

# Inflationary Expectatives: Predicting, Persistence, and Disagreement\*

Sidney Martins Caetano\*\*  
Geraldo Edmundo Silva Júnior\*\*\*

**Abstract:** Because the expectations of inflation are considered a key-variable in the specification of macroeconomic models and due to the current discussion in literature on the disagreement concerning expectations, the present work aimed to study the divergence problem, starting from the statistical verification of the superior predicting ability of institutions that take part in the research (Focus- Market Report) of the Gerin/Bacen. The application of the SPA test, in short and long term analyses, suggested strength in the disagreement in the expectations of inflation, in which some models/institutions exert superior predicting ability. It is also possible to observe the presence of some degree of inertia in the dynamics of the expectations, in which the measures of persistence suggest sharp shock effects that diminish as time passes, in most of the cases analyzed.

**Keywords:** Inflation. Expectations. Monetary policy.

**Resumo:** Devido às expectativas de inflação serem consideradas uma variável-chave na especificação de modelos macroeconômicos e à nova discussão sobre desacordo em relação às mesmas, o presente trabalho teve o objetivo de analisar o problema de divergência, a partir da verificação estatística da habilidade preditiva superior das instituições que participam da pesquisa da Gerin/Bacen (Focus-Relatório de Mercado). A aplicação do teste SPA, em períodos considerados de curto e longo prazo, sugerem que alguns modelos/instituições exercem uma habilidade preditiva superior. Também é possível observar a presença de algum grau de inércia na dinâmica das expectativas, em que as medidas de persistência sugerem efeitos de choque acentuado que diminuem com o passar do tempo, na maioria dos casos analisados.

**Palavras-chave:** Inflação. Expectativas. Política monetária.

**JEL Classification:** D84; E31; E52.

---

\* The authors thank the Executive Management of Relationship with Investors (Gerin) of the Central Bank of Brazil for dataset and attention.

\*\* Department of Economics, Federal University of Juiz de Fora. E-mail: sidneymcaetano@gmail.com

\*\*\* Department of Economics, Federal University of Sao Carlos. E-mail: gedmundos@gmail.com

# 1 Introduction

In Brazil, after some years of high economic growth, the 80s were marked by the conjunction of two factors: increase in the inflation rate and decrease in the growth rate. This situation continued until the beginning of the 90s, leading the country to adopt seven stabilization plans in less than ten years. Nowadays, there is a certain consensus that price stability brings several advantages for self-sustained growth. For such, it is necessary to convince society that, in medium and long terms, higher inflation does not generate more growth. On the contrary, it creates an unfavorable environment for investments and penalizes the poorest population, promoting income concentration.

To avoid losing the price stability reached with the Plano Real in 1994, it was implemented the current inflation targeting regime. In this regime, the Central Bank of Brazil (Bacen) committed itself to act in a way to guarantee that the inflation observed be in line with a pre-established goal, guided by some basic principles implemented in the realm of the Monetary Policy Committee (Copom). Among these principles, it is pointed out the evaluation of inflation future trend. Besides, the Bacen and a greater number of central banks in the world (New Zealand, USA, Canada, Chile, United Kingdom, Sweden, Peru, Mexico, Brazil and Israel, among others) have dedicated special attention to the regular monitoring of the private sector expectations. With that, the monetary authorities have a crucial instrument of monetary policy management, since the expectations play a relevant role in the inflationary dynamics itself. It is worth to add that the main central banks of the world adopt price stability as the main objective of the monetary policy, either explicitly or implicitly, as is the case of the United States<sup>1</sup> (BANK OF ENGLAND, 1999; SVENSSON, 1997; WOODFORD, 2003).

Expectations are important because they are a component in the formation of prices, wages and economy contracts, besides determining the term structure of *ex-ante* real interest rates associated to a determined nominal interest curve, which, by its turn, affects the private decisions of consumption and investments. The monitoring of expectations, especially inflation expectations, is even more important in countries which follow the inflation targeting regimes. According to this regime, when conducting the monetary policy, the Central Bank searches to coordinate the formation of the private expectations.

