

Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon

Thiago Carlomagno Carlo & Emerson Fernandes Marçal

To cite this article: Thiago Carlomagno Carlo & Emerson Fernandes Marçal (2016): Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon, Applied Economics, DOI: [10.1080/00036846.2016.1167824](https://doi.org/10.1080/00036846.2016.1167824)

To link to this article: <http://dx.doi.org/10.1080/00036846.2016.1167824>



Published online: 09 Jun 2016.



Submit your article to this journal [↗](#)



Article views: 11



View related articles [↗](#)



View Crossmark data [↗](#)

Forecasting Brazilian inflation by its aggregate and disaggregated data: a test of predictive power by forecast horizon

Thiago Carlomagno Carlo ^a and Emerson Fernandes Marçal ^{b,c}

^aEconomics, UBS Brazil Broker Dealer, São Paulo, Brazil; ^bEconomics, São Paulo School of Economics, São Paulo, Brazil; ^cBusiness Department, Mackenzie Presbyterian University, São Paulo, Brazil

ABSTRACT

This work aims to compare the forecast efficiency of different types of methodologies applied to Brazilian consumer inflation (*Índice de Preços ao Consumidor Amplo*; IPCA). We will compare forecasting models using disaggregated and aggregated data from IPCA over 12 months ahead. We used IPCA in a monthly basis, over the period between January 1996 and March 2012. Out-of-sample analysis will be made through the period of January 2008 to March 2012. The disaggregated models were estimated by Seasonal Autoregressive Integrated Moving Average (SARIMA) and will have different levels of disaggregation from IPCA as groups and items, as well as disaggregation with more economic sense used by Brazilian Central Bank as: (1) services, monitored prices, food and industrials and (2) durables, non-durables, semi-durables, services and monitored prices. Aggregated models will be estimated by time series techniques as SARIMA, state-space structural models and Markov-switching. The forecasting accuracy among models will be made by the selection model procedure known as Model Confidence Set developed by Peter Hansen, Asger Lunde and James Nason. We were able to find evidence of forecast accuracy gains in models using more disaggregated rather than aggregate data.

KEYWORDS

Inflation; forecasting; ARIMA; state-space model; Markov-switching; model confidence set

JEL CLASSIFICATION

C53; E31; C52

1. Introduction

Inflation forecasting exercises play an important role within the empirical work of macroeconometrics. In the Brazilian case, the period of sharply rising inflation in the 1980s and first half of the 1990s led to the adoption of several stabilization plans, creating awareness of the benefits of price stability, which finally occurred with the Real Plan in 1994. Economic stabilization in recent historical terms also creates problems for the analysis of inflation dynamics in Brazil due to a relatively short time series, a common problem in developing countries. Both the problems of data availability and of abundance of institutional changes make it difficult to project the future through the past information, as Schwartzman (2006) observes. A second important step in economic stabilization came from the adoption of an inflation-targeting regime in June 1999. The adoption of inflation targeting became a fundamental difference in the development of econometric inflation models in Brazil. According to Chauvet and Morais (2010), to make the inflation-targeting

regime operational, the comparison of inflation forecasts with the stated goal becomes an important tool in monetary policy decision-making, given the lag of transmission mechanisms.

Inflation forecasting is not only important for the monetary authority but also of great importance to private agents, who try to understand and react to central bank behaviour. A more accurate inflation forecast is also very useful in fiscal policy, wage negotiations and also in the financial markets.

In this context, it is crucial to obtain inflation forecasts even more precise for several time horizons. Seeking the best forecast techniques, several approaches have been studied over time, including aggregated and disaggregated models, with linearity and non-linearity; structural models for time series; and the Phillips curve-based models are among some of the techniques studied.

In this work, we seek to identify the best estimation technique of Brazilian inflation among the univariate time series models. We use linear and non-linear models and aggregated and disaggregated

data, evaluating out-of-sample forecast performance for periods of 1–12 months ahead using the *Model Confidence Set* (MCS) by Hansen, Lunde, and Nason (2011).

Our results suggest that forecasts obtained from models using disaggregated data tend to outperform time series models based on aggregated data to forecast Brazilian consumer price index (*Índice de Preços ao Consumidor Amplo*; IPCA) for almost all horizons up to 12 months. These gains appear to be higher for short term horizons. Forecast combination does not improve forecast accuracy by a relevant amount. Brazilian Central Bank (BCB) official forecasts are outperformed by almost all econometric models in all horizons in our exercise.

The work is divided into five parts, besides this introduction. Section II is a literature review of inflation forecasting methodologies and model comparison methods. Section III will provide a brief description of the methodology and models that will be estimated, as well as a breakdown of the MCS. It will also describe how the data were treated, their peculiarities, and the breakdowns used. Section IV will present the results of the estimated models and comparisons of models. In Section V, we discuss the limitations and possible extensions of the work. Finally, the conclusions will be presented.

II. Literature review

Forecasting methods

One of the possibilities that could result in more accurate forecasts involves breakdown of the data. This alternative focuses on the decomposition of the index into several subcomponents, each with its associated weight; instead of forecasting the main variable, each of the subcomponents is projected individually to be reaggregated later. In this approach, the breakdown of the data can increase the accuracy of the forecast, insofar as each of its subcomponents may be modelled by their individual characteristics. Lütkepohl (1984) argued that if the data generating process is known, it is preferable to design multiple disaggregated time series and then aggregate them, rather than forecasting the aggregated series directly. However, in practice, the data generating process is unknown, and Lütkepohl presents evidence that given the variability of the model

specification and estimation, it may be preferable to design the aggregate variable directly. The loss of information due to aggregate series is also addressed by Kohn (1982), Palm and Nijman (1984), Rose (1977), and Tiao and Guttman (1980). It is not clear whether the aggregation of forecasts actually improves their accuracy, as the result also depends on the level of disaggregation of the series. According to Hendry and Hubrich (2006), the potential specification error in the forecast model with disaggregated variables is due to model selection, uncertainty in the estimation, as well as measurement errors and structural breaks. This explains why the theoretical results of the expected forecast are not confirmed in empirical studies. For example, Hubrich (2005) indicated that the aggregation of the forecasts of the components of inflation does not necessarily help with 12-month forecasts for the Eurozone. Duarte and Rua (2007) found an indirect relationship between forecast horizon and amount of information, that is, more disaggregated models are more efficient in short-term forecasts, while aggregate models are more efficient for longer horizons.

Another way to incorporate richer structures for modelling economic series is the use of forecast models of aggregate series with non-linear structure. Economic series, in general, are subject to changes in economic policies, a recurring problem in emerging countries, which may generate changes of regime. Possible structural changes in economic series have motivated the development of models that incorporate this particular feature. Many models of non-linear time series have been proposed in the literature, such as the bilinear model (Granger and Anderson (1978)); threshold autoregressive model (Tong (1978)); the state-space model used in engineering to represent a variety of physical processes, and its introduction and comparison of applications in engineering and econometrics performed by Mehra (1974); and the Markov-switching model proposed by Hamilton (1989).

A large number of studies on non-linearity in inflation investigate Friedman's hypothesis that high inflation rates lead to a higher variance of future inflation. Friedman (1977) suggests that uncertainty about the inflation regime may be the fundamental source of the positive relationship between inflation and volatility. The Markov-switching models have been especially designed to capture

this kind of phenomenon in the series. Studies in this regard are found in Evans and Wachtel (1993), Kim (1993), Bidarkota and McCulloch (1998), and Bidarkota (2001).

Another alternative study involves a combination of the models forecasts so that a forecast derived from the combination of several other forecasts may result in greater accuracy in the estimation. Various combinations are possible, but in many cases, a simple average of the forecasts can significantly improve the results, as noted by Clemen (1989). The first to study the combination of forecasts on economic situations were Reid (1968, 1969) and Bates and Granger (1969). Earlier studies in forecast combination were founded in the fields of statistics and psychology. To go deeper on the subject, Clemen (1989) made an extensive literature review of forecast combination in his work.

