

Estudios de Economía Aplicada
Nº 16, 2000. Págs. 63-92

Estudio de la migración interregional en España, a través de la Ecuación Master

FAURA MARTÍNEZ, U.
GÓMEZ GARCÍA, J.
ARANDA GALLEGO, J.
Universidad de Murcia

RESUMEN

En este trabajo modelizamos mediante la ecuación master los flujos migratorios entre los estados de un sistema formado por varias regiones. El modelo nos permite conocer cuál es la evolución temporal de la distribución de probabilidad de la población del sistema. Distinguiendo entre crecimiento natural, migración interior y exterior, desarrollamos la ecuación master para cada una de estas componentes. Establecemos las ecuaciones del estado medio y de la covarianza, así como la solución estacionaria. Estimamos con datos sobre migraciones entre Comunidades Autónomas de España para el periodo 1986-1995 los parámetros de la ecuación master mediante tres procedimientos diferentes: Mínimos Cuadrados Ordinarios, regresión log-normal y regresión de Poisson.

Palabras clave: Ecuación Master, Migración, Log-normal, Regresión de Poisson.

ABSTRACT

In this paper, we modeling migration flows between the states of a system formed by different regions using the equation master. The model allows us to represent the temporal evolution of the probability distribution of the population. As we distinguish within natural growth, internal and external migration, we could figure the mean value equations, the covariance equations and the stationary solution. Finally, we have estimated the equation master parameters for migration between Spanish regions for the period 1986-1995 on three ways: ordinary least squares, log-normal regression and Poisson regression.

Key words: Master Equation, Migration, Log-normal and Poisson regression.

Artículo recibido el 15 de noviembre de 1999. Aceptado 15 de marzo de 2000.

1. Introducción

Está extensamente reconocido que la variable población interviene, cada vez más, en los modelos económicos explicativos, tanto para interpretar cambios de tendencia sufridos por la economía a largo plazo, como para predecir comportamientos futuros. Difícilmente pueden entenderse los cambios que ocurren en la economía fuera del contexto demográfico en el que tienen lugar.

En la variación de las tasas de población, intervienen dos factores, el crecimiento natural y las migraciones, siendo la información que aporta el estudio de esta última, mucho mayor que la que pueda dar el estudio de cualquier otra componente del cambio demográfico. Cabré *et al.* (1985), afirman que los movimientos migratorios son el principal factor diferenciador de la población, por su efecto directo sobre el volumen y la estructura de la población, y por su efecto indirecto sobre el crecimiento natural. En esta misma línea encontramos el trabajo de Leguina Herrán y Fernández Cordón (1991), quienes consideran que las migraciones interiores constituyen la variable clave (superior al crecimiento vegetativo) que explica las diferencias de dinamismo demográfico de una región a otra.

Han sido muchos los modelos utilizados para estudiar las migraciones (modelos basados en las cadenas de Markov¹, modelos gravitacionales² y modelos que intentan unir las ideas básicas de los anteriores³), pero gran parte de ellos son modelos estáticos, o, modelos dinámicos que no tienen en cuenta la naturaleza estocástica del proceso migratorio (no se conoce con exactitud qué induce a un individuo a realizar una migración) ni la posibilidad de que el sistema objeto de estudio no esté en equilibrio. Por ello, nosotros estudiamos los flujos migratorios a partir de la ECUACIÓN MASTER, una ecuación diferencial que nos indica como evoluciona la distribución de probabilidad de la población de un sistema formado por varias regiones. En la sección 2, explicamos cuáles son las características más importantes de la ecuación master y establecemos las ecuaciones del valor medio y de la covarianza, de una manera más sencilla que en los trabajos originales sobre la ecuación master. En la

1. McGinnis (1968), Ginsberg (1971, 1979), Mcfarland (1970), Rowe y Krishnan (1983), Salkin *et al* (1975), Geweke *et al* (1986), Sampson (1990), etc.

2. Zipf (1946), Stouffer (1940), Sjaastad (1962), Todaro (1969), Lowry (1966), Anselin e Isard (1979), Gordon (1991), etc.

3. Isserman *et al* (1985), Rogerson (1984), Greenwood y Hunt (1984), Bentolila y Dolado (1990), González Pérez (1991), Santiago Hernando (1994), etc.

sección 3, modelizamos en España los flujos migratorios entre Comunidades Autónomas, CC.AA., durante el período 1986-1995 y estimamos los parámetros de la ecuación master mediante tres métodos diferentes⁴ (**Regresión log-normal**, entendiendo por tal, Mínimos Cuadrados Ordinarios (MCO) una vez linealizada la ecuación a estimar, **Regresión MCO** directamente y **Regresión de Poisson**).

2. Ecuación master migratoria

2.1 Conceptos generales

Denotando por $n = (n_1, n_2, \dots, n_L)^T$ la configuración poblacional del sistema, con n_i la población de la región i , queremos estudiar la dinámica de $n(t)$, en concreto, la probabilidad de que se dé el estado n en el período t .

La expresión de la ecuación master viene dada por⁵

$$\frac{d}{dt} P(n; t) = \sum_{m \neq n} [W(n/m; t) P(m; t) - W(m/n; t) P(n; t)] \quad (1)$$

siendo

$$W(n/m; t) = \frac{\partial}{\partial \Delta t} P(n; t + \Delta t / m; t) |_{\Delta t = 0} \quad (2)$$

Las $W(n/m; t)$ indican tasas de transición por unidad de tiempo desde el estado n al estado m , y coinciden, para n/m , con lo que se conoce en los procesos de Markov como *Intensidad de Transición desde m a n , en el instante t* .

La ecuación master (1), puede extenderse para todo m , pues el caso concreto que falta, $m=n$, se anula en el sumando global, admitiendo la forma

$$\frac{d}{dt} P(n; t) = \sum_m [W(n/m; t) P(m; t) - W(m/n; t) P(n; t)] \quad (3)$$

Esta ecuación puede interpretarse como una ecuación de tasas de transición y la probabilidad de cambio del estado n se puede descomponer como la suma de dos términos con efectos opuestos:

1. Aumento debido a cambios en la tasa de transición del estado m al n .
2. Disminución debido a cambios en la tasa de transición del estado n al m .

4. En todos los trabajos sobre la ecuación master consultados, siempre se ha utilizado el método log-normal, con algún breve comentario sobre el método de MCO y nada sobre la regresión de Poisson.

5. Ver Weidlich y Haag (1988).

Es decir, llegadas al estado n y salidas del estado n . Además, ofrece el mayor conocimiento posible de la evolución del sistema en condiciones de incertidumbre o restricción de información. Esta restricción de información está representada por $W(n/m;t)$.

Como los cambios producidos en el estado del sistema pueden ser debidos tanto al crecimiento natural como a las migraciones, se puede descomponer la ecuación master en varios sumandos, cada uno de los cuales refleja la variación debida, o, al crecimiento natural, o, a la migración. En el sumando referente a la migración se puede distinguir entre migración interna y externa (este enfoque no se ha tenido en cuenta en los diversos estudios realizados). La migración interna es aquella que ocurre entre las regiones que forman el sistema, y migración externa sería la migración entre cualquier región del sistema y el exterior del mismo. Ahora bien, este planteamiento se podría modificar si en lugar de tener L regiones, tomamos $L + 1$, siendo esa nueva región lo que hay fuera del sistema. Pero, para estudiar la evolución temporal de las $(L + 1)$ regiones, necesitamos conocer el crecimiento natural del resto del sistema, y esto es prácticamente imposible; por ello, es necesario distinguir por separado entre migración interior y exterior. La ecuación master, viene dada entonces por

$$\frac{d}{dt} P(n;t) = \left[\frac{d}{dt} P(n;t) \right]_{\text{MIGR}} + \left[\frac{d}{dt} P(n;t) \right]_{\text{CN}} + \left[\frac{d}{dt} P(n;t) \right]_{\text{MIG-EXT}} \quad (4)$$

reflejando cada uno de estos sumandos la parte debida a las migraciones interiores (MIGR), al crecimiento natural (CN) y a las migraciones exteriores (MIG-EXT).

