Estadística Española • Volumen 56, número 185 / 2014, pp. 381-399

Incorporating demographic knowledge in parametric models for forecasting households size*

Silvia Bermúdez Parrado
Instituto de Estadística y Cartografía de Andalucía

Abstract

In this paper we show how expert knowledge, such as hypothesis of convergence, can be included in a parametric model for forecasting households size. To do this we fit the parameters of a generalized multinomial logistic model by solving an optimization problem, in which the demographic knowledge is expressed as parameters constraints. The parameters are tuned by means of standard and user-friendly optimization software.

The methodology is tested on Spain and one of its regions, Andalusia.

Keywords: Projections, Households size, Households shares, Hypothesis of convergence, Mathematical Optimization.

AMS Classification: 91D20, 62P25, 65K05, 90C05

Modelos paramétricos para la proyección del tamaño de los hogares incorporando conocimiento demográfico

Resumen

En este artículo se demuestra cómo el conocimiento experto puede ser incorporado al modelo paramétrico en forma de Hipótesis de Convergencia para proyectar la variable tamaño del hogar. Para esto, se han ajustado los parámetros de un modelo logístico multinomial generalizado mediante la resolución de un problema de optimización, en el que el conocimiento demográfico se integra en las restricciones del problema. Los parámetros se obtienen utilizando software de optimización muy extendido y de fácil manejo.

* Research partially supported by project MTM2012-36163, Spain, FQM-329, Andalucía, Spain, both co-fded by EU ERD funds.

La metodología que aquí se presenta se prueba con datos de España y una de sus comunidades autónomas, Andalucía.

Palabras clave: Proyecciones, Tamaño del hogar, Distribución de los hogares, Hipótesis de convergencia, Optimización Matemática.

Clasificación AMS: 91D20, 62P25, 65K05, 90C05

1 Introduction

Households are basic units of cohabitation and consumption, and for this reason they have notorious relevance in public planning. It is then critical to know the volume and patterns of the households in a given territory, as well as to have accurate forecasts to support mid-term and long-term decisions which concern households and their inhabitants, such as social and assistance services or housing procurement, see e.g. (2, 11, 14, 15, 18) and the references therein.

Two features associated with households size call for a rigorous forecast analysis: the number of households of each size class, as well as their distribution (shares).

When the number of households of the different size classes are to be forecasted, a good amount of information is usually required as input, (1, 3, 9, 16, 17, 19, 23, 24), and sophisticated computational tools, such as LIPRO, (25), may be required. Once such forecast is done, shares forecast can easily be computed. Two main drawbacks are found in this two-phase approach. First, complex, not always easily accessible and error-pruning, information is needed; second, possible, if not likely, errors in the population forecast, are inherited by shares forecast, which are then derived forecasts.

In this paper, household shares are forecasted directly, as in e.g. (2, 14), without forecasting first the number of households of different sizes. Our main contribution is to show that, this way, expert knowledge, (5), can be easily accommodated. Indeed, if hypotheses of convergence (6, 7, 8, 26) are imposed, so that different territories expected to behave similarly in the future are forced to have convergent forecasts for their shares, one only needs to add nonlinear constraints to the optimization problem. To solve such optimization problems, user-friendly algorithmic tools, such as the Solver tool in Microsoft Excel, (12, 10), are shown to be sufficient.

The paper is structured as follows. After this introduction, Section 2 describes a parametric model due to (14), which is here improved at the expense of introducing more parameters. Then, hypotheses of convergence are included in both models. All these models are tested and compared on household data from Spain and one of its regions, namely, Andalusia. Conclusions and future research lines are given in Section 3.

2 The models

2.1 HOR model

In (14), Haupt, Oberhofer and Reichsthaler introduce a multinomial logistic model, called hereafter HOR model, which will be seen here as the basic model. Its main features are briefly discussed.

A population H of households is given, whose features are evolving in time. Let us assume in this population the variable household size is measured, taking the values $1,2,...,I-1$, and aggregating to I the tail of the distribution, i.e., those households with I or more members.

For the year t and for each possible value $i = 1,2,...,I$ of the household size, let f_{it} denote the fraction of households of size i in year t.

In the HOR model, the fraction f_{it} of households which, in years $t = T+1, T+2,...,T+r$, take the value i, $i = 1,2,...,I$ are forecasted by a multinomial logistic model, e.g. (13, 22),

$$\hat{f}_{it} = \frac{e^{\alpha_i+\beta_i t}}{\sum_{j=1}^{1} e^{\alpha_j+\beta_j t}} \quad \forall i = 1,...,I \quad \forall t = 1,... \qquad [1]$$

for some parameters α_i, β_i, $\forall i = 1,...,I$, to be tuned.

In (14) such parameters are fitted by maximizing the sample likelihood. In this paper we propose to minimize the sum of the squares of the errors in the model for periods $1,2,...,T$, i.e., by solving the nonlinear optimization problem

$$\min \sum_{t=1}^{T} \sum_{i=1}^{I} \left(f_{it} - \hat{f}_{it}\right)^2$$

$$\hat{f}_{it} = \frac{e^{\alpha_i+\beta_i t}}{\sum_{j=1}^{1} e^{\alpha_j+\beta_j t}} \quad \forall i = 1,...,I \quad \forall t = 1,...,T \qquad [2]$$

$$\alpha_i, \beta_i \in \mathbb{R} \ \forall_i = 1, \dots, I$$

Observe that the model is over-parameterized, and then there is no loss of generality in imposing $\alpha_1 = \beta_1 = 0$.

As an illustration, we have fitted the model to the data of households size in Andalusia, Spain, in the period 1987 − 2011, split into 5 categories, as done in (29). Data are obtained from the so-called Encuesta de Población Activa (EPA) (28).

ESTADÍSTICA ESPAÑOLA Vol. 56 Núm. 185 / 2014

Table 1

Optimal parameters (HOR model). Andalusia, 1987 − 2011

Parameter	α_1	β_1	α_2	β_2	α_3	β_3	α_4	β_4	α_5	β_5
Value	0.000	0.000	0.810	-0.013	0.861	-0.020	1.222	-0.033	1.617	-0.089

Problem [2] was solved with the optimization tool Solver, (10, 12), which allows one to solve in a user-friendly environment nonlinear optimization problems like [2]. We should point out that, although Solver presents some limitations in terms of the number of variables and constraints to be used, it successfully handles problems of the size considered here.