In the Brazilian literature, most alternative inflation forecasting models use the Phillips curve. Bogdanski, Tombini, and Werlang (2000) presented a small-size structural model to represent the monetary policy transmission channels, estimating a Phillips curve as part of the model. In an attempt to build a leading indicator to anticipate Brazilian inflation, Chauvet and Morais (2010) created a dynamic factor model using the Kalman Filter. Among the attempts to estimate the Phillips curve for Brazil, Schwartzman (2006) estimated a disaggregated Phillips curve, using the three-stage least square method and quarterly data for different samples. Areosa and Medeiros (2007) derived a structural model for inflation in an open economy, using a representation of the neo-Keynesian Phillips curve and a hybrid curve. Sachsida, Ribeiro, and Santos (2009) estimated a Phillips curve with quarterly data, in which they adopted a Markov-switching model, noting that the estimated coefficients are highly sensitive, and suggesting the inadequacy of the Phillips curve for explaining inflation dynamics in the Brazilian economy. Arruda, Ferreira, and Castelar (2011) compared inflation forecasts in Brazil with monthly data using linear and non-linear time series and the Phillips curve.

Comparison of predictive power of models

Not only are estimation techniques important, but so are comparisons of results from different forecast

methods. Several studies in the literature have addressed the problem of selecting the best model from a set of forecasting models. It is well known that increasing the number of variables in a regression improves the model fit within the sample – reducing, for example, the mean squared error – but the worst forecasts end up being from those models with the greatest number of variables. Engle and Brown (1986) compared the model selection procedures based on six information criteria with two testing procedures, showing that the criteria that more heavily penalizes the overparametrization of the models showed the best performance. Inoue and Kilian (2006) compared the selection procedures based on out-of-sample information criteria and evaluation by the mean squared forecast error (MSFE). A similar problem, known as multi-comparison with control, occurs when all objects are compared to a benchmark model, that is, when a base model is chosen whose performance is compared with all other models. Along these lines, a test for equal predictive ability was proposed by Diebold and Mariano (1995), consisting of a formal test between two competing forecast models under the null hypothesis of equivalence of accuracy. Harvey, Leybourne, and Newbold (1997) suggested a modification to Diebold and Mariano's test to best treat cases of small samples. Instead of testing equivalent forecast accuracy, White (2000) proposed a procedure to test the superior predictive ability (SPA) of the models. While the equivalence test has a single null hypothesis to be tested, the SPA test has multiple hypotheses. Hansen (2005) presented a new SPA test, following White's construction, but with a different statistic test, with a distribution dependent on the sample for the test of the null hypothesis. Finally, Hansen, Lunde, and Nason (2011) introduced the MCS, which is a multi-model comparison procedure without the need to set a benchmark model. The MCS is a methodology that determines the 'best' model or set of 'best' models within a confidence interval and has several advantages over previously proposed methods.

III. Modelling and forecasts comparison techniques

Model confidence set

The MCS is a methodology for comparing forecasts from a set of models, determining the 'best' model or

set of 'best' models within a confidence interval. The methodology that we follow was introduced by Hansen, Lunde, and Nason (2011).

According to Hansen, an attractive feature of the MCS is that it recognizes the limitations of the data, evaluating the sample information about the relative performance of the collection of models under consideration. If the sample is informative, the MCS will result in a single 'best' model. For a sample offering less informative data, it is difficult to distinguish the models, which may result in an MCS containing several or even all models. This feature differs from the existing selection methods that choose a single model without considering the existing information in the data.

Another advantage of MCS is that it is possible to comment on the significance that is valid in the traditional manner, a property that is not satisfied in the approach typically used to report the p -value of a multiple comparison of pairs of models. Thus, the advantage of MCS is that the procedure allows more than one model to be considered the 'best', resulting in a set of 'best' models.

Let us suppose that for a collection of models that compete with each other, M_0 , the procedure determines a set, M^* , which consists of the best set of forecast models. The MCS procedure is based on an equivalence test, δ_M , and a rule of elimination, e_M . The equivalence test, δ_M , is applied to the set of objects $M = M_0$. If δ_M is rejected, there is evidence that the models in M are not equally good, and thus the rule of elimination, e_M , is used to remove the object with the worst performance in M . This procedure is repeated sequentially until δ_M is accepted, and the MCS will be defined as the set of 'survival' models. The same level of significance is applied in all tests, which asymptotically ensures a significance level of α , $P(M^* \subset \hat{M}_{1-\alpha}^*) \geq 1 - \alpha$, and in the case where M^* consists of only one object, we have the stronger results that $\lim_{n \rightarrow \infty} P(M^* \subset \hat{M}_{1-\alpha}^*) = 1$, that is, a single object 'survives' on tests when n tends to infinity and is never removed, so its probability is equal to 1. The MCS also determines a p -value for each model (MCS p -value), which follows an intuitive form related to the p -value of the individual tests of equivalence between models. Briefly, for a given model

$i \in M_0$, the p -value of MCS, \hat{p}_i , is on the border where $i \in \hat{M}_{1-\alpha}^*$, if and only if $p_i \geq \alpha$. Thus, a model with low p -value is unlikely to be one of the models selected for the set of best models M^* , while a model with a high p -value probably will be part of M^* .

In this sense, there is no need to assume a particular model as being the 'best', nor is the null hypothesis test defined by a single model. Each model is treated equally in comparing and assessing only its out-of-sample predictive power. This is one of the most interesting features of the MCS that will be used in this work. The MCS was implemented using the OxMetrics 6.0 and MULCOM 2.0 package of Hansen and Lunde (2010).

SARIMA models¹

ARIMA model is a method widely discussed in the literature, initially introduced by Box and Jenkins (1976). The estimated model is as follows:

$$\begin{aligned} \phi(L)\varphi(L^s)(\Delta^d y_t - \beta_1 D_1 - \dots - \beta_{12} D_{12}) \\ = \theta(L)\delta(L^s)\varepsilon_t \end{aligned} \quad (1)$$

These models were estimated by the X-12 ARIMA software provided by the United States Census Bureau, using the procedure of automatic model selection derived from that used by TRAMO (Gomez and Maravall (2001)).

Structural time series models

According to Watson and Engle (1983), the use of unobservable variables in economics is widely accepted as a useful approach to describe economic phenomena. The general idea of the structural model for time series is that the series are a sum of components, not necessarily observed, such as tendency, seasonality and cycles, wherein each component evolves according to a particular dynamic. Structural models in a state-space representations offer a very interesting approach to forecasting, by allowing parameters to be stochastic. One can define a state-space model as follows:

$$y_t = Z_t \alpha_t + \varepsilon_t \quad \varepsilon_t \sim NI(0, H_t) \quad (2)$$

¹SARIMA means Seasonal Autoregressive Integrated Moving Average Model.

$$\alpha_{t+1} = T_t \alpha_t + R_t \eta_t \quad \eta_t \sim NI(0, Q_t) \quad (3)$$

where y_t is a vector of observations $p \times 1$; α_t is called the state vector, which is unobserved, and has dimension $m \times 1$; ε_t and η_t are independent error terms. The model estimation is done using the Kalman filter combined with maximum likelihood, in which the forecast error is minimized. The Kalman filter is composed by a set of equations to estimate recursively in time the mean and conditional variance of the state vector. The estimated state-space structural model was decomposed into trend, seasonality and stochastic cycle.

