Banco Central de Chile
Documentos de Trabajo

Central Bank of Chile
Working Papers

N° 545

Diciembre 2009

FORECASTING CHILEAN INFLATION FROM DISAGGREGATE COMPONENTS

Marcus Cobb

BANCO CENTRAL DE CHILE

CENTRAL BANK OF CHILE

Documentos de Trabajo del Banco Central de Chile
Working Papers of the Central Bank of Chile
Agustinas 1180
Teléfono: (56-2) 6702475; Fax: (56-2) 6702231

Documento de Trabajo
N° 545

Working Paper
N° 545

FORECASTING CHILEAN INFLATION FROM DISAGGREGATE COMPONENTS

Marcus Cobb
Gerencia de Análisis Macroeconómico
Banco Central de Chile

Resumen
En este documento se realiza un ejercicio para determinar si incorporar información desagregada de precios mejora la precisión de la proyección de inflación total. Se utilizan diferentes métodos autorregresivos univariados y multivariados para distintos niveles de agregación a fin de proyectar la inflación en un periodo de inflación estable y uno de aceleración inflacionaria. Los resultados muestran que un cierto nivel de desagregación podría ser beneficioso cuando la inflación no es baja y estable, sugiriendo que bajo ciertas circunstancias el enfoque desagregado logra capturar de forma más eficiente la dinámica inflacionaria subyacente. Los beneficios son notorios en un horizonte de tres, seis y doce meses, mientras que en un horizonte de un mes la mejora parece insignificante.

Abstract
In this paper an exercise is performed to determine the usefulness of utilizing disaggregated price data to forecast headline inflation more accurately. A number of methods based on univariate and multivariate autoregressive models are used for different levels of disaggregation for a period of stable inflation and a period of accelerating inflation. The results show that a certain level of disaggregation could be beneficial when inflation is not low and stable, suggesting that under certain circumstances the disaggregate approach captures the underlying dynamics of inflation more efficiently. The benefits are noticeable for the three-, six- and twelve-month horizons, as opposed to the one-month horizon, where improvements seem negligible.

I would like to thank Ana Galvão, Michael Pedersen and an anonymous referee for useful discussions and comments on the ideas presented in the paper.

I - Introduction:

Inflation is constantly being monitored by private economic agents and policy-makers alike, as there are a considerable number of ways in which it affects them. From the return on private investment to revisions in economic policy, few are in a position to remain indifferent to its evolution and the monetary authorities therefore normally take a great deal of trouble to try to maintain inflation at a relatively low and stable rate. In this context, it is obvious that the future evolution of prices is a major concern, allowing those who can forecast inflation most accurately to have an advantage over the rest.

When it comes to forecasting, there are an extensive number of methods and approaches available and their relative success or failure to outperform each other is in general conditional to the problem at hand. Additionally, the restrictions set by the availability of data and the forecasting horizon makes it harder when deciding on an approach. To complicate matters, there has been some concern recently about the value of some approaches in a context of low volatility and persistent inflation, something that seems to be the situation in many countries at present. As a result, there have been a number of papers exploring the usefulness of using or incorporating disaggregate data to forecast the aggregate Consumer Price Index (CPI) as a way of capturing the underlying forces affecting the overall inflation that might not be immediately apparent when dealing with the aggregate alone.

In this context, I take the disaggregate components of Chile's CPI to see if forecasts based on the broader information set can provide an estimate of headline inflation that is more accurate than forecasts that rely solely on the aggregate CPI. To achieve this, I perform a forecasting exercise using the CPI information that spans from December 1998 to December 2008. In particular, I generate forecasts one, three, six and twelve months ahead and compare the accuracy of those obtained from disaggregate data relative to that of directly forecasting headline inflation. To account for the uncertainty regarding the most appropriate modelling procedure, the exercise is performed for a number of univariate and multivariate autoregressive specifications. The period used to evaluate the relative accuracy of the models, the last three years of the sample, is especially interesting as it begins with approximately a year of stable inflation and then suffers the effects of a sudden increase of food and oil prices that takes overall inflation near to ten percent.

The rest of the document is arranged as follows. In section 2 the theoretical framework on which the empirical exercise relies are discussed. Section 3 presents some recent studies and a brief summary of their findings. In section 4 the data is presented and its main features are discussed. Section 5 deals with the empirical framework by presenting the different models and explaining the methodology used for evaluation purposes. Section 6 explains the outcomes of the different methods and the general results; and section 7 summarizes the findings and proposes topics for further research.

II - Theoretical Framework:

In the forecaster's toolbox there is an array of methods and approaches. From the simple single equation regression models to the more complicated dynamic optimizing ones, there is growing consensus that one model alone is unlikely to fill the bill in any situation (Pagan and Robertson, 2002). In particular some have argued that, in the recent decades, standard approaches to inflation forecasting, like the Phillips Curve in its original and generalized versions, have become less useful when it comes to accurately predicting changes (Atkeson and Ohanian, 2001. and Fisher et al., 2002, regarding U.S. inflation) and attribute this poor performance to the low volatility of recent periods and changes in monetary policy regime[1]. Stock and Watson (2007) argue on the one hand that, due to the decrease in volatility, inflation has become more predictable in a sense that large changes in inflation are infrequent, but on the other hand, it has become harder to forecast because models with economic and statistical content are unable to identify the driving forces underlying inflation and therefore hardly outperform the naïve counterparts.

[1] In a recent study, Pincheira (2009) finds that persistence in Chilean inflation increased significantly mid 2007 and for 2008.

An interesting question that arises from the previous paragraph is whether this low-volatility at an aggregate level is the product of a similar behaviour at the micro-level. Altissimo et al. (2009) find that this is not necessarily true. They show empirically the importance of heterogeneity at the disaggregate level and how the subsequent aggregation procedure determines the dynamics of the aggregate data. They state a case for the compatibility of both disaggregate flexibility and aggregate stickiness. This finding, in any case, is in line with the earlier theoretical findings of Granger (1990), where the aggregation on a large scale of simple stationary autoregressive processes with different parameters could produce a fractional integrated process or, in other words, a near unit root.

With this in mind, a natural question to ask is whether resorting to the potentially more volatile disaggregate information could increase the forecasting accuracy of an aggregate measure. The discussion on such matters is far from being new but, given that both the aggregate and disaggregate approaches have known benefits and drawbacks that occur to different extents depending on the setting, the answer still remains unclear. As Barker and Pesaran (1990) point out:

> “When the disaggregate model is correctly specified and the available data are free from measurement errors, then the investigator could do no worse by adopting a disaggregate approach as compared to an aggregate one; and he or she may do better.” (p. 4)

Of course in the previous sentence there are at least two big, and certainly naive, assumptions. Both the collection of error-free data and the finding of the “real” data generating processes at any level of aggregation are highly unlikely. Without even considering the quality of the data, a model in an economics environment is at its best a good approximation to the behaviour of a probably complex system and therefore the desirability of following a full-scale micro approach, as advised in the previous statement, is far from being obvious.

This less-than-perfect setting calls for a careful review of the respective advantages and disadvantages of any approach being considered. Relying on the exposition by Barker and Pesaran

(1990), the main benefits of disaggregation are that, by allowing for different specifications across disaggregate variables to capture the individual dynamics, the better understanding of the underlying behaviour would suggest that a better overall prediction could be expected. The more accurate disaggregate information should also contribute to this better understanding. In a similar way, the additional information should allow for more powerful tests to be applied to the formulated hypothesis. Finally, the parameter estimates should allow for a better understanding of the mechanism at work, as opposed to those from an aggregate equation that rely on a particular aggregation structure.

Nevertheless, there are a number of reasons which suggest that following the aggregate approach may outweigh the potential benefits from the disaggregate approach. Probably the most important of these is that the aggregate model may be less subject to specification error (Grunfeld and Griliches, 1960), since the data generating process from the micro relationships are rarely known. Additionally, potential errors in variables at a micro level may cancel out when they are added together (Aigner and Goldfeld, 1974). A third point is that the disaggregate variables may have an unobserved influence on each other, but one that cancels out in the aggregate. If this continues to be true over the prediction period, the aggregate might give a better forecast.

Ultimately the decision on whether to pursue one approach or the other comes down to a trade-off between the specification errors in the disaggregate model and the magnitude of the aggregation error. However, even at a theoretical level things do not become any easier for the researcher, since Pesaran et al. (1989) provide a theoretical framework where the gain in terms of specification error may or may not compensate for the aggregation loss. All of this again suggests that the gains of disaggregation will depend of the problem at hand.

From the above evidence it is clear that the extra effort of disaggregation will only pay off if the behaviour from the heterogeneous sub-components can be sufficiently well approximated to produce a better forecast than that of a weighted average of them. The importance of an adequate description of the underlying micro dynamics therefore sets the further question as to how far to disaggregate. Given that the different levels of disaggregation are still a set of weighted averages of relatively less heterogeneous sub-components, one could expect the “real” dynamics to be less obvious with every aggregation step or to be lost completely fairly early in the aggregation process. If this were true, few improvements could be achieved from utilizing mid-level indexes to forecast the aggregate, but again this would depend on the relative weakness of the aggregation error and specification error effects.

III –Literature Review:

In the context of increasing forecasting accuracy[2], there have been a number of recent papers that use disaggregated data to forecast headline inflation in a univariate and multivariate autoregressive set-up. Most of them perform a forecasting competition by confronting a series of forecasts based on a direct approach, meaning that the aggregate series is forecasted directly, paying no attention to the available disaggregate information, with other forecasts based on an indirect approach, where the forecasts of the individual sub-components are combined to obtain an aggregate one. Then their relative accuracy is contrasted to try to determine the usefulness of the micro-data.

Examples of these forecasting competitions are Espasa et al. (2002), where factor and cointegration analysis are initially used to evaluate the sources of non-stationarity in the data. Their analysis reveals different trending behaviour in the components and therefore they conclude that a disaggregate estimation could provide more accurate forecasts and so proceed to utilize a number of approaches to forecast inflation. They find that, for Non-Energy US CPI, univariate models outperform the multivariate and provide a more accurate forecast than the direct aggregate approach when forecasting five to twelve steps ahead.

[2] It is worth mentioning that the focus here is solely to increase forecasting accuracy, not to give an explanation for the development of inflation and a particular forecast. Obviously, there must be a story behind a certain development, but for the purposes of this paper it will remain untold and probably unknown.

In a different paper, Hubrich (2005) uses various AR and VAR specifications for the Euro area. She finds no significant differences between methods for one-step-ahead forecast, but, unlike the previous case, she also finds that none of them outperform a random walk on a 12-month horizon. Benalal et al. (2004) on the other hand, also for the Euro area and using different univariate and multivariate specifications, find that forecasting the aggregate CPI index directly is better than the sub-component approach for horizons beyond one year, while getting mixed results for shorter horizons. They also find that the sub-component approach excels at forecasting core inflation, this is CPI excluding oil and unprocessed food.

Taking a different approach, Hendry and Hubrich (2006) develop a theoretical framework which shows that a forecast based on a joint estimation of the aggregate measure and its sub-components should asymptotically outperform those of a wholly direct or indirect approach. Empirically however, they get mixed results. For the Euro area they find that the joint procedure does not perform as well as the direct aggregate approach and that the indirect forecasts usually perform worst. For the U.S. they find that the disaggregate information does help forecasting, but underline that model selection is very important to whether the disaggregate information is useful. Based on their findings, they conclude that the extent to which the disaggregate information contributes to a more accurate forecast is a matter to be settled empirically.