Trabajando con estados adyacentes, definidos como aquellos que permiten la transición de un solo individuo, el cambio en el sistema va a estar representado por una secuencia de estados adyacentes. Para estudiar estados adyacentes se introduce el operador $E_i^{\pm k}$, definido por $E_i^{\pm k} f(n) = f(n_1, n_2, \dots, n_i \pm k, \dots, n_L)$, lo que nos va permitir desarrollar cada uno de los sumandos de la ecuación master. Siguiendo a Haag (1989), el primer término correspondiente a las migraciones interiores está dado por

$$\begin{aligned} \left[\frac{d}{dt} P(n;t) \right]_{\text{MIG}} &= \sum_{i,j} W_{ij}(n^{(ji)}) P(n^{(ji)}; t) - \sum_{i,j} W_{ji}(n) P(n;t) \\ &= \sum_{i,j} W(n; E_i^{+1} E_j^{-1} n) P(E_i^{+1} E_j^{-1} n; t) - \sum_{i,j} W(E_i^{+1} E_j^{-1} n; n) P(n;t) \\ &= \sum_{i,j} E_i^{+1} E_j^{-1} W(E_j^{+1} E_i^{-1} n; n) P(n;t) - \sum_{i,j} W(E_j^{+1} E_i^{-1} n; n) P(n;t) \\ &= \sum_{i,j} (E_i^{+1} E_j^{-1} - 1) W(E_j^{+1} E_i^{-1} n; n) P(n;t) \end{aligned}$$

Por otro lado, hemos obtenido que la variación debida al crecimiento natural es:

$$\begin{aligned} \left[\frac{d}{dt} P(n;t) \right]_{CN} &= \sum_i (E_i^{-1} - 1) W(E_i^{+1} n; n) P(n;t) \\ &\quad + \sum_i (E_i^{+1} - 1) W(E_i^{-1} n; n) P(n;t) \end{aligned}$$

Y para la migración exterior

$$\begin{aligned} \left[\frac{d}{dt} P(n;t) \right]_{MIG-EXT} &= \sum_i \left[W_{Ei} P(E_i^{-1} n; t) - W_{Ei} P(n; t) \right. \\ &\quad \left. + W_{iE} P(E_i^{+1} n; t) - W_{iE} P(n; t) \right] \\ &= \sum_i \left[(E_i^{-1} - 1) W_{Ei} P(n; t) + (E_i^{+1} - 1) W_{iE} P(n; t) \right] \end{aligned}$$

con las correspondientes tasas de transición, denotadas por W_{iE} y W_{Ei} para transiciones de una región i del sistema hacia el exterior, E , y del exterior hacia i , respectivamente.

Para obtener las tasas de transición correspondiente a las migraciones interiores, hay que tener en cuenta que si $p_{ij}(t, \Delta t)$ es la probabilidad de transición de la región i a la región j durante el intervalo de tiempo $[t, t + \Delta t]$ para un individuo particular de i , tenemos que *la intensidad de transición por unidad de tiempo para dicho individuo desde el estado i al j* , está dada por

$$q_{ij}(t) = \lim_{\Delta t \rightarrow 0} \frac{p_{ij}(t, \Delta t)}{\Delta t}$$

y para intervalos infinitesimales, $p_{ij}(t, \Delta t) = q_{ij}(t) \Delta t$. Suponiendo que los individuos deciden cambiar de estado de forma independiente y que sólo hay una transición por unidad de tiempo (esto se puede hacer pues estamos tomando intervalos infinitesimales), la tasa de transición se puede expresar como⁶

$$W(E_j^{+1} E_i^{-1} n; n) \equiv W_{ij}(n) \equiv W_{ij}(n_i, n_j) = q_{ij}(t) n_i \quad (5)$$

6. Demostración en Kanaroglou, Liaw y Papageorgiou (1986).

Es decir, la tasa de transición desde i a j , es n_i veces la tasa de transición individual $q_{ij}(t)$, pues cualquiera de las n_i personas en la región i puede emigrar independientemente a la región j con una tasa $q_{ij}(t)$.

Para el proceso de crecimiento natural, tenemos de forma análoga a (5) que las tasas de transición de nacimientos y muertes, denotadas por $W_i^+(n)$ y $W_i^-(n)$ vienen dadas por

$$\begin{cases} W_i^+(n) \equiv W(E_i^{+1}n/n) = \beta_i(t)n_i \\ W_i^-(n) \equiv W(E_i^{-1}n/n) = \mu_i(t)n_i \end{cases} \quad (6)$$

siendo $\beta_i(t)$ y $\mu_i(t)$ las tasas de nacimiento y muerte (por unidad de tiempo) respectivamente⁷. En las migraciones exteriores, las tasas de transición, W_{iE} y W_{Ei} , son más difíciles de determinar⁸.

2.2 Tasas de transición como función de utilidades y factores de motivación

Las tasas de transición individuales, q_{ij} , se definen en términos de las funciones de utilidad que caracterizan la atracción de cada área, u_i . Además, deben verificar unos requisitos mínimos:

1. Ser positivas por definición.
2. La tasa de transición individual q_{ij} tiene que ser mayor que q_{ji} , si la utilidad de la región j , u_j , es mayor que la de la región i , u_i .
3. q_{ij} ha de ser una función monótona no decreciente de la diferencia $(u_j - u_i)$ ⁹, pues una mayor diferencia no puede inducir una menor tasa de transición.

Una expresión válida para $q_{ij}(t)$ sería, $q_{ij} = \exp(u_j - u_i)$. Como también hay otros factores que reflejan los efectos de distancia tanto geográfica, como social y económica, la tasa individual podría tener la expresión:

7. Pumain (1982) confirma empíricamente estos resultados.

8. Como en (5) y (6), se podría contrastar que

$$\begin{aligned} W_{iE} &= n_i^* q_{iE}(t) \\ W_{Ei} &= n_i^{**} q_{Ei}(t) \end{aligned}$$

siendo n_i^* y n_i^{**} el número de personas de la región i que hay en E y el número de personas de E que hay en i , respectivamente. En ambos casos, n_i^* y n_i^{**} , se toman como aproximación al número de movimientos que pueden realizarse.

9. También puede depender del cociente de utilidades y dar lugar a un nuevo conjunto de hipótesis, y posteriormente decidir con cual de ellas se obtiene mejores estimaciones.

$$q_{ij} = v_{ij}(t) \exp(u_j - u_i) \quad (7)$$

con un factor de movilidad $v_{ij}(t)$ que normalmente es simétrico, y un factor de atracción/repulsión dependiendo de las funciones de utilidad de la región de origen, u_i y de la región de destino, u_j .

Teniendo en cuenta la ponderación que se puede dar a las ventajas o desventajas de la región de origen frente a la de destino, tenemos

$$q_{ij} = v_{ij}(t) \exp(u_j - \alpha u_i) \quad (8)$$

siendo α un *factor de sensibilidad*, que refleja tal ponderación.

Así, las tasas de transición vienen dadas por:

$$W_{ij}(n_j, n_i) = n_i v_{ij}(t) \exp(u_j - u_i) \quad (9)$$

$$W_{ij}(n_j, n_i) = n_i v_{ij}(t) \exp(u_j - \alpha u_i) \quad (10)$$

según se utilice la ecuación 7 u 8.

Para los factores de movilidad se admiten diversas hipótesis:

$$v_{ij}(t) = v_0(t) f_{ij}$$

siendo $v_0(t)$ un factor de movilidad global, que caracteriza la movilidad media de la población bajo estudio, con f_{ij} dependiente de las regiones de origen y destino. Es posible que el término f_{ij} también dependa del tiempo (a diferencia de los supuestos tradicionales):

$$v_{ij}(t) = v_0(t) f_{ij}(t)$$

y se puede relacionar este factor con una serie de variables explicativas como la distancia económica, geográfica o psicológica:

$$f_{ij}(t) = a(t) \exp(-\beta_t d_{ij}^\theta) \quad (11)$$

$$f_{ij}(t) = a(t)d_{ij}^{-\theta t} \quad (12)$$

$$f_{ij}(t) = a(t) \exp\left(\frac{-\beta_t d_{ij}^{\theta_1}}{1 + \gamma_t d_{ij}^{\theta_2}}\right) \quad (13)$$

$$f_{ij}(t) = a(t) \frac{\exp(\beta_t d_{ij}^{\theta})}{\exp(\beta_t d_{ij}^{\theta}) + \gamma} \quad \text{con } \gamma > 0 \quad (14)$$

Las hipótesis (11) y (12) son las hipótesis clásicas de dependencia exponencial y potencial, donde el parámetro θ puede tomar cualquier valor¹⁰. Como distancias grandes o pequeñas pueden infravalorar esas medidas, otra posibilidad es la dada por (13), en la que γ es un parámetro asociado con el efecto de saturación, y θ_1 y θ_2 nos indican la fuerza con la que depende de la distancia¹¹. En (14) expresamos una dependencia logística¹², con $\theta > 0$.

La función de utilidad, u_i , se descompone en dos partes:

$$u_i(n_i; t) = s_i(n_i; t) + \delta_i(t)$$

una, dependiente del tamaño (s_i), y otra, que indique las preferencias por una u otra región independientemente del tamaño (δ_i).