Markov-switching model

Markovian regime switching models are non-linear models of time series analysis in which there is a probability of transition between regimes connected to the immediately preceding period. These models are very useful in modelling series that present periods with distinct behaviours. For example, one should not expect a recessionary economy to behave in the same way as an economy in a period of expansion. Hamilton (1989) highlights this point. This non-linearity in economic series can be modelled from regime-switching models. We will use a Markov-switching autoregressive model, allowing two regimes, one for periods of high inflation and the other for periods of low inflation. In addition, our model will include dummies for deterministic seasonality. The estimated model will have the following form:

$$y_t = c_1 + \phi_1 y_{t-1} + \sum_{i=1}^{11} \theta_{1i} D_i + \varepsilon_{1t} \quad \text{if } s_t = 1 \quad (4)$$

and

$$y_t = c_2 + \phi_2 y_{t-1} + \sum_{i=1}^{11} \theta_{2i} D_i + \varepsilon_{2t} \quad \text{if } s_t = 2 \quad (5)$$

where s_t takes values $\{1, 2\}$, being the following transition probabilities between regimes defined:

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

in which p_{ij} can be defined as the probability of the regime i being followed by the regime j . The phenomenon is governed by a non-observed process in

which the probabilities model the transition from one conditional function to another.

IV. Data

The series used for the models are from IPCA – the official consumer price inflation measure in Brazil, released by the Brazilian Institute of Geography and Statistics (*Instituto Brasileiro de Geografia e Estatística*; IBGE) – starting in January 1996 and ending in March 2012. Currently the IPCA has the following subdivisions, in hierarchical order: Groups (9), Subgroups (19), Items (52) and Sub-items (365). However, the subdivisions of IPCA change over time, following the disclosure of the Consumer Expenditure Survey (*Pesquisa de Orçamentos Familiares*; POF) by IBGE, to more adequately reflect the changes in consumption patterns of the target population over the years. Altogether, throughout the IPCA reporting period, there were five changes in the basket of goods used, and within the review period of the study, there were three changes: in August 1999, July 2006 and January 2012.²

The most recent change to the weighting structure took place in January 2012, resulting in a reduction in the number of sub-items, from 384 to the present 365, due to the inclusion of 31 new sub-items and exclusion of 50, without changes in the other subdivisions: the 9 groups and 52 items remained unchanged. However, previous alterations showed more significant changes. It was decided to make the sets of groups and items of the IPCA compatible with the more current structure, using the IBGE's structure converter. Thus, we worked with 9 groups and 52 items in the forecast in more disaggregated levels, insofar as there were fewer changes over the period under review, making its compatibilization easier. Only two of the current items lack a structure corresponding to the period between January 1991 and July 1999, and therefore begin in August 1999. Another six items have changed more significantly but could be reconciled to complete the series. Table 1 contains more details.

Another way to analyse the IPCA in a disaggregated form is by using the item classification system of the BCB, which follows international standards recommended by United Nations. To this end, the BCB announces the items that comprise the following classifications: services, durable goods, nondurable goods,

²See IBGE (2013) available at Consumer Price Methodology.

Table 1. Compatibilization of structures of IPCA items.

Code	Item	January 1991 to July 1999	August/1999 to June 2006	From July 2006
3301	Maintenance and repair services	n.d.	Item 3301. Maintenance and repair services	Item 3301. Maintenance and repair services
6203	Health insurance	n.d.	Item 6203. Health insurance	Item 6203. Health insurance
7203	Photography and video recording	Aggregation of sub-items 7,201,011 Camera and 7,201,011 Photography accessories	Item 7203. Photography and video	Item 7203. Photography and video
7201	Recreation	Item Recreation less sub-items 7,201,011 Camera e 7,201,011 Photography accessories	Item 7201. Recreation	Item 7201. Recreation
8101	Courses	Aggregation of sub-items: 7,301,004. Textbooks; 7,301,006. Formal courses; 7,301,020. Books and technical journals; 7,301,021. Nursery	Item 8101. Courses less sub-item 8,101,014. Courses in general	Item 8101. Cursos
8104	Courses in general	Sub-item 7,301,007. Courses in general	Sub-item 8,101,014. Courses in general	Item 8104. Courses in general
8103	Stationery articles	Aggregation of sub-items: 7301002. Notebooks; 7,301,003. Stationery articles	Item 8103. Stationery articles	Item 8103. Stationery articles
9101	Communication	Item 5201. Communication	Item 9101. Communication	Item 9101. Communication

Source: Author.

semi-durable goods, tradable goods, non-tradable goods and monitored prices. With the change in the POF in January 2012, the BCB also made some changes to the previous classifications, and because the objective of this work is to test the best breakdowns for projected inflation, we decided to make the forecasts in both the old and the new classifications.

The main differences between the BCB's current and previous classification are in the groups of services and monitored prices. In the new classification, services include the subgroup 'Food away from home' (weight of 7.97% of the IPCA in January 2012) and the sub-item 'Airline tickets' (weight of 0.57% of the IPCA in January 2012). The group of monitored prices excluded the item 'Airline tickets' and also the sub-item 'Cell Phone' (weight of 1.52% of the IPCA in January 12). Thus, the new classification gives the Services group a weight of 33.72% for January 2012, whereas the old classification would have given 23.66% for the same period. Table 2 presents the changes of weighting for each instance of change in the structures of the IPCA because of POF changes.³ One of the classifications that draws the most attention is that of monitored prices, given the great weighting change over the period. At the beginning of the review period, in January 1996, the share of monitored prices was only 13.25%, reaching 31.29% in July 2006, but being greatly reduced in the last weighting change in January 2012 to 26.70%.

Table 2. IPCA weights of the BCB classifications in changes to the POF.

	Jan-96 ¹	Aug-99	Jul-06	Jan-12
Food and beverages	26.54	22.30	20.29	23.12
Services	30.43	24.33	23.06	23.66
Monitored prices	13.25	24.40	31.19	26.70
Industrials	29.79	28.97	25.46	26.52
IPCA	100.00	100.00	100.00	100.00
Food at home	18.88	17.15	13.36	15.15
Monitored prices new	13.22	26.61	29.31	24.51
Services new	37.53	29.48	31.38	33.72
Industrials new	30.37	26.76	25.94	26.62
IPCA	100.00	100.00	100.00	100.00
Services	30.43	24.33	23.06	23.66
Monitored prices	13.25	24.40	31.19	26.70
Durable goods	8.25	11.55	10.67	11.56
Non-durable goods	33.56	31.99	26.04	29.51
Semi-durable goods	14.51	7.72	9.03	8.56
IPCA	100.00	100.00	100.00	100.00
Tradables	43.92	39.97	33.44	36.13
Non-tradables	42.83	35.63	35.37	37.16
Monitored prices	13.25	24.40	31.19	26.70
IPCA	100.00	100.00	100.00	100.00

Source: IBGE; Central Bank of Brazil.

We will thus employ the IPCA forecasts with the following breakdowns, in which we will use the names defined in Table 3. With models at the aggregate level, for the purpose of comparison we will use the state-space methodology and Markov-switching, as discussed in the previous sections. An advantage of working with disaggregated prices, in these cases, is the ability to better control for them due to changes of weighting in the structures.

³The details of the BCBs classifications and the differences with the new classification can be obtained from the Box 'Updates of the Weighting Structures IPCA and INPC and of IPCA Classifications of the Quarterly Inflation' Report of December 2011. Available at Inflation Report.

Table 3. Estimated models and their breakdowns.

Name	Composition	Model
DD	Overall Index	Double Difference
Seasonal DD	Overall Index	Seasonal Double Difference
SARIMA_1	Overall Index	SARIMA
SARIMA_9	9 groups	SARIMA
SARIMA_52	52 items	SARIMA
SARIMA_4	Industrials old, Services old, Monitored old and group Food and Beverages	SARIMA
SARIMA_3	Tradables, Non-tradables, Monitored old	SARIMA
SARIMA_5	Durables, Semi-durables, Non-durables, Services old and Monitored old	SARIMA
SARIMA_4N	Industrials new, Services new, Monitored new and subgroup Food at home	SARIMA
BCB_focus	Brazilian Central Bank Survey – FOCUS	-
BCB_QIR_mkt	Brazilian Central Bank Forecast using FOCUS scenario for interest rate	-
BCB_QIR_ref	Brazilian Central Bank Forecast using BCB scenario	-
MS	Overall Index	Markov-Switching
STSM	Overall Index	Structural Time Series Model

Source: Authors.