The optimal values of the parameters, as provided by the software, for household size in Andalusia are given in Table 1, and the optimal value of the problem is 0.0045.

Figure 1 represents the observed data, and the data fitted and projected to the period 2011−2016. Remark that the forecast horizon is very short, which makes such projections rather reliable. We see that the fit is good, though higher deviations are observed in the tails. See for instance the plots for $i = 2$ and $i = 4$.

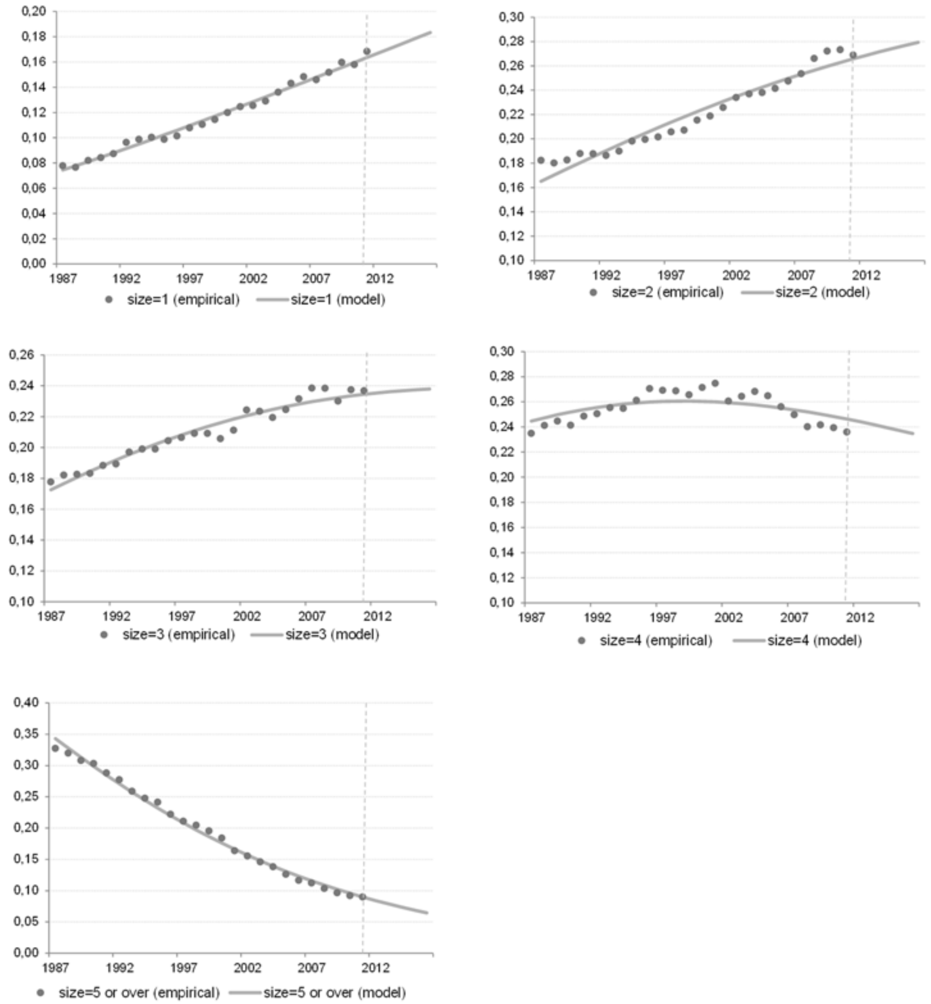
The very same analysis, with the very same software, has been performed for the household size in Spain. However, the data series available is shorter (2005 − 2011, instead of 1987 − 2011), (27). In spite of the much smaller amount of information available, the fit for Spain with the HOR model is rather good, see Figure 2, with a very small optimal value of the problem, namely, 0.000057.

2.2 An improved model

It was observed in the figures presented that the HOR model, though providing a reasonable fit, yields larger errors in the tails when data present a strong nonlinear behavior. At the expense of introducing more parameters, thus making the model less parsimonious, we propose to replace [2] by what we call Improved HOR (I-HOR)

Figure 1

Data fit and projections (HOR model). Andalusia, 1987 − 2016



model,

$$\hat{f}_{it} = \frac{t^{\mu_i} e^{\alpha_i+\beta_i t}}{\sum_{j=1}^{I} t^{\mu_j} e^{\alpha_j+\beta_j t}} \quad \forall i = 1,....I \quad \forall t = 1,...,T \qquad [3]$$

We have then I more parameters to tune, generated by replacing the exponential- type numerators in [1] by gamma-type numerators, popular in Demography, e.g. (4, 20). However, the resulting problem can still be successfully handled by solver, without a remarkable increase of running times. As for model [2], [3] is over-parameterized, and thus we can set, without loss of generality, $\alpha_1 = \beta_1 = \mu_1 = 0$.

