Forecasting Spanish Inflation Using Information from Different Sectors and Geographical Areas

Juan de Dios Tena[1], Antoni Espasa [2] and Gabriel Pino[3]

Abstract

This paper evaluates different strategies to forecast Spanish inflation using information of price series for 57 products and 19 regions in Spain. We consider vector equilibrium correction (VeqC) models that include cointegration relationships between sectoral Spanish prices and sectoral prices in the regions of Valencia, Andalucía, Madrid, Cataluña and the Basque Country. This approach is consistent with economic intuition and is shown to be of tangible importance after suitable econometric evaluation. It is found that inflation forecast can always be improved by aggregating projections from different sectors and geographical areas. Moreover, both levels of disaggregation are required in order to obtain a significantly better inflation forecast.

Key Words: inflation forecast, vector equilibrium correction models, relative prices.

JEL Codes: C2, C5.

[1] Juan de Dios Tena, Universidad Carlos III, Departamento de Estadística and Laboratorio de Predicción y Análisis Macroeconómico (Instituto Flores de Lemus), C/Madrid 126. 28903 Getafe (Madrid), España, e-mail: jtena@est-econ.uc3m.es and Universidad de Concepción, Departamento de Economía, Victoria 471 - Oficina 242 - Concepción, Chile, email: juande@udec.cl .

[2] Antoni Espasa, Universidad Carlos III, Departamento de Estadística, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: espasa@est-econ.uc3m.es

[3] Gabriel Pino, Universidad de Concepción, Departamento de Economía, Victoria 471, Chile, e-mail: gapino@udec.cl.

1. Introduction

This paper evaluates the advantages of forecasting Spanish inflation by aggregating projections from different sectors and geographical areas. Although forecasting inflation using information from their disaggregate components has become a recurrent topic in the economic literature, [4] however this literature typically focuses solely on the use of sectoral disaggregate series in order to forecast inflation for a given country.

This lack of concern on regional prices to forecast inflation can be explained as sectoral price heterogeneity is, in principle, more evident than divergences in the geographical prices in a country. Thus, sectoral prices could be differently affected by economic shocks as they can have a different degree of access to the financial markets, some of them are more affected by international market operations or simply because of the different demand price elasticities for the different products in the market. However, it is typically argued that prices in the different regions of a given country should be similar otherwise these differences would be eliminated by arbitrage.

In spite of this hypothesis, in practice it is possible to observe important divergences in the inflation rate of the different Spanish regions. For example, in January 2007 the annual rate of inflation in the Canary Island was 1,5% while it was 2,7% in La Rioja. These divergences can be explained because even in the case that there is a unique equilibrium for regional prices, they could be differently affected by economic policies of regional scope or show different speed of adjustment to this equilibrium after an economic shock.

Here, we forecast Spanish inflation using the information provided by price series from 57 different products and 19 Spanish regions. Concretely, we compare the inflation forecast obtained from a benchmark specification, an ARIMA model for the Spanish Consumer Price Index, with those obtained from three alternative strategies. In the first strategy, we obtain Spanish

[4] See Hendry and Hubrich (2005), Espasa et al. (2002) and Zellner and Tobias (2000) for some examples.

inflation forecast by aggregating sectoral prices from different sectors. We do this by using univariate ARIMA models for each of the components. A second strategy considers the heterogeneity of the different geographical areas and obtains inflation forecast by aggregating projections from 19 ARIMA models applied to aggregate price series in the different Spanish regions. Our final strategy is based on the use of sectoral and regional information. Concretely, we follow a two-steps procedure. In the first step, we obtain price projections in the regions of Valencia, Andalucía, Madrid, Cataluña and Basque Country by specifying vector equilibrium corrections (VeqC) models in which each of the 57 sectoral prices in the aforementioned regions is allowed to work in a cointegration relationship with the national price for the same sector. This approach is justified as it is plausible to assume that regional and national price are cointegrated however their speed of adjustment to the equilibrium could be different in each case. Then, in a second step, we aggregate the projections obtained for the regions of Valencia, Andalucía, Madrid, Cataluña and Basque Country in and those obtained from the remaining Spanish regions using ARIMA model from previous step in order to obtain a forecast of the Spanish inflation.

The structure of this paper is as follows. The next section presents an analysis of each the individual series and test for the presence of regular and seasonal differences. Section 3 discusses model specification used by the different strategies. Inflation forecasts under different strategies are compared in Section 4. Section 5 presents a sectoral analysis for the relative prices in Valencia, Andalucía, Madrid, Cataluña and Basque Country. Some concluding remarks follow in Section 6.

2. Data Description

This section describes the features of the individual price series included in the analyis. We use both aggregate information as well as information related to different sectors and geographical areas. Concretely, the following series are considered in our analysis: 1) the aggregate Spanish Consumer Price Index; 2) price series for 57 different articles in the Spanish economy; 3) aggregate price series for each of the 19

Spanish regions; and 4) price series for 57 different articles in the Spanish regions of Valencia, Andalucía, Madrid, Cataluña and the Basque Country.[5]

The series are available from the Spanish Office for National Statistics at the following URL: http://www.ine.es. Also, all the series are in natural logarithm and cover the period 1993:01-2006:12. However, given that the purpose of the paper is to evaluate the forecast performance of different disaggregation strategies, in the econometric analysis we only use information for the period 1993:01-2002:10 in order to evaluate how useful the different strategies are in order to forecast the annual rate of inflation in the period 2002:11-2006:12.