In a recent paper for México, Capistran et al. (2009) explore the use of seasonal models to forecast short-term inflation considering both models with stochastic and deterministic seasonality. They find that models with deterministic seasonal patterns perform best and that the aggregation of individual forecasts provides a more accurate aggregate forecast.

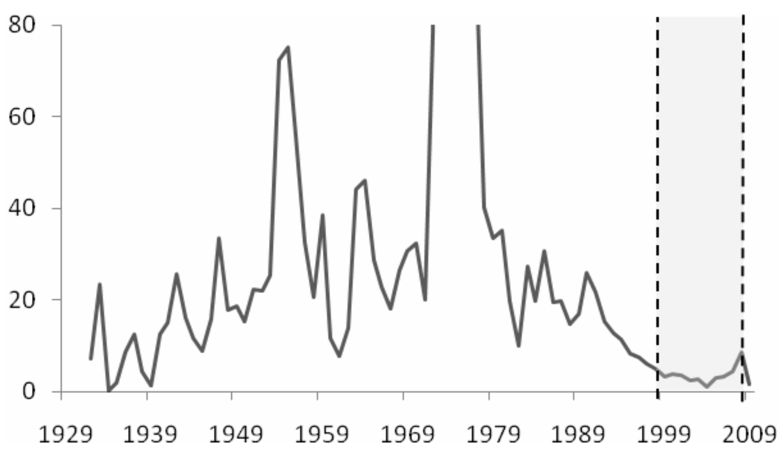
Regarding Chilean inflation, Diaz and Leyva (2008) seek to exploit the joint dynamic of disaggregated data in a multivariate setting in order to improve headline CPI forecast. By using an ad hoc general grouping of the CPI components to account for the idiosyncratic effects of food and energy prices, they hypothesise that the predictive ability of multivariate models in high inflation environments should be superior to the univariate approaches. They focus both on point and density forecast ability tests, but only prove their hypothesis for point forecasts.

IV - Data:

Chile's history as a country with low stable inflation does not go back very far. According to Morande (2001) and as illustrated in Figure 1, inflation exhibited a volatile behaviour starting in the 1930's, and this increased steadily, on average, to a peak in 1973, when it reached 606% per year. After the disinflationary efforts from the seventies the yearly inflation during the eighties stayed within a range between 10 and 25% but still showed the same volatility as before.

The Central Bank of Chile announced and followed an inflation targeting framework in 1990 that was fully implemented it in September 1999 when it jointly confirmed its indefinite inflation target (centred on 3% with a tolerance degree of +-1%, comparable with industrialized countries) and established a free-floating exchange rate (Banco Central de Chile, 2007).

Figure 1: Chile's Consumers Price Index 1932 - 2009
(yearly change, percentage)



Note: 2009 inflation only includes the first semester
Source: Instituto Nacional de Estadisticas (INE)

From 1990 onwards inflation fell gradually and steadily, meaning that since 1999, inflation had always remained under 5% until September 2007 when it started growing, literally fuelled by the surge in oil and food prices, to a peak of 9.9% in October 2008, since when it has fallen to 1.9% in June 2009.

The sample considered, delimited in Figure 1 by the dotted lines, consists of Chilean CPI information starting in December 1998 and finishing in December 2008, published and available from the National Statistics Institute (Instituto Nacional de Estadisticas – INE) on a monthly basis. It is worth mentioning that the complete sample is subject to the same general methodology and is therefore free of any generalized break due to framework revisions and other such factors. This would not be true if the sample had been lengthened either way, as the relevant framework starts in December 1998. In January 2009 the framework was updated by the Instituto Nacional de Estadísticas and the groupings were changed (Instituto Nacional de Estadísticas, 2009).

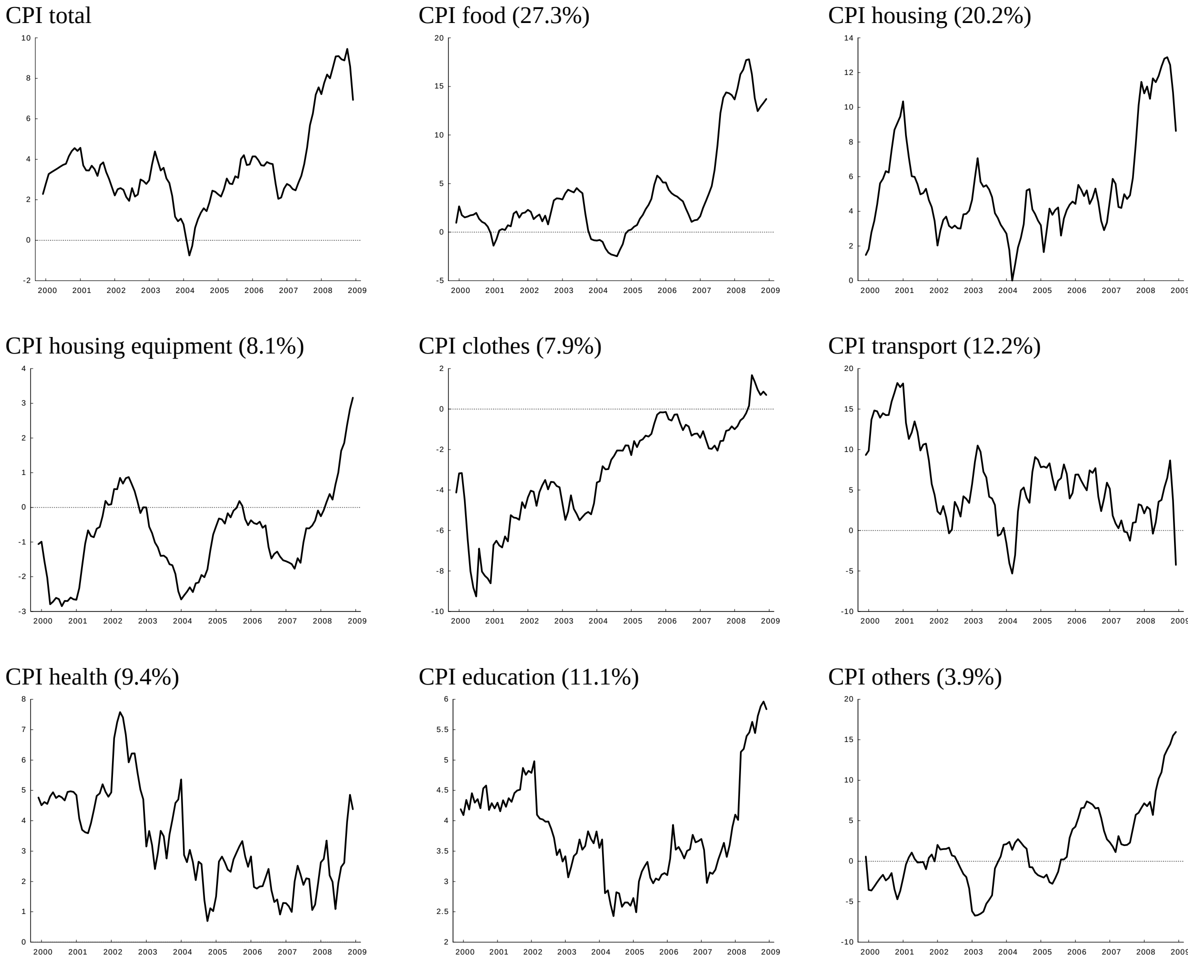
Additional to the headline inflation, the available information also covers, for the purpose of this paper, the following disaggregation levels: groups, sub-groups and articles, being the number of components at each level, 8, 41 and 156 respectively and the relative importance of each component at the first level of disaggregation: food (27.25%), housing (20.15%), housing equipment (8.11%), clothes (7.9%), transport (12.18%), health (9.39%), education (11.12%) and others (3.9%). The weights at this, and the other levels considered, are fixed and the aggregation is done by means of an arithmetic weighted mean[3] (Instituto Nacional de Estadísticas, 1999). The year-on-year inflation of each group and the aggregate is shown in Figure 2.

Given that the sample is relatively short, in line with the approach followed by Díaz and Leyva (2008), for the purposes of estimating multivariate models I additionally generated an ad hoc grouping of four sub-components in order to reduce parameter estimation uncertainty. As opposed to their approach, which means aggregating the original groups directly, I kept food (27.2%), merged housing and housing

[3] At further disaggregation levels, the use of geometric weighted mean appears.

equipment (24.8%), merged the remaining five components (42.3%) and removed the sub-groups corresponding to energy, namely: transport fuel, home fuel and electricity, from the previous regroupings, and grouped them in a separate index (5.7%).

Fig. 2: Year on year CPI Inflation (in %) of aggregate and subcomponents
(relative wheights)



Source: Instituto Nacional de Estadisticas (INE)

When needed, the series were seasonally adjusted utilizing the U.S. Census Bureau X12-ARIMA program (U.S. Census Bureau, 2007), which is probably the most used methodology in governments around the world in performing seasonal adjustments (Pedregal and Young, 2002), and is the methodology used by the Central Bank of Chile to adjust inflation data where the procedure is required (Bravo et al., 2002).

To make a detailed description of all the series would take more space than is justified, given the number of them, but it is worth pointing out the general characteristics regarding integration orders and seasonal patterns.

Using the untransformed indexes (in levels) to perform the Augmented Dickey Fuller tests (ADF)[4], headline inflation and all but one of the groups are I(1) under all specifications. Only clothes is I(0) or I(1) depending on whether a trend is included in the test. On the other hand, the Automatic Identification of Differencing Orders procedure from X12-ARIMA[5] classifies all of the above series as I(1) except for House Equipment that appeares as I(2). For the sub-group and articles level, the automatic procedure showed that at the sub-group level the vast majority of series were integrated of order one, but a couple were found to be I(2) and one to be stationary. At an Articles level all but four series were found to have a unit-root.

Finding that one of the groups is I(2) is at odds with a I(1) headline inflation given that the absence of a second I(2) group eliminates the possibility of any sort of cointegration. This could be evidence of a weakness in the identification procedure given that theoretically, under these circumstances, headline CPI should be I(2). However, the low power of Unit Root tests in cases of borderline behaviour is well documented in the literature particularly in a context of seasonal series (Phillips and Xiao, 1998) and this could be an example of it. In any case, according to Dickey and Pantula (1987), overdifferentiation would not be a relevant problem in a forecasting context, while Gomez and Maravall(1998) add that for a model based seasonal adjustment, the moderate overdifferentiation would be compensated by the MA parameters, and therefore wouldn't be a problem either. Another interesting point is made by Pincheira and Garcia (2009) that argue that including a stochastic trend on inflation would increase forecasting performance in periods where inflation follows trending patterns.

[4] Tables IV.1 and IV.2 present the statistics and asymptotic p-values for the core simple 1998.12 – 2005.12 and for the complete simple.

[5] Due to the low power of conventional tests when regular and seasonal Unit-roots may be present, particularly in presence of the large MA roots, the procedure basically compares the modulus of the AR roots of the model estimated with a critical value fixed “a priori”, taking care of the possibility of large MA roots (Gomez and Maravall, 1998 ; Maravall, 2003).

Turning to the seasonality, it would not be uncommon to find some seasonal patterns in a number of series given the monthly frequency of the data and this becomes particularly relevant considering that in a recent paper Capistran et al. (2009) find that, for Mexico at least, the dominant component in inflation is the seasonal one as opposed to the trend component and therefore the identification and estimation of the seasonal effects could be non trivial for a forecasting exercise.