Sobre la forma que puede tener la parte referente al tamaño, existen diversas posibilidades, ver Haag y Max (1995, pp. 246), como

$$u_i(n_i; t) = \delta_i(t) + \kappa n_i(t) + \sigma n_i^2(t) \quad (15)$$

$$u_i(n_i; t) = \delta_i(t) + \kappa \ln(n_i(t)) \quad (16)$$

$$u_i(n_i; t) = \delta_i(t) \quad (17)$$

siendo $\kappa > 0$ un parámetro que representa la tendencia de la población a concentrarse. El parámetro σ , llamado parámetro de saturación, recoge todos aquellos factores que dependan de las interacciones entre cada par de individuos del área i (congestión del tráfico, aumento de la renta, etc.). Los parámetros $\{\delta_i, \kappa, \sigma\}$ reciben el nombre de *parámetros de tendencia* del sistema. Existen trabajos empíricos que apoyan

10. Siempre se ha tomado $\theta = 1$, ver Haag, Munz, Pumain, Sanders y Saint-Julien (1992).

11. Al igual que en el caso anterior, sólo ha sido considerado el caso $\theta_1 = \theta_2 = 1$.

12. Este caso no ha sido aplicado nunca en este contexto.

el uso de la hipótesis (15)¹³. Pero, también cabe la posibilidad de que los parámetros dependan del tiempo o de la región, y dar lugar a un nuevo conjunto de hipótesis para las utilidades.

Se puede considerar el efecto de la variable tamaño sin dimensión, mediante la transformación

$$\frac{n_i(t) - n^*(t)}{n^*(t)} \quad (18)$$

que describe la desviación relativa de la población regional respecto a su valor medio espacial

$$n^*(t) = \frac{1}{L} \sum_{i=1}^L n_i \quad (19)$$

2.3. Ecuaciones del valor medio y covarianza

En la mayoría de los casos, toda la información contenida en la distribución de probabilidad no puede ser utilizada en un análisis empírico. Aunque en un momento determinado sólo se da un estado, la distribución nos informa sobre la probabilidad en t de todos los posibles estados. Lo más general es que $P(n;t)$ sea una distribución unimodal y es de esperar que el valor medio de n , coincida prácticamente con el estado más probable. Por ello, es importante obtener ecuaciones para el valor medio de los componentes de la configuración poblacional.

El valor medio de la población del estado j en el período t , está definido por

$$\bar{n}_j(t) = \sum_n n_j P(n;t) \quad (20)$$

Es la suma sobre todos los posibles estados ponderado por la probabilidad de que se dé ese estado. Las ecuaciones para el valor medio se pueden obtener directamente a partir de la ecuación master (1) o (3):

$$\frac{d\bar{n}_j(t)}{dt} = \overline{G_j(n;t)} \quad (21)$$

13. Ver Weidlich y Haag (1988).

con

$$G_j(n; t) = \sum_m (m_j - n_j) W(m/n; t) \quad (22)$$

llamado momento de primer orden. Teniendo en cuenta que la tasa de transición total del sistema viene dada por

$$W(m/n; t) = \sum_{i,j} W(E_j^{+1} E_i^{-1} n/n) + \sum_i W(E_i^{+1} n/n) + \sum_i W(E_i^{-1} n/n)$$

y sustituyendo en (22), obtenemos las ecuaciones del valor medio¹⁴,

$$\frac{d\bar{n}_j(t)}{dt} = \sum_i \overline{W_{ij}(n; t)} - \sum_i \overline{W_{ji}(n; t)} + \overline{W_j^{+1}(n; t)} - \overline{W_j^{-1}(n; t)} + \overline{W_{Ej}} - \overline{W_{jE}} \quad (23)$$

La interpretación intuitiva de la ecuación (23) es clara. El cambio en el tiempo del valor medio $\bar{n}_j(t)$, es debido a la tasa media de inmigración desde cualquier región (incluido el exterior) hasta j , menos la tasa media de emigración desde j a cualquier otra región (incluido el exterior), más la tasa media de nacimiento en j menos la de defunción.

Suponiendo que estamos en la hipótesis del modelo (9), las ecuaciones correspondientes son:

$$\begin{aligned} \frac{d\bar{n}_j(t)}{dt} &= \sum_{i=1}^L v_{ij} \bar{n}_i \exp[u_j(\bar{n}_j) - u_i(\bar{n}_i)] - \sum_{i=1}^L v_{ji} \bar{n}_j \exp[u_i(\bar{n}_i) - u_j(\bar{n}_j)] \\ &+ \rho_j \bar{n}_j + \overline{W_{Ej}} - \overline{W_{jE}} \end{aligned} \quad (24)$$

Siempre que la tasa de crecimiento natural junto con la migración externa, puedan expresarse como

$$\rho_j(t) \bar{n}_j + (\overline{W_{Ej}} - \overline{W_{jE}}) \approx \rho(t) \bar{n}_j(t) \quad (25)$$

se puede separar el proceso de crecimiento natural y migración externa de la redistribución migratoria interna de la población. Para ello, introducimos el porcentaje de población relativa a través de

14. Otra demostración más compleja e incluyendo sólo migraciones interiores, se puede encontrar en Kanaroglou *et al* (1986), pp. 920-921.

$$x_j(t) = \frac{\bar{n}_j(t)}{N(t)} \quad \text{con} \quad 0 \leq x_j(t) \leq 1 \quad (26)$$

con evolución temporal dada por

$$\frac{dx_j(t)}{dt} = \sum_{i=1}^L v_{ij}(t) x_i \exp[u_j(t) - u_i(t)] - \sum_{i=1}^L v_{ji}(t) x_j \exp[u_i(t) - u_j(t)] \quad (27)$$

La solución estacionaria del estado medio, corresponde a la solución de las ecuaciones (27), junto con (26), es

$$\hat{n}_i = \frac{N \exp[2u_i(\hat{n}_i)]}{\sum_{j=1}^L \exp[2u_j(\hat{n}_j)]} \quad (28)$$

Para la covarianza, definida mediante la siguiente expresión

$$\sigma_{ij}(t) = \overline{[(n_i - (\bar{n}_i)_t)(n_j - (\bar{n}_j)_t)]_t} = \overline{n_i n_j} - \bar{n}_i \bar{n}_j$$

obtenemos (mediante una demostración propia), las **ecuaciones de la covarianza:**

$$\begin{aligned} \frac{d\sigma_{ij}(t)}{dt} &= \frac{d}{dt} \overline{n_i n_j} - \frac{d\bar{n}_i}{dt} \bar{n}_j - \bar{n}_i \frac{d\bar{n}_j}{dt} \\ &= \sum_n \sum_m (m_i m_j - n_i n_j) W(m/n; t) P(n; t) - \bar{n}_j \overline{G_i(n; t)} - \bar{n}_i \overline{G_j(n; t)} \end{aligned}$$

y, teniendo en cuenta (21) y (22)

$$\begin{aligned} \frac{d\sigma_{ij}(t)}{dt} &= \sum_n \sum_m (m_i m_j - n_i n_j) W(m/n; t) P(n; t) \\ &\quad - \sum_n G_i(n; t) \bar{n}_j P(n; t) - \sum_n G_j(n; t) \bar{n}_i P(n; t) \\ &= \sum_n \sum_m (m_i m_j - n_i n_j) W(m/n; t) P(n; t) \\ &\quad - \sum_n \sum_m (m_i - n_i) \bar{n}_j W(n/m; t) P(n; t) \\ &\quad - \sum_n \sum_m (m_j - n_j) \bar{n}_i W(n/m; t) P(n; t) \end{aligned}$$

Sumando y restando a la expresión anterior $(m_i n_j + n_i m_j + n_i n_j) W(m/n; t) P(n; t)$ y agrupando

$$\begin{aligned} \frac{d\sigma_{ij}(t)}{dt} &= \sum_m \sum_n (m_i - n_i)(m_j - n_j) W(m/n; t) P(n; t) \\ &\quad + \sum_n (n_i - \bar{n}_i) G_j(n; t) P(n; t) + \sum_n (n_j - \bar{n}_j) G_i(n; t) P(n; t) \end{aligned}$$

Si definimos como momentos de segundo orden

$$G_{ij}(n; t) = \sum_m (m_i - n_i)(m_j - n_j) W(m/n; t)$$

tenemos finalmente que

$$\frac{d\sigma_{ij}(t)}{dt} = \overline{G_{ij}(n; t)} + \overline{(n_i - \bar{n}_i) G_j(n; t)} + \overline{(n_j - \bar{n}_j) G_i(n; t)}$$

o de forma aproximada

$$\frac{d\sigma_{ij}(t)}{dt} \approx G_{ij}(\bar{n}; t) + \sum_k \left(\sigma_{ik}(t) \frac{\partial G_j(\bar{n}; t)}{\partial \bar{n}_k} + \sigma_{jk}(t) \frac{\partial G_i(\bar{n}; t)}{\partial \bar{n}_k} \right)$$