An important point to mention at this stage is that some of the series show a structural break in their evolution in the year 2002 because of the methodological change in the way that price series for each of the 57 disaggregate articles were constructed. This break can be clearly observed, for example, in men and women clothes for Spain and the different regions. The presence of this break is used in our particular context to check whether or not there are advantages in the use of disaggregate models to forecast an aggregate variable even when the specification of disaggregate models is a hard task because of the observed structural breaks in the disaggregate time series.[6]

Notice also that we study 57 products as this is the highest disaggregate level that can be considered at the regional level. Some papers in the literature use an intermediate level of analysis in five groups of products for the Spanish economy: 1) processed food, 2) non-energy industrial goods, 3) services, 4) unprocessed food, and 5) energy; see Espasa et al. (1987). Here, we could not use this set-up because, at the moment, these five series are only available from 2002:01 at the regional level.

Figures of the series are not shown to save space, however the inspection reveals that practically all of them growth

[5] A description of the sectors and regions is confined to the Appendix.
[6] See Espasa and Albacete (2004) and Hubrich (2003) *interalia* for a discussion on this issue.

smoothly during the period under consideration.[7] Two exceptions to this general pattern are the price of mail and communications that grows until the end of the 90s and then decrease and the price of recreational objects that remains stable until 2001 and then shows a negative trend. However, even these two series cannot be considered as stationary as they show little tendency to return to the mean.

Series in first differences, on the other hand, show regular crossing points and no obvious trend. Additionally, some series such as for example meats, transport and tourism exhibit a clear seasonal behaviour.

For a formal test on the number of unit roots in the series we employed the methodology proposed by Osborn, Chui, Smith y Birchenhall (1988) (OCSB henceforth) who extended the procedure of Hazsa and Fuller (1982) to seasonal time series for monthly data. Although we are aware of other more sophisticated procedures to investigate the presence of seasonal unit roots such as the tests proposed by Franses (1991) and Beaulieu and Miron (1993), we choose the DHF test because of its simplicity that allows us to determine whether or not to take seasonal differences instead of testing for unit roots one by one at each of the harmonic frequencies of the seasonal cycle.

Following OCSB, our test regression for a given y_t variable takes the form:

$$\Delta\Delta_{12} y_t = c + \sum_{k=1}^{11} \delta_k D_{kt} + \beta_1 \Delta_{12} y_{t-1} + \beta_2 \Delta y_{t-12} + \sum_{i=1}^{p} \phi_i \Delta\Delta_{12} y_{t-i} + \varepsilon_t \qquad (2.1)$$

where c is a constant term; D_{kt} is a centered seasonal dummy variable for the *kth* month; Δ and Δ_{12} denote respectively the regular and seasonal difference operator; p is the number of augmentation lags that in our case is chosen using the sequential approach by Ng and Perron (1995); and $\varepsilon_t \sim \mathbf{iid}(\mathbf{0}, \sigma^2)$ is the disturbance term.

[7] All the information not explicitly reported in this paper can be obtained from the authors upon request.

If y_t is non stationary but $\Delta\Delta_{12}y_t$ is a stationary invertible process then, following OCSB y_t is denoted as being I(1,1). The I(1,1) null hypothesis, $\beta_1 = \beta_2 = 0$, can be tested by using a F-type statistic.

One alternative to the I(1,1) null hypothesis is that stationarity is obtained after first differences. This alternative hypothesis, denoted as I(1,0), can be represented in equation (2.1) by $\beta_1 = 0$ and $\beta_2 < 0$. A second alternative is that the process requires annual differencing to be stationary. This alternative hypothesis I(0,1) is captured in equation (2.1) by $\beta_1 < 0$ and $\beta_2 = 0$. Following DHF, separate t-type statistics for $\beta_1 = 0$ and $\beta_2 = 0$ can be used to distinguish between the two possible alternatives.

We do not report the test results for each of the individual time series in order to save space but show instead only the results for each of the 57 Spanish products. They indicate that most of the series are stationary after one regular difference. Moreover, some of the I(1,0) have seasonal components that can be captured with dummy variables. This is the case for example of the different type of meats, potatoes, home and medical services, transport and secondary school.

[INSERT TABLE 1]
[INSERT TABLE 2]
[INSERT TABLE 3]

The null hypothesis can not be rejected at the conventional levels in a few exceptional cases. This is the situation for example of price of clothes and footwear, furniture and floor covering and textile and home accessories.

When the test is applied to the aggregate series we find that the null cannot be rejected at the 5% level what is a surprising result given that the null hypothesis is rejected in the vast majority of cases at the disaggregate level. This also happens with the aggregate series of the different regions with the only exception of Madrid. The result suggest the convenience of using a disaggregate approach for analysing the properties of the time series as these features could be masked in the aggregate counterpart.

3. Model Specification.

We now turn to the central point of the paper that is the comparison of different approaches to forecast the Spanish inflation for the period 2002:11-2006:12. More specifically, we compare the projections obtained under our benchmark specification: a simple ARIMA model for the aggregate series, with those obtained under three alternative strategies.

The first alternative strategy relates to the sectoral disaggregation level and considers 57 ARIMA models for each of the dissagregate Spanish products. Our second strategy is based on the projections obtained from ARIMA models for each of the 19 Spanish regions. Finally, the third strategy considers both sectoral and geographic disaggregation. More specifically, we estimate a vector equilibrium correction (VeqC) model for each of the 57 different products in the Spanish regions of Valencia, Andalucía, Madrid, Cataluña and the Basque Country that incorporate a potential long-run equilibrium between each of the products in the different regions with the same product in Spain. Using this approach, we forecast inflation in the 5 aforementioned regions by aggregating the projections obtained from the vector equilibrium correction (VeqC) models for each of the products. Projections in the remaining Spanish regions are obtained in a similar way as we do in the second alternative strategy. Then, we aggregate these prices and those obtained from the remaining Spanish regions using ARIMA model in order to obtain a forecast of the Spanish inflation.