Relying on the X12-ARIMA program and its Combined Test for the Presence of Identifiable Seasonality, which is basically a combination of a non-standard one-factor ANOVA test for testing stable seasonality and a non-standard two-factor ANOVA test for testing moving seasonality (Lothian and Morry, 1978b), just over 25% of the series showed identifiable seasonality at an articles level. This rose to 32% in the sub-group level and to half at group level. Up to this point, the test was seldom on the borderline between acceptance and rejection, but at the aggregate CPI level the result depended on the sample considered, resulting in identifiable seasonality being rejected a bit more often than not.

In light of these findings and in relation to the need to deal with the seasonality present in some of the series, I reviewed the Monitor and the Quality Assessment Statistics (Lothian and Morry, 1978a) that the X12-ARIMA provides when identifying the underlying seasonal patterns. Essentially these statistics try to assess the viability of separating the irregular component from the seasonal, trend and cycle components by comparing the relative contribution with the variance of each one of them. In particular if the variance of the seasonal component is too small when compared to the irregular variance, it will not be possible to separate them and therefore the procedure will result in a poor seasonal adjustment. These measures should be related somehow to the identifiability of seasonality, since one would expect the statistics to be poor in the series without seasonality, and relatively good in those with a clear seasonal pattern.

The calculated statistics showed an acceptable figure, meaning that the irregular and seasonal components of a series are found to be separable, for just over half of the series at an article level, over

60% at a sub-group level and about three quarters at a group level. The statistic for the headline CPI suggested, as opposed to the Combined Test for the Presence of Identifiable Seasonality, a clear seasonal pattern.

All of the above, in line with Altissimo et al. (2009), suggests how determinant aggregation is in the subsequent dynamics of the aggregated data, as both the integration order and seasonal patterns "creep up" the aggregation levels and could therefore potentially affect the model identification.

V - Empirical Framework:

The main objective of this exercise is to determine whether forecasting headline inflation using its components results in a more accurate forecast than that of forecasting it directly. However, the potential improvement of including disaggregate information, if any, could depend on the forecasting method and the forecasting exercise is therefore preformed not only for different levels of aggregation but also for a number of univariate and multivariate autoregressive methods. This should shed some light on whether disaggregation is generally helpful or not, but also permits to compare the overall performance of the different methods and particularly if any method consistently outperforms or is outperformed by the rest.

V.1 – Recursive forecasting procedure:

To evaluate the different methods a simulated out-of-sample forecast exercise was performed to determine the relative forecasting accuracy. The exercise was set up to replicate a real-time environment, meaning that only information that would have been available at that point in time was used for the identification and estimation of each model[6]. The exercise starts in December 2005, meaning that the first forecast is for January 2006. This leaves a minimum span of seven years for identification and estimation purposes. Using this core sample (1998:12 – 2005-12), each one of the

[6] This means that for every period, model identification, model estimation, seasonal adjustment and seasonal pattern forecasts were run using information up to that period.

models detailed in section V.2 were identified accordingly and estimated for all series (that is headline inflation, the eight groups, the 41 sub-groups and the 156 articles). Using these models, forecasts were calculated one, three, six and twelve steps ahead. Thereafter the sample was extended sequentially a month at a time and all the models were re-estimated including the new observations.

V.2 – Forecasting Models

V.2.1 - Univariate Autoregressive Models:

Regardless of the numerous developments in econometric modelling the ARMA approach continues to provide an often strong benchmark, against which other models can be compared (Stock and Watson, 1999) and, given the relatively short sample, the univariate models seemed to be a place to start. Therefore, a univariate autoregressive model AR(*p*,*d*) was defined as

$$\left(1-\sum_{i=1}^{p}\phi_i L^p\right)\nabla^d x_t = \delta + \mu_t \tag{1}$$

with x_t : log of the variable in period *t*
μ_t : white noise
∇ : difference operator
L : lag operator

where the differencing order *d* chosen to render the series stationary was selected according to the empirical unit root tests of section IV.

For the first model, the lag order *p* was chosen using the Schwartz Bayesian Information Criterion (BIC) and for the second, the Akaike Information Criterion (AIC) was utilized, both employing the unmodified series;

$$BIC = \ln\left(\frac{1}{T}\sum_{i=1}^{T} e_i^{\,2}\right) + \frac{k}{T}\ln(T)$$

(2)

$$AIC = \ln\left(\frac{1}{T}\sum_{i=1}^{T} e_i^{\,2}\right) + \frac{2k}{T}$$

With $\sum_{i=1}^{T} e_i^{\,2}$: the sum of squared residuals

k : number parameters being estimated

T : number of observations

Acknowledging the fact that the presence of identifiable seasonality in at least some series raises the need to address the issue, the two previous AR specifications were run utilizing seasonally adjusted data. The two models, the first of which used the BIC to choose the lag order while the second relied on the AIC, were used to produce forecasts for the seasonally adjusted series after which the estimated seasonal pattern was reintroduced to generate the forecasts for the unmodified series. It is worth mentioning .

The fifth model is a seasonal autoregressive moving average (SARIMA) (p,d,q)x(P,D,Q), defined as;

$$\left(1-\sum_{i=1}^{p}\phi_i L^p\right)\left(1-\sum_{i=1}^{P}\Phi_i L^P\right)\nabla^d \nabla^D x_t = \delta + \left(1-\sum_{i=1}^{q}\theta_i L^q\right)\left(1-\sum_{i=1}^{Q}\Theta_i L^Q\right)\mu_t \qquad (3)$$

with x_t : log of the variable in period t

μ_t : white noise

∇ : difference operator

L : lag operator

where the regular differencing order d and the seasonal differencing order D chosen to render the series stationary were, again, selected according to the empirical unit root tests of section IV. The AR lag orders - regular (p) and seasonal (P)- and the MA lag orders - regular (q) and seasonal (Q)- were chosen using BIC. This method was estimated using the original data given that this specification incorporates the seasonal pattern in the model and does not require identifying the seasonal and non seasonal components separately. This means it does not assume that this separation is possible.

V.2.2 - Unobserved Component Models:

As noted by Ghysels et al. (2006), seasonality is often neglected when it comes to modelling, even when it is a dominant feature for the economic agent, and they therefore argue that this feature of economic series should be fully integrated in the forecasting process.

The Unobserved Component models (also known as Structural Models) assume that a time series can be separated into a number of relatively simple components, that normally include a low frequency stochastic trend, a periodic cycle, a seasonal component and an irregular component (Pedregal and Young, 2002). As stated by Proietti (2002), structural models have a reduced form ARIMA representation that is subject to parameter restrictions. These are important for signal extraction and forecasting, given that they allow for more complex lag structures than the direct ARIMA approach would permit.

Regarding the accuracy of these models, Harvey (2006) recognizes that few studies have compared the forecasting performance of structural time-series models with other time-series methods in a broad set-up, but cites Andrews (1994) as saying that the unobserved components approach performs relatively well on monthly data and particularly for seasonal data and longer horizons.

An example of this sort of models is the X12-ARIMA that is widely used for seasonal adjustment, not necessarily for forecasting. Pedregal and Young (2002) classify it as an *ad hoc* approach as it uses a number of centred moving-average filters to extract the components from the original series as opposed to relying on explicit models. The decomposition proceeds assuming that the original series can be represented as the product (or sum, depending on the series) of its components:

$$X_t = TC_t + S_t + I_t \qquad (4)$$

with X_t : log of the original variable in period t
TC_t:log of the trend-cycle component in period t
S_t: log of the seasonal component in period t
I_t: log of the irregular component in period t

As described by Ladiray and Quenneville (1999) the trend-cycle component is modelled using a 13-term 2x12 moving average on the original series[7]. From the de-trended series, that is $X_t^{S+I} = X_t - TC_t$, the seasonal component is extracted using a 3x3 moving average one month at a time. The final seasonally adjusted series results from the sum of the trend-cycle and irregular components.

In order to apply the symmetric moving average filters to the complete sample, the time series are extended using forecasts and backcasts from a SARIMA model before the seasonal adjustment (Dagum, 1980). This procedure would result in smaller future revisions as opposed to alternative methods (Findley et al.,1998). Taking this into account, Ghysels et al. (2006) argue that, given that seasonal adjustment procedures involve forecasting for the identification of the different components, the out of sample forecast should be linked to the forecasting embedded in the seasonal adjustment process.

This leads to a second set of models where I used two automatic selection procedures of the X12-ARIMA program and used the model identified in the filtering process for the seasonal adjustment to produce a forecast beyond the end of sample. This means that in this group all methods were applied to the unmodified series as the SARIMA model identification and seasonal adjustment was embedded in the process.

The first method, within the Unobserved Components models, selects the respective underlying SARIMA model by selecting from a pool of models the specification that has the lowest in-sample

[7] These are the default options. The actual program follows a much more complex procedure involving outlier detection and regression effects and allows customizing many features, a complete description may be found in Ladiray and Quenneville (1999).

average mean error for the last three years of the sample[8]. Regarding this procedure Inoue and Kilian (2006) find that this sort of criteria tends to select overparameterized models, resulting in excessive finite sample prediction mean squared error when compared to the information criterion methods that consistently choose the best approximating model. This means that this method could play an important role in the accuracy of the seasonal adjustment process and therefore on forecasts based on these seasonally adjusted series.

Taking this information into consideration, the second method is based on the procedure in the TRAMO ("Time series Regression with ARIMA noise, Missing values and Outliers") time-series modelling program (Gomez and Maravall, 1997)[9]. The procedure is extensively documented in U.S. Census Bureau (2007), but basically, it relies on estimating the SARIMA “airline model” as the default option, then identifying the differencing orders with empirical unit root tests, identifying the SARMA model orders through an iterative procedure involving the BIC criterion and then comparing it with the airline model to quantify its improvements and evaluate which of them is more adequate. If the improvements of the statistics are not significant, the airline model is used. Maravall (2003) points out that empirically the airline model approximates well a large number of economic series.

The third method is an obvious extension of the previous one; instead of comparing the significance of the increase in performance of the chosen SARIMA over the airline model, it utilizes the model identified as having the lowest BIC value.

[8] This corresponds to the standard and only method available in the previous version of the program, updated in 2007 (Monsell, 2007). Originally the programs pool was composed of only five models with the airline model (0 1 1) x (0 1 1) as the default. For this exercise, the pool size was increased dramatically by including all possible combinations of models up to the maximum lag and differentiation orders of the new version of the program. These maximums are: up to four AR and MA lags and differencing twice. This makes the available models for each version of the program to be the same.

[9] It is the default routine the newest version of X12-ARIMA (Monsell, 2007).