V. Results

Evaluation of the models' performance was conducted using the out-of-sample results for forecasts for up to 12 months forward. All series and models were estimated recursively, starting in January 2008 and ending in March 2012. This means that with every new piece of information added, the models and all their parameters were re-estimated to always obtain the 'best' model, conditional on information at that time. In this sense, the sample periods for each estimation varies according to the final sample, that is, for evaluating the forecasts of the models one step ahead the sample is 51 months, while for forecast 12 months ahead the sample is 40. We preferred to use the full sample for each forecast horizon, because the more informative is the series, the better the results found in MCS.

For the disaggregated models, their subcomponents were reaggregated using the weightings corresponding to the IPCA estimated for the first month, while for the remaining months ahead, these were estimated based on the result of the forecasts in the previous period.

The loss function used to compare models was the MSFE which can be defined as the average difference between the estimated value and the actual squared. The advantage of this measure is that the larger projection deviations receive the highest penalty. It should be reiterated that the MSFE was calculated

upon the number index and not in relation to rate of change so that the results which will be presented later have relatively larger absolute values than if compared in percentage terms.

Initial analysis suggests better MSFE performance on more disaggregated models, while aggregated models showed the worst results. Lower MSFEs were found in the models with the highest level of disaggregation; the model made by forecasting the 52 IPCA items (SARIMA_52) is the one with the lowest MSFE, followed by 9 group subdivisions (SARIMA_9). The third smallest MSFE was the IPCA was divided into four groups: Food and beverages, industrials, services and monitored prices (SARIMA_4). The largest MSFE's were found in the aggregate models in the three forecast methods used (SARIMA, Markov-switching and structural model). Graphically, the lower performance of the aggregated models compared disaggregated models becomes even clearer (Figures 1 and 2), but the model estimated by state-space methodology presents the worst results among all models.

We also performed combinations of estimated models, whenever it proved possible that a combination of forecasts resulted in increased accuracy in terms of the loss function. To select the possible combinations, we tested using regressions, following the idea formalized and extended by Chong and Hendry (1986), wherein the effective value is regressed in function of the forecasts obtained on the basis of the estimated models plus a constant, and the estimated coefficients cannot be zero nor one, so that one model overlaps the other. Any value other than that shows that the joint forecasts provide useful information about the actual value. The weightings for the combination were obtained by regression of the estimates of the values achieved without constant, under the restriction that the sum of the weightings equals one. As shown by Granger and Ramanathan (1984), this method returns the optimal variance-covariance outcome. The weights were estimated for all time horizons in order to always obtain the best combination of templates in each period.

We found five possible combinations that resulted in improved performance of the combined model compared to the results for each of the individual models. Among the aggregate models, the combination of the projections of the model estimated by

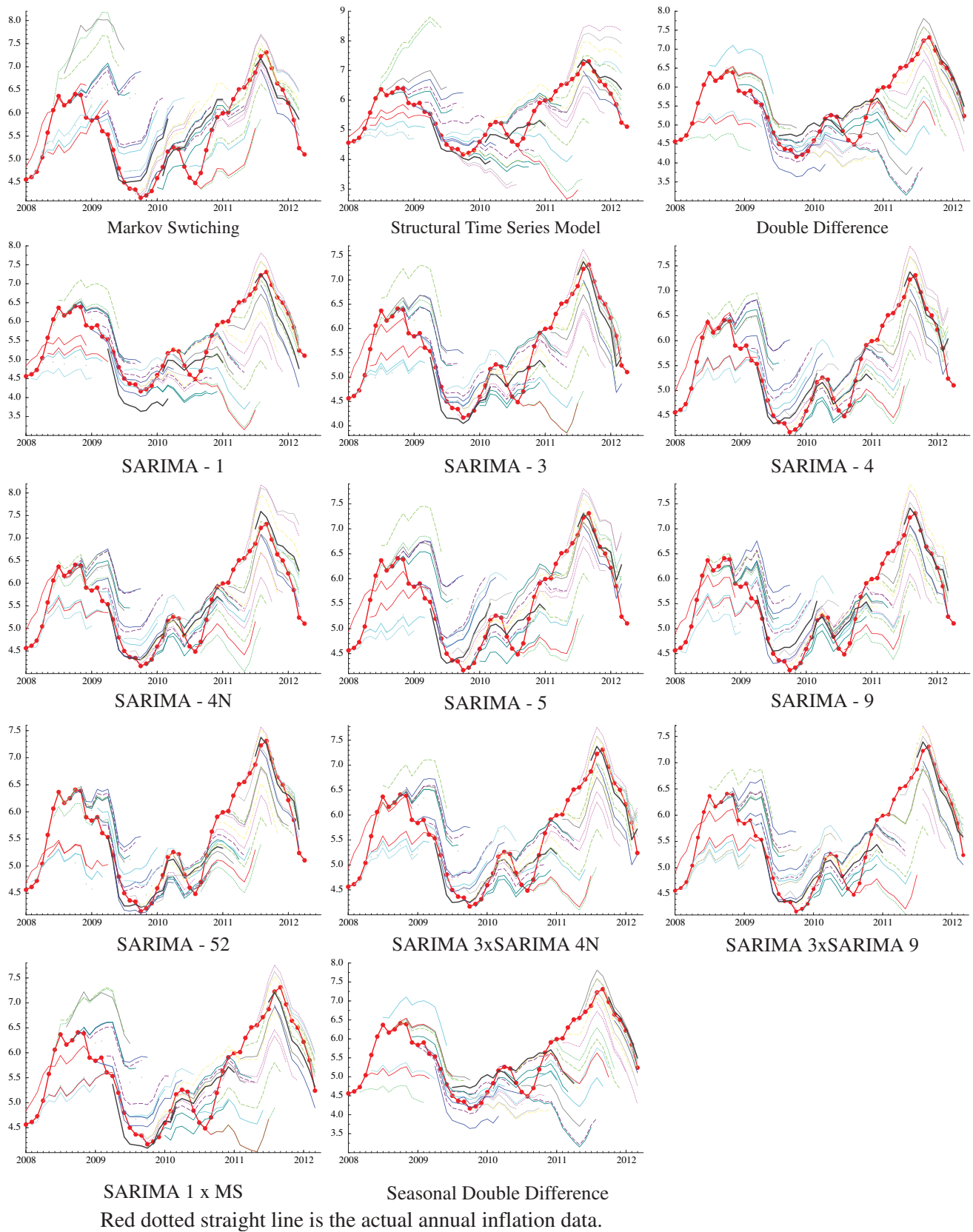


Figure 1. Consumer price index – actual versus predicted values – month over same month of previous year.

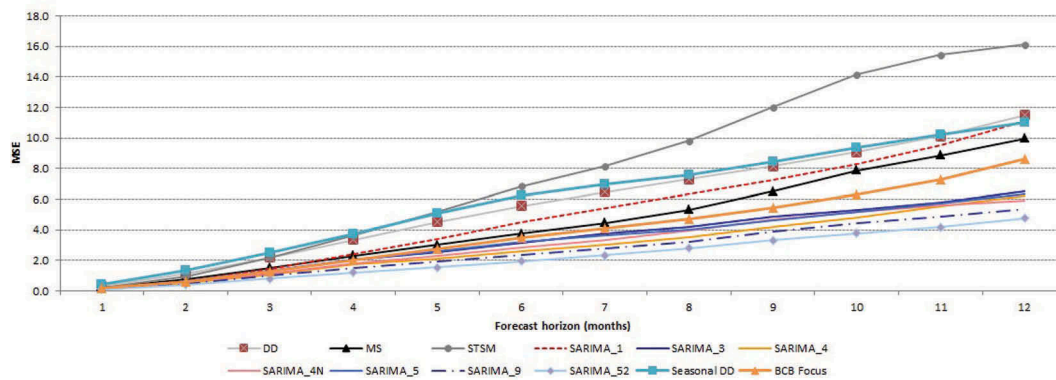


Figure 2. Mean square forecast errors of the estimated models by forecast horizon.