For the first alternative our ARIMA models are specified using the TRAMO/SEATS automatic procedure as it is completely aseptic and not affected by the subjective annalist criterion. Tables 4, 5 and 6 describe the univariate specifications for each of the 57 price series in Spain under the benchmark and the first alternative strategy. Notice that, in most of the cases, the growth pattern of most of the series is either captured by including a regular and a seasonal difference in the model or by specifying a model with a regular difference and a constant term.

This is in contradiction with the unit root tests results shown in the previous section that suggest the presence of only one unit root in most cases. In fact, there is an open debate in the literature on whether inflation is stationary or generated by a unit root process; see for example Culver and Papell (1997) for a discussion on this issue. Assuming that there is not an equilibrium rate for inflation is a quite strong hypothesis. However, it is also true that the inflation level is typically affected by different breaks on the levels of the series and therefore, for empirical purposes, it is reasonable to capture these changes with a unit root. Hence, we use here the specifications proposed by TRAMO that in most of the cases do not consider inflation as a stationary process. However, there are not material change in the results when we base the analysis on ARIMA models specified considering the results about the cointegration order of the series in the previous section and based on the interpretation of figures and correlograms using the methodology proposed by Box and Jekins; see Box et al. (1994).

[INSERT TABLE 4]
[INSERT TABLE 5]
[INSERT TABLE 6]

Regarding the second alternative, we follow a similar approach to the previous case using TRAMO/SEATS in the model specification of the ARIMA model for the 19 regional series. However, weights to aggregate the regional series in order to obtain the Spanish Consumer Price Index do not exist. We solve this problem using as weights the share of expenditure of each region in the Spanish expenditure. Proceeding in this way, we obtain an aggregate series that is very similar to the Spanish Consumer Price Index.

As to the third alternative, we depart from the following specification:

$$\Delta y_{ijt} = \gamma_{ij} + \alpha_{ij}\left[\beta_{ij}^{'}, \delta_{ij}, \theta_{ij}\right]\begin{bmatrix} y_{ijt-1} \\ 1 \\ t \end{bmatrix} + \Phi_{ij}\Delta y_{ijt-1} + \Gamma_{ij} D_t + \varepsilon_{ijt} \qquad (3.1)$$

where y_{ijt} is a (2x1) vector containing price levels in the ith product of the jth region and the price for the same product in Spain; α_{ij} and β_{ij} are respectively the (2x1) adjustment and cointegration vectors; Φ_{ij} is a (2x2) matrix of parameters; γ_{ij} is a (2x1) vector of intercept parameters; δ_{ij} and θ_{ij} are scalars; D_t includes intervention dummies and Γ_{ij} is the matrix of parameters associated to the interventions; and ε_{ijt} is a (2x1) vector of serially uncorrelated errors.

Notice that expression (3.1) considers that there is a long-run equilibrium for sectoral prices between Spain and each of the different Spanish regions. This is consistent with economic theory as, at least for transable products, arbitrage should eliminate regional price differentials in the long-run. However, dynamic adjustments to the equilibrium level are allowed to be different across regions and products.

Specification (3.1) allow for two deterministic intercepts, one of them belong to the α_{ij} space (δ_{ij}) and the other to the orthogonal space (γ_{ij}). However, the trend component is forced to belong only to the α_{ij} space (θ_{ij}) as imposing a quadratic deterministic trend in y_{ijt} is an implausible assumption; see for example Johansen (1995).

In practice, there are four possible cases of interest: 1) $\delta_{ij}= \gamma_{ij} = \theta_{ij} = 0$; 2) $\delta_{ij} \neq 0$, $\gamma_{ij} = \theta_{ij} = 0$; 3) $\delta_{ij} \neq 0$, $\gamma_{ij} \neq 0$ and $\theta_{ij} = 0$; and 4) $\delta_{ij} \neq 0$, $\gamma_{ij} \neq 0$ and $\theta_{ij} \neq 0$. Notice that only the last two cases allows for deterministic linear trend in y_{ijt} .

We specify VeqCM models selecting for each of the 57 models the case that minimize the Akaike information criterium. In most of the cases we select cases 3 and 4 that allows for linear trend in the variables what is consistent with the results found in the previous section. Also, the trace test indicates that the null of no cointegration can be rejected at the conventional levels for many of the sectors. Moreover, in most of the cases where the null was not rejected, the trace test statistic is close to the critical values for rejection. This is consistent with the visual inspection of the series.

Results in the previous section suggest that some of the series could need one regular and one seasonal difference to become stationary. However, when we test cointegration for the series in differences instead of levels , we find that the null of cero and one cointegration relationship can be rejected in all the cases at the conventional levels. This is an indication that series in annual differences could be considered as stationary and therefore we perform for all the cases cointegration tests for series in levels.

4. Forecast Evaluation

Using the four strategies mentioned in the previous section, we forecast the Spanish annual rate of inflation. Here, we consider both static and dynamic forecast. The first one assumes that for a forecast horizon k , the practitioner knows all the information at $t+h-1$ ($h \leq k$) in order to forecast the value of the relevant variable at $t+h$. Dynamic forecast, on the other hand, assumes that values of the variables are not updated with the new information during the forecast horizon.

Static forecast is especially consistent with the forecasting activity as the analyst has to adjust their predictions after the new information arrives. Dynamic forecast on the other hand is a good indicator of how successful the different strategies are to predict the evolution of the series in the medium and long run.

We also use both the sum of the square forecast errors and the sum of the absolute values of the forecast errors to evaluate the accuracy of the different approaches.

Table 7 shows the sum of the square and absolute forecast errors for the aggregate inflation computed using the four approaches discussed in the previous section. Results indicate the advantages of using a disaggregate approach regardless of the loss function considered in both the static and dynamic evaluation. Indeed, forecasting inflation by aggregation of sectors and geographic areas outperform the benchmark case but the best strategy is always to consider both disaggregation levels (by sectors and regions) at the same time.