Figure 2

Data fit and projections (HOR model). Spain, 2005 − 2016

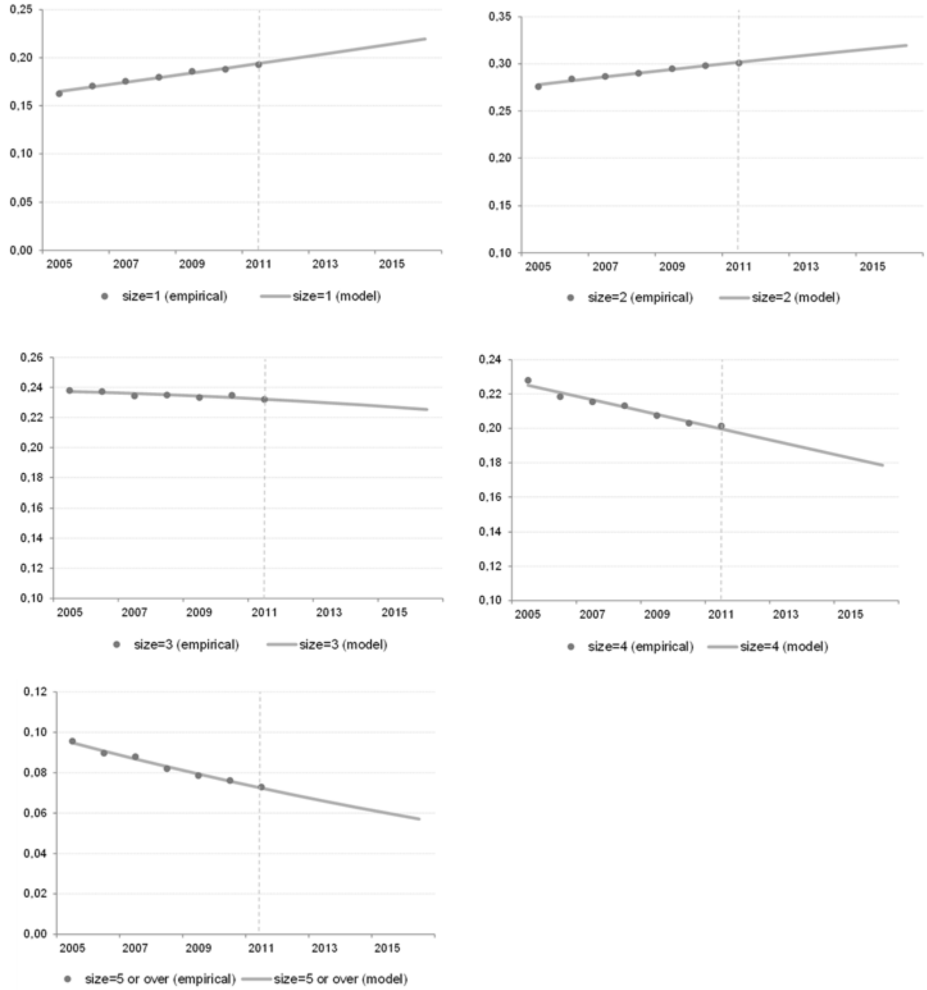


Table 2

Optimal parameters (I-HOR model). Andalusia, 1987 − 2011

Parameter	α_1	α_2	α_3	α_4	α_5
Value	0.000	0.879	0.857	1.110	1.522
Parameter	β_1	β_2	β_3	β_4	β_5
Value	0.000	-0.007	-0.020	-0.044	-0.103
Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
Value	0.000	-0.063	0.003	0.110	0.116

Figure 3 shows the fit with our I-HOR model for Andalusia. The results are much sharper than those given by the HOR model: The optimal value as given by solver was 0.00196. The parameters values are presented in Table 2.

Doing the same for the data of Spain, we have obtained the fit presented in Figure 4. The error, 0.000025, is slightly smaller than the error given by the HOR model. However, the fit looks much better in our improved model.

Figure 3

Data fit and projections (I-HOR model). Andalusia, 1987 – 2016

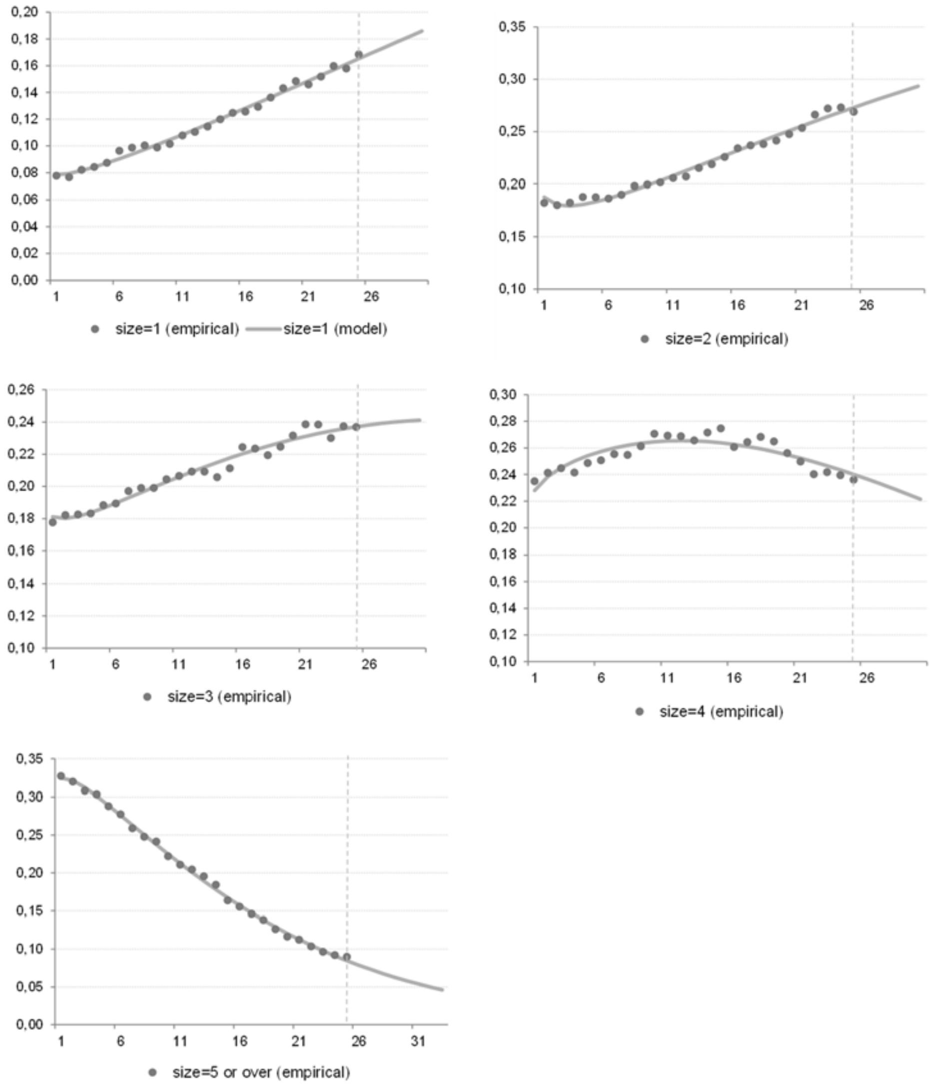
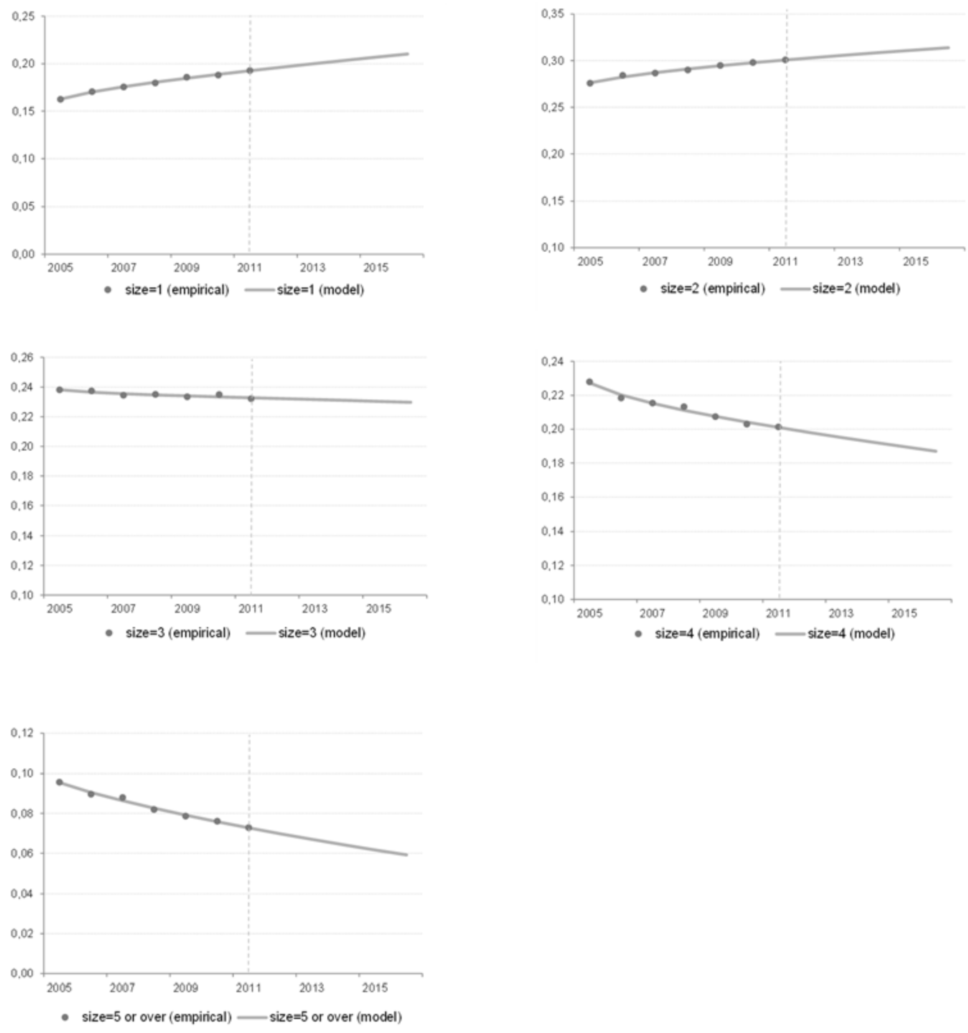