Y sabiendo que los momentos de segundo orden, vienen dados por

$$\begin{cases} G_{ij}(n; t) = -(W_{ij}(n; t) + W_{ji}(n; t)) & i \neq j \\ G_{jj}(n; t) = \sum_{i \neq j} (W_{ij}(n; t) + W_{ji}(n; t)) + W_j^+(n; t) + W_j^-(n; t) \end{cases}$$

la ecuación aproximada para la covarianza es

$$\begin{aligned} \frac{d\sigma_{ij}(t)}{dt} &= -W_{ij}(\bar{n}; t) - W_{ji}(\bar{n}; t) \\ &\quad + \sum_{k \neq j} \sigma_{ik}(t) \frac{\partial W_{kj}(\bar{n}; t)}{\partial \bar{n}_k} + \sum_{k \neq i} \sigma_{jk}(t) \frac{\partial W_{ki}(\bar{n}; t)}{\partial \bar{n}_k} \\ &\quad + \sigma_{ij}(t) (\rho_j(t) + \rho_i(t)) \end{aligned}$$

Las ecuaciones de la covarianza son importantes porque (Helbing (1995, pp.69):

1. Reflejan la naturaleza estocástica del cambio en el tiempo y permiten distinguir entre distintos sistemas que tengan iguales ecuaciones aproximadas del valor medio.
2. Son necesarias para determinar la fiabilidad de las predicciones realizadas a partir de las ecuaciones del valor medio y para evaluar el intervalo de tiempo después del cual son menos útiles las ecuaciones del valor medio.
3. Permiten correcciones de las ecuaciones aproximadas del valor medio.

2.4. *Estimación de parámetros*

Es evidente que la dinámica del sistema va a depender de la función de utilidad escogida, es decir, de los parámetros de tendencia (δ_i , κ , σ , dependientes o no del tiempo o de la región), cualquiera que sea su composición en términos de factores de motivación. Por ello, es conveniente hacer el estudio por separado de los factores de tendencia y de los factores de motivación. Considerando un sistema con L regiones en el que se tiene información, para cada año, sobre el tamaño poblacional de cada región, $n_i^e(t)$, y el número de transiciones en el período $[t, t+1]$ de una región a otra, $W_{ij}^e(t)$; la matriz de migración teórica puede venir dada por distintas expresiones (9,10). Sea cual sea su expresión, siempre W_{ij} es una función de $v_{ij}(t)$ y de las utilidades, $u_i(t)$, y se va a determinar a partir de las $T(L^2-L)$ observaciones de $W_{ij}^e(t)$, suponiendo que el período de observación tiene T intervalos. Esta estimación va a depender además de que se considere factores de movilidad simétricos o no simétricos, de que $f_{ij}(t)$ sea o no dependiente del tiempo. Utilizaremos tres métodos para estimar los parámetros¹⁵:

1. **Regresión log-normal**, entendiéndose por tal, MCO una vez linealizada la ecuación a estimar.

2. **Regresión MCO**, directamente.

3. **Regresión de Poisson**, pues la variable a estimar, las migraciones por unidad de tiempo (tasas de transición), es una variable discreta, y, por tanto, sería conveniente realizar la regresión utilizando técnicas de regresión de variables discretas, como la regresión de Poisson (que es una estimación Máximo-Verosímil).

La comparación entre las distintas técnicas de regresión, debe hacerse desde el punto de vista de la acuracidad de los valores pronosticados. La medición de la bondad de ajuste se realiza en la estimación por mínimos cuadrados, a través del coeficiente de determinación, R^2 . La medida de R^2 no es comparable directamente

15. Los detalles de los distintos métodos serán objeto de otro artículo posterior.

con la medida de la desviación, D , usada en la regresión de Poisson. En el modelo de Poisson se toma como base para determinar el ajuste, el modelo nulo (Grand Mean Model), en el que la migración de i a j está estimada simplemente como el flujo medio. Si llamamos D_0 a la desviación del modelo nulo, la bondad del ajuste se puede calcular mediante

$$\phi = \frac{D_0 - D}{D_0} \quad (29)$$

siendo D la desviación de nuestro modelo de trabajo. Cuanto más próxima esté ϕ de 1 mejor será el ajuste. En el caso de la regresión MCO, el coeficiente de determinación R^2 coincide con ϕ , pues D_0 es la suma de los cuadrados totales, SCT, y D es la suma de los cuadrados de los errores, SCE. Por tanto, mediante ϕ podemos comparar la bondad de ajuste de los distintos métodos desarrollados. Otra posibilidad para comparar los distintos métodos es, siguiendo el trabajo de Knudsen y Fotheringham (1986), el estadístico ϕ^{16} , definido por

$$\phi = \sum_i \sum_j p_{ij} \ln \left(\frac{p_{ij}}{q_{ij}} \right) \quad (30)$$

donde

$$p_{ij} = \frac{W_{ij}}{\sum_i \sum_j W_{ij}} \quad y \quad q_{ij} = \frac{\hat{W}_{ij}}{\sum_i \sum_j \hat{W}_{ij}}$$

siendo \hat{W}_{ij} el valor estimado. El estadístico ϕ aumenta conforme la bondad del ajuste disminuye.

Una vez estimados los parámetros de tendencia, podemos obtener cuáles son las variables socioeconómicas (factores de motivación) que explican dichos parámetros.

3. Caso empírico: Migraciones interregionales en España, período 1986-1995

3.1. *Las migraciones interiores, principal factor del crecimiento demográfico*

En primer lugar, comprobamos mediante un procedimiento residual que las migraciones interiores¹⁷ son el principal factor de diferenciación poblacional entre

16. Realizan un estudio de distintas medidas de bondad de ajuste y recomiendan utilizar el estadístico ϕ .

17. Los datos proceden de las Estadísticas de Variaciones Residenciales, EVR, publicada por el INE, en la serie Migraciones. Para los años en los que se realiza un censo o padrón, los datos ofrecidos por EVR

CC.AA.¹⁸, durante el período 1986-1995. Calculamos la tasa de crecimiento total de cada CC.AA., y la tasa de crecimiento de la migración interior, dadas respectivamente por

$$g_i = \frac{1}{n_i} \cdot \frac{dn_i}{dt} \quad \text{y} \quad m_i = \frac{1}{n_i} \cdot \sum_{j \neq i} (w_{ji} - w_{ij})$$

y definimos el crecimiento residual (debido al crecimiento natural y a las migraciones externas) mediante

$$\rho_i = g_i - m_i$$

La tasa de crecimiento total de España, vendrá dada por

$$g = \sum_i \frac{n_i}{N} g_i = \frac{1}{N} \sum_i \frac{dn_i}{dt}$$

y el índice de crecimiento residual es

$$\hat{\rho}_i = \frac{g_i - m_i}{g}$$

Mediante la varianza de dicho índice (la media vale 1)

$$g = \sum_i \frac{n_i}{N} g_i = \frac{1}{N} \sum_i \frac{dn_i}{dt}$$

podemos saber si hay diferencias entre las CC.AA. Estas varianzas anuales están alrededor de 0.055, salvo para los 1991 y 1995, en los que aumenta un par de milésimas. Por tanto, no es un factor diferenciador entre CC.AA, el crecimiento natural y la migración externa.

están infravalorados, pues durante los meses en los que se efectúa el censo o padrón, los nuevos residentes son dados de alta directamente en el padrón municipal de su nueva residencia y, a la vez de baja en el municipio de origen, con lo que no quedan registrados en las cifras mensuales de variaciones municipales, de forma que el número de migrantes anual ofrecido por la EVR es más bajo. Es necesario ajustar los datos en estos años y para ello realizamos una interpolación con el año anterior y posterior (este método ha sido utilizado por Bentolila y Dolado (1990) y Ródenas Calatayud (1997)).

18. No tenemos en cuenta en nuestro estudio Ceuta y Melilla, por ser las migraciones con estas regiones muy pequeñas, y porque no tenemos suficiente información estadística sobre las variables económicas necesarias para explicar los flujos.

3.2. Estimación de parámetros

Queremos estimar las utilidades y el factor de movilidad global, junto con f_{ij} dependiente del tiempo y simétrico. Para ello, compararemos los residuos entre los valores observados y los teóricos, por los tres métodos comentados. La regresión de Poisson y la MCO, ha sido realizada mediante el paquete informático SAS, mientras que la estimación log-normal ha sido llevada a cabo mediante un programa elaborado en FORTRAN. En la tabla 1, tenemos la bondad del ajuste, para cada año y para cada uno de los métodos. Los resultados indican que el mejor ajuste se produce cuando aplicamos el método MCO, seguido del de Poisson y el log-normal; si bien, estos dos últimos también son buenos.