In Brazil, the Bacen weekly issues the FOCUS - Market Report by Executive Management of Relationship with Investors (Gerin). This document brings market projections and the evolution of important macroeconomic variables, besides pointing out the expectations of the institutions with the greatest predicting ability,

---

1 Mankiw (2006).

the so-called Top 5 institutions<sup>2</sup>. The Gerin, created in April 1999, is also part of the structure of the monetary regime of inflation targeting, aiming to improve the communication between the Bacen and the private sector, focusing on domestic and foreign investors<sup>3</sup>. Then, in May 1999, the Research on Market Expectations began. Although it was created in May 1999, only with the increasing relevance of the expectations monitoring for the monetary policy, in November 2001, the *internet homepage* of the Market Expectation System was created, with the objective of speeding the collection process and minimizing the occurrence of misinformation (MARQUES; FACHADA; CAVALCANTI, 2003).

Currently, the research monitors the market expectations for different prices indexes, the Gross Domestic Product growth and industrial production, exchange rate, Selic interest rate, tax variables and the indicators of the external sector. However, a recurrent criticism to the Gerin expectation research refers to the fact that it is based in a sample basically composed of financial sector specialists. Therefore, it should not be a *proxy* of the economic agents expectations, but instead, of a small part out of the whole. In fact, according to the Central Bank of Brazil (BACEN, 2004), out of 104 informants, 84 were financial institutions (52 commercial banks, 23 investment funds, and 9 brokerage firms and securities), 14 consultancy companies and 6 non-financial companies and class entities. The Bacen argues that, in practice, companies and entities, regardless of the sector to which belong, present smaller capacity to respond to macroeconomic changes than the participants of the financial sector and consultancy companies, attributable to the nature of their professional activity, mainly in the presence of data of high frequency and in an environment of volatilities<sup>4</sup>.

If some economic agents linked to companies and entities do not respond in the same speed as the financial analysts do, then it is reasonable to imagine the possibility of the existence of divergences in expectations about macroeconomic variables; in the present case, the economic prices.

Some works in Brazil had the objective of verifying whether or not the Bacen prediction is similar to that of the market, (BACEN, 2003), and if time series models perform do better prediction than the Bacen does (LIMA; CÉSPEDES, 2003). The results of these authors showed that some simple autoregressive models are able to generate predictions more exact than those from the Bacen, and that the mar-

---

2 More information about the Central Bank System of Market Expectations can be found in Marques, Fachada, and Cavalcanti (2003).

3 The Gerin manages a list of thousands of e-mails, to which it sends the main records of the Bacen. Besides, it produces its own reports, all available at [www.bcb.gov.br](http://www.bcb.gov.br), including: Focus Market Report; Time Series of the Market Expectations: weekly updated spreadsheets, containing the daily data of the market expectations for the main economy variables since 2001; Top 5; "Frequently Asked Questions" Series; Weekly Review of Open Market.

4 Companies and organizations, independently of the sector to which they belong, must meet certain criteria of admission and regularity so that they can take part in the research on market expectations.

ket expectations are inflation predictions statistically similar to those of the Bacen. So, it is undeniable that the comparison among the estimates or expectations and the results should be an important determinant of the prediction capacity of the models considered, even if their structures are not specified. So, with the instruments used in the financial area, discussed in section 2, for the comparison of the predicting capacity of the models, it would be possible to verify important results in the supposition that some model would present superior predictive ability to the others. Among them:

- a) even with some consideration among the sector expectations, the superior predictive ability would restrain the capacity of the Bacen even in the specification of the Selic interest rate-target through the expectation divergence as Mankiw, Reis and Wolfers (2003);
- b) distortions caused by the phases of the economic cycles and the inflation path, either ascending or descending, would be observed;
- c) the consequences of prediction errors about variables, initially mentioned by Lima and Céspedes (2003), would be unknown, mainly the long term horizon.

In this case, it is important to verify statistically the superior predictive ability of the institutions that take part of the “Focus-Market Report”, spreading their expectations concerning the Brazilian official inflation, measured by the IPCA variation, as well as the alternatives available to the economic agents. For such, it is used the new *Superior Predictive Ability* (SPA) test of Hansen (2005), in which “models” are proposed as *benchmark*, for both short and long terms.