Figure 4

Data fit and projections (I-HOR model). Spain, 2005 − 2016



Summarizing, Figures 1-4 show that the I-HOR model yields a better fit than HOR. This is confirmed by the optimal errors obtained by the two models. However, since the I-HOR model has more parameters, it is not demonstrated in the previous analysis whether such better fit may cause overfitting. We have made the following test: the data in Andalusia for years 1987 − 2008 have been used to estimate the parameters in models HOR and I-HOR, and the obtained forecasts are compared against the true values in years 2009 − 2011. Tables 3 and 4 show respectively the optimal parameters for HOR and I-HOR models in this new experiment. If we compare such parameters with those given in Tables 1 and 2 for period 1987 − 2011, we see that the optimal parameters only suffer very small changes.

Table 3

Optimal parameters (HOR model). Andalusia, 1987 – 2008

Parameter	α_1	β_1	α_2	β_2	α_3	β_3	α_4	β_4	α_5	β_5
Value	0.000	0.000	0.819	-0.014	0.854	-0.019	1.200	-0.030	1.604	-0.087

Table 4

Optimal parameters (I-HOR model). Andalusia, 1987 − 2008

Parameter	α_1	α_2	α_3	α_4	α_5
Value	0.000	0.881	0.864	1.122	1.528
Parameter	β_1	β_2	β_3	β_4	β_5
Value	0.000	-0.007	-0.018	-0.040	-0.102
Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
Value	0.000	-0.066	-0.012	0.086	0.108

Now we analyze how HOR and I-HOR models are able to forecast when their optimal parameters, as given in Tables 3 and 4 are used. Table 5 shows the ratio between mean squared errors of HOR and I-HOR models in the 5 categories along the forecasting period 2009−2011, showing that HOR and I-HOR have a similar behavior in category i = 3, but I-HOR clearly outperforms HOR in the remaining categories. In other words, with the data of Andalusia, I-HOR model gives a better fit without overfit.

An extension of this improved model would consist of including weights, to penalize in a different way the different years. We did not observe in our experiments a significant improvement with such a model (which would involve the tuning of the weights) and thus the results are not reproduced here.

Table 5

Ratios of HOR vs. I-HOR errors. 2009 − 2011. Andalusia

Size	*1*	*2*	*3*	*4*	*5*
$\frac{HOR}{I-HOR}$	1.63	4.33	0.93	3.58	2.14

2.3 Incorporating convergence hypotheses in the model

So far two models have been described for households size, namely, the HOR model of (14), and our improvement, I-HOR, showing that the fit for Andalusia and Spain are very good, see Figures 1-4. The two territories, Andalusia and Spain, are hierarchically nested, and thus it is not sensible to perform independent analyses. Indeed, projections should keep some degree of consistency for the different units in the hierarchy. To do that we propose to perform the projections simultaneously for all the units in the nested hierarchy, allowing one to accommodate in the model hypothesis on the joint behavior of the different units. More precisely, given a set of territories, $n = 1,2,...,N,$ we want to fit and forecast f_{nit}, the fraction of households in territory n having size i in year t, and this fit is done simultaneously for the different territories n, accommodating in the model hypotheses of convergence. In other words, we seek values of the parameters defining f_{nit} imposing that, at time T', the distribution of households is identical in the different territories n, $n = 1,...,N$, i.e., $f_{niT'} = f_{n*iT'}$ for all n, $n*$. This convergence hypothesis may be seen as too restrictive, and an approximate convergence can be used instead, by assuming that, at time T', for all households size i, and all territories $n, n*,$ the discrepancy between f_{niT} and $f_{n*iT'}$ should be small, i.e.,

$$\left| f_{niT'} - f_{n*iT'} \right| \leq \varepsilon \quad \forall n, n* = 1,...,N,$$

where ε is a small positive number. This leads to the following nonlinear optimization problem, which we will call I-HOR model with hypotheses of convergence, I-HOR-C.