[INSERT TABLE 7]

However, results in Table 7 do not provide any indication on how important are forecast differences. Thus, a more formal analysis requires showing if the projections obtained under the different strategies are statistically different. We do this by using the standard test proposed by Diebold and Mariano (1995) (DM henceforth). This is an asymptotic test to compare the forecast obtained under different strategies. For a brief description of this procedure, we define $e_{i,t}$ and $e_{j,t}$ as the forecast errors using procedures *i* and *j* respectively. We also define $p(\bullet)$ as the loss function.

The DM statistic is an asymptotic test for the null hypothesis $E(d_t)=0$, where $d_t = p(e_{jt}) - p(e_{it})$. Based on these definitions, DM consider the $\bar{d}$ statistic which denotes the sample mean of d_t , $t = 1,...,T$, $\bar{d} = T^{-1}\sum_{t=1}^{T} d_t$.

DM show that, under the null hypothesis, the standardized statistic

$$ST = \frac{\bar{d}}{\sqrt{T^{-1}2\pi\hat{f}_d(0)}} \qquad (4.1)$$

converges to a normal distribution when $T \rightarrow \infty$. $\hat{f}_d(0)$ is a consistent estimator of the spectral density of d_t at the zero frequency such that

$$\hat{f}_d(0) = (2\pi)^{-1} \sum_{k=-T+1}^{T-1} w(k/S(T))\hat{\gamma}_d(k) \qquad (4.2)$$

where $w(\bullet)$ is a lagged window function, $S(\bullet)$ is the truncated lag and $\hat{\gamma}_d$ is the estimated autocorrelation function.

We show in Tables 8 and 9 the results of the DM test using the sum of the square forecast errors as the loss function in the static and dynamic forecast exercise.[8] In the static forecast case, notice that improvements in the accuracy of inflation forecast can only be obtained from a geographical disaggregation. In the dynamic case, results are even more compelling showing that the inflation forecast obtained from geographic disaggregation significantly outperform the other strategies at the conventional levels.

[INSERT TABLE 8]
[INSERT TABLE 9]

5. Sectoral Analysis

A disaggregate analysis by regions and sectors is not only useful to forecast the aggregate rate of inflation but also to study all the disaggregate information available in order to detect the sources of differences in the regional rate of inflation. Here, we study the evolution of price differentials restricting our analysis to the behaviour of each of the 57 sectoral prices in the regions of Valencia, Andalucía, Madrid, Basque Country and Cataluña.

Divergences of the regional inflation rates could be due to: 1) different patterns of preferences across regions that are reflected in the weights given to each of the 57 different sectoral prices; 2) the heterogeneity inherent in the evolution of each of the individual time series. While different patterns of preferences in different regions can explain structural divergences that are likely to persist in time, diferences in the evolution of disaggregate prices can explain transitory divergences.

We compute the differences between the 2007 weights of the different regions and Spain, it can be observed that the regions

[8] Results based on the sum of the absolute values of the forecast errors do not change the conclusions of this analysis.

of Madrid and Valencia give relatively more importance to the consumption of service products. In the case of Madrid, this is because the expenditure in tourism and hotels is higher than in Spain while in Valencia the reason is the important role played by expenditure in personal transportation. The region of Andalucía gives less importance to service in the basket of consumption. However, Andalucía gives relatively more weight to food consumption. Cataluña and the Basque Country show patterns of consumption that are very similar to the national average.

Regarding the evolution of the individual series, we pay attention to the tests of weak exogeneity obtained form the sectoral vector equilibrium correction (VeqC) models in Section 2. Results of this test indicate whether regional prices are exogenously determined or react to shocks to the long-term equilibrium in the previous period. It is found that the hypothesis of weak exogeneity for food products is rejected in 80% of the cases in the Basque country at the conventional levels. However, the percentage of rejection of the null hypothesis for the remaining regions is around the 50%. Also, the hull hypothesis for services is rejected in Madrid for practically all the cases at the conventional levels but it is only rejected in less that 60% of the cases in Andalucía and the Basque Country.

Additionally we compute product by product the difference between the annual rate of inflation in each of the regions and Spain. We also compute the standard deviation of these differences and the first order autocorrelation of these differences. Figures 1 and 2 show the most important cases of historical price divergences between Spain and the different regions. We find that in general prices in Valencia grow less than in the rest of Spain in most of the products. This is especially true for bread, rented apartments, own apartments, furniture and floor covering, major appliances and nondurable household items. Andalucía shows historically lower inflation rates in other meats, footwear for babies and children, medical services, personal transportation and recreation and its inflation is historically higher in secondary school. For Cataluña, it must be mentioned the high inflation rates compared to Spain in many items such as footwear, furniture and floor covering, textiles and home accessories, recreation and primary school. Prices in Madrid historically grow less than in the rest of Spain. This is especially true for pork, heater, lighting and water distribution and recreation. Prices in the Basque Country grows more than in the rest of Spain in coffee, cacao and infusions, recreation and primary school while inflation in the Basque country is significantly lower than in Spain in Secondary School.

[INSERT FIGURE 1]

[INSERT FIGURE 2]

6. Concluding remarks

This paper evaluates different strategies to forecast Spanish inflation using information of price series for 57 products and 19 regions in Spain. In this procedure we use simple statistical devices based on the use of univariate ARIMA models and vector equilibrium correction (VeqC) models that allows for potential cointegration relationships between sectoral prices in Spain and in the regions of Valencia, Cataluña, Madrid, Cataluña and the Basque country. It is found that forecasting Spanish inflation by aggregating projections for different sectors and geographical areas always result in a lower sum of the square and absolute forecast errors. Moreover, in order to obtain substantial forecast improvements it is necessary to consider the highest disaggregation level.

Bibliography

Anderson, R. G. , D. Hoffman and R. H. Rasche, 1998, "A vector error-correction forecasting model for the US economy", Working Paper 1998-008C, Federal Reserve Bank of St. Louis.