With the *random walk* model, representing the *benchmark*, it was investigated the inferiority of the prediction of the institutions selected in the presence of a singular model, as well as the problem of the disagreement in the inflation expectations, in case some agents adopt this model. As the others, it is verified if it is reasonable to consider, for example, that the Bacen Focus/Top5 daily surveys Bacen are good forecasters of the inflation rate in short and long terms and if corroborate results in literature. In part, the contribution of this article follows the reasoning of evaluating the superior predictive ability of the institutions in face of simple *benchmark* models, such as *random walk* and AR(p), but it is distinguished because it is a statistical evaluation of the problem and it contributes to the current discussion on the disagreement in the inflationary expectations, something that is not done by Lima and Céspedes (2003). It is also proposed a new TOP 5 (TOP 5\*), compared with the other models, including the Gerin/Bacen TOP 5, also, in other part, evalu-



ating the inertia degree in the dynamics of the inflation expectations from the ratio of variances, systematically used to measure shock persistence in the long run.

The present work presents the following steps: besides the present introduction, in Section 2, some theoretical foundations are discussed on the relevance of the expectations in economy; Section 3 presents the models of superior predictive ability in *snooping data* and the new test of superior predictive ability proposed by Hansen (2005); Section 4 shows the empiric results achieved, their interpretations and a discussion of the disagreement problem in expectations; and, finally, Section 5 closes with the conclusions.

## 2 Theoretical Foundations

For almost half a century, inflation expectations were as a key-variable in the specification of macroeconomic models. Started by Friedman (1968) and Phelps (1969), the expectations should explain the phenomenon of the stagflation. Years later, Lucas (1973) gave more attention to this variable as rational expectations in an expectations-augmented Phillips curve. Such models solve the problem of the perfect prediction in the models of the neoclassical intertemporal equilibrium. An important postulate on the role of the expectation divergence has been considered in the economic literature since the contribution of Lucas (1973). In the eighties, based on the work of Taylor (1980), Roberts (1995) suggested a New Keynesian Phillips Curve as a model, with the dominance of price rigidity, in which the current inflation would depend on the product and the expectations of the future inflation. However, in such model, some puzzles were detected, namely: in disinflation, there shall be expansion instead of recession; the persistence of inflation shall not be explained; and, finally, the gradual response to real variables shocks should not be supported by this model. Recently, however, Mankiw and Reis (2002) structured a model of the New Keynesian Phillips Curve, in which the rigidity would be of information, instead of prices, causing a slow adjustment in prices.

Based on the informational rigidity, Mankiw, Reis and Wolfers (2003) suggested that the inflation expectation disagreement as a key-variable in the macroeconomic dynamics and, for such, they considered three different sources of expectation statistics, that are: *Michigan Survey Consumer Attitudes and Behavior*, *Livingston Survey* and *Survey of Professional Forecasters*. Based on the analysis of the information, they concluded that the divergence would enlarge with the increase of the inflation and that this divergence would tend to augment with a sudden inflation change, regardless of the movement direction. Still, the expectation dispersion of the intra-group inflation should enlarge the variance; and finally, that there would not be any clear relation between the divergence and real economic variables. In other words, none of the models clarifies or structures the expectation divergence

as a macroeconomic variable. Mankiw, Reis and Wolfers (2003) have proved the diversity and the expectation divergence for the United States, but in the economic literature, the adoption of an instrument that statistically proves the expectation divergence has not been proved studied.

So, the debate was open around the predictive capacity of the models, without a specific goal. Important results have been observed in literature. For Branch (2004), there is much inflation predicting models with the traditional approach of limited rationality that considers agents would use an adaptive rule of their predictions. In other words, the predicting ability of a specific model upon the set of alternative models would make it possible to corroborate the hypotheses of divergence and expectational heterogeneity, given different costs of achievement of relevant information on the prediction of such models, since the agents or groups do not have any econometrists on duty. Besides, in the Adaptively Rational Equilibrium Dynamics (ARED) models, the agents would choose a prediction function among a set of alternatives, with an increasing cost of ordination due to the sophistication of the models. Then, the probability of a predictor selection would depend on the net benefits in comparison to the respective models. Branch (2004) widened the analysis of such models when he considered inflation adaptive expectations 12 steps ahead, in which the optimality of a predictor would depend on the agent that chooses it, thus suggesting a heterogeneous rationality in the selection process and, with that, an expectational divergence.