Tabla 1: Bondad de ajuste

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
<i>REGRESIÓN LOG-NORMAL</i>										
R^2	0.9962	0.9926	0.9899	0.9923	0.9932	0.9927	0.9890	0.9933	0.9913	0.9813
R^*	0.9913	0.9830	0.9769	0.9823	0.9843	0.9832	0.9747	0.9847	0.9800	0.9570
<i>REGRESIÓN MCO</i>										
	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
ϕ	0.9978	0.9970	0.9969	0.9977	0.9973	0.9975	0.9958	0.9964	0.9959	0.9961
ϕ^*	0.9950	0.9931	0.9931	0.9947	0.9939	0.9943	0.9905	0.9919	0.9908	0.9912
<i>REGRESIÓN DE POISSON</i>										
	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
ϕ	0.9954	0.9940	0.9940	0.9957	0.9957	0.9962	0.9938	0.9953	0.9948	0.9940
ϕ^*	0.9897	0.9864	0.9865	0.9904	0.9904	0.9914	0.9861	0.9894	0.9883	0.9864

El * indica el coeficiente corregido por los grados de libertad.

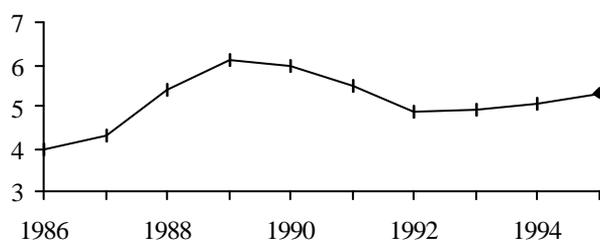
Por tanto, a la hora de estudiar los factores de movilidad, así como las utilidades en función de las variables socioeconómicas, utilizaremos los resultados de la estimación mínimo cuadrática, MCO, para f_{ij} dependiente de t ¹⁹.

19. Si en lugar del estadístico ϕ , utilizamos ϕ dado por (30), se obtienen las mismas conclusiones, tomando dicho estadístico valores comprendidos entre 0.002 y 0.005, siendo menores para la estimación MCO, seguida de la Poisson y de la log-normal.

3.3. Movilidad global, utilidades y preferencias

El factor de movilidad global, $v_0(t)$, es una medida de la frecuencia de cambios de residencia entre las distintas CC.AA., sin tener en cuenta entre qué regiones se realiza. La evolución de la movilidad global dependiente del tiempo, $v_0(t)$, gráfico 1, nos indica que la movilidad aumenta con el tiempo, salvo para el período 1989-92, en el que disminuye. La predisposición de la población a moverse desde una región a otra, ha ido generalmente en aumento.

Gráfico 1: Movilidad global, periodo 1986-95, (x1000)



Para la utilidad, hemos impuesto la condición de que para cada t , la suma de las utilidades sea cero, pues las tasas de transición dependen de la diferencia de utilidades y por tanto estarán estimadas salvo una constante (resultados en tabla 2).

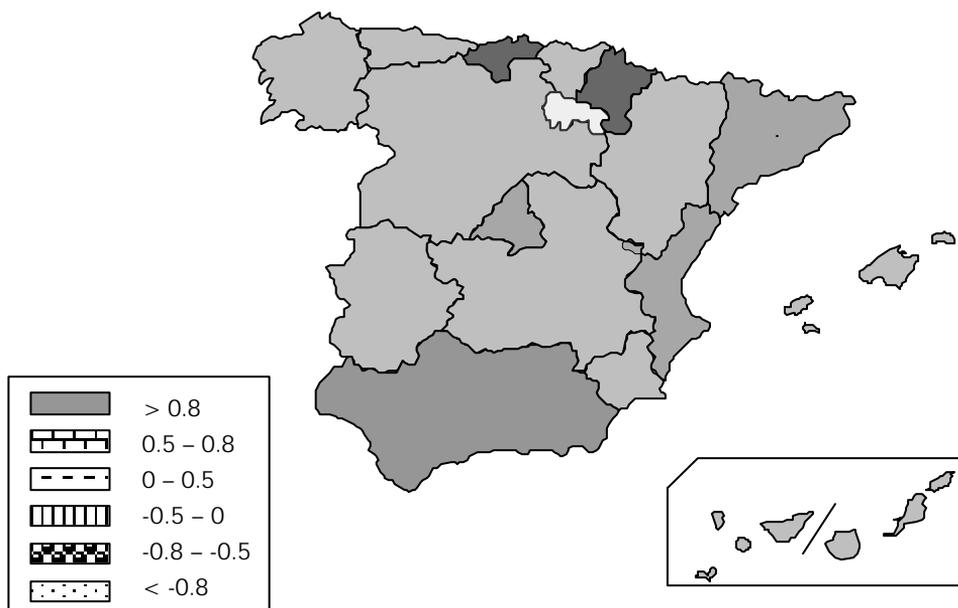
Tabla 2: Utilidades por CC.AA.

CC.AA	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995
And.	0.8019	0.8004	0.789	0.8277	0.8376	0.8307	0.8212	0.835	0.7966	0.7444
Arg.	-0.1647	-0.1669	-0.0304	-0.0415	-0.0354	-0.0535	-0.0967	-0.1541	-0.1767	-0.1675
Ast.	-0.3310	-0.3454	-0.3549	-0.2937	-0.2699	-0.262	-0.2475	-0.2537	-0.3557	-0.3469
Bal.	-0.1896	0.0913	0.1644	0.0364	-0.2165	-0.2817	-0.3170	-0.3622	-0.2146	-0.0568
Can.	0.3447	0.4549	0.3968	0.2397	-0.0437	0.0331	0.1576	0.1669	0.2491	0.2762
Cant.	-0.6793	-0.7214	-0.7188	-0.6448	-0.6507	-0.6234	-0.6026	-0.5573	-0.5207	-0.5327
C.L.	0.0882	0.0257	0.0349	-0.0355	-0.0109	0.0057	0.0240	0.0741	0.1012	0.0533
C.M.	0.0269	-0.0263	-0.0535	-0.0426	0.0679	0.1352	0.2091	0.1791	0.1715	0.1859
Cat.	0.5664	0.6856	0.7709	0.7888	0.777	0.7127	0.6273	0.5765	0.5486	0.6180
C.V.	0.6012	0.6113	0.6699	0.7200	0.6916	0.6610	0.6145	0.6199	0.6138	0.6506
Ext.	-0.2389	-0.3108	-0.4510	-0.4841	-0.4088	-0.3535	-0.2803	-0.2154	-0.2049	-0.2810
Gal.	0.1649	0.1020	0.1059	0.1659	0.2343	0.2436	0.2768	0.2862	0.2617	0.2175
Mad.	0.6449	0.6945	0.6448	0.639	0.6551	0.6318	0.5958	0.5220	0.4611	0.4617
Mur.	0.0520	0.0199	-0.0121	0.0795	0.0098	0.0413	0.0648	0.0521	-0.0245	-0.0610
Nav.	-0.6015	-0.6873	-0.6171	-0.5965	-0.5445	-0.5407	-0.5498	-0.5664	-0.5434	-0.5713
PV.	-0.2085	-0.3119	-0.3647	-0.2946	-0.1859	-0.1982	-0.2159	-0.1580	-0.1914	-0.1712
Rioja	-0.8776	-0.9156	-0.9741	-1.0637	-0.907	-0.9821	-1.0813	-1.0447	-0.9717	-1.0192

A partir de los valores medios temporales de la utilidad, gráfico 2, podemos clasificar las regiones en seis grupos:

1. Un grupo de regiones con utilidad alta: Andalucía.
2. Un grupo de regiones con utilidad moderadamente alta: Cataluña, C. Valenciana y Madrid.
3. Un grupo de regiones con utilidad positiva entre 0 y 0.5: Canarias, Castilla-León, Castilla la Mancha, Galicia y Murcia.
4. Un grupo de regiones con utilidad negativa entre -0.5 y 0: Aragón, Asturias, Baleares, Extremadura y País Vasco.
5. Un grupo de regiones con utilidad baja: Navarra y Cantabria.
6. Un grupo de regiones con utilidad muy baja: La Rioja.

Gráfico 2: Valores medios temporales de las utilidades.



Si realizamos para cada CC.AA., una regresión de la función de utilidad frente al tiempo, tabla 3, podemos comprobar como desde un punto de vista estadístico (según el valor de la F de Snedecor), para Cantabria, Castilla la Mancha, Galicia y Madrid, existe una relación lineal entre la función de utilidad del modelo y el tiempo, positiva para las tres primeras CCAA. y negativa para Madrid. Al realizar una estimación cuadrática, obtenemos resultados significativos para las mismas CC.AA., junto con Andalucía y Aragón, aunque sólo se produce una mejora importante para la región de Madrid. En Andalucía, Aragón, Galicia y Madrid, la función de utilidad

presenta forma de U invertida; y para Cantabria y Castilla la Mancha forma de U. En el resto de las CC.AA. no se observa ninguna dependencia temporal concreta.