V.2.3 - Vector Autoregression Models:

The third set of methods, given the relatively short sample, was simple multivariate autoregressive models, where a VAR(*p*) is of the form:

$$\mathbf{y}_t = \mathbf{m} + \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_2\mathbf{y}_{t-2} + \boldsymbol{\varepsilon}_t \tag{5}$$

or alternatively

$$\begin{bmatrix} y_{1t} \\ \vdots \\ y_{Nt} \end{bmatrix} = \begin{bmatrix} m_1 \\ \vdots \\ m_N \end{bmatrix} + \begin{pmatrix} a_{11}^1 & \dots & a_{1N}^1 \\ \vdots & \ddots & \vdots \\ a_{N1}^1 & \cdots & a_{NN}^1 \end{pmatrix} \begin{bmatrix} y_{1t-1} \\ \vdots \\ y_{Nt-1} \end{bmatrix} + \cdots + \begin{pmatrix} a_{11}^p & \dots & a_{1N}^p \\ \vdots & \ddots & \vdots \\ a_{N1}^p & \cdots & a_{NN}^p \end{pmatrix} \begin{bmatrix} y_{1t-p} \\ \vdots \\ y_{Nt-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{Nt} \end{bmatrix}$$

Regarding the estimation of the VARs, both the Engle-Granger (Engle and Granger, 1987) and Johansen (Johansen, 1991) tests suggested that there was no cointegration between the levels of the variables considered for both the sets of sub-components. Therefore, bearing in mind that, according to Hjalmarsson and Osterholm (2007), in a system with near-integrated variables, Johansen's tests for cointegration (both maximum eigenvalue and trace) have a higher-than-normal probability of reaching an erroneous conclusion about the cointegrating rank of the system, I followed the recommendations in Hamilton (1994) and disregarded the possibility of cointegration in levels and proceeded to estimate the VARs in differences and therefore imposed Unit Root in all the series. It is worth mentioning that all the models in this sub-section used seasonally adjusted series only and the patterns were therefore reintroduced into the respective forecasts accordingly to generate the forecasts for the unmodified series.

The first VAR, following Hubrich (2005) who cites Stock and Watson (1999), was a simple Phillips Curve model using inflation and the change in unemployment, considering twelve lags. That is $\mathbf{y}_t^{\mathrm{T}} = \left[\Delta \log CPI_t \quad \Delta \log Unemp_t\right]$ and $p = 12$.

The second, in order to reduce parameter estimation uncertainty but still retain the dynamics of food and energy (Díaz and Leyva, 2008), was a VAR using the four ad hoc groupings mentioned in the section dealing with data, that is $\mathbf{y}_t^{VAR.Adhoc} = \mathbf{y}_t^{\mathrm{T}} = \left[\Delta \log Food_t \quad \cdots \quad \Delta \log Energy_t\right]$[10]. The third was a VAR using the eight original groups from the first level of disaggregation, that is $\mathbf{y}_t^{VAR.eight} = \mathbf{y}_t^{\mathrm{T}} = \left[\Delta \log Food_t \quad \cdots \quad \Delta \log Others_t\right]$[11].

Taking Hendry and Hubrich's (2006) theoretical findings regarding the joint estimation of aggregate and disaggregates into account, the fourth and fifth models were VARs that additionally included headline CPI in the second and third models, that is $\mathbf{y}_t^{\mathrm{T}} = \left[\Delta \log CPI_t \quad \mathbf{y}_t^{VAR.Adhoc}\right]$ for the fourth model and $\mathbf{y}_t^{\mathrm{T}} = \left[\Delta \log CPI_t \quad \mathbf{y}_t^{VAR.eight}\right]$ for the fifth.

Regarding the lag structure, given that the number of parameters to be estimated in a VAR grows exponentially with the number of lags and variables, the relatively short sample does not allow for a very long lag structure. The VARs lag selection, with exception of the Phillips Curve model, relied on the generalization of the BIC provided in the software package for econometric analysis Gretl (Gnu Regression, Econometrics and Time-series Library, Cottrell and Lucchetti, 2009), that was consistently one lag for both specifications and successive samples. Just for comparison purposes, I also estimated the VARs imposing three and six lags to allow for a richer lag structure without losing too many degrees of freedom, acknowledging the fact that both the Akaike and Hannan-Quinn information criteria suggested a higher lag order (though both would generally suggest adding more lags even after running out of degrees of freedom).

V.3 – Headline Inflation Rate Forecast

The outcome of the previous process was a number of forecasts for headline inflation and its components (groups, sub-groups and articles). From those methods that produced a forecast for headline inflation and relied solely on the past values of CPI, termed as direct approaches, the inflation rate forecast was constructed according to;

[10] The four ad hoc groups as mentioned in Section IV are: food, housing composite, remainder and energy.
[11] The eight original groups are: food, housing, housing equipment, clothes, transport, health, education and others.

$$\pi_{t+h}^{direct} = \left(\frac{\mathrm{E}\left[P_{t+h} \mid P_1 ... P_t \right]}{P_t} \right) - 1 \tag{6}$$

Where:

π_{t+h}^{direct} : forecasted *h*-inflation rate with direct method,

P_t : CPI index in *t* in levels,

$\mathrm{E}\left[P_{t+h} \mid P_1 ... P_t \right]$: forecasted CPI index for *t+h*, conditional to its realized past values,

In order to construct the indirect forecasts, all the relevant individual sub-component forecasts were summed up using the fixed weights that are used to construct the original aggregate CPI index, according to:

$$\pi_{t+h}^{indirect} = \left(\frac{\sum_{j=1}^{J} \omega_j \mathrm{E}\left[P_{t+h}^j \mid P_1^j ... P_t^j \right]}{\sum_{j=1}^{J} \omega_j P_t^j} \right) - 1 \tag{7}$$

Where:

$\pi_{t+h}^{indirect}$: forecasted *h*-inflation rate with indirect method,

P_t^j : index of component *j* in *t* in levels,

ω_j : fixed weight for component *j*,

$\mathrm{E}\left[P_{t+h}^j \mid P_1^j ... P_t^j \right]$: forecasted index of component *j* for *t+h*, conditional to its realized past values.

It is worth mentioning that the aggregation is performed for the levels of the series (and forecasts) so that the sum of components is equal to the headline IPC. Pedersen (2009) shows that headline inflation calculated from the sum of the components (in levels) and that of a weighted sum of the year on year variation of the components is seldom equal. For the considered sample, the latter consistently underestimates the annual Inflation.

V.4 - Forecast combinations:

It might be presumptuous to claim that a single forecasting method incorporates all relevant information and the concept of forecasting combination therefore seems appealing. The concept and methodology have existed for some time since Bates and Granger (1969) introduced them, claiming that a combined forecast should be considered when multiple forecasts were to hand and, in particular, that the weights for the different sources should be assigned according to their recent relative performance (Newbold and Harvey, 2002). Additionally, Timmermann (2005) suggests that, conditional to the success of assigning the right weights, the diversification gains offered by a combined forecast may still make it an attractive strategy, even when the best single model can be determined for each point in time.

One could argue, however, that the gains stemming from combination would only be sufficient if the pooled forecasts were from methodologies of very different natures, but Newbold and Granger (1974) find that combining different similar univariate forecasts tends to result in a positive outcome.

Considering all the previous information, the number of combination approaches one could follow appears to be endless, so I examined only a few of them. The first of them is in line with Bates and Granger (1969) and involved combining the best forecasts at an aggregate level, either from a direct or indirect univariate approach or from a multivariate approach. I took all the aggregate forecasts and evaluated their root mean squared forecast error (RMSFE) over an out-of-sample comparison window for each forecasting horizon, where for a given lead time l and n errors $e_{t+i}(l)$ that denote the l-step-ahead forecast error for the forecast at time $t+i$;

$$RMSFE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} e_{t+i-1}^{2}(l)} \qquad (8)$$

Regarding the comparison window, Newbold and Harvey (2002) argue that, for evaluation purposes, if the relative performance of the different forecasts is suspected to be varying over time, only a few, most recent forecast errors should be considered. With this in mind, and the fact of the sample being relatively short, the comparison window was composed of to the three most recent months for each of the forecasting horizons. For example, this means that to produce a one step-ahead forecast for January 2008, the one step-ahead forecasts of all the previous methods were evaluated in an out-of-sample comparison over the months of October, November and December 2007.

The first combined forecast for this approach involved using the forecast from the best single models at each point in time (the one with the lowest RMSFE over the comparison window). The second and third were the simple and weighted averages of the ten best forecasts, where the weights for each individual forecast were inversely proportional to their RMSFE in the evaluation period. The number of forecasts included in the combinations was clearly arbitrary and subject to discussion, and probably to optimization, but it was chosen in the spirit of Newbold and Harvey (2002). They point out that especially in cases with a relatively large number of forecasts, the fact of necessarily assigning a weight to every forecast, no matter how bad its recent performance may have been, might be troublesome and a previous elimination stage might therefore be beneficial.

A second scope for forecast combination is to make use of the more accurate methods for the disaggregate series. Relying on the same three-month comparison window, I indirectly generate an aggregate forecast, utilising the individual sub-component forecasts stemming from the best single model at each point in time for every disaggregation level, that is groups, sub-groups and articles.

V.5 - Forecast performance comparison:

Once the corresponding forecasted annual inflation was calculated, following (6) or (7) depending on whether the forecast was direct or indirect, they were compared to the actual realization. The sample considered allowed for thirty-six one-month-ahead evaluation periods and twenty-four twelve-month-ahead evaluation periods.

To assess the differences between stable and accelerating inflation, the same exercise was conducted, but with forecasting comparison samples restricted to 2006.01 to 2007.06 for stable inflation, and to 2007.07 to 2008.12 for accelerating inflation. It is important to note that, given that the combined forecast methods require an evaluation period to rank the single methods, this results in a shrinking of the out-of-sample evaluation period to 2006.4-2008.12, and therefore the low stable inflation period is rendered too small for a significant twelve-month-ahead comparison.

VI - Results:

The main aspect that was being tested in this particular experiment was the benefits of disaggregation for the forecasting accuracy. However the exercise also allowed contrasting a number of other aspects. A second one was the consistency with which one method performed over different forecasting horizons. A third aspect was the relative performance of the different types of forecasting methods, and finally, how did the results compare for two different scenarios, one of low stable inflation and another of accelerating overall inflation.

Tables VI.1 to VI.3 present the resulting RMSFE for the complete sample and both sub-periods. The methods are shown for each of the forecasting horizons in the same order as they are presented in the empirical framework section[12], preceded by a forecast using a random walk with drift. It is worth remembering that, due to the fact that the combination of forecasts required a comparison window, the sample included in the results does not include the first three months of 2006. However, the results for the single method forecasts including these three months, do not alter their relative RMSFE accuracy ranking.

VI.1.1 - Performance over the period with low stable inflation:

Table VI.1 presents the RMSFE for the April 2006 – June 2007 period for a one, three and six-step-ahead forecasting horizon (the twelve-month horizon is left out, given that the sample only contains a few observations). The first thing to be noticed is that the benefits of disaggregation are relatively mild for the univariate methods, especially in the one-month-ahead horizon. On the three and six month horizons, the benefits appear to be more significant, but not consistent throughout the different methods.

Another thing is that the performance of the different methods is fairly erratic over the different circumstances between and within the different sets of methods, meaning that methods that show a relatively low RMSFE in some circumstances appear with a relatively high one in others. An example of this is seen in the unobserved components methods significantly reduced their RMSFE by disaggregating

[12] The glossary for Tables VI.1 to VI.3 is presented in Table VI.0

to the sub-group level for the six-month horizon, but increased it for the one-month-ahead and remained practically the same for the three-month-ahead forecast. In contrast, both of the VAR methods performed reasonably well over the different horizons, particularly the joint estimation procedure, making a case for the use of the disaggregate information.

Overall, the methods with the lowest RMSFE for this period were the VAR(JE-3), the VAR(JE-6) and the AR(BIC-SA), all of them for group level disaggregation, for the one, three and six month horizon respectively. If it were necessary to choose a particular method based on the results, it would be VAR(JE-3), which performed particularly well when using the group level disaggregation.