Table 6

Optimal parameters (I-HOR-C model). Andalusia and Spain

Spain	Parameter	α_1	α_2	α_3	α_4	α_5
	Value	0.000	0.539	0.400	0.358	-0.478
	Parameter	β_1	β_2	β_3	β_4	β_5
	Value	0.000	-0.017	-0.027	-0.024	-0.062
	Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
	Value	-0.000	0.010	-0.017	-0.074	-0.036
Andalusia	Parameter	α_1	α_2	α_1	α_1	α_5
	Value	0.000	0.541	0.494	0.648	-0.052
	Parameter	β_1	β_2	β_3	β_4	β_5
	Value	0.000	-0.023	-0.037	-0.033	-0.072
	Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
	Value	-0.000	0.067	-0.064	-0.034	-0.037

$$\min \sum_{n=1}^{N} \sum_{t=1}^{T} \sum_{i=1}^{I} \left(f_{nit} - \hat{f}_{nit} \right)^2$$

s.t.

$$\hat{f}_{nit} = \frac{t^{\mu_{ni}} e^{\alpha_{ni} + \beta_{ni} t}}{\sum_{j=1}^{I} t^{\mu_{nj}} e^{\alpha_{nj} + \beta_{nj} t}} \; \forall n = 1,...,N \; \forall i = 1,...,N \; \forall t = 1,...,N \qquad [4]$$

$$\left| \hat{f}_{niT'} - \hat{f}_{n*iT'} \right| \leq \varepsilon \quad \forall n, n* = 1,...,N, \forall i = 1,...,I$$

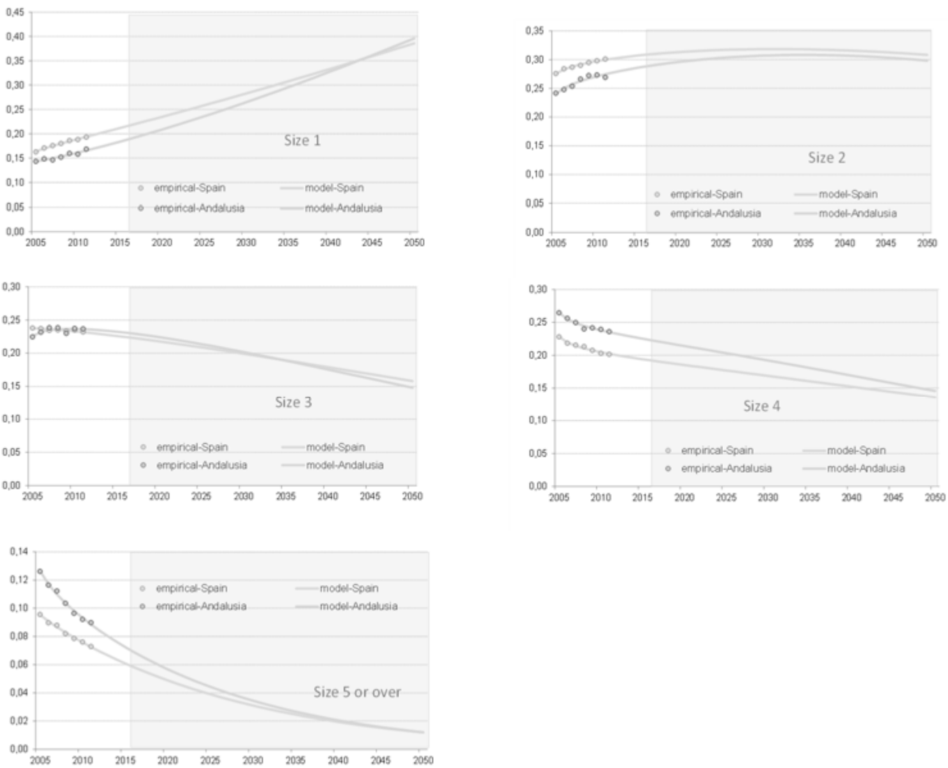
$$\alpha_{ni}, \beta_{ni}, \mu_{ni} \in \mathbb{R} \;\; \forall i = 1,...,I \;\; \forall n = 1,...,N$$

The I-HOR-C model is tested on N = 2 territories (Andalusia and Spain), with convergence at time T' = 2050, much further than the forecasting horizon considered. The optimal parameters in the I-HOR-C model [4] using a hypothesis of approximate convergence, with $\varepsilon = 0.01$, are given in Table 6.

Figure 5 shows the resulting forecasts for Andalusia and Spain as territories, and $\varepsilon = 0.01$. Data are available for the period 2005 − 2011, (27, 28), and the forecasting horizon is 5 years, i.e., the forecasting interval is 2011−2016. We have shaded in Figure 5 the time period outside our forecasting horizon, but depicted to show the fulfillment of the hypothesis of convergence.

Figure 5

Data fit and projections (I-HOR-C model). Spain and Andalusia, 2005−2050



The forecasts obtained this way may be, as expected, different than those obtained without hypothesis of convergence. The forecasts without and with convergence for Andalusia and Spain are shown respectively in Figures 6 and 7.