Tabla 3: Estimación lineal y cuadrática de las utilidades regionales en función del tiempo

	Regresión Lineal			Regresión Cuadrática			
	t	F	R ²	t	t ²	F	R ²
Andalucía	-0.0021	0.40 (0.543)	-0.071	0.0301	-0.0029	6.97 (0.022)	0.5702
Aragón	-0.0054	0.58 (0.467)	-0.049	0.0665	-0.0065	8.04 (0.015)	0.6100
Asturias	0.0026	0.26 (0.622)	-0.089	0.0489	-0.0042	4.15 (0.065)	0.4121
Baleares	-0.0285	2.32 (0.166)	0.128	-0.0830	0.0049	1.28 (0.336)	0.0589
Canarias	-0.0205	1.51 (0.253)	0.054	-0.1630	0.0129	4.59 (0.053)	0.4440
Cantabria	0.0223	60.84 (0.000)	0.869	0.0074	0.0014	33.22 (0.000)	0.8775
Cataluña	-0.0122	1.53 (0.251)	0.056	0.0703	-0.0075	4.26 (0.061)	0.4204
C.-León	0.0037	0.55 (0.479)	-0.052	-0.0400	0.0040	4.26 (0.062)	0.4203
C. Mancha	0.0291	21.40 (0.002)	0.694	0.0237	0.0005	9.43 (0.010)	0.6519
C. Valenc.	-0.0008	0.03 (0.865)	-0.081	0.0301	-0.0028	1.47 (0.292)	0.0952
Extrema.	0.0134	1.64 (0.237)	0.066	-0.0640	0.0070	3.11 (0.108)	0.3192
Galicia	0.0172	11.68 (0.009)	0.543	0.0426	-0.0023	6.86 (0.022)	0.5658
Madrid	-0.0244	30.33 (0.001)	0.765	0.0200	-0.0040	45.71 (0.000)	0.9085
Murcia	-0.0062	1.73 (0.224)	0.075	0.0247	-0.0028	2.46 (0.155)	0.2453
Navarra	0.0102	6.56 (0.034)	0.382	0.0325	-0.0021	4.55 (0.054)	0.4409
Pais Vasco	0.0148	5.81 (0.043)	0.348	0.0020	0.0016	2.72 (0.134)	0.2763
Rioja	-0.0130	3.86 (0.085)	0.241	-0.0486	0.0032	2.92 (0.120)	0.2989

Entre paréntesis indicamos el nivel de significación.

R² es el coeficiente de determinación corregido.

Para saber si a lo largo del tiempo, las diferencias entre las utilidades aumentan o disminuyen, calculamos la varianza²⁰ de las utilidades entre regiones, mediante

$$\sigma_u^2 = \frac{1}{L} \sum_{i=1}^L u_i^2(t) - \left(\frac{1}{L} \sum_{i=1}^L u_i(t) \right)^2$$

20. Basta con comparar las varianzas de cada año, pues las medias temporales coinciden, al imponer en la restricción que la suma de las utilidades sea cero.

y obtenemos que existe una tendencia creciente en los cuatro primeros años, con máximo en 1989. A partir de este año, la varianza disminuye con mínimo en 1994 (cuando hay más diferencias entre las utilidades regionales, hay más posibilidad de movimientos migratorios que cuando las diferencias son mínimas).

Realizada la regresión de las utilidades en función del tamaño poblacional²¹, obtenemos las preferencias. La dispersión espacial de las preferencias, calculada de forma similar a la de las utilidades, nos indica la importancia del efecto tamaño para explicar las diferencias entre las utilidades. Si comparamos la varianza de las utilidades con la de las preferencias, podemos comprobar cómo la varianza de las preferencias es casi 20 veces más pequeña que la de las utilidades. Esto nos indica que las diferencias entre las utilidades regionales son debidas principalmente al efecto de su tamaño, siendo las CC.AA. más homogéneas respecto a las preferencias independientes del tamaño. Estudiamos la atractividad de cada CC.AA., separando las variaciones temporales de las preferencias, tabla 4. Siguen la misma estructura que las utilidades, tabla 3, si bien el ajuste es mejor para las preferencias.

Tabla 4: Estimación lineal y cuadrática de las preferencias regionales en función del tiempo

	<i>Regresión Lineal</i>			<i>Regresión Cuadrática</i>			
	t	F	R ²	t	t ²	F	R ²
Andalucía	-0.0041	1.76 (0.221)	0.077	0.0267	-0.0028	8.61 (0.013)	0.628
Aragón	-0.0042	0.35 (0.571)	-0.078	0.0672	-0.0064	7.18 (0.020)	0.578
Asturias	0.0026	0.26 (0.622)	-0.019	0.0489	-0.0042	4.15 (0.065)	0.452
Baleares	-0.0311	2.85 (0.130)	0.171	-0.0875	0.0051	1.56 (0.275)	0.110
Canarias	-0.0233	1.94 (0.201)	0.095	-0.1650	0.0129	4.82 (0.048)	0.459
Cantabria	0.0226	61.96 (0.000)	0.872	0.0076	0.0014	33.79 (0.000)	0.879
Cataluña	-0.0145	2.14 (0.182)	0.112	0.0670	-0.0074	4.49 (0.056)	0.437
C. León	0.0071	2.10 (0.185)	0.109	-0.0369	0.0040	5.88 (0.032)	0.520
C. Mancha	0.0300	23.31 (0.001)	0.712	0.0286	0.0001	10.20 (0.008)	0.672
C. Valenc.	-0.0043	0.79 (0.400)	-0.024	0.0255	-0.0027	1.63 (0.262)	0.123
Extremad.	0.0154	2.16 (0.180)	0.115	-0.0602	0.0069	3.29 (0.098)	0.338
Galicia	0.0225	16.93 (0.003)	0.639	0.0490	-0.0024	9.42 (0.010)	0.652
Madrid	-0.0279	41.23 (0.000)	0.818	0.0153	-0.0039	58.88 (0.000)	0.928
Murcia	-0.0086	3.33 (0.105)	0.206	0.0227	-0.0028	3.51 (0.088)	0.358
Navarra	0.0103	6.76 (0.032)	0.390	0.0329	-0.0021	4.73 (0.050)	0.454
Pais Vasco	0.0183	8.84 (0.018)	0.466	0.0046	0.0012	4.12 (0.066)	0.410
Rioja	-0.0128	3.72 (0.090)	0.232	-0.0488	0.0033	2.86 (0.124)	0.293

21. Para ello, hemos utilizado la transformación (18), de forma que se verifica que la suma de las utilidades coincide con la suma de las preferencias y es igual a cero.

Los valores $f_{ij}(t)$ que representan la parte simétrica de los flujos migratorios, describen la influencia de la "distancia" sobre la frecuencia de la migración. Para comprobar si $f_{ij}(t)$ depende de la distancia geográfica entre i y j , d_{ij} , y determinar cuál de las hipótesis (11, 12, 13, 14) se verifica en nuestro caso, hemos escogido como distancia la que hay entre las capitales de las Comunidades Autónomas, y para las islas, al no existir datos, le hemos aplicado en el caso de Baleares, la suma de la distancia Palma de Mallorca-Madrid y la de Madrid a la CC.AA. en cuestión. Para Canarias, hemos procedido de igual forma, con la distancia Santa Cruz de Tenerife-Madrid.

Ninguna de las funciones propuestas es válida para estos datos, indicando, por tanto, que existen otros factores además de la distancia geográfica que influyen en la migración. Llegamos a idéntico resultado, si utilizamos como distancia el número de fronteras a cruzar en el movimiento. Quizás, si trabajáramos con unidades regionales más pequeñas, como provincias o ciudades, para las cuales el concepto de distancia geográfica es más evidente, se obtendrían otros resultados.

Pero, analizando uno a uno los datos por regiones, podemos comprobar cómo generalmente, todas presentan un factor de movilidad mayor para regiones contiguas y cuando se refiere a Cataluña, Comunidad Valenciana y Madrid. Por ejemplo, Andalucía, tiene con Cataluña, un factor de movilidad de 9.64, seguido de 7.29 con Madrid y 4.86 con Valencia. También es alto con Canarias, Baleares y Murcia, y muy pequeño con las zonas del norte. Para Aragón, se tiene 2.40 con Cataluña, seguido de 1.36 con la Comunidad Valenciana y 1.18 con Madrid. Por tanto, de alguna forma, no es la distancia geográfica la que va a determinar los factores de movilidad entre regiones, sino que influye bastante la accesibilidad de las regiones, es decir, que estén bien comunicadas entre sí.

3.4. *Elección de las variables socioeconómicas*

Las variables se han escogido teniendo en cuenta que deben abarcar los diferentes campos socioeconómicos (mercado de trabajo, vivienda, etc.) y que deben estar disponibles a nivel de CC.AA. para el periodo 1985-1995. Necesitamos los datos referentes a 1985 pues admitimos que las variables pueden actuar con retardo sobre la migración, y sólo consideramos un año anterior porque de cualquier otra forma se podrían mezclar las consecuencias de la migración con sus causas²². Según estos criterios, las variables seleccionadas para la regresión de las utilidades y de la movilidad global han sido las siguientes:

22. Ver Bover y Antolín (1993).

Variables de Tamaño:

- Población, $n_i(t)$.
- Población al cuadrado, $n_i^2(t)$.
- Densidad de población, $Dn_i(t)$.