Beaulieu, J.J. and Miron, J.A. (1993), "Seasonal unit roots in aggregate U.S data", Journal of Econometrics, 55, 305-328.

Box, G.E.P., Jenkins, G.M. and Reinsel, G.C., 1994, <u>Time Series Analysis: Forecasting and Control</u>, (Third Edition), Englewood Cliffs, HJ: Prentice Hall.

Culver, S.E. and Papell, D.H,, 1997, "Is there a unit root in the inflation rate? Evidence from sequential break and panel data models, Journal of Applied Econometrics, 12, 436-44.

Dickey D.A. and Fuller, W.A., (1979), "Distribution of the estimators for autoregressive time series with a unit root", Journal of the American Statistical Association, 74, 427-431.

Espasa, A. and Albacete, R., (2004), "Econometric modelling of short-term inflation forecasting in the EMU", Working paper 03-43, Statistics and Econometric Series 09, Universidad Carlos III.

Espasa, A., Matea, M. L., Manzano, M. C. & Catasus, V. (1987), "La inflación subyacente en la economía española: estimación y metodología", Boletín Económico del Banco de España, marzo, 32-51.

Espasa, A., E. Senra and R. Albacete, (2002), "Forecasting inflation in the European Monetary Union: a disaggregate approach by countries and by sectors", European Journal of Finance, 8, 402-421.

Franses, P.H., (1991), "Seasonality, nonstationarity and the forecasting of monthly time series", International Journal of Forecasting, 199-208.

Hasza, D.P. and W.A. Fuller, (1982), "Testing for Nonstationary Parameter Specifications in Seasonal Time Series Models", The Annals of Statistics, 10, 1209-1216.

Hendry, D. (2001), "Modelling UK Inflation, 1875-1991", Journal of Applied Econometrics, 16, 255-275.

Hendry, D.F. and K. Hubrich, (2005), "Forecasting aggregates by disaggregates", manuscript.

Hubrich, K., (2003), "Forecasting euro area inflation: does aggregating forecasts by HICP components improve forecasts accuracy?", ECB Working Paper, No.247.

Johansen, S., (1995), <u>Likelihood-based inference in cointegrated vector autoregressive models</u>, Oxford University Press.

Ng, S. and Perron, P., (1995), "Unit root tests in ARMA models with data dependent methods for selection of the truncation lag", Journal of the American Statistical Association.

Osborn, D.R., Chui, A.P.L., Smith, J.P. and Birchenhall, C.R., (1988), "Seasonality and the order of integration for consumption", Oxford Bulletin of Economics and Statistics, 50, 361-377.

Stock, J. H. and Watson, M. W. (1999), "Forecasting Inflation". Journal of Monetary Economics, 44, 293-335.

Zellner, A. and Tobias, J., (2000), "A note on aggregation, disaggregation and forecasting performance", Journal of Forecasting, 19, 457-469.

Appendix

Time Series

The time series considered in the analyses can be freely obtained from the INE (Instituto Nacional de Estadística). We use time series for the following disaggregate products (rúbricas):

Food Products

R1: Cereals.
R2: Bread.
R3: Beef.
R4: Ovine.
R5: Pork.
R6: Bird.
R7: Other meats.
R8: Fish.
R9: crustaceans, mollusks and prepared of fish.
R10: Eggs.
R11: Milk.
R12: Milk products.
R13: Oil and fats.
R14: Fresh fruits.
R15: Preserved fruit.
R16: Vegetables.
R17: Preserved vegetables.
R18: Potatoes.
R19: Coffee, cacao and infusions.
R20: Sugar.
R21: Other food products.
R22: Non alcoholic drinks.
R23: Alcoholic drinks.
R24: Tobacco.

Other products

R25: Men clothes.
R26: Women clothes.
R27: Clothes for babies and children.
R28: Complements and Repairs.
R29: Men footwear.
R30: Women footwear.
R31: Footwear for babies and children.
R32: Repairs of footwear.
R33: Rented apartments.
R34: Heater, lighting and water distribution.
R35: Own apartments.
R36: Furnitures and floor covering.
R37: Textile and home accessories.
R38: Major appliances.
R39: Household items.
R40: Non durable household items.

Services

R41: Home services.
R42: Medical services.
R43: Medicines and other chemical products.
R44: Personal transportation.
R45: Public urban transportation.
R46: Public intercity transportation.
R47: Mail and communications.

R48: Recreational objects.
R49: Publications.
R50: Recreation.
R51: Primary school.
R52: Secondary school.
R53: University.
R54: Other expenditures in education.
R55: Personal items.
R56: Tourism and hotels.
R57: Other goods and services.

Also, the Spanish regions consider and their respective weights considered are:

Andalucia	*153.10*
Aragón	*32.12*
Asturias	*25.26*
Baleares	*24.15*
Canarias	*39.89*
Cantabria	*12.46*
Castilla y león	*55.08*
Castilla la mancha	*36.08*
Cataluña	*173.79*
Comunidad Valenciana	*105.41*
Extremadura	*16.34*
Galicia	*55.97*
Madrid	*162.66*
Murcia	*26.03*
Navarra	*13.84*
Basque Country	*58.19*
Rioja	*6.92*
Ceuta y Melilla	*2.71*

All the series are in monthly basis and cover the period 1993:01-2006:10.

Table 1. OCSB test for the aggregate Consumer Price Index and prices of the food sector in Spain.