Figure 6

Data fit and projections (I-HOR vs I-HOR-C). Andalusia, 2005 − 2016

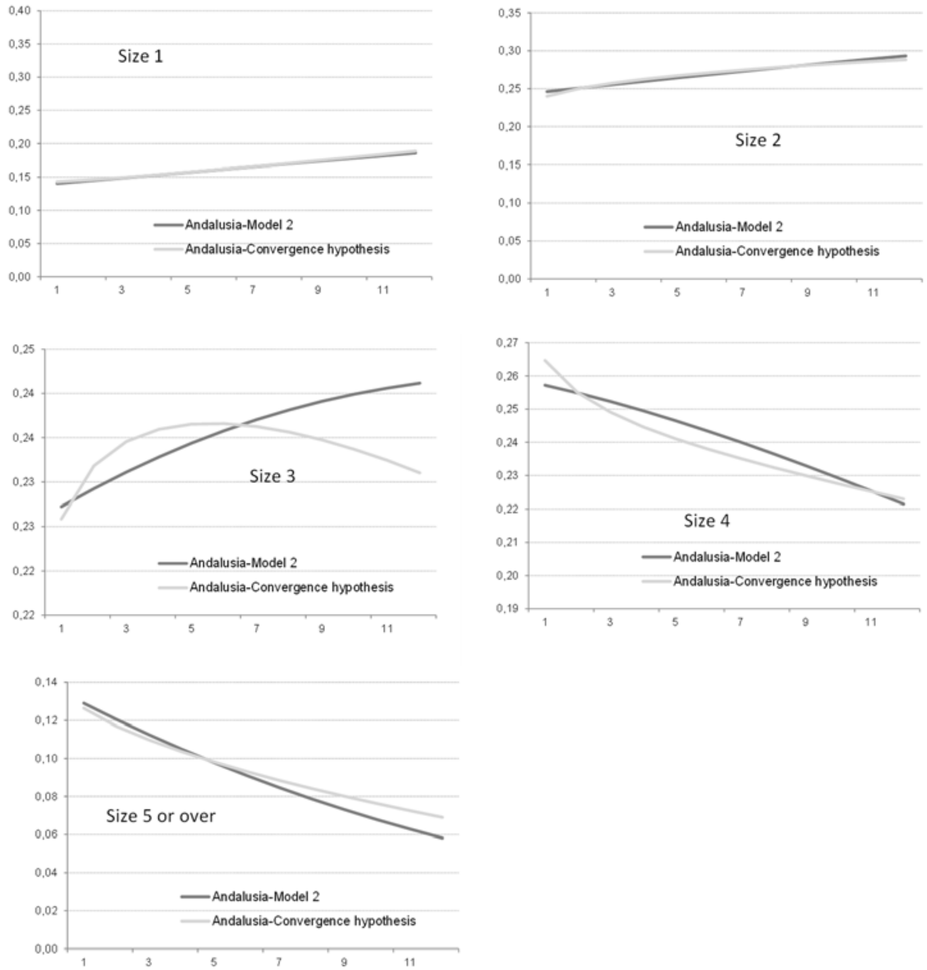
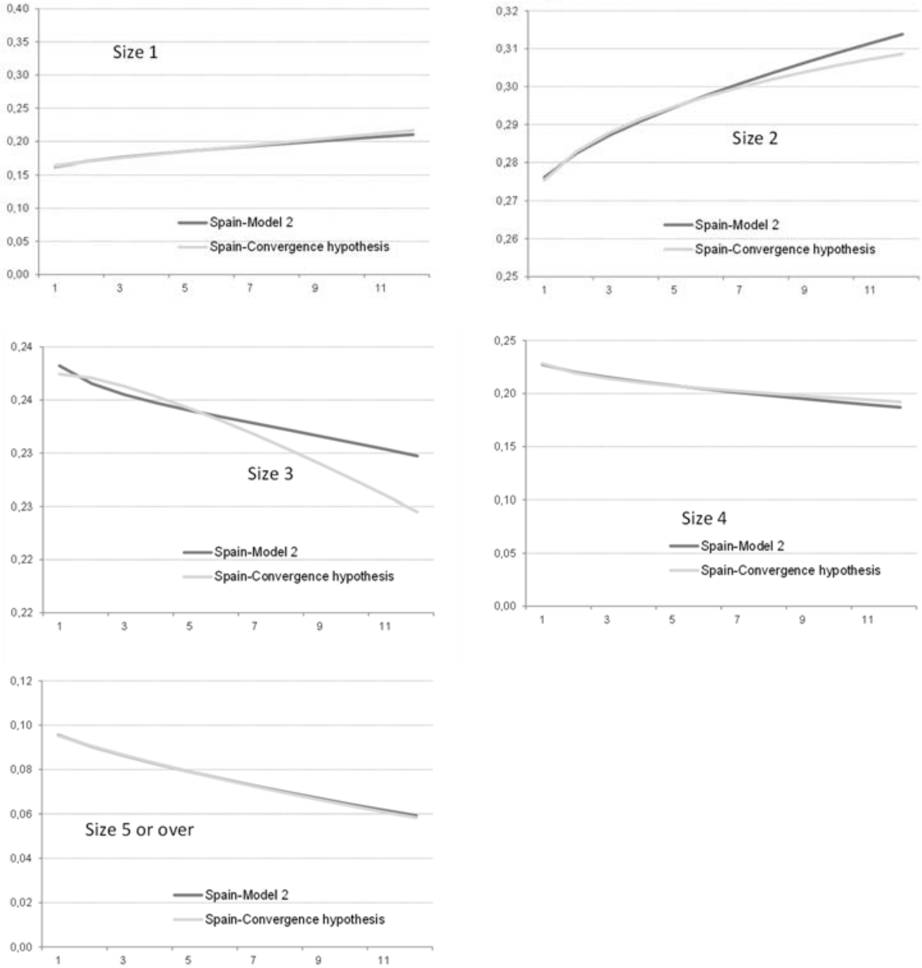


Figure 7

Data fit and projections (I-HOR vs I-HOR-C). Spain, 2005 − 2016



2.4 Hypotheses of convergence and data series of different lengths

The time series available for the different territories may not necessarily have the same length, since the different statistical offices may provide different types of information for different territories, due to a number of reasons. For instance, IECA, the office in charge of the diffusion of data on households for Andalusia, uses longer time series than those used by INE, the statistical office providing the data for the series of Spain.