Mercado Laboral:

- Tasa de actividad, $TA_i(t)$.
- Tasa de ocupación, $TO_i(t)$.
- Tasa de paro, $TP_i(t)$.
- Porcentaje regional de activos en la agricultura, (población activa de la región en el sector agrícola dividido por la población activa total de la región), $Aagr_i(t)$.
- Porcentaje regional de activos en la industria, $Aind_i(t)$.
- Porcentaje regional de activos en el sector servicios, $Aserv_i(t)$.
- Porcentaje regional de activos en la construcción, $Aconst_i(t)$.
- Porcentaje regional de ocupados en la agricultura, (población ocupada de la región en el sector agrícola dividido por la población ocupada total de la región), $Oagr_i(t)$.
- Porcentaje regional de ocupados en la industria, $Oind_i(t)$.
- Porcentaje regional de ocupados en el sector servicios, $Oserv_i(t)$.
- Porcentaje regional de ocupados en la construcción, $Oconst_i(t)$.

Vivienda:

- Precio del metro cuadrado de la vivienda, $V_i(t)$.

Otras variables:

- Renta per cápita, $R_i(t)$.
- Producto Interior Bruto per cápita, $PIB_i(t)$.
- Prestaciones por desempleo por habitante, $PD_i(t)$.

Las series de variables empleadas en este trabajo, han sido obtenidas principalmente de las publicaciones del INE, salvo para la renta, que hemos consultado la publicación del BBV²³. Para la vivienda sólo existen datos a partir de 1987, y no para todas las CC.AA.. Para completarlos, hemos utilizado el Índice de Precios de Consumos, IPC, calculando el precio en t como el producto del precio en $(t+1)$ por el IPC del año t con base en $(t+1)$. Todas las variables que indican valor monetario, están en pesetas constantes (base 1985).

23. Como sólo existen datos de renta para años impares, realizamos una interpolación para obtener los datos que faltan.

Para evitar problemas con las diferentes unidades de medida, hemos transformado todas las variables según (18). Las estimaciones han sido realizadas con el programa SPSS.

Representación de la movilidad global en función de variables socioeconómicas

Debido a su propia estructura, para estudiar la movilidad global debemos utilizar variables a nivel nacional, no por CC.AA.. Con sólo 4 variables, renta per cápita, activos en el sector servicios y en el sector agrícola, junto con los ocupados en agricultura, podemos explicar casi el 98% de la misma (detalles en tabla 5). Todas las variables están con un retardo.

Tabla 5: Resultados del análisis de regresión sobre la movilidad global

<i>Variables</i>	<i>Coefficientes no estandarizados</i>	<i>Coefficientes estandarizados</i>	<i>t</i>
Constante	-3.248E-03		-6.995
A _{serv} (t-1)	2.4556E-05	1.285	5,707
R(t-1)	2.612E-05	2.008	5.782
O _{agr} (t-1)	1.101E-04	4.481	4.672
A _{agr} (t-1)	-5.351E-05	-2.036	-3.020
R ² =0.987			
R ^{2*} =0.976			

t es el valor de la t de student para cada parámetro ajustado.

Representación de las utilidades en función de variables socioeconómicas

La estructura espacio-temporal de las utilidades, viene explicada por 9 variables, tabla 6, con un coeficiente de determinación corregido de 0.920, teniendo la variable población los signos correctos, siendo estos los factores que más influyen en la utilidad (basta fijarse en el valor de los coeficientes estandarizados). Cuanto mayor sea la tasa de actividad de la región y las prestaciones por desempleo, mayor será la utilidad de la región. También tienen signo positivo, la renta per cápita, el porcentaje de activos en el sector de la construcción y de ocupados en los servicios (todos con un retardo). Por el contrario, conforme aumenta el precio medio de la vivienda (con un retardo) y la tasa de paro, la CC.AA. es menos atractiva.

Si para cada CC.AA., comparamos las utilidades estimadas con las teóricas, se mantiene la ordenación espacial de las utilidades regionales, si bien la calidad de

Tabla 6: Resultados del análisis de regresión sobre las utilidades.

Variables	Coefficientes no estandarizados	Coefficientes estandarizados	t
$n_i(t)$	1.130	1.977	20.447
	-0.372	-1.158	-11.982
$Aconst_i(t-1)$	0.350	0.143	5.293
$Oserv_i(t-1)$	0.915	0.280	7.762
$TA_i(t)$	0.987	0.108	3.802
$V_i(t-1)$	-0.410	-0.416	-8.379
$R_i(t-1)$	0.167	0.141	2.007
$PD_i(t)$	0.286	0.234	3.980
$TP_i(t)$	-0.160	-0.089	-2.436

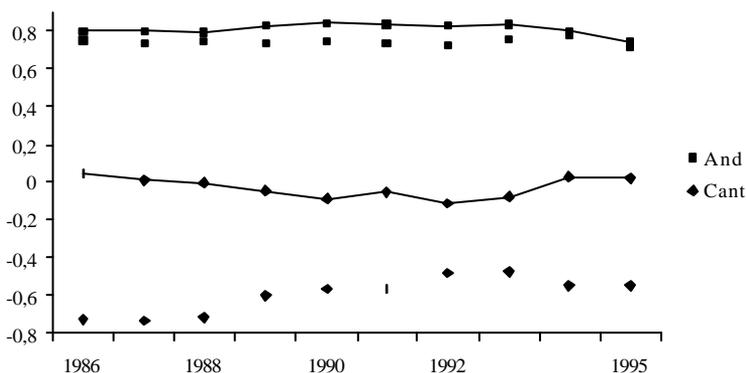
$R^2 = 0.924$

$R^{2*} = 0.920$

Durbin-Watson = 1.883

t es el valor de la t de student para cada parámetro ajustado.

ajuste para cada CC.AA. es más baja. De hecho, la función de utilidad dada por la tabla 3, no sirve para explicar por separado la utilidad de todas y cada una de las CC.AA.. Así, en regiones como por ejemplo Andalucía, podría ser válida la estructura dada anteriormente para la utilidad, pero no en otras, como Cantabria, gráfico 3. Por ello, analizamos cada una de las CC.AA. por separado, y establecemos cuál de las hipótesis formuladas se verifica.

Gráfico 3: Valores estimados (línea sólida) y observados (puntos) de las utilidades

La regresión para cada CC.AA., nos conduce a los resultados ofrecidos en la tabla 7. Sólo en Canarias y la C. Valenciana, aparece el tamaño poblacional como variable explicativa en las utilidades, a través de n_i , en ningún caso es significativa la población al cuadrado. Mientras hay un grupo de CC.AA. (Canarias, Cataluña, Madrid, Navarra, etc.) en las que la utilidad se puede explicar mediante las variables que tenemos, hay otras (Rioja, Murcia, Aragón, etc.) en las que haría falta tener en cuenta otras variables.

Tabla 7: Regresión de las utilidades de cada CC.AA.

CC.AA.	Variables explicativas	R^{2*}
Andalucía	Aconst(t-1)	0.642
Aragón	TA(t)	0.659
Asturias	Aconst(t)	
Baleares	Oagr(t) Aind(t-1) R(t-1) Aserv(t)	0.831 0.838
Canarias	n_i TP(t) Oserv(t-1)	 0.994
Cataluña	Aserv(t) TA(t-1) V(t)	 0.986
Cantabria	Aind(t-1) Oconst(t-1)	 0.896
Castilla-León	Aagr(t-1)	0.718
C. la Mancha	Aserv(t) D(t-1)	 0.871
C. Valenciana	n_i TP(t)	 0.724
Extremadura	Oagr(t) Oind(t-1)	 0.887
Galicia	Aserv(t)	0.880
Madrid	Oserv(t) TP(t) Oagr(t)	 0.957

Tabla 7: Regresión de las utilidades de cada CC.AA. (cont.)

CC.AA.	Variables explicativas	R^{2*}
Murcia	PD(t) Aind(t)	0.642
Navarra	Aagr(t) Oind(t-1) TA(t) D(t-1)	0.953
País Vasco	Aserv(t) Oagr(t) Oconst(t)	0.893
Rioja ^a		

^a No puede estimarse con las variables actuales.

Si realizamos la regresión año a año, independientemente de las CC.AA., obtenemos los datos reflejados en la tabla 8. En todos los años, la bondad del ajuste es muy buena.

Tabla 8: Regresión de las utilidades por año.