SARIMA and Markov-switching showed great improvement in performance. Among the disaggregated models, four possible combinations were found, but two of them showed no improvement on forecasts for all horizons. The combination of the model disaggregated into four groups by the new classification with those disaggregated into five or three groups showed improvement in all

horizons, while the combination of the model disaggregated in nine groups showed gains at longer horizons, albeit slight. The combination of the model disaggregated in nine groups also showed improvement when combined with the model disaggregated into three groups for the first 3 months and in the last 5 months. The improvements in the combined results can be better observed in Figure 3.

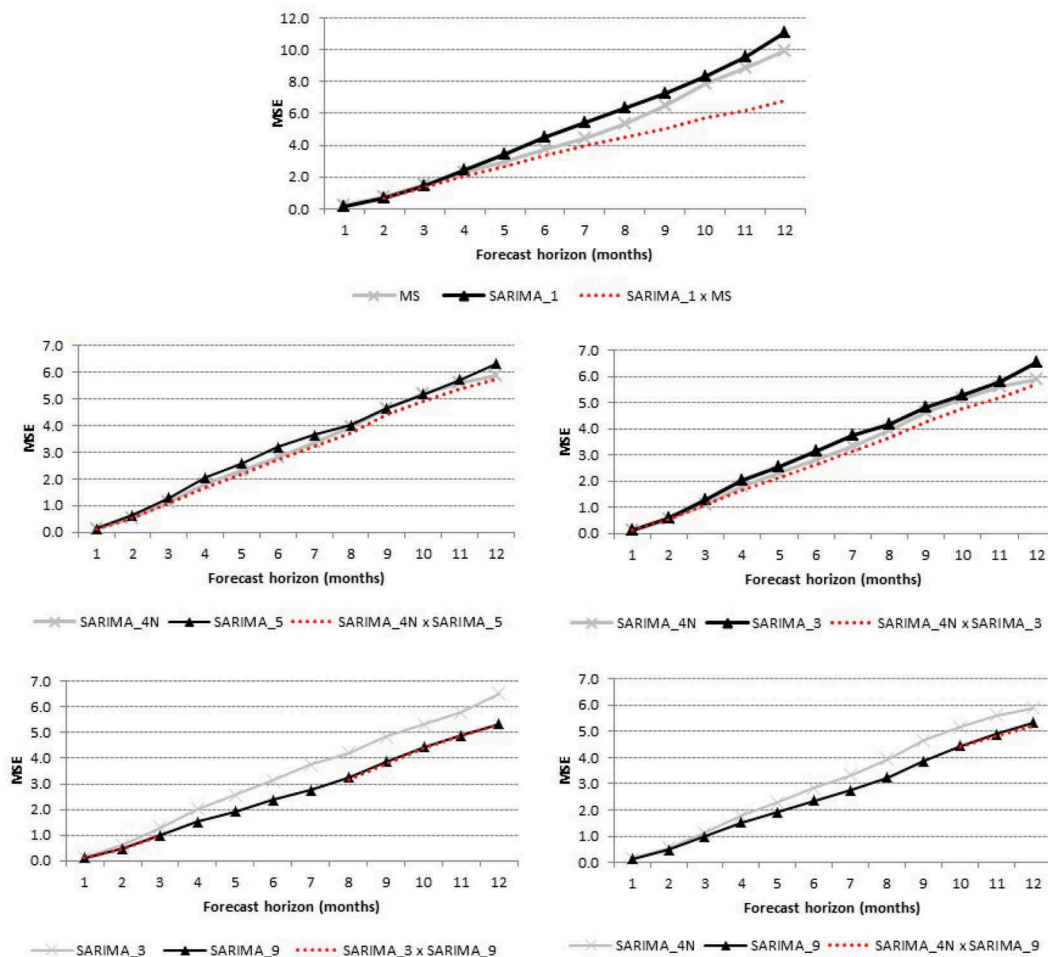


Figure 3. Mean square forecast errors of the combined models estimated by forecast horizon.

Table 4. Mean square forecast error of the models for the forecast horizon.

	Forecast horizon (months)											
	1	2	3	4	5	6	7	8	9	10	11	12
DD	0.330	1.128	2.185	3.360	4.511	5.560	6.466	7.294	8.144	9.069	10.103	11.485
MS	0.219	0.782	1.516	2.307	3.007	3.755	4.439	5.310	6.508	7.881	8.867	9.964
STSM	0.216	0.977	2.183	3.625	5.170	6.856	8.166	9.821	12.010	14.162	15.438	16.121
Seasonal DD	0.412	1.331	2.489	3.738	5.068	6.229	7.003	7.620	8.471	9.370	10.256	11.025
BCB Focus	0.151	0.622	1.279	1.999	2.764	3.496	4.116	4.702	5.449	6.308	7.309	8.620
SARIMA_1	0.146	0.662	1.461	2.422	3.409	4.469	5.423	6.342	7.231	8.298	9.529	11.089
SARIMA_3	0.140	0.613	1.306	2.021	2.559	3.132	3.742	4.193	4.840	5.311	5.803	6.526
SARIMA_4	0.134	0.541	1.050	1.732	2.085	2.582	3.055	3.497	4.172	4.798	5.535	6.237
SARIMA_4N	0.143	0.553	1.127	1.790	2.293	2.838	3.341	3.941	4.655	5.171	5.601	5.893
SARIMA_5	0.143	0.633	1.303	2.062	2.583	3.200	3.646	4.010	4.641	5.170	5.725	6.307
SARIMA_9	0.128	0.482	0.988	1.517	1.928	2.365	2.764	3.240	3.876	4.439	4.880	5.331
SARIMA_52	0.098	0.408	0.811	1.201	1.541	1.953	2.352	2.788	3.312	3.766	4.173	4.763
SARIMA_1 × MS	0.164	0.652	1.331	2.035	2.661	3.385	3.936	4.465	5.008	5.691	6.152	6.793
SARIMA_4N × SARIMA_3	0.134	0.545	1.081	1.678	2.135	2.660	3.148	3.641	4.276	4.767	5.211	5.703
SARIMA_4N × SARIMA_5	0.136	0.550	1.093	1.697	2.171	2.722	3.212	3.717	4.397	4.924	5.382	5.734
SARIMA_4N × SARIMA_9	0.128	0.503	1.007	1.532	1.950	2.428	2.861	3.359	3.999	4.393	4.815	5.207
SARIMA_3 × SARIMA_9	0.121	0.480	0.965	1.607	2.030	2.508	2.971	3.151	3.799	4.403	4.866	5.291

The shaded areas and in bold denotes de minimum EQM.

Table 4 shows the MSFE for all models in forecasts made up to 12 months ahead. The models are ordered from the lowest level of disaggregation to greatest possible. At the end of the table are the results obtained by combining the projections which showed some improvement in terms of the loss function.

Results of MCS

However, only the MSFE analysis alone does not tell us whether the differences in the models are statistically significant or if the sample period is informative enough to define a better model. Thus, we perform statistical tests to see which models can statistically be considered the best based on the

MCS selection criteria. Table 5 shows all p -values resulting from the MCS and those models considered part of the set of 'best' models (M^*), with 90% and 75% probability, indicated with one and two asterisks respectively, as in Hansen, Lunde, and Nason (2011).

