lambda}_\pi - \hat{\pi}\hat{\pi}')\hat{\mathfrak{D}}'_\pi\right]^{-1} \begin{pmatrix} \hat{e}_1 \\ \hat{e}_2 \\ \hat{e}_3 \end{pmatrix} \tag{A.9}$$

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.

Endnotes

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

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

3 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.

4 Data sources and the information obtained from each of them are described in detail in the next section and in Table A1.3.

5 The form of Ω_j^{-1} results from the matrix inversion lemma:

$$(\Lambda_j - \phi_j\phi_j')^{-1} = \Lambda_j^{-1} + \frac{\Lambda_j^{-1}\phi_j\phi_j'\Lambda_j^{-1}}{(1-\phi_j'\Lambda_j^{-1}\phi_j)}$$

and the fact that $\Lambda_j^{-1}\phi_j = \imath$ and $\phi_{jq} = 1 - \imath'\phi_j$. As for the MD criterion note that

$$[\hat{\phi}_j - \phi_j(\alpha,\beta)]'\hat{\Lambda}_j^{-1}[\hat{\phi}_j - \phi_j(\alpha,\beta)] + \frac{1}{\hat{\phi}_{jq}}[\hat{\phi}_j - \phi_j(\alpha,\beta)]'\imath\imath'[\hat{\phi}_j - \phi_j(\alpha,\beta)]$$

$$= \sum_{\ell=1}^{q-1}\frac{1}{\hat{\phi}_{j\ell}}(\hat{\phi}_{j\ell} - \phi_{j\ell}(\alpha,\beta))^2 + \frac{1}{\hat{\phi}_{jq}}[\imath'\hat{\phi}_j - \imath'\phi_j(\alpha,\beta)]^2,$$

and that $[\imath'\hat{\phi}_j - \imath'\phi_j(\alpha,\beta)]^2 = [\hat{\phi}_{jq} - \phi_{jq}(\alpha,\beta)]^2$.

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.

8 See Bover and Arellano (1999) for a more extensive calculation of predicted probabilities.

9 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.

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

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

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

12 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.

13 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.

14 Alternatively, substituting $\phi_{jq} = 1 - \sum_{\ell=1}^{q-1}\phi_{j\ell}$ in (A.1) we obtain

$$L = \sum_{j=1}^{3} \left[n_j \sum_{\ell=1}^{q-1} \hat{\phi}_{j\ell} \ln \phi_{j\ell} + n_j \hat{\phi}_{jq} \ln \left(1 - \sum_{\ell=1}^{q-1} \phi_{j\ell} \right) \right]$$

$$\propto \sum_{j=1}^{3} \left[n_j \sum_{\ell=1}^{q-1} \hat{\phi}_{j\ell} \ln G_j(z(\xi_\ell); \alpha, \beta) + n_j \hat{\phi}_{jq} \ln \left(1 - \frac{1}{p_j} \sum_{\ell=1}^{q-1} \pi_\ell G_j(z(\xi_\ell); \alpha, \beta) \right) \right].$$

The problem with this way of enforcing the restrictions (A.2) is that the expression whose log is taken in the last term could be negative for some values of α and β.

References

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

Angrist J, Krueger A (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, Bover O (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, Meghir C (1992) Female Labour Supply and On-the-Job Search: An Empirical Model Estimated Using Complementary Data Sets. *Review of Economic Studies* 59:537–559

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

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

Bover O, Arellano M (1999) Learning About Migration Decisions from the Migrants: An Exercise in Endogenous Sampling and Complementary Datasets. Working Paper 9908, Research Department, Bank of Spain

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

Ferguson TS (1958) A Method of Generating Best Asymptotically Normal Estimates with Application to the Estimation of Bacterial Densities. *Annals of Marthematical Statistics* 29:1046–1062

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

Imbens GW, Lancaster T (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 CF, McFadden D (eds) *Structural Analysis of Discrete Data with Econometric Applications*. MIT Press, Cambridge, Chapt. 5, 198–272

Manski CF, Lerman SR (1977) The Estimation of Choice Probabilities from Choice Based Samples. *Econometrica* 45:1977–1988

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, Martí M (1997) Son Bajos los Flujos Migratorios en España? *Revista de Economía Aplicada* 5:155–171