Product	β_1	β_2	$F_{1,2}$	F_s (1)	**Lags**
Aggregate	0.59	-4.01	9.78	0.81	10
Food Prices					
Product	β_1	β_2	$F_{1,2}$	F_s (1)	**Lags**

R1	3.24(**)	-8.39(**)	38.53(**)	3.06(**)	0
R2	2.53(*)	-8.46(**)	36.82(**)	2.36(*)	4
R3	1.73	-9.98(**)	54.34(**)	5.30(**)	1
R4	-1.28	-10.73(**)	74.54(**)	8.32(**)	9
R5	0.89	-10.06(**)	55.69(**)	4.07(**)	1
R6	-1.90	-8.87(**)	47.08(**)	2.49(*)	12
R7	2.10	-8.27(**)	37.54(**)	2.36(*)	1
R8	-1.10	-7.08(**)	29.04(**)	2.86(**)	10
R9	0.16	-10.50(**)	59.39(**)	5.80(**)	0
R10	0.46	-7.18(**)	33.71(**)	2.27(*)	9
R11	2.13(*)	-8.95(**)	44.70(**)	2.47(*)	1
R12	4.18(**)	-8.54(**)	37.93(**)	2.78(**)	12
R13	2.08	-5.93(*)	20.56(*)	1.01	12
R14	3.61(**)	-8.19(**)	36.27(**)	0.44	12
R15	5.36(**)	-7.35(**)	28.11(**)	1.93	12
R16	3.58(**)	-12.54(**)	92.99(**)	1.13	8
R17	2.56(*)	-8.52(**)	40.30(**)	1.31	1
R18	-1.50	-7.14(**)	31.93(**)	2.98(**)	12
R19	-1.85	-11.53(**)	82.67(**)	0.82	12
R20	3.73(**)	-7.46(**)	30.41(**)	1.11	12
R21	4.58(**)	-10.17(**)	54.42(**)	2.55(*)	1
R22	1.57	-7.76(**)	32.17(**)	1.80	1
R23	3.00(**)	-7.89(**)	35.33(**)	2.82(**)	2
R24	-0.82	-7.76(**)	34.40(**)	0.52	0

Notes: (1) is the F statistic to test for the significance of seasonal dummies. * indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 5% significance nivel; ** indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 1% significance nivel. The critical values of Rodrigues and Osborn (1997) are used.

Table 2. OCSB test for prices of other products in Spain.

Product	β_1	β_2	$F_{1,2}$	F_s (1)	**Lags**
R25	-1.04	0.22	0.58	0.85	10
R26	-0.33	-1.17	0.73	1.09	6
R27	1.25	-3.01	4.64	1.62	9
R28	0.23	-2.37	3.88	0.68	7

R29	-0.51	-2.22	4.15	1.23	7
R30	-1.60	0.34	1.40	0.78	9
R31	-0.67	-1.12	1.60	0.78	9
R32	3.74(**)	-9.11(**)	41.64(**)	4.61(**)	0
R33	8.96(**)	-19.18(**)	197.76(**)	10.42(**)	0
R34	-0.03	-7.50(**)	28.72(**)	0.05	1
R35	3.54(**)	-10.10(**)	53.55(**)	3.78(**)	1
R36	0.81	-4.14	12.76	1.63	7
R37	-0.11	-1.76	2.31	0.98	6
R38	2.95(**)	-7.06(**)	25.58(**)	2.58(*)	1
R39	1.53	-6.43(**)	23.26(**)	1.71	2
R40	1.47	-6.69(**)	27.66(**)	1.56	1

Notes: (1) is the F statistic to test for the significance of seasonal dummies. * indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 5% significance nivel; ** indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 1% significance nivel. The critical values of Rodrigues and Osborn (1997) are used.

Table 3. OCSB test for prices of services in Spain.

Product	β_1	β_2	$F_{1,2}$	$F_s^{(1)}$	**Lags**
R41	-0.35	-8.24(**)	40.61(**)	5.08(**)	3
R42	0.03	-5.03	15.23	3.35(**)	11
R43	-0.87	-6.30(*)	21.29(*)	0.66	12
R44	2.32(*)	-11.33(**)	69.84(**)	2.26(*)	1
R45	1.18	-5.67	16.94	3.10(**)	12
R46	-0.96	-8.25(**)	39.26(**)	2.40(*)	0
R47	1.04	-12.01(**)	78.78(**)	2.74(**)	0
R48	2.42(*)	-5.07	13.18	0.79	2
R49	1.47	-8.49(**)	39.04(**)	1.01	3
R50	1.45	-10.04(**)	55.96(**)	2.58(*)	0
R51	1.63	-7.85(**)	32.43(**)	2.39(*)	0
R52	1.27	-9.46(**)	49.05(**)	5.68(**)	0
R53	0.06	-21.42(**)	256.20(**)	17.63(**)	0
R54	3.53(**)	-10.96(**)	60.76(**)	6.69(**)	0
R55	1.59	-7.35(**)	30.70(**)	3.33(**)	1
R56	-1.59	-4.68	13.81	1.95	0
R57	4.23(**)	-8.27(**)	35.24(**)	5.31(**)	0

Notes: (1) is the F statistic to test for the significance of seasonal dummies. * indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 5% significance nivel; ** indicates the null hypothesis of zero coefficient(s) in equation (1) is rejected at the 1% significance nivel. The critical values of Rodrigues and Osborn (1997) are used.

Table 4. Univariate ARIMA for the aggregate Consumer Price and prices of the food products in Spain.