The results suggest that there is a gain in accuracy in forecasting models with a higher level of disaggregation because the IPCA projected in the 52 items appears on $\hat{M}_{90\%}^*$ in all the forecast horizons. In forecast horizons from 1 up to 7 months ahead, only models with disaggregated data appear to belong to the final set (M^*). Other evidence that also suggests a gain from disaggregating is that the model with the second highest degree of disaggregation, the IPCA projected in all nine groups,

Table 5. Model confidence set and p -value of the estimated models.

	Forecast horizon (months)											
	1	2	3	4	5	6	7	8	9	10	11	12
Sample (n)	51	50	49	48	47	46	45	44	43	42	41	40
Double difference	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Focus – Brazilian Central Bank	0.011	0.051	0.035	0.005	0.005	0.010	0.026	0.080	0.116*	0.105*	0.075	0.002
Seasonal double difference	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
MS	0.011	0.002	0.005	0.001	0.002	0.005	0.022	0.049	0.063	0.063	0.044	0.046
STSM	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
SARIMA_1	0.095	0.027	0.016	0.001	0.000	0.001	0.000	0.000	0.001	0.003	0.005	0.001
SARIMA_3	0.131*	0.042	0.022	0.003	0.038	0.062	0.126*	0.228*	0.275**	0.270**	0.233*	0.303**
SARIMA_4	0.149*	0.223*	0.216*	0.072	0.156*	0.160*	0.243*	0.341**	0.322**	0.308**	0.233*	0.216*
SARIMA_4N	0.046	0.223*	0.197*	0.072	0.099	0.106*	0.163*	0.273**	0.260**	0.263**	0.265**	0.517**
SARIMA_5	0.011	0.004	0.020	0.001	0.002	0.005	0.092	0.185*	0.227*	0.248*	0.233*	0.420**
SARIMA_9	0.149*	0.271**	0.216*	0.147*	0.184*	0.223*	0.294**	0.341**	0.322**	0.308**	0.307**	0.591**
SARIMA_52	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**	1.000**
SARIMA_1 × MS	0.011	0.006	0.023	0.002	0.001	0.005	0.022	0.178*	0.227*	0.238*	0.225*	0.216*
SARIMA_3 × SARIMA_9	0.153*	0.247*	0.197*	0.073	0.126*	0.160*	0.243*	0.341**	0.322**	0.308**	0.290**	0.541**
SARIMA_4N × SARIMA_3	0.149*	0.181*	0.159*	0.004	0.078	0.123*	0.199*	0.311**	0.322**	0.308**	0.270**	0.373**
SARIMA_4N × SARIMA_5	0.149*	0.051	0.144*	0.004	0.063	0.088	0.177*	0.309**	0.322**	0.308**	0.265**	0.303**
SARIMA_4N × SARIMA_9	0.149*	0.271**	0.216*	0.133*	0.158*	0.187*	0.264**	0.341**	0.322**	0.308**	0.307**	0.591**

The forecasts in 90% and 75% are identified by one and two asterisks, respectively.

The forecasts in $\hat{M}_{90\%}^*$ is identified by one asterisk and p -value in bold in the shaded area, while those in $\hat{M}_{75\%}^*$ identified by two asterisks and p -value in the shaded area.

Table 6. MCS results comparing SARIMA-52 with Brazilian Central Bank Forecasts.

	Forecast Horizon (quarters)							
	1	2	3	4	1	2	3	4
	MCS p -values				MSFE* 10^3			
BCB_focus	0.016	0.022	0.012	0.011	1.28923	3.57158	5.17361	8.35339
BCB_QIR_mkt	0.010	0.022	0.012	0.011	1.68938	5.02961	7.97065	11.45024
BCB_QIR_ref	0.009	0.022	0.011	0.011	1.70869	5.15959	8.0914	10.67439
SARIMA_52	1.000**	1.000**	1.000**	1.000**	0.68785	1.7274	2.76742	4.46715

The forecasts in 90% and 75% are identified by one and two asterisks, respectively.

appears on M^* in all forecast horizons. The aggregate models (MS, STSM and SARIMA_1) fail to appear in MCS, but the combination of the projections made by the MS and SARIMA_1 models resulted in better performance, making up part of the MCS in projections between 8 and 12 months ahead. The combinations of forecasts from models SARIMA_3 and SARIMA_4N showed an increase in p -value of MCS in relation to models that make up almost all forecast horizons, except for few horizons. The two combinations performed with the model SARIMA_9, the second most disaggregated level, showed p -values of MCS nearly equal to SARIMA_9. Finally, the double difference model and seasonal double difference models fail to appear in MCS for any forecast horizon as expected. Clements and Hendry (2001) discuss why these benchmarks may be difficult to beat in a context of structural change.

It is quite interesting to note that the gains made in the disaggregated models occur at all forecast horizons both in the short-term and in the longer spans (Table 5). It should be noted that the p -values found in the disaggregated models are higher, and the forecast on the highest level of disaggregation appears with a p -value equal to 1 in all 12 forecast horizons.

Forecasts from market analysts collected by BCB, FOCUS Survey, are outperformed by almost all time series models in all horizons.

Remember that all estimated models are univariate and multivariate models that can provide an additional gain. Analysis of multivariate models is one suggestion for future studies on the topic.

Comparing econometric forecasts with BCB official forecasts

Brazil is an inflation target country since 1999. BCB is obliged to release official forecasts for inflation over a time horizon. These inflation forecasts are available at a quarterly frequency, but our forecast

exercise was performed using monthly data. We use end of period data for each quarter to compare BCB forecasts with our best model (SARIMA_52) due to frequency differences.

The initial set of models for running MCS algorithm is composed by SARIMA_52 forecasts, FOCUS Survey forecasts and two different official forecasts released by BCB (reference scenario and market scenario).⁴

Table 6 shows the results of MCS procedure. For all forecast horizons up to four quarters, the final set contains only SARIMA_52 model suggesting that BCB forecasts can be improved, not only in terms of MSFE but also to be, at least, part of the final set. The performance of BCB forecasts is worse in average when compared to our SARIMA_52 benchmark. Even FOCUS Survey shows a better performance in terms of MSFE. Figure 4 shows time evolution of actual and fitted values for these models.

Our results contrast with Groen, Kapetanios, and Price (2009). Their results suggest that inflation forecasts from Bank of England outperform an ARIMA and an Exponential Threshold Autoregressive Model. Joutz and Stekler (2000) compare forecasts made by Federal Reserve Bank (FED) with private forecasts and an ARIMA econometric model. Their results suggest that, in general, FED's forecasts have lower MFSE compared to econometric model-based and private forecasts, but these differences are not statistically different.