One option, as done in Section 2.3, is to cut the longest series and to unify their length to those of the territories with shortest series. Instead, series with different length can be handled by modifying slightly I-HOR-C, yielding what we call I-HOR-C'.

Suppose that, for territory *n, n* = 1,...,*N*, the time series starts at year t_{0_n}, but all series finish at year *T*. Let φ_n be a weight associated to errors for territory *n*. We consider the following problem:

$$\min \sum_{n=1}^{N}\sum_{t_n=t_{0_n}}^{T}\sum_{i=1}^{I}\phi_n\left(f_{nit_n}-\hat{f}_{nit_n}\right)^2$$

s.t.

$$\hat{f}_{nit_n}=\frac{t_n^{\mu_{ni}}e^{\alpha_{ni}+\beta_{ni}t_n}}{\sum_{j=1}^{I}t_n^{\mu_{nj}}e^{\alpha_{nj}+\beta_{nj}t_n}}\ \forall_n=1,\dots,N\ \forall_i=1,\dots,I\ \forall_{t_n}=t_{0_n},\dots,\mathrm{T} \qquad [5]$$

$$\left|\hat{f}_{niT'}-\hat{f}_{(n+1)iT'}\right|\le\varepsilon \quad \forall n=1,\dots,N,\forall i=1,\dots,I$$

$$\alpha_{ni},\beta_{ni},\mu_{ni}\in\mathbb{R}\ \ \forall_n=1,\dots,N\ \ \forall_i=1,\dots,I$$

The weights φ_n are to be fixed in advance. A possible approach is to chose φ_n by weighing each territory according to its associated time series length, so that the objective function of [5] reflects the global error:

$$\phi_n\alpha\frac{1}{T-t_{0_n}+1}$$

$$\phi_n=\frac{\frac{1}{T-t_{0_n}+1}}{\sum_m\frac{1}{T-t_{0_m}+1}} \qquad [6]$$

For instance, for the case analysed in this paper for *N* = 2, with Spain and Andalusia with territories, t_{0_1} = 2005 and t_{0_2} = 1987, i.e., the series available for Spain is 2005 − 2011 whilst for Andalusia the series covers the period 1987 – 2011. Applying formula [6] one obtains:

$$\phi_1=\frac{\frac{1}{7}}{\frac{1}{7}+\frac{1}{25}}=0.78125$$

$$\phi_2=1-\phi_1=0.21875$$

The optimal parameters for model [5] are presented in Table 7.

Table 7

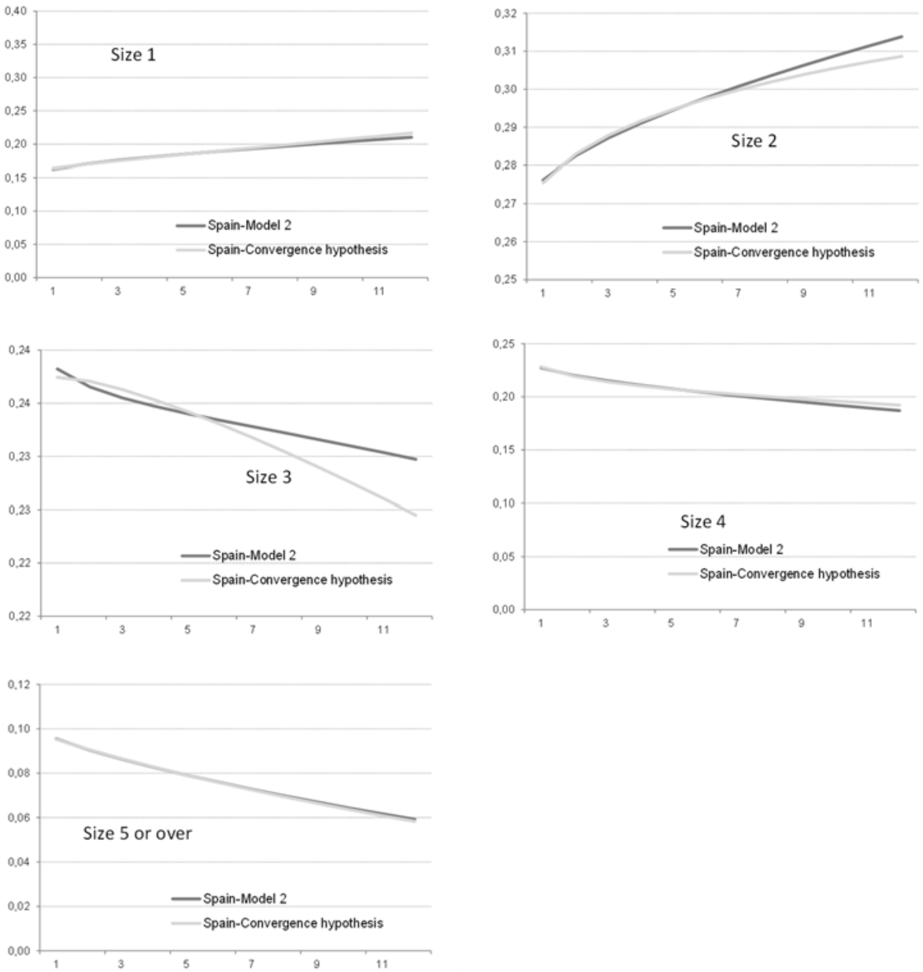
Optimal parameters (I-HOR-C model). Andalusia and Spain

Spain	Parameter	α_1	β_1	α_2	β_2	α_3
	Value	0.000	0.086	0.536	0.081	0.395
	Parameter	β_3	α_4	β_4	α_5	β_5
	Value	0.073	0.359	0.061	-0.481	0.034
	Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
	Value	0.012	-0.016	-0.045	-0.060	-0.053
Andalusia	Parameter	α_1	β_1	α_2	β_2	α_3
	Value	0.000	0.023	0.879	0.016	0.858
	Parameter	β_3	α_4	β_4	α_5	β_5
	Value	0.003	1.112	-0.021	1.523	-0.080
	Parameter	μ_1	μ_2	μ_3	μ_4	μ_5
	Value	-0.068	-0.132	-0.066	0.040	0.048

For the analysis in Spain and Andalusia, the optimal value is 0.0015, which can be expressed as the sum of 0.000034 (the objective value for Spain), and 0.001972 (same for Andalusia), multiplied by their weights, 0.7813 and 0.2188 respectively. Such values coincide with those obtained with model I-HOR, but treating the fitting problem for the two territories as independent problems, Figures 3 and 4. In other words, jointly analyzing the two territories with I-HOR-C' gives results as good as those which would have been obtained by separate analyses. The fit and forecast for Spain and Andalusia in $2011-2016$ is represented in Figure 8.