Product	Δ	Δ_{12}	C	**Model**	S_{1t}	S_{2t}
Aggregate	1	1	No	AR(1), MA(12)	No	No
Food Prices						
Product	$\Delta^{(1)}$	$\Delta_{12}^{(2)}$	$C^{(3)}$	**Model**	$S_{1t}^{(4)}$	$S_{2t}^{(5)}$
R1	1	0	Yes	AR(12), MA(12)	Yes	No
R2	1	0	Yes	AR(1), AR(12)	Yes	No

R3	1	1	No	AR(1), MA(12)	Yes	Yes
R4	1	0	Yes	AR(1), AR(2), AR(12), MA(1)	Yes	No
R5	1	1	No	AR(1), AR(12), MA(1)	Yes	No
R6	1	1	No	MA(12)	Yes	No
R7	1	0	Yes	AR(1), AR(12)	Yes	No
R8	1	1	No	AR(1), AR(2), MA(12)	Yes	No
R9	1	1	No	MA(1), MA(12)	Yes	No
R10	1	1	No	AR(1), AR(12), MA(12)	Yes	No
R11	1	1	No	AR(1), MA(12)	Yes	Yes
R12	1	1	No	AR(1), MA(12)	Yes	Yes
R13	1	1	No	AR(1), MA(1), MA(12)	No	No
R14	1	0	Yes	AR(1)	No	No
R15	1	1	No	AR(1), AR(2), MA(12)	No	No
R16	1	0	Yes	AR(1), AR(12), MA(1), MA(12)	No	No
R17	1	0	No	AR(1)	No	No
R18	1	1	No	MA(1), MA(12)	Yes	No
R19	0	0	Yes	AR(1), AR(2)	No	No
R20	1	0	No	AR(1)	No	No
R21	2	1	No	MA(1), MA(12)	Yes	No
R22	1	1	No	MA(1), MA(12)	No	No
R23	1	0	Yes	AR(1), AR(12)	Yes	No
R24	1	1	No	MA(12)	No	No

Notes: (1) is the number of regular differences; (2) is the number of seasonal differences; (3) indicate the presence of a constant term in the estimation; (4) seasonal dummies; (5) dummy for the methodological change.

Table 5. Univariate ARIMA for other products in Spain.

Product	Δ	Δ_{12}	C	**Model**	S_{1t}	S_{2t}
R25	0	1	Yes	MA(1), MA(2)	No	Yes
R26	1	0	Yes	MA(1), MA(2)	No	Yes
R27	1	0	Yes	MA(1), MA(2)	No	Yes
R28	1	0	Yes	AR(1), AR(2), MA(1)	No	No
R29	1	0	Yes	AR(2), AR(3)	No	No

R30	1	1	No	AR(6), MA(2), MA(3), MA(6)	No	Yes
R31	2	0	No	AR(1), AR(2), AR(3)	No	No
R32	2	1	No	MA(1), MA(12)	Yes	No
R33	1	1	No	AR(1), AR(2), MA(12)	Yes	No
R34	1	0	No	MA(1), MA(12)	No	No
R35	2	0	No	MA(1), MA(2), MA(3)	Yes	No
R36	1	1	No	AR(1), AR(2), MA(1), MA(2), MA(12)	No	No
R37	1	1	No	AR(2), AR(3)	No	Yes
R38	1	1	No	AR(1), MA(12)	Yes	Yes
R39	1	0	Yes	AR(12), MA(1), MA(2)	No	No
R40	1	0	No	AR(1), AR(12)	No	No

Notes: (1) is the number of regular differences; (2) is the number of seasonal differences; (3) indicate the presence of a constant term in the estimation; (4) seasonal dummies; (5) dummy for the methodological change.

Table 6. Univariate ARIMA for services in Spain.

Product	Δ	Δ_{12}	C	**Model**	S_{1t}	S_{2t}
R41	1	1	No	MA(12)	Yes	No
R42	1	1	No	MA(1)	Yes	No
R43	1	0	Yes	MA(1)	No	No
R44	1	1	No	MA(1), MA(12)	No	No
R45	1	1	No	AR(11), MA(11)	Yes	No
R46	1	1	No	MA(1), MA(12)	Yes	No
R47	1	1	No	MA(1)	No	No
R48	2	0	No	MA(1)	No	Yes
R49	1	0	Yes	MA(1), MA(2)	No	No
R50	1	0	Yes	AR(1)	Yes	No
R51	1	1	No	MA(12)	Yes	No
R52	1	1	No	MA(12)	No	No
R53	1	1	No	MA(1)	Yes	No

R54	1	1	No	MA(1), MA(12)	Yes	Yes
R55	2	1	No	MA(1), MA(12)	Yes	No
R56	1	1	No	MA(1), MA(12)	No	No
R57	1	1	No	MA(1), MA(12)	Yes	No

Notes: (1) is the number of regular differences; (2) is the number of seasonal differences; (3) indicate the presence of a constant term in the estimation; (4) seasonal dummies; (5) dummy for the methodological change.

Table 7. Sum of square and absolute forecast errors of the annual rate of inflation in Spain using different strategies.

Sum of square forecast errors				
	S1(1)	**S2(2)**	**S3(3)**	**S4(4)**
Static forecast	6.0	5.8	5.2	4.2
Dynamic forecast	27.6	45.0	21.3	16.0
Sum of absolute forecast errors				
	S1(1)	**S2(2)**	**S3(3)**	**S4(4)**
Static forecast	13.5	13.6	12.5	11.5
Dynamic forecast	29.9	41.1	26.2	22.1

Notes: (1) S1 refers to the strategy based on an ARIMA model for the aggregate consumer price index series; (2) S2 is the forecast based on the specification of univariate ARIMA models for each of the 57 disaggregate series; (3) S3 is the forecast strategy based in specification of the ARIMA model for the 19 regional series; (4) S4 is the forecast strategy based in specification of the ARIMA model for the 19 regional series based on the use of VeqC models the exploits cointegration relationships between disaggregate series in Spain and Andalucia, Cataluña, Madrid, the Basque Country and Valencia.

Table 8. DM test for the different strategies of the sum of the square forecast errors for the static forecast.