VI.1.2 - Performance over the period with accelerating inflation:

Table VI.2 presents the RMSFE for the July 2007 – December 2008 period for the one, three, six and twelve-step-ahead forecasting horizon. Just as in the case of low and stable inflations, at the one month-ahead forecast there is no substantial difference between the better methods at any level of aggregation or family of methods. However, for the univariate autoregressive methods, the initial disaggregation to group level made almost every method improve. The unobserved components methods, on the other hand, did better at an aggregate level.

As regards the multivariate methods, the group level VAR(6), ad hoc VAR(3) and group level VAR(JE-6) performed relatively well, as did the simple and weighted mean combined forecasts, but were nevertheless marginally worse than the best univariate methods.

Concerning the three and six month ahead forecast, the benefits of disaggregation are more noticeable, but again the first level of disaggregation proved to be the most rewarding for the univariate autoregressive methods. The multivariate autoregressive methods, based on the three-month lag length, also performed relatively well for both aggregation levels. Again, the unobserved component methods hardly benefited from the disaggregation while the aggregate combined forecasts performed well. For the twelve-month horizon, disaggregation to the group level also resulted in a lower RMSFE for most of the methods.

For the period of accelerating inflation, the univariate autoregressive methods at group level seem to have the edge, given that the AR(BIC), AR(AIC-SA) and ARIMA exhibit the lowest overall RMSFE for the one, six and twelve-month-ahead forecast horizon. Only for the three-step-ahead forecast did the ad hoc disaggregation level VAR(3) register the lowest overall RMSFE. In any case, if it were necessary to choose a particular method over the others, it would have to be the AR(AIC-SA), which performed relatively well for all the horizons. It is worth mentioning that the simple and weighted mean combined forecasts also showed good results, while the VAR(3) at group level exhibited comparable results for all but the twelve-month horizon.

VI.1.3 - Performance over the complete sample:

Table VI.3 presents the results for the complete sample. First of all, it is worth noting that, as the sample is composed of a low and an accelerating inflation period, the latter will tend to feature more predominantly in the figures than the former, due to the fact that the RMSFE penalizes larger forecasting errors more heavily. This means that the outcome may not necessarily be equivalent to averaging the results of the two sub-samples. Furthermore, the low and stable sample is considerably smaller for the longer forecasting horizons.

As in the case of the previous two sub-sections, the benefit of disaggregation is not immediately apparent for the one-month forecast, but is noticeable for longer horizons. It also becomes noticeable that the group level of disaggregation is the one that offers most improvement. Taking into account all forecasting horizons, it is the univariate autoregressive methods that use group level information, the standard and joint estimation VARs that use a three month lag length and the aggregate combined forecasts that consistently perform better than the rest. In particular the most accurate methods for each of the one, three, six and twelve month forecasting horizons are AR(BIC) at a group level, VAR(3) at an ad hoc level, AR(AIC-SA) at a group level and ARIMA at a group level respectively. These results confirm the value of utilizing disaggregate data but also leave in evidence the lack of a single best method.

VI.2 - Significance of the difference between forecasts:

While examining the results presented in the previous section and in the tables at the end of the document, it is natural to ask to what extent the forecasts at different level of aggregation are significantly different from a statistical perspective. The same question arises regarding whether the best method is significantly different from the runner-up or even from the worst. In an attempt to assess the sampling uncertainty involved in comparing point estimates, Diebold and Mariano (1995) present a general asymptotic test to compare two competing forecasts. Formally, it consists in testing whether the difference between two forecast error loss functions $d_t=g(e_1)-g(e_2)$ is significantly different from zero. The test statistic is $DM = \bar{d} \big/ \sqrt{\hat{V}(\bar{d})}$ where $\bar{d}$ is the sample mean of the difference between the forecast error loss functions (d_t) and $\hat{V}(\bar{d})$ is the sample means estimated variance. Under the null hypothesis of no difference between forecasts, the statistic has an asymptotic standard normal distribution and is therefore fairly easy to calculate. However, considering that the original test possesses proven undesired small-sample properties in situations that may occur frequently in economic contexts, the modified statistic proposed by Harvey et al. (1997) which circumvents the problem, is utilized. The square of the forecast error is used as the loss function for the test and a separate analysis is conducted for both the low stable inflation period and the accelerating inflation period.

From the last paragraph it is clear that the proposed test only allows two forecasts to be compared at a time, so a pair-wise comparison is performed for all forecasts. Taking the results of this exercise into account, a forecast is considered to be inferior if, in a comparison with any other forecast (not only those of the same method), the modified Diebold-Mariano test (MDM) rejects the null hypothesis of the difference between them to be zero in a way that means that the latter has a lower forecasting error. In other words, a forecast is inferior if there is at least one other forecast that is significantly better according to the MDM. It is worth making the distinction that, unlike the RMSFE criterion that ranks each one of the methods, the MDM only separates the methods into those considered to be inferior and those that are not. The results of the application of the MDM are shown in the Tables 1 to 3 by enclosing in a shaded area the RMSFE of those forecasts found to be inferior.

For the period of low and stable inflation, at a one-month horizon, the MDM classifies a number of forecasts as inferior, particularly those that use a high level of disaggregation, but does not reject equal accuracy for most of the multivariate methods, the combined forecasts and the aggregate univariate forecast. It does not rule out the random walk either. For the three and six-month horizon, a similar thing occurs, but the article level unobserved components method also appears to be as accurate as the best methods.

For the accelerating inflation period and the one-month horizon, most of the univariate AR group level and multivariate methods remain in the group of non-inferior methods, as do the combined forecasts and only the best aggregate methods. The three-month horizon, is similar, but only one of the aggregate models is not considered as inferior. At six and twelve months, all the aggregate measures and unobserved components methods are seen to be inferior.

VI.3 - The overall picture:

The results presented in the preceding section, both the ranking based on RMSFE and the further significance analysis using the Modified Diebold-Mariano test, suggest that, even though no particular method is seen to outperform the others, disaggregation to a certain level could be beneficial for forecasting accuracy when inflation is not low and stable, though when inflation is low and stable, the better disaggregate methods seem to perform at a level that is not significantly worse than the aggregate ones, making a case for disaggregation. On the other hand, disaggregating beyond a certain point does not seem to pay off, suggesting that after the initial improvement the misspecification error becomes more important than the reduction in aggregation error. Regarding the forecasting horizons, even though the benefits seem negligible for the one-month horizon, these become more apparent at three, six and twelve months ahead.

The results stemming from the unobserved components models suggest that linking the seasonal adjustment model with the forecasting model does not have the desired effect on forecasting accuracy. This can be seen in that the methods that chose the model independently from the one chosen in the seasonal adjustment process consistently outperformed those that linked the processes.

Regarding the multivariate models, although the joint estimation procedure generated forecasts comparable with the standard VARs over the whole sample, the standard disaggregate approach performed slightly better in the accelerating inflation period, once again making a case for the indirect approach in these circumstances. However, the indirect approach does not seem to work favourably for the combined forecast methods, given that the disaggregated combined methods seldom performed better then the aggregate ones, and normally performed worse.

An interesting outcome from this exercise is that the multivariate methods with a longer imposed lag structure generally performed better than the ones with a lag structure chosen by the BIC. The same was true when comparing the univariate AR methods chosen according to BIC and AIC meaning that, at least for this particular case, parsimony would not have been beneficial for forecasting accuracy.

Concerning the consistency of the competing methods over the different circumstances, it is worth noting that, even when the simple and weighted mean combined forecasts were never the best in the different cases, they were never far from it and, according to the MDM test, were not inferior to the other forecasts. Therefore, in line with Fildes and Ord (2002), who point out that forecast combination on average outperforms the methods being combined, it does seem worth considering them among the candidates to be used as the preferred method. Another point made by Fildes and Ord (2002) is that more sophisticated methods do not typically produce more accurate forecasts than simple ones, something that seem to be corroborated by the fact that simple AR models generally performed just as well or better than the other methods.

VII - Conclusion:

In this paper I carried out an exercise to find out whether using disaggregated CPI data to forecast headline inflation had a positive effect on its accuracy, all of this in the context of Chile. A number of methods – univariate, multivariate and forecast combinations, were used for different levels of disaggregation, subsequently ranked according to their root mean square forecasting error; and compared, using the modified Diebold-Mariano test, for both a period of both stable and accelerating inflation. The results suggested that a certain level of disaggregation could be beneficial in a period where inflation is not low and stable, but is not necessarily harmful otherwise. This therefore leads us to conclude that, at least under some circumstances, in times of relative price instability, the indirect approach could manage to capture the underlying dynamics more efficiently than the direct approach. The benefits are appreciable both for univariate and multivariate methods, and are more noticeable at the three, six and twelve month horizons. This statement must be taken in context, given that it relies on one inflationary episode. In particular, this episode was one where inflation was driven nearly exclusively by a small number of components. Whether these improvements in forecasting accuracy are still observed in episodes of a different nature (generalized inflation or any sort of deflation) is a question that remains unanswered.

An interesting feature is that longer lag structures seem to perform better in terms of forecasting accuracy than shorter ones for both the univariate and the multivariate methods, something that is at odds with the common understanding that parsimony is generally beneficial when choosing models that involve forecasting.

Regarding the performance of the competing methods over the different forecasting horizons, it is clear that no single method consistently outperforms the rest. But, taking this into account, it should be noted that the simple and weighted mean combined forecasts were never far removed from the better methods, making a case for the risk diversification implied by these methods. This seems especially appealing given that the performance of these forecasts could potentially show further improvements if the imposed parameters –comparison window length and number of pooled forecasts- were subject to some sort of optimization.

The research presented in this paper could be continued and extended in a number of ways. The more immediate are, for example, to perform the same exercise for core inflation measures, like CPI excluding energy and unprocessed foods, or to measure the performance on predicting inflation direction changes as opposed to magnitudes. A less obvious extension could be to combine different levels of disaggregation, to try to benefit from the gains in terms of specification from both the direct and indirect approach. The idea behind this is that by trying to model the disaggregated series, some of the series will be approximated relatively well, while others may not be, hence the lack of improvement from the full disaggregate approach. If these two groups could somehow be separated, one could forecast the "predictable" series individually and therefore produce better forecasts than those implied by a completely aggregate approach. One could then utilize a more aggregate approach to forecast the remaining series, in the hope that the aggregation will cancel out the individual errors to some extent.

The results in this paper show that using the disaggregate CPI information may significantly increase forecasting accuracy under certain conditions, but it does not explore the qualitative nature of the accuracy achieved. In order to address this, the results should be contrasted with other more sophisticated methods that have not been covered in this document but, ultimately, the final assessment as to whether the best forecast is good enough or not will depend on the problem in hand and, specifically, what is considered to be an acceptable margin of error.

References:

Aigner, D. J. and S. M. Goldfeld (1974) "Estimation and Prediction from Aggregate Data when Aggregates are Measured More Accurately than Their Components", Econometrica, Vol. 42, No. 1 (Jan., 1974), pp. 113-134

Altissimo, F., B. Mojon, and P. Zaffaroni (2009) "Can aggregation explain the persistence of inflation?" Journal of Monetary Economics, Volume 56, Issue 2, March 2009, Pages 231-241

Andrews, R.L. (1994) "Forecasting Performance of Structural Time Series Models", Journal of Business & Economic Statistics, American Statistical Association, vol. 12(1), pages 129-33, January.