Año	Variables explicativas	R^{2*}
1986	$n_i(t)$ $n_i^2(t)$ Oind(t-1) Oserv(t-1) V(t)	0.965
1987	$n_i(t)$ $n_i^2(t)$ Oconst(t) Aconst(t-1) Aserv(t-1) TP(t) V(t-1)	0.992
1988	$n_i(t)$ Oconst(t-1) TP(t-1)	0.840
1989	$n_i(t)$	0.704
1990	$n_i(t)$ $n_i^2(t)$	0.869

Tabla 8: Regresión de las utilidades por año (cont.)

Año	Variables explicativas	R ² *
1991	$n_i(t)$ $n_i^2(t)$	0.851
1992	$n_i(t)$ $n_i^2(t)$ Aconst(t-1)	0.866
1993	$n_i(t)$ $n_i^2(t)$ Oserv(t) V(t)	0.946
1994	$n_i(t)$ $n_i^2(t)$ Oserv(t-1) V(t-1)	0.965
1995	$n_i(t)$	0.961

Aparece siempre el tamaño poblacional y la población al cuadrado, salvo para 1988 y 1989. Cuando las variables económicas no aportan información para explicar las utilidades (nos estaría indicando que con estas variables no existe diferencias económicas en estos años), la bondad disminuye.

Bibliografía

- ANSELIN, L.; ISARD, W. (1979): "On Alonso's General Theory of Movement". *Man, Environment, Space and Time*, Vol.1, nº 1, pp. 52-63.
- BENTOLILA, S.; DOLADO, J.J. (1990): "Mismatch and Internal Migration in Spain". *Banco de España, Documento de trabajo, nº 9006*.
- BOVER, O.; ANTOLÍN, P. (1993): "Migraciones Regionales en España". *Banco de España, Documento de Trabajo nº 9318*.
- CABRÉ, A.; MORENO, J.; PUJADAS, I. (1985): "Cambio Migratorio y Reconversión Territorial en España". *Revista Española de Investigaciones Sociológicas*, nº 32, pp. 43-65.
- GEWEKE, J.; MARSHALL, C.; ZARKIN, G.A. (1986): "Mobility Indices in Continuous Time Markov Chains". *Econometrica*, Vol.54, nº 6, pp. 1407-1423.
- GINSBERG, R.B. (1971): "Semi-Markov Processes and Mobility". *Journal of Mathematical Sociology*, Vol.1, pp. 233-262.
- GINSBERG, R.B. (1979a): "Timing and Duration Effects in Residence Histories and Other Longitudinal Data. I.- Stochastic and Statistical Models". *Regional Science and Urban Economics*, Vol. 9, pp. 311-331.

- GINSBERG, R.B. (1979b): "Timing and Duration Effects in Residence Histories and Other Longitudinal Data. II.- Studies of Duration Effects in Norway, 1965-1971". *Regional Science and Urban Economics*, Vol. 9, pp. 369-392.
- GONZÁLEZ PÉREZ, J.M. (1991): "Modelo Explicativo de los Flujos Migratorios en España: Incidencia en la dispersión del desempleo interregional 1960-85". *Documento de Trabajo*, nº 33, Facultad de Ciencias Económicas y Empresariales, Universidad de la Laguna, Diciembre 1991.
- GORDON, I. (1991): "Multi-Stream Migration Modelling". *Migration Models*, pp. 73-91; Ed. Stillwell-Congdon.
- GREENWOOD, M.J.; HUNT, G.L. (1984): "Econometrically Accounting for Identities and Restrictions in Models of Interregional Migration". *Regional Science and Urban Economics*, Vol.14, pp. 113-128.
- HAAG, G. (1989): *Dynamic Decision Theory: Applications to Urban and Regional Topics*. Studies in Operational Regional Science, Kluwer Academic Publishers.
- HAAG, G.; MAX, H. (1995): "Rank-Size Distribution of Settlement Systems: A Stable Attractor in Urban Growth". *Papers in Regional Science: The Journal of the RSAI*, Vol.74, nº 3, pp. 243-258.
- HAAG, G.; MUNZ, M.; PUMAIN, D., SANDERS, L.; SAINT-JULIEN, TH. (1992): "Interurban Migration and the Dynamics of a System of Cities: 1.The Stochastic Framework with Application to the French Urban System". *Environment and Planning A* Vol. 24, pp. 181-198.
- HAAG, G.; WEIDLICH, W. (1984): "A Stochastic Theory of Interregional Migration". *Geographical Analysis*, Vol.16, nº 4, pp. 331-357.
- HAAG, G.; WEIDLICH, W. (1986): "A Dynamic Migration Theory and its Evaluation for Concrete Systems". *Regional Science and Urban Economics*, Vol. 16, nº 1, pp. 57-80.
- HAKEN, H. (1974): «The Stationary Solution of the Master Equation for Detailed Balance». *Physical Letters*, 46A, 7.
- HELBING, D. (1995): *Quantitative Sociodynamics. Stochastic Methods and Models of Social Interaction Processes*, Ed. Kluwer Academic Publishers.
- ISSERMAN, A.M.; PLANE, D.A.; ROGERSON, P.A.; BEAUMONT, P.M. (1985): "Forecasting Interstate Migration with Limited Data: A Demographic-Economic Approach". *Journal of the American Statistical Association*, Vol. 80, nº 390, pp. 277-285.
- KANAROGLOU, P.; LIAW, K-L.; PAPAGEORGIOU, Y.Y. (1986): "An Analysis of Migration Systems: 1. Theory". *Environment and Planning A* Vol. 18, pp. 913-928.
- KNUDSEN, D.C.; FOTHERINGAM, A.S. (1992): "Matrix Comparison, Goodness of Fit, and Spatial Interaction Modelling". *International Regional Science Review*, Vol. 10, nº 2, pp. 127-147.

- LEGUINA HERRÁN, J.; FERNÁNDEZ CORDÓN, A. (1991): "Situación y Perspectivas de la Población Española". *Economía y Sociología del Trabajo*, nº 11, pp. 52-60.
- LOWRY, I. (1966): *Migration and Metropolitan Growth: Two Analytic Models*. San Francisco, Ed. Chandler.
- MCFARLAND, D.D. (1970): "Intragenerational Social Mobility as a Markov Process: Including a Time-Stationary Markovian Model that Explains Observed Declines in Mobility Rates over Time". *American Sociological Review*, Vol. 35, pp. 463-475.
- MCGINNIS, R. (1968): "A Stochastic Model Social Mobility". *American Sociological Review*, Vol. 33, nº 5, pp. 712-721.
- PUMAIN, D. (1982): *La Dynamique des Villes*, París.
- RÓDENAS CALATAYUD, C. (1997): "¿Son Bajos los Flujos Migratorios en España?". *Revista de Economía Aplicada*, nº 15 (Vol. V), pp. 155-171.
- ROGERSON, P.A. (1984): "New Directions in the Modelling of Interregional Migration". *Economic Geography*, nº 60, pp. 111-121.
- ROWE, G.T.; KRISHNAN, P. (1983): "Internal Migration in India. Results from Three Stochastic Process Models". *Population Geography*, Vol. 5, pp. 70-80.
- SALKIN, M.S.; LIANOS, T.P.; PARIS, Q. (1975): "Population Predictions for the Western United States: A Markov Chain Approach". *Journal of Regional Science*, Vol. 15, nº 1, pp. 52-60.
- SAMPSON, M. (1990): "A Markov Chain Model for Unskilled Workers and the Highly Mobile". *Journal of the American Statistical Association*, Vol. 85, nº 409, pp. 177-180.
- SANTIAGO HERNANDO, RAFAEL DE (1994): *Migraciones, Salarios y Desempleo. Un Modelo para la Economía Española*. Secretariado de Publicaciones, Universidad de Valladolid.
- SJAASTAD, L.A. (1962): "The Cost and Return of Human Migration". *Journal of Political Economy*, Vol. 70, nº 5, Supplement, pp. 80-93.
- STOUFFER, S.A. (1940): "Intervening Opportunities: A Theory Relating Mobility and Distance". *American Sociological Review*, Vol. 5, nº 6, pp. 845-867.
- TODARO, M.P. (1969): "A Model of Labor Migration and Urban Unemployment in Less Developed Countries". *The American Economic Review*, pp. 138-148.
- WEIDLICH, W. (1988): "Stochastic Migration Theory and Migratory Phase Transitions". *Journal of Computational and Applied Mathematics*, Vol. 22, nº 2-3, pp. 359-368.
- WEIDLICH, W; HAAG, G. (1987): "A Dynamic Phase Transition Model Spatial Agglomeration Processes". *Journal of Regional Science*, Vol. 27, nº 4, pp. 529-569.
- WEIDLICH, W; HAAG, G. (1988): *Interregional Migration: Dynamic Theory and Comparative Analysis*. Ed. Springer.

ZIPF, G.K. (1946): "The P_1P_2/D Hypothesis: On the Intercity Movement of Persons". *American Economic Review*, Vol .11, pp. 677-686.