	Static forecast
S1(1) versus **S2(2)**	0.23
S1(1) versus **S3(3)**	3.84(**)
S1(1) versus **S4(4)**	2.53(*)
S2(2) versus **S3(3)**	0.57
S2(2) versus **S4(4)**	1.81
S3(3) versus **S4(4)**	1.47

Notes: (1) S1 refers to the strategy based on an ARIMA model for the aggregate consumer price index series; (2) S2 is the forecast based on the specification of univariate ARIMA models for each of the 57 disaggregate series; (3) S3 is the forecast strategy based in specification of the ARIMA model for the 19 regional series; (4) S4 is the forecast strategy based in specification of the ARIMA model for the 19 regional series based on the use of VeqC models the exploits cointegration relationships between disaggregate series in Spain and Andalucia, Cataluña, Madrid, Basque Country and Valencia. **,* indicate respectively rejection at the 1% and 5% significant levels.

Table 9. DM test for the different strategies of the sum of the square forecast errors for the dynamic forecast.

	Dynamic forecast
S1$^{(1)}$ versus **S2**$^{(2)}$	-1.11
S1$^{(1)}$ versus **S3**$^{(3)}$	1.99(*)
S1$^{(1)}$ versus **S4**$^{(4)}$	1.74
S2$^{(2)}$ versus **S3**$^{(3)}$	1.86
S2$^{(2)}$ versus **S4**$^{(4)}$	2.8(**)
S3$^{(3)}$ versus **S4**$^{(4)}$	1.53

Notes: (1) S1 refers to the strategy based on an ARIMA model for the aggregate consumer price index series; (2) S2 is the forecast based on the specification of univariate ARIMA models for each of the 57 disaggregate series; (3) S3 is the forecast strategy based in specification of the ARIMA model for the 19 regional series; (4) S4 is the forecast strategy based in specification of the ARIMA model for the 19 regional series based on the use of VeqC models the exploits cointegration relationships between disaggregate series in Spain and Andalucia, Cataluña, Madrid, Basque Country and Valencia. **,* indicate respectively rejection at the 1% and 5% significant levels.

Figure 1. Historical price divergences between Spain and the regions of Cataluña, Madrid and the Basque Country.

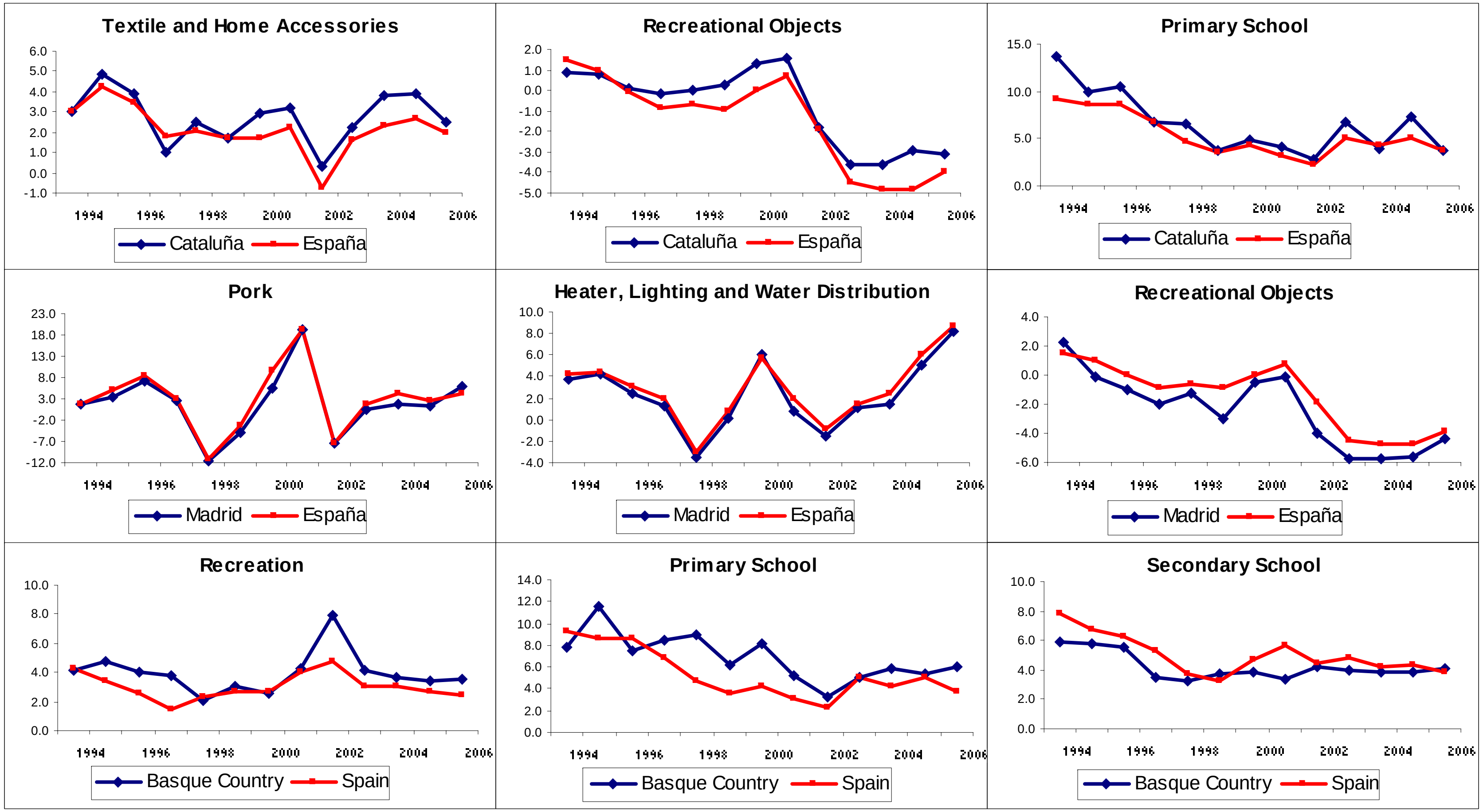

Figure 2. Historical price divergences between Spain and the regions of Valencia and Andalucía.