Atkeson, A., and L.E. Ohanian (2001) "Are Phillips Curves useful for forecasting inflation?", FRB Minneapolis Quarterly Review (Winter) pp. 2–11

Banco Central de Chile (2007), "La Política Monetaria del Banco Central de Chile en el Marco de Metas de Inflación", Banco Central de Chile.

Barker, T.S. and M.H. Pesaran (1990) "Disaggregation in econometric modelling: an introduction", In Disaggregation in Econometric Modelling, Chapter 1, Ed. T.S Barker and M.H. Pesaran, published by Routledge.

Bates, J.M. and C.W.J. Granger (1969) "The combination of forecasts", Operations Research Quarterly 20, 451-468.

Benalal, N., Diaz del Hoyo, J. L., Landau, B., Roma, M. & Skudelny, F. (2004) "To aggregate or not to aggregate? euro area inflation forecasting", Working Paper 374, European Central Bank.

Bravo, H.F., V. Correa, L. Luna and F. Ruiz (2002) "Desestacionalización de series económicas: el procedimiento usado por el Banco Central de Chile", Central Bank of Chile, Working Papers N° 177, August 2002

Capistrán, C., C. Constandse and M.R. Francia (2009) "Using Seasonal Models to Forecast Short-Run Inflation in Mexico", Working Papers 2009-05, Banco de México.

Cottrell, A. and R. Lucchetti (2009) "Gretl User's Guide", http://gretl.sourceforge.net/

Dagum, E.B. (1980), "The X-11-ARIMA Seasonal Adjustment Method", Statistics Canada, Catalogue 12-564E.

Díaz, J. and G. Leyva (2008) "Forecasting Chilean inflation in difficult times", Central Bank of Chile Working Papers N° 511

Dickey, D. and S. Pantula (1987), "Determining the Ordering of Differencing in Autoregressive Processes", Journal of Business & Economic Statistics, American Statistical Association, vol. 5(4), pages 455-61, October.

Diebold, F.X. and R.S. Mariano (1995) "Comparing Predictive Accuracy" Journal of Business & Economic Statistics, American Statistical Association, vol. 13(3), pages 253-63, July.

Dolado, J., J. Gonzalo y F. Marmol (2001). "Cointegration", in Companion to Theoretical Econometrics, ed. by B.H.Baltagi , 634-654. Blackwell, New York.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error-Correction: Representation, Estimation, and Testing", Econometrica 55, pp. 251-276

Espasa, A., Senra, E. and Albacete, R. (2002) "Forecasting inflation in the European Monetary Union: A disaggregated approach by countries and by sectors", European Journal of Finance 8(4): 402–421.

Fildes, R. and Ord, K. (2002) "Forecasting competitions – their role in improving forecasting practice and research". In Clements, M. P. and Hendry, D. F. (Eds.), A Companion to Economic Forecasting. Oxford: Blackwells.

Findley, D.F., Monsell, B.C., Bell, W.R., Otto, M.C. and Chen B. (1998), "New Capabilities and Methods of the X-12-ARIMA Seasonal-Adjustment Program", Journal of Business and Economic Statistics, Vol. 16, No. 2., 127-177.

Fisher, J.D.M., C.T. Liu, and R. Zhou (2002) "When can we forecast inflation?", FRB Chicago Economic Perspectives (1Q) pp. 30-42.

Ghysels, E., D.R. Osborn and P.M.M. Rodrigues (2006) "Forecasting Seasonal Time Series", Handbook of Economic Forecasting, Chapter 13, Ed: G. Elliott, C.W.J. Granger and A. Timmermann, Elsevier, 2006, Volume 1, Pages 659-711,

Gomez, V. and A. Maravall (1997) "Programs TRAMO and SEATS : Instructions for the user". Working Paper 97001, Ministerio de Economıa y Hacienda, Direction General de Analisis y Programacion Presupuestaria, Madrid.

Gomez, V. and A. Maravall (1998) "Automatic Modeling Methods for Univariate Series", Banco de España Working Papers 9808, Banco de España.

Granger, C. W. J. (1990) "Aggregation of time series variables: a survey", In Disaggregation in Econometric Modelling, Chapter 2, Ed. T.S Barker and M.H. Pesaran, published by Routledge.

Grunfeld Y.and, Z. Griliches (1960) "Is aggregation necessarily bad?", The Review of Economics and Statistics, Vol. 42, No. 1 (Feb., 1960), pp. 1-13

Hamilton, J.D. (1994) "Time Series Analysis", Princeton University Press

Harvey, A.(2006) "Forecasting with Unobserved Components Time Series Models", In: Handbook of Economic Forecasting, Chapter 7, Editors: G. Elliott, C.W.J. Granger and A. Timmermann, Elsevier, 2006, Volume 1, Pages 327-412

Harvey, D., S. Leybourne and P. Newbold (1997) "Testing the equality of prediction mean squared errors" International Journal of Forecasting, Elsevier, vol. 13(2), pages 281-291, June.

Hendry, D.F. and K. Hubrich (2006) "Forecasting economic aggregates by disaggregates" Working Paper Series 589, European Central Bank

Hjalmarsson, E. and P. Osterholm (2007) "Testing for cointegration using the Johansen methodology when variables are near-integrated", International Finance Discussion Papers 915, Board of Governors of the Federal Reserve System (U.S.).

Hubrich, K. (2005) "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?", International Journal of Forecasting 21(1): 119–136.

Inoue, A. and L. Kilian (2006) “On the selection of forecasting models”, Journal of Econometrics, Volume 130, Issue 2, February 2006, Pages 273-306,

Instituto Nacional de Estadísticas (1999), “Índice de Precios al Consumidor (IPC), base diciembre 1998=100, Aspectos Metodológicos”, Instituto Nacional de Estadísticas.

Instituto Nacional de Estadísticas (2009), “Metodología del Índice de Precios al Consumidor (IPC) Base diciembre de 2008=100”, Instituto Nacional de Estadísticas

Johansen, S. (1991) “Cointegration and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models”, Econometrica, Vol.59, No.6.

Johnston, J. and J. DiNardo (1997) “Econometric Methods”, Fourth Edition, McGraw-Hill

Ladiray, D. and Quenneville, B. (1999) "Understanding the X11 Method", Time Series Research and Analysis Centre, Statistics Canada, http://www.census.gov/pub/ts/papers.

Lothian, J. and Morry, M. (1978a) “A Set of Quality Control Statistics for the X-11-ARIMA Seasonal Adjustment Method” Staff Paper STC1788E, Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa.

Lothian, J. and Morry, M. (1978b) “A Test for the Presence of Identifiable Seasonality When Using the X-11-ARIMA Program”, Staff Paper STC2118, Seasonal Adjustment and Time Series Analysis Staff, Statistics Canada, Ottawa.

Maravall, A. (2003), “Notes on Programs TRAMO and SEATS, Part II, Time Series Regression with ARIMA Noise, Missing Observations and Outliers”.

Monsell, B.C. (2007) “Release Notes for Version 0.3 of X-12 ARIMA”, Research Report Series, Statistical Research Division, U.S. Census Bureau, June 2007

Morande, F. (2001) “Una decada de metas de inflacion en Chile: desarrollos, lecciones y desafios”, Economia chilena, Vol. 4, Nº. 1, 2001 , pags. 35-62

Newbold, P. and C. W. J. Granger (1974) “Experience with Forecasting Univariate Time Series and the Combination of Forecasts”, Journal of the Royal Statistical Society, Series A (General), Vol. 137, No. 2 (1974), pp. 131-165

Newbold, P. and D.I. Harvey (2002) “Forecasting Combination and Encompassing”, in A Companion to Economic Forecasting 2002, Chapter 7, edited by M.P. Clements and D.F. Hendry

Pagan, A.R. and J. Robertson (2002) “Forecasting for Policy”, in A Companion to Economic Forecasting 2002, Chapter 7, edited by M.P. Clements and D.F. Hendry

Pedregal D.J. and P.C. Young (2002) “Statistical approaches to modelling and forecasting time series”, In A companion to economic forecasting, Chapter 4, edited by M.P. Clements and D.F. Hendry

Pesaran, M. H., R. G. Pierse, and M. S. Kumar (1989) “Econometric analysis of aggregation in the context of linear prediction models”, Econometrica, Vol. 57, No. 4 (Jul., 1989), pp. 861-888

Phillips P.C.B. and Z. Xiao (1998), “A Primer on Unit Root Testing”, Cowles Foundation Discussion Papers 1189, Cowles Foundation, Yale University.

Pincheira, P (2009), “La Dinámica de la Persistencia Inflacionaria en Chile”, Economía Chilena, Vol. 12, Nº. 1, 2009 , pags. 97-104

Pincheira, P. and A. Garcia (2009), “En búsqueda de un buen benchmark predictivo para la inflación”, Documentos de Trabajo [BC], Nº 514, Banco Central de Chile

Pollock, D.S.G. (1999) “A Handbook of Time Series Analysis, Signal Processing, and Dynamics” Academic Press Inc.

Proietti, T. (2002) “Forecasting with Structural Time-Series Models”, in A Companion to Economic Forecasting 2002, Chapter 5, edited by M.P. Clements and D.F. Hendry

Ramchander, S., M. W. Simpson and M. K. Chaudhry (2005) “The influence of macroeconomic news on term and quality spreads”, The Quarterly Review of Economics and Finance, Volume 45, Issue 1, February 2005, Pages 84-102

Stock, J. and M. W. Watson (1999) “Forecasting inflation”, Journal of Monetary Economics, 1999, 44(2,Oct), 293-335.

Stock, J. and M. W. Watson (2007) "Why Has U.S. Inflation Become Harder to Forecast?," Journal of Money, Credit and Banking, Blackwell Publishing, vol. 39(s1), pages 3-33, 02.

Timmermann, A. G. (2005) “Forecast Combinations”, CEPR Discussion Paper No. 5361.

U.S. Census Bureau (2007) “X-12-ARIMA Reference Manual: Version 0.3”, Time Series Staff, Statistical Research Division, Washington.