VI. Possible extensions and limitations of the work

This study examines forecasts from a broader level of disaggregation of Brazilian inflation models, and for comparing the forecasting models we used a recent technique presented by Hansen, Lunde, and Nason (2011). Nevertheless, there remain some

⁴More information about Brazilian inflation target regime can be found at <http://www.bcb.gov.br/?INFLATION>

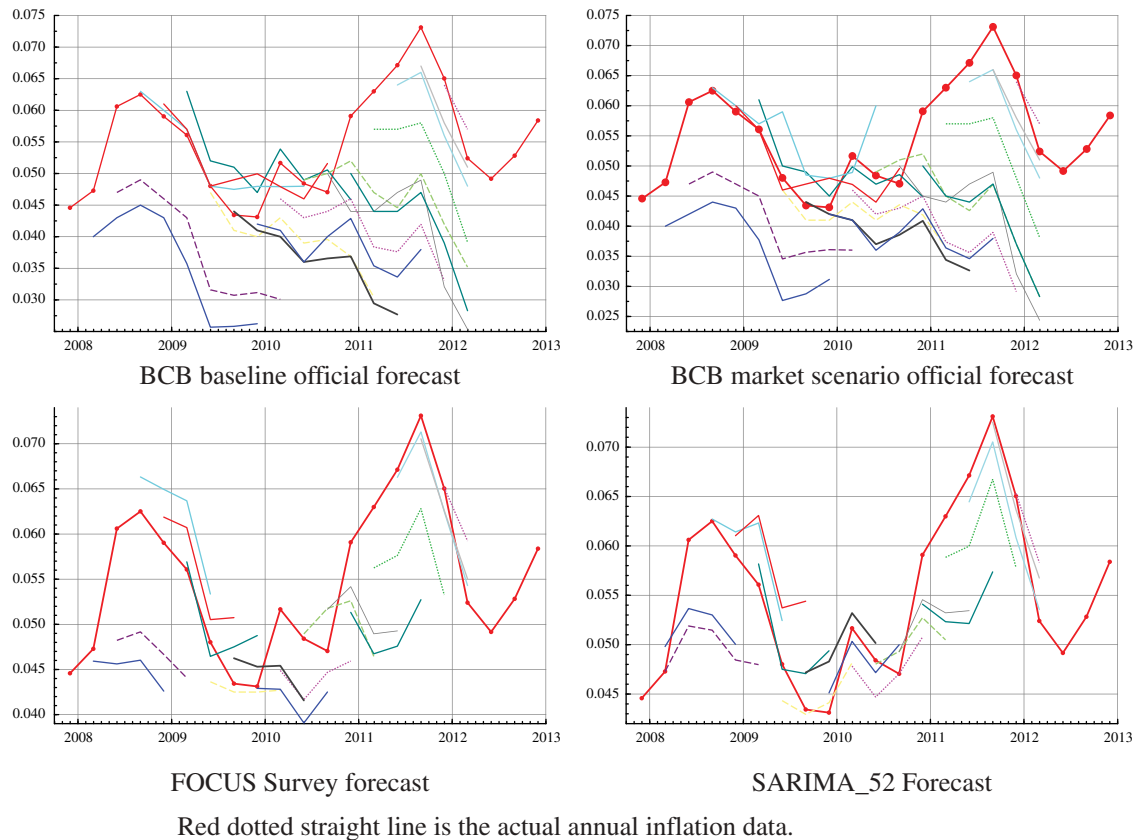


Figure 4. Actual and fitted – BCB official forecast, focus survey forecast and SARIMA-52 forecasts.

limitations as well as possibilities for various extensions by future studies. One limitation of the study lies in the relatively short sample in historical terms, especially when one considers that the forecasts began in January 2008, losing 51 data points, and also the series experienced some turbulence due to the change of exchange rate regime in January 1999 and the relevant currency devaluation in 2002. Another observation should be made regarding the forecast period, which incorporates several exogenous shocks such as the financial crisis marked by the bankruptcy of Lehman Brothers in September 2008; the sharp rise in commodity prices in 2010, reflecting the accommodative monetary policies in developed countries and the European debt crisis, with the first request for the IMF bailout of Greece in April 2010. The projection period also generated a relatively short sample MSFEs, 51 data points for forecasts 1 month ahead and 40 for data 12 months ahead, which could compromise the

evaluation due to the real difference in accuracy of the models studied. However, one of the advantages of MCS is to recognize the limitations of the data, and if these are not adequately explanatory a final set in MCS with all models would be generated. Thus, a possible extension of the work would accomplish the same forecasting exercise for different periods of the IPCA series in order to see whether similar results are obtained in other periods.

Another possible extension to the work would be to incorporate a wider range of forecasting models, including multivariate VAR models, Phillips curve aggregated and disaggregated models, and factorial models, for example. Regarding the model with the highest level of disaggregation, SARIMA_52, it would be possible to estimate each item with the model that best fits the series, not just the SARIMA class of models, which could lead to an even greater gain in disaggregation data.

There is a very large range of possible extensions to the work that could result in additional gains for research methodologies on Brazilian inflation forecasts.

VII. Final remarks

This paper presents a comparison of inflation forecasts for up to 12 months ahead, as measured by the IPCA, using univariate linear and non-linear time series, projecting rates from aggregated and disaggregated information. A comparison of the performance of the forecasts was conducted by MCS, introduced by Hansen, Lunde, and Nason (2011), by which it was found that there are significant gains in disaggregated inflation forecasts, more clearly for shorter forecasting horizons and less significantly for longer horizons. The analysis by the MSFE and p -value of MCS suggests that disaggregation offers a performance gain even for longer time horizons, although these have not been shown to be statistically significant. One factor that may have limited the results showing significance for longer horizons is the relatively short sample, both in-sample and for the out-of-sample forecasts. The non-linear models estimated for the aggregated IPCA, Markov-switching and structural model showed the worst results among the models. The worst forecast performance was presented by the structural model, suggesting that models with stochastic coefficients can generate results within the sample that are not repeated out-of-sample. The combination of the forecasts of aggregate models showed significant improvement in the performance of forecasts, something that was not as evident in the combination of the forecasts of disaggregated models.

The results presented in the previous section are somewhat similar to those found in some international work, which showed gains through disaggregation of forecasts, such as Duarte and Rua (2007), who found an inverse relationship between amount of information and forecast horizon: that is, more disaggregated models had better predictive power in shorter terms. Hendry and Hubrich (2012) also found gains through disaggregation for forecasting inflation in the Eurozone. Sachsida, Ribeiro, and Santos (2009) found the Markov-switching model poorly suited to forecast Brazilian quarterly inflation.

The work aims to contribute to the literature emphasizing that the disaggregated analysis of price indices can generate more accurate forecasts. The results are stronger for shorter forecasting horizons.

Acknowledgements

The authors acknowledge the comments and suggestions made by Paulo Picchetti, Pedro Luiz Valls Pereira, Marislei Nishijima and participants of 67th European Meeting of Econometric Society. All remaining errors are ours.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Thiago Carlomagno Carlo  <http://orcid.org/0000-0002-2914-5313>

Emerson Fernandes Marçal  <http://orcid.org/0000-0002-0841-5644>

References

- Areosa, W. D., and M. Medeiros. 2007. "Inflation Dynamics in Brazil: The Case of a Small Open Economy." *Brazilian Review of Econometrics* 27 (1): 131–166.
- Arruda, E. F., R. T. Ferreira, and I. Castelar. 2011. "Modelos Lineares e Não Lineares da Curva de Phillips Para Previsão da Taxa de Inflação no Brasil." *Revista Brasileira de Economia* 65 (3): 237–252.
- Bates, J. M., and C. W. Granger. 1969. "The Combination of Forecasts." *Operational Research Society* 20 (4): 451–468.
- Bidarkota, P. V. 2001. "Alternative Regime Switching Models for Forecasting Inflation." *Journal of Forecasting* 20 (1): 21–35. doi:10.1002/(ISSN)1099-131X.
- Bidarkota, P. V., and J. H. McCulloch. 1998. "Optimal Univariate Inflation Forecasting with Symmetric Stable Shocks." *Journal of Applied Econometrics* 13 (6): 659–670. doi:10.1002/(ISSN)1099-1255.
- Bogdanski, J., A. A. Tombini, and S. R. da C. Werlang. 2000. *Implementing Inflation Targeting in Brazil*. Working Paper Series 1. Brasília: Brazilian Central Bank.
- Box, G. E., and G. M. Jenkins. 1976. "Time Series Analysis, Control, and Forecasting." *San Francisco, CA: Holden Day* 3226 (3228): 10.
- Chauvet, M., and I. Morais. 2010. "Predicting Recessions in Brazil." 6th Colloquium on Modern Tools for Business Cycle Analysis: The Lessons from Global Economic Crisis, Luxembourg, September 26–29. EuroStat; European Commission.