Figure 8

Data fit and projections (I-HOR-C' model). Andalusia and Spain, 1987 −2050



3. Concluding remarks

In this paper we have shown how to accommodate hypotheses of convergence in parametric models for forecasting household shares. We have enriched the basic multinomial logistic model of (14), which is later integrated in the joint analysis of different territories under hypotheses of convergence. The so-obtained models are handled with the very same software.

We conclude from the results in Section 2 on a state, Spain, and a region, Andalusia, that very good fits are obtained with user-friendly optimization software. For simplicity only two territories have been analyzed, but the methodology can be applied to more complex contexts, with for instance, many more interrelated territories.

References

(1) AKKERMAN, A. (1996), «A problem in household composition», *Mathematical Population Studies*, 6, 3–18.

(2) ALHO, J. AND KEILMAN, N. (2012), «On future household structure», *Journal of the Royal Statistical Society*, 173, 117–143.

(3) BELL, M., COOPER, J. AND LES, M. (1995), *Household and Family Forecasting Modelsls: A Review*. Department of Housing and Regional Development, Camberra.

(4) BERMÚDEZ, S., BLANQUERO, R., HERNÁNDEZ, J. AND PLANELLES, J. (2012) , «A new parametric model for fitting fertility curves», *Population Studies*, 66, 297–310.

(5) BILLARI, F.C., GRAZIANI, R. AND MELILLI, E. (2012), «Stochastic population forecasts based on conditional expert opinions», *Journal of the Royal Statistical Society*, 175, 491–511.

(6) BOWLES, P. AND WAGMAN, B. (1997), «Globalization and the welfare state: Four hypotheses and some empirical evidence», *Eastern Economic Journal*, 23, 3: 317–336.

(7) COLEMAN, D.A. (2002), «Populations of the industrial world - A convergent demographic community?», *International Journal of Population Geography*, 8, 319–344.

(8) DORIUSM S.F. (2008), «Global Demographic Convergence? A Reconsideration of Changing Intercountry Inequality in Fertility», *Population and Development Review*, 34, 3: 519–537.

(9) EDIEV, D.M., YAVUZ, S. AND YÜCEŞAHIN, M.M. (2012), «Private Households in Turkey: Big Changes Ahead», *Population Review*, 51, 28–49.

(10) FRONTLINE SYSTEMS (2010), *Excel Solver, Optimization Software, Monte Carlo Simulation (electronic resource)*. Incline Village, Nevada: Frontline Systems Inc. http://www.solver.com/.

(11) FUKAWA, T. (2011), «Household projection and its application to health/longterm care expenditures in Japan using INAHSIM-II», *Social Science Computer Review*, 29, 1: 52–66.

(12) FYLSTRA, D., LASDON, L., WATSON, J. AND WAREN, A. (1998), «Design and use of the Microsoft Excel Solver», *Interfaces*, 28, 5: 29–55.

(13) HOFFMAN, S. AND G. DUNCAN (1988), «Multinomial and Conditional Logit Discrete-Choice Models in Demography», *Demography*, 25, 3: 415–427.

(14) HAUPT, H., OBERHOFER, W. AND REICHSTHALER, T. (2003), "A varyingcoefficient approach to estimation and extrapolation of household size", *Mathematical Population Studies*, 10, 249–273.

(15) JIANG, L. AND O'NEILL, B.C. (2007), «Impacts of demographic trends on US household size and structure», *Population and Development Review*, 33, 3: 567 – 591.

(16) KEILMAN, N. AND CHRISTIANSEN, S. (2010), «Norwegian Elderly Less Likely to Live Alone in the Future», *European Journal of Population*, 26, 47–72.

(17) KONO, S. (1987), «The headship rate method for projecting households», in *Family Demography: Methods and their Application*, Clarendon Press, Oxford, 287–308.

(18) LEIWEN, J. AND O'NEILL, B. (2004), «Toward a New Model for Probabilistic Household Forecasts», *International Statistical Review*, 72, 51–64.

(19) MASON, A. AND RACELIS, R. (1992), «A comparison of four methods for projecting households», *International Journal of Forecasting*, 8, 509–527.

(20) PERISTERA, P. AND KOSTAKI, A. (2007), «Modelling fertility in modern populations», *Demographic Research*, 16, 141–194.

(21) RINNOOY K., A. AND TIMMER, G. (1987), «Stochastic Global Optimization Methods», *Mathematical Programming*, 39, 57–78.

(22) ROBINS, P. AND DICKINSON, K. (1985), «Child Support and Welfare Dependence: A Multinomial Logit Analysis», *Demography*, 22, 367–380.

(23) UNITED STATES NATIONAL RESOURCES PLANNING COMMITTEE (1938), *The problems of a changing population*. Government Printing Office, Washington, D.C..

(24) UNITED NATIONS (1973), *Methods of Projecting Household and Families*, Population Studies. United Nations, New York.

(25) VAN IMHOFF, E. (2005), *LIPRO 4.0. Tutorial*. The Hague: Netherlands Interdisciplinary Demographic Institute.

(26) WILSON, C. (2001), «On the Scale of Global Demographic Convergence 19502000», *Population and Development Review*, 27, 1: 155–171.

(27) *www.ine.es/jaxi/menu.do?type=pcaxis&path=/t22/e308_mnu&file=inebase&N=&L=0 Encuesta de Población Activa. Instituto Nacional de Estadística.*

(28) *www.juntadeandalucia.es/institutodeestadisticaycartografia/epa/index.htm Encuesta de Población Activa. Instituto de Estadística y Cartografía de Andalucía.*

(29) *www.juntadeandalucia.es/institutodeestadisticaycartografia/merctrab/index.htm Mercado de Trabajo en Andalucía, datos estructurales. Instituto de Estadística y Cartografía de Andalucía.*