Tables:

Table IV.1

Augmented Dickey-Fuller Unit-root test
Lag Length based on SIC
Sample: 1998M12 2005M12

	Exogenous Variable								
	none			constant			constant, trend		
	N° Lags	t-Statistic	p-value	N° Lags	t-Statistic	p-value	N° Lags	t-Statistic	p-value
Levels									
CPI	2	4.47	1.0000	2	-0.52	0.8812	2	-2.12	0.5265
Groups									
food	1	1.36	0.9558	1	-1.11	0.7068	1	-3.27	0.0788
housing	1	3.45	0.9998	1	-0.46	0.8920	1	-3.01	0.1347
housing equipment	1	-2.61	0.0096	1	-1.15	0.6926	1	-1.88	0.6578
clothes	4	-5.74	0.0000	4	-3.75	0.0050	4	0.41	0.9989
transport	2	3.00	0.9993	2	-1.54	0.5080	2	-1.97	0.6110
health	0	5.67	1.0000	0	-1.67	0.4448	0	-1.31	0.8798
education	0	3.39	0.9998	0	-1.03	0.7393	0	-4.21	0.0067
others	1	-0.83	0.3532	1	-3.05	0.0345	1	-2.58	0.2895
1st Differences									
CPI	4	-1.85	0.0613	1	-7.35	0.0000	1	-7.29	0.0000
Groups									
food	0	-5.42	0.0000	0	-5.63	0.0000	0	-5.54	0.0001
housing	0	-5.45	0.0000	0	-6.80	0.0000	0	-6.75	0.0000
housing equipment	0	-4.99	0.0000	0	-5.76	0.0000	0	-5.76	0.0000
clothes	0	-7.78	0.0000	3	-8.04	0.0000	3	-9.64	0.0000
transport	1	-6.25	0.0000	1	-7.39	0.0000	1	-7.46	0.0000
health	0	-6.69	0.0000	0	-8.93	0.0000	0	-9.10	0.0000
education	0	-8.70	0.0000	0	-9.96	0.0000	0	-9.94	0.0000
others	0	-8.01	0.0000	0	-8.02	0.0000	0	-8.27	0.0000
2nd Differences									
CPI	3	-9.32	0.0000	3	-9.26	0.0000	3	-9.20	0.0000
Groups									
food	0	-10.61	0.0000	0	-10.53	0.0001	0	-10.50	0.0000
housing	1	-10.10	0.0000	1	-10.04	0.0000	1	-9.98	0.0000
housing equipment	1	-9.22	0.0000	1	-9.17	0.0000	1	-9.11	0.0000
clothes	4	-16.38	0.0000	4	-16.28	0.0001	4	-16.49	0.0001
transport	3	-9.69	0.0000	3	-9.63	0.0000	3	-9.57	0.0000
health	4	-8.36	0.0000	4	-8.31	0.0000	4	-8.25	0.0000
education	1	-12.38	0.0000	1	-12.31	0.0001	1	-12.24	0.0000
others	1	-10.16	0.0000	1	-10.10	0.0000	1	-10.04	0.0000

Table IV.2

Augmented Dickey-Fuller Unit-root test
Lag Length based on SIC
Sample: 1998M12 2008M12

	Exogenous Variable								
	none			**constant**			**constant, trend**		
	N° Lags	t-Statistic	p-value	N° Lags	t-Statistic	p-value	N° Lags	t-Statistic	p-value
Levels									
CPI	2	3.40	0.9998	2	1.29	0.9985	2	-0.69	0.9713
Groups									
food	1	1.95	0.9875	2	1.81	0.9997	1	-0.46	0.9841
housing	1	4.04	1.0000	1	1.65	0.9996	1	-0.53	0.9809
housing equipment	1	-1.33	0.1689	1	-2.13	0.2316	1	0.42	0.9990
clothes	4	-5.45	0.0000	4	-5.79	0.0000	4	-0.53	0.9807
transport	2	1.95	0.9876	2	-2.00	0.2874	2	-2.36	0.3964
health	0	5.86	1.0000	0	-1.59	0.4829	0	-1.98	0.6055
education	0	4.34	1.0000	0	-0.01	0.9552	0	-3.78	0.0207
others	3	1.25	0.9459	1	3.26	1.0000	1	2.06	1.0000
1st Differences									
CPI	0	-3.94	0.0001	1	-5.51	0.0000	1	-5.68	0.0000
Groups									
food	0	-4.19	0.0000	0	-4.63	0.0002	1	-5.44	0.0001
housing	0	-5.27	0.0000	0	-6.73	0.0000	0	-7.02	0.0000
housing equipment	0	-5.78	0.0000	0	-5.93	0.0000	0	-6.49	0.0000
clothes	4	-3.78	0.0002	4	-5.19	0.0000	3	-11.95	0.0000
transport	1	-6.88	0.0000	1	-7.36	0.0000	1	-7.50	0.0000
health	0	-8.77	0.0000	0	-11.11	0.0000	0	-11.20	0.0000
education	0	-10.50	0.0000	0	-12.23	0.0000	0	-12.19	0.0000
others	2	-3.05	0.0025	2	-3.25	0.0194	0	-8.87	0.0000
2nd Differences									
CPI	1	-9.33	0.0000	1	-9.28	0.0000	1	-9.27	0.0000
Groups									
food	0	-10.40	0.0000	0	-10.34	0.0000	0	-10.31	0.0000
housing	0	-14.96	0.0000	0	-14.90	0.0000	0	-14.85	0.0000
housing equipment	2	-10.20	0.0000	2	-10.17	0.0000	2	-10.20	0.0000
clothes	4	-20.37	0.0000	4	-20.28	0.0000	4	-20.38	0.0000
transport	3	-7.96	0.0000	3	-7.94	0.0000	3	-7.93	0.0000
health	4	-9.36	0.0000	4	-9.32	0.0000	4	-9.28	0.0000
education	1	-14.87	0.0000	1	-14.81	0.0000	1	-14.75	0.0000
others	1	-13.16	0.0000	1	-13.11	0.0000	1	-13.05	0.0000

Table VI.0:
Glossary for Tables VI.1 to VI.3

Model	Description
RW	random walk with drift
1 - Univariate Autoregressive Models	
AR(BIC) AR(AIC)	univariate autoregressive methods that use unmodified series and BIC and AIC as the lag selection criteria
AR(BIC-SA) AR(AIC-SA)	univariate autoregressive methods with seasonally adjusted series, using BIC and AIC as lag selection criterion
SARIMA	SARIMA that uses unmodified series and BIC as lag selection criteria
2 - Unobserved Component Models	
UC(AFE)	unobserved components method that chooses the underlying model based on the average forecast error over the last three months of the sample
UC(X12)	unobserved components method that chooses the underlying model based on the embedded procedure in the X12-ARIMA program
UC(BIC)	unobserved components method that chooses the underlying model based on the lowest BIC
3 - Vector Autoregression Models	
VAR(Ph)	simple Philips Curve with 12 lags
VAR(BIC)	standard VAR that uses the BIC as the lag selection criterion
VAR(3)	standard VAR with a imposed lag structure of three lags
VAR(6)	standard VAR with a imposed lag structure of six lags
VAR(JE-BIC)	joint estimation VAR that uses the BIC as the lag selection criterion
VAR(JE-3)	joint estimation VAR with a imposed lag structure of three lags
VAR(JE-6)	joint estimation VAR with a imposed lag structure of six lags
4 - Forecast combinations	
Best Single	combination of the single best aggregate methods
Simple Mean	simple average of the ten best single methods
Weighted Mean	weighted average of the ten best single methods
Group	combination of the best single method for each group component
Sub-group	combination of the best single method for each sub-group component
Articles	combination of the best single method for each article component

Table VI.1:
Root Mean Squared Forecast Error of year on year inflation in percentage points, for periods 2006(4) to 2007(6) - one, three and six steps ahead.

	direct	indirect								
		group (8)	sub-group (41)	article (156)			ad hoc (4)	group (8)		
One month horizon										
RW	0.390									
Univariate Methods:					Multivariate Methods:				Combined Methods:	
AR(BIC)	0.336	0.302	0.389	0.334	VAR(Ph)	0.243			Aggregate:	
AR(AIC)	0.336	0.302	0.373	0.327					Best Single	0.415
AR(BIC-SA)	0.265	0.305	0.341	0.327	VAR(BIC)		0.294	0.283	Simple Mean	**0.298**
AR(AIC-SA)	0.249	0.300	**0.330**	**0.308**	VAR(3)		0.253	0.249	Weighted Mean	0.318
SARIMA	**0.234**	0.293	0.408	0.361	VAR(6)		0.340	0.266		
									Disaggregate:	
UC(AFE)	0.281	0.293	0.376	0.383	VAR(JE-BIC)		0.257	0.246	Group	0.286
UC(X12)	0.308	0.304	0.415	0.338	VAR(JE-3)		**0.244**	$\boxed{\mathbf{0.218}}$	Sub-group	0.393
UC(BIC)	0.307	0.321	0.411	0.380	VAR(JE-6)		0.303	0.240	Articles	0.374
Three month horizon										
RW	0.830									
AR(BIC)	0.898	0.772	0.955	0.776	VAR(Ph)	0.675			Aggregate:	
AR(AIC)	0.898	0.774	0.923	0.779					Best Single	0.701
AR(BIC-SA)	0.701	0.661	0.800	0.728	VAR(BIC)		0.703	0.707	Simple Mean	**0.659**
AR(AIC-SA)	0.679	0.632	0.777	0.712	VAR(3)		0.598	0.609	Weighted Mean	0.691
SARIMA	0.730	0.828	0.968	0.835	VAR(6)		0.677	0.598		
									Disaggregate:	
UC(AFE)	**0.673**	0.725	**0.730**	**0.674**	VAR(JE-BIC)		0.682	0.697	Group	0.792
UC(X12)	0.822	0.804	0.758	0.694	VAR(JE-3)		**0.578**	0.533	Sub-group	0.759
UC(BIC)	0.824	0.774	0.753	0.705	VAR(JE-6)		0.674	$\boxed{\mathbf{0.528}}$	Articles	0.635
Six month horizon										
RW	1.112									
AR(BIC)	1.187	1.033	1.148	1.015	VAR(Ph)	**0.678**			Aggregate:	
AR(AIC)	1.187	1.042	1.121	1.057					Best Single	0.713
AR(BIC-SA)	0.738	$\boxed{\mathbf{0.536}}$	0.804	0.789	VAR(BIC)		0.565	0.618	Simple Mean	**0.674**
AR(AIC-SA)	0.722	0.639	0.817	0.809	VAR(3)		**0.539**	0.596	Weighted Mean	0.680
SARIMA	0.959	1.088	1.144	1.151	VAR(6)		0.666	0.698		
									Disaggregate:	
UC(AFE)	0.878	0.741	**0.572**	0.667	VAR(JE-BIC)		0.585	0.570	Group	0.791
UC(X12)	0.735	0.887	0.643	**0.649**	VAR(JE-3)		0.566	0.547	Sub-group	0.661
UC(BIC)	0.731	0.827	0.824	0.861	VAR(JE-6)		0.592	0.633	Articles	0.602

Note: (1) given that the base sample for estimation runs from 1998(12) to 2005(12) the actual periods being forecasted when including combined forecasts for the respective horizons start at: 2006(4) for one month, 2006(6) for three months and 2006(9) for six months.
(2) the bolded numbers indicate the lowest RMSFE for the respective disaggregation level for a specific forecast horizon.
(3) the shaded area indicate those forecasts that are considered inferior, as defined in the text, by the Diebold-Mariano statistic.
(4) the lowest overall RMSFE for a specific horizon is enclosed in a rectangle.
(5) the twelve month horizon is not shown given the few observations.