- Chong, Y. Y., and D. F. Hendry. 1986. "Econometric Evaluation of Linear Macroeconomic Models." *The Review of Economic Studies* 53 (4): 671–690. doi:10.2307/2297611.
- Clemen, R. T. 1989. "Combining Forecasts: A Review and Annotated Bibliography." *International Journal of Forecasting* 5 (4): 559–583. doi:10.1016/0169-2070(89)90012-5.
- Clements, M. P., and D. F. Hendry. 2001. *Forecasting Non-Stationary Economic Time Series*. Cambridge, MA: MIT Press.
- Diebold, F. X., and R. S. Mariano. 1995. "Comparing Predictive Accuracy." *Journal of Business & Economic Statistics* 13 (3): 253–263.
- Duarte, C., and A. Rua. 2007. "Forecasting Inflation through a Bottom-Up Approach: How Bottom Is Bottom?" *Economic Modelling* 24 (6): 941–953. doi:10.1016/j.econmod.2007.03.004.
- Engle, R. F., and S. J. Brown. 1986. "Model Selection for Forecasting." *Applied Mathematics and Computation* 20 (3–4): 313–327. doi:10.1016/0096-3003(86)90009-3.
- Evans, M., and P. Wachtel. 1993. "Inflation Regimes and the Sources of Inflation Uncertainty." *Journal of Money, Credit and Banking* 25 (3): 475–511. doi:10.2307/2077719.
- Friedman, M. 1977. "Nobel Lecture: Inflation and Unemployment." *The Journal of Political Economy* 85: 451–472. doi:10.1086/260579.
- Gomez, V., and A. Maravall. 2001. "Automatic Modeling Methods for Univariate Series." In *A Course in Time Series Analysis*, 171–201. New York: John Wiley & Sons. Wiley Series in Probability and Statistics.
- Granger, C., and A. Anderson. 1978. *An Introduction to Bilinear Time Series Models*. Vandenhoeck und Ruprecht. Munich: Angewandte Statistik und Okonometrie.
- Granger, C. W., and R. Ramanathan. 1984. "Improved Methods of Combining Forecasts." *Journal of Forecasting* 3 (2): 197–204. doi:10.1002/for.3980030207.
- Groen, J. J., G. Kapetanios, and S. Price. 2009. "A Real Time Evaluation of Bank of England Forecasts of Inflation and Growth." *International Journal of Forecasting* 25 (1): 74–80. doi:10.1016/j.ijforecast.2008.09.005.
- Hamilton, J. D. 1989. "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle." *Econometrica* 57 (2): 357–384. doi:10.2307/1912559.
- Hansen, P. R. 2005. "A Test for Superior Predictive Ability." *Journal of Business & Economic Statistics* 23 (4): 365–380. doi:10.1198/073500105000000063.
- Hansen, P., and A. Lunde. 2010. *Mulcom 2.0 Manual – Econometric Toolkit for Multiple Comparison*. Aarhus: University of Aarhus.
- Hansen, P. R., A. Lunde, and J. M. Nason. 2011. "The Model Confidence Set." *Econometrica* 79 (2): 453–497. doi:10.3982/ECTA5771.
- Harvey, D., S. Leybourne, and P. Newbold. 1997. "Testing the Equality of Prediction Mean Squared Errors." *International Journal of Forecasting* 13 (2): 281–291. doi:10.1016/S0169-2070(96)00719-4.
- Hendry, D. F., and K. Hubrich. 2006. *Forecasting Economic Aggregates by Disaggregates*. Working Paper Series 586. Frankfurt: European Central Bank.
- Hendry, D. F., and K. Hubrich. 2012. "Combining Disaggregate Forecasts or Combining Disaggregate Information to Forecast an Aggregate." *Journal of Business & Economic Statistics* 29 (2): 216–227. doi:10.1198/jbes.2009.07112.
- Hubrich, K. 2005. "Forecasting Euro Area Inflation: Does Aggregating Forecasts by HICP Component Improve Forecast Accuracy?" *International Journal of Forecasting* 21 (1): 119–136. doi:10.1016/j.ijforecast.2004.04.005.
- IBGE. 2013. *Methodological Report Series*. Rio de Janeiro: Instituto Brasileiro de Geografia e Estatísticas.
- Inoue, A., and L. Kilian. 2006. "On the Selection of Forecasting Models." *Journal of Econometrics* 130 (2): 273–306. doi:10.1016/j.jeconom.2005.03.003.
- Joutz, F., and H. Stekler. 2000. "An Evaluation of the Predictions of the Federal Reserve." *International Journal of Forecasting* 16 (1): 17–38. doi:10.1016/S0169-2070(99)00046-1.
- Kim, C.-J. 1993. "Unobserved-Component Time Series Models with Markov-Switching Heteroscedasticity: Changes in Regime and the Link between Inflation Rates and Inflation Uncertainty." *Journal of Business & Economic Statistics* 11 (3): 341–349.
- Kohn, R. 1982. "When Is an Aggregate of a Time Series Efficiently Forecast by Its Past?" *Journal of Econometrics* 18 (3): 337–349. doi:10.1016/0304-4076(82)90087-2.
- Lütkepohl, H. 1984. "Forecasting Contemporaneously Aggregated Vector ARMA Processes." *Journal of Business & Economic Statistics* 2 (3): 201–214.
- Mehra, R. 1974. "Topics in Stochastic Control Theory Identification in Control and Econometrics: Similarities and Differences." In *Annals of Economic and Social Measurement*, Vol. 3, 21–48. Cambridge, MA: NBER.
- Palm, F. C., and Th. E. Nijman. 1984. "Missing Observations in the Dynamic Regression Model." *Econometrica* 52 (6): 1415–1435. doi:10.2307/1913513.
- Reid, D. J. 1968. "Combining Three Estimates of Gross Domestic Product." *Economica* 35 (140): 431–444. doi:10.2307/2552350.
- Reid, D. J. 1969. *A Comparative Study of Time Series Prediction Techniques on Economic Data*. Nottingham: University of Nottingham, Library Photographic Unit.
- Rose, D. E. 1977. "Forecasting Aggregates of Independent ARIMA Processes." *Journal of Econometrics* 5 (3): 323–345. doi:10.1016/0304-4076(77)90043-4.
- Sachsida, A., M. Ribeiro, and C. H. Santos. 2009. *A Curva de Phillips e a Experiência Brasileira*. IPEA Working Paper Series – 1429. Brasília: Instituto de Pesquisa Econômicas Aplicadas.

- Schwartzman, F. F. 2006. "Estimativa de Curva de Phillips Para o Brasil Com Preços Desagregados." *Economia Aplicada* 10 (1): 137–155. doi:[10.1590/S1413-80502006000100008](https://doi.org/10.1590/S1413-80502006000100008).
- Tiao, G. C., and I. Guttman. 1980. "Forecasting Contemporaneous Aggregates of Multiple Time Series." *Journal of Econometrics* 12 (2): 219–230. doi:[10.1016/0304-4076\(80\)90008-1](https://doi.org/10.1016/0304-4076(80)90008-1).
- Tong, H. 1978. *On a Threshold Model*. In *Pattern Recognition and Signal Processing*, edited by C. Chen, 575–586. Amsterdam: Sijthoff & Noordhoff. NATO ASI Series E: Applied Sc. (29), ISBN 9789028609785.
- Watson, M. W., and R. F. Engle. 1983. "Alternative Algorithms for the Estimation of Dynamic Factor, Mimic and Varying Coefficient Regression Models." *Journal of Econometrics* 23 (3): 385–400. doi:[10.1016/0304-4076\(83\)90066-0](https://doi.org/10.1016/0304-4076(83)90066-0).
- White, H. 2000. "A Reality Check for Data Snooping." *Econometrica* 68 (5): 1097–1126.