Table VI.2:
Root Mean Squared Forecast Error of year on year inflation in percentage points, for periods 2007(7) to 2008(12) - one, three, six and twelve steps ahead.

	direct	indirect group (8)	indirect sub-group (41)	indirect article (156)			ad hoc (4)	group (8)		
One month horizon										
RW	0.743									
Univariate Methods:					Multivariate Methods:				Combined Methods:	
AR(BIC)	0.575	**0.478**	**0.522**	0.554	VAR(Ph)	0.687			Aggregate:	
AR(AIC)	**0.526**	0.485	0.527	0.578					Best Single	0.636
AR(BIC-SA)	0.569	0.524	0.582	0.596	VAR(BIC)		0.581	0.604	Simple Mean	**0.522**
AR(AIC-SA)	0.527	0.519	0.592	0.612	VAR(3)		**0.527**	0.655	Weighted Mean	0.526
SARIMA	0.708	0.564	0.552	**0.533**	VAR(6)		0.607	0.523		
									Disaggregate:	
UC(AFE)	0.608	0.610	0.671	0.781	VAR(JE-BIC)		0.610	0.630	Group	0.547
UC(X12)	0.552	0.592	0.604	0.834	VAR(JE-3)		0.547	0.557	Sub-group	0.573
UC(BIC)	0.507	0.597	0.608	0.817	VAR(JE-6)		0.577	0.531	Articles	0.803
Three month horizon										
RW	1.512									
AR(BIC)	1.613	1.089	1.207	1.300	VAR(Ph)	1.919			Aggregate:	
AR(AIC)	1.587	1.100	1.188	1.272					Best Single	**0.990**
AR(BIC-SA)	**1.412**	1.061	1.186	1.202	VAR(BIC)		1.226	1.282	Simple Mean	1.040
AR(AIC-SA)	1.415	**1.038**	**1.154**	**1.158**	VAR(3)		**0.985**	1.109	Weighted Mean	1.038
SARIMA	1.495	1.077	1.269	1.307	VAR(6)		1.094	1.144		
									Disaggregate:	
UC(AFE)	1.594	1.404	1.435	1.693	VAR(JE-BIC)		1.329	1.366	Group	1.149
UC(X12)	1.611	1.394	1.514	1.745	VAR(JE-3)		1.046	1.056	Sub-group	1.232
UC(BIC)	1.469	1.429	1.449	1.562	VAR(JE-6)		1.117	1.099	Articles	1.392
Six month horizon										
RW	2.671									
AR(BIC)	2.939	2.061	2.382	2.406	VAR(Ph)	3.757			Aggregate:	
AR(AIC)	2.933	2.069	2.273	2.287					Best Single	2.220
AR(BIC-SA)	**2.642**	1.994	2.371	2.365	VAR(BIC)		2.459	2.525	Simple Mean	**2.050**
AR(AIC-SA)	2.671	**1.953**	**2.264**	**2.229**	VAR(3)		**2.181**	2.043	Weighted Mean	2.091
SARIMA	2.915	2.060	2.424	2.516	VAR(6)		2.341	2.204		
									Disaggregate:	
UC(AFE)	2.919	2.721	2.610	2.920	VAR(JE-BIC)		2.670	2.726	Group	2.193
UC(X12)	3.044	2.664	2.893	3.268	VAR(JE-3)		2.318	2.197	Sub-group	2.171
UC(BIC)	2.694	2.706	2.636	2.714	VAR(JE-6)		2.393	2.233	Articles	2.151
Twelve month horizon										
RW	4.483									
AR(BIC)	4.778	3.555	4.398	3.917	VAR(Ph)	6.004			Aggregate:	
AR(AIC)	4.776	3.564	4.128	3.666					Best Single	**3.542**
AR(BIC-SA)	4.719	3.537	4.265	3.925	VAR(BIC)		4.545	4.574	Simple Mean	3.555
AR(AIC-SA)	4.747	3.474	**4.121**	**3.653**	VAR(3)		**4.199**	3.692	Weighted Mean	3.543
SARIMA	**4.547**	**3.299**	4.227	4.085	VAR(6)		4.442	3.877		
									Disaggregate:	
UC(AFE)	4.835	4.443	4.573	4.717	VAR(JE-BIC)		4.784	4.847	Group	3.982
UC(X12)	5.006	4.502	5.102	5.240	VAR(JE-3)		4.287	4.078	Sub-group	3.996
UC(BIC)	4.714	4.371	4.402	4.070	VAR(JE-6)		4.367	3.917	Articles	3.715

Note: (1) the bolded numbers indicate the lowest RMSFE for the respective disaggregation level for a specific forecast horizon.
(2) the shaded area indicate those forecasts that are considered inferior, as defined in the text, by the Diebold-Mariano statistic.
(3) the lowest overall RMSFE for a specific horizon is enclosed in a rectangle.

Table VI.3:
Root Mean Squared Forecast Error of year on year inflation in percentage points, for periods 2006(4) to 2008(12) - one, three, six and twelve steps ahead.

One month horizon

	direct	indirect								
		group (8)	sub-group (41)	article (156)			ad hoc (4)	group (8)		
RW	0.608									
Univariate Methods:					Multivariate Methods:				Combined Methods:	
AR(BIC)	0.481	**0.408**	0.466	0.467	VAR(Ph)	0.533			Aggregate:	
AR(AIC)	0.450	0.412	**0.464**	0.481					Best Single	0.546
AR(BIC-SA)	0.457	0.438	0.488	0.493	VAR(BIC)		0.472	0.485	Simple Mean	**0.435**
AR(AIC-SA)	**0.424**	0.434	0.490	0.497	VAR(3)		**0.425**	0.512	Weighted Mean	0.444
SARIMA	0.546	0.461	0.492	**0.463**	VAR(6)		0.504	0.426		
							0.000	0.000	Disaggregate:	
UC(AFE)	0.487	0.492	0.557	0.632	VAR(JE-BIC)		0.482	0.494	Group	0.448
UC(X12)	0.458	0.483	0.526	0.656	VAR(JE-3)		0.436	0.437	Sub-group	0.499
UC(BIC)	0.428	0.491	0.527	0.655	VAR(JE-6)		0.472	0.424	Articles	0.645

Three month horizon

RW	1.272									
AR(BIC)	1.360	0.968	1.109	1.111	VAR(Ph)	1.526			Aggregate:	
AR(AIC)	1.342	0.977	1.085	1.093					Best Single	**0.880**
AR(BIC-SA)	1.168	0.915	1.042	1.030	VAR(BIC)		1.040	1.079	Simple Mean	0.900
AR(AIC-SA)	**1.165**	0.891	**1.013**	**0.996**	VAR(3)		**0.845**	0.933	Weighted Mean	0.909
SARIMA	1.234	0.980	1.153	1.133	VAR(6)		0.942	0.954		
									Disaggregate:	
UC(AFE)	1.291	1.169	1.191	1.362	VAR(JE-BIC)		1.105	1.135	Group	1.015
UC(X12)	1.338	1.183	1.253	1.403	VAR(JE-3)		0.881	**0.876**	Sub-group	1.060
UC(BIC)	1.240	1.199	1.207	1.275	VAR(JE-6)		0.956	0.905	Articles	1.138

Six month horizon

RW	2.242									
AR(BIC)	2.461	1.764	2.030	2.023	VAR(Ph)	3.039			Aggregate:	
AR(AIC)	2.456	1.771	1.942	1.939					Best Single	1.830
AR(BIC-SA)	**2.163**	1.630	1.961	1.954	VAR(BIC)		2.001	2.058	Simple Mean	**1.692**
AR(AIC-SA)	2.185	**1.612**	**1.880**	**1.852**	VAR(3)		**1.778**	1.677	Weighted Mean	1.726
SARIMA	2.407	1.775	2.060	2.132	VAR(6)		1.919	1.816		
									Disaggregate:	
UC(AFE)	2.399	2.226	2.121	2.375	VAR(JE-BIC)		2.169	2.212	Group	1.821
UC(X12)	2.480	2.201	2.352	2.649	VAR(JE-3)		1.889	1.792	Sub-group	1.785
UC(BIC)	2.204	2.225	2.170	2.236	VAR(JE-6)		1.951	1.830	Articles	1.762

Twelve month horizon

RW	4.083									
AR(BIC)	4.324	3.217	3.985	3.555	VAR(Ph)	5.434			Aggregate:	
AR(AIC)	4.323	3.225	3.743	3.334					Best Single	**3.205**
AR(BIC-SA)	4.269	3.200	3.868	3.566	VAR(BIC)		4.114	4.141	Simple Mean	3.217
AR(AIC-SA)	4.295	3.148	**3.742**	**3.326**	VAR(3)		**3.802**	3.343	Weighted Mean	3.206
SARIMA	**4.115**	**2.994**	3.829	3.714	VAR(6)		4.019	3.517		
									Disaggregate:	
UC(AFE)	4.377	4.029	4.148	4.278	VAR(JE-BIC)		4.328	4.385	Group	3.612
UC(X12)	4.529	4.076	4.620	4.742	VAR(JE-3)		3.879	3.690	Sub-group	3.626
UC(BIC)	4.265	3.955	3.988	3.692	VAR(JE-6)		3.953	3.548	Articles	3.374

Note: (1) given that the base sample for estimation runs from 1998(12) to 2005(12) the actual periods being forecasted when including combined forecasts for the respective horizons start at: 2006(4) for one month, 2006(6) for three months, 2006(9) for six months and 2007(3) for twelve months.
(2) the bolded numbers indicate the lowest RMSFE for the respective disaggregation level for a specific forecast horizon.
(3) the shaded area indicate those forecasts that are considered inferior, as defined in the text, by the Diebold-Mariano statistic.
(4) the lowest overall RMSFE for a specific horizon is enclosed in a rectangle.

Documentos de Trabajo Banco Central de Chile	Working Papers Central Bank of Chile
NÚMEROS ANTERIORES	PAST ISSUES

La serie de Documentos de Trabajo en versión PDF puede obtenerse gratis en la dirección electrónica: www.bcentral.cl/esp/estpub/estudios/dtbc. Existe la posibilidad de solicitar una copia impresa con un costo de $500 si es dentro de Chile y US$12 si es para fuera de Chile. Las solicitudes se pueden hacer por fax: (56-2) 6702231 o a través de correo electrónico: bcch@bcentral.cl.

Working Papers in PDF format can be downloaded free of charge from: www.bcentral.cl/eng/stdpub/studies/workingpaper. Printed versions can be ordered individually for US$12 per copy (for orders inside Chile the charge is Ch$500.) Orders can be placed by fax: (56-2) 6702231 or e-mail: bcch@bcentral.cl.

DTBC-544 Diciembre 2009
Towards an operational framework for financial stability: "fuzzy" measurement and its consequences
Claudio Borio y Mathias Drehmann

DTBC-543 Diciembre 2009
Institutional Bias Towards The Status Quo
Dalibor Eterovic

DTBC-542 Diciembre 2009
A Network Model of Super-Systemic Crises
Prasanna Gai y Sujit Kapadia

DTBC-541 Diciembre 2009
New Framework for Measuring and Managing Macrofinancial Risk and Financial Stability
Dale F. Gray, Robert C. Merton y Zvi Bodie

DTBC-540 Diciembre 2009
Households' Financial Vulnerability
Marcelo Fuenzalida y Jaime Ruiz-Tagle

DTBC-539 Diciembre 2009
Exchange Rate Volatility and Export Margins
Roberto Álvarez, Michael Doyle y Ricardo A. López

DTBC-538 Diciembre 2009
On The Role and Effects of IMF Seniority
Diego Saravia

DTBC-537 Diciembre 2009
The Balance Sheet Channel
Ethan Cohen-Cole y Enrique Martínez-García

DTBC-536 Diciembre 2009
Entry into Export Markets and Product Quality Differences
Roberto Álvarez y J. Rodrigo Fuentes

DTBC-535 Diciembre 2009
Análisis de Derechos Contingentes: Aplicación a Casas Comerciales
Rodrigo Alfaro y Natalia Gallardo

DTBC-534 Diciembre 2009
Efectos del Calentamiento Global sobre la Industria Pesquera
Carlos Medel

DTBC-533 Diciembre 2009
Hybrid Inflation Targeting Regimes
Carlos García, Jorge Restrepo y Scott Roger

DTBC-532 Diciembre 2009
Bank Credit and the 2008 Financial Crisis: A Cross-Country Comparison
Ari Aisen y Michael Franken

DTBC-531 Octubre 2009
La Curva De Rendimiento Bajo Nelson-Siegel
Rodrigo Alfaro

DTBC-530 Octubre 2009
The Long And The Short Of Emerging Market Debt
Luis Opazo, Claudio Raddatz y Sergio Schmukler