According to Döpke *et al.* (2006), the informational rigidity would be observed for countries like Germany, France, Italy and the United Kingdom. For Italy, the readjustment of the expectations would occur in only six months and one year gap for the other countries, thus characterizing the informational rigidity. So, the superiority and the consistency of the informational rigidity model would have solved the puzzles around the inflation persistence and the reductions costs. However, the comparison of the predictive power of the price rigidity models, informational rigidity and price rigidity with an indexation mechanism, implemented by Korenok (2008) and Korenok and Swanson (2007), restricts the results of Döpke *et al.* (2006), because the dominance would be of the price rigidity models, in periods of increasing inflation.

Due to such possibility, the consideration of the comparison of the inflation prediction models is important, due to the expectational divergence phenomena, the predictive ability, and the problem of inflation persistence as well. In the scope of the Corporate Finances, the empiric literature has presented relevant methods and results likely to be used in the Macroeconomics field.

Using the instrument started by Lo and Mackinlay (1990), who sought to quantify the bias of the structures of the *Capital Asset Price Model* (CAPM) prediction models and the *Asset Price Theory* (APT) model for the *snooping data* of

the North American financial market, Hansen and Lunde (2005) confronted 330 models of the ARCH structure for the exchange rate German Mark/Dollar and the return of the actions of the IBM and established the GARCH (1, 1) model as *benchmark*. After carrying out the superior predictive ability test of Hansen (2005) and the White (2000) test, they concluded that the GARCH (1, 1) model is superior, at least for the exchange rate. Koopman, Jungbacker, and Hol (2005) compared the models of unobserved components (UC), ARFIMA long memory models, stochastic volatility models (SV) and autoregressive conditional heteroskedasticity models (GARCH) and, based on the daily data of approximately 7 years for the index of Standard & Poor's 100 (S&P 100), the long memory models were considered superior to the others.

In Brazil, using the superior predictive ability test of White (2000), the *Reality Check for Data Snooping*, Saffi (2003) demonstrated that none of the 14630 strategies of technical analysis would generate statistically significant returns because of the consideration of the *snooping* data effects. So, the weak hypothesis of market efficiency should be proven. Concerns with the Bacen predicting capacity was suggested by Lima and Céspedes (2003) when the authors compared with the following prediction models: AR(2), AR(4), ARIMA (0,1,2) and ARIMA (2,1,4) to the FOCUS and TOP 5 predictions and concluded that the accuracy of the Bacen about long horizons of prediction should be questioned.

Although the authors question the authority capacity, especially concerning long term horizons, since they would be considered in the specification of the alterations in the Selic interest rate-target, the authors do not consider the structural diversity of the FOCUS results. They do not presented any test with enough statistical rigor so that a superior prediction model could be presented.

### 3 Superior Predictive Ability Models in Snoop Data

The origin of the concern about the use of the same financial dataset for the generation of prediction models, although distinguished in their estimates, reminds the work of Cowles (1933), entitled "Can Stock Market Forecasters Forecast?".

This author analyzed the proposals spread by 45 brokerages from January 1928 to July 1932, during the Great Depression period. His motivation was to demonstrate that the same dataset on prices and variations of the assets would generate different proposals for clients of those agencies. Then, the genesis of superior predictive ability issue would be the joint treatment of prediction models, the use of enclosing prediction models to test efficiency, as well as equal predictive ability (WEST, 1996).

### 3.1 The Bias Problem in *Snooping Data*

A treatment based on a similar idea was suggested by White (2000) and defined as bias in *snooping data*. The event of *snooping data*, according to this author, would take place when a set of data was used more than once with the objective of inferring or selecting a model among a set of alternatives. Such practice would be dangerous for researchers and for market analysts, if careful comparisons were not performed. In the scope of the equal predictive ability test, there would be concerns on the binary comparison between the *benchmark* model, considered the most efficient, and an alternative model, represented by a null (simple) hypothesis of equal predictive ability. In the superior predictive ability (SPA) test, there would be a multiple comparison between the alternative models, composed in a nullity hypothesis that represents inferiority to the *benchmark* model, and the alternative hypothesis, superior to the *benchmark*.

For SPA test, the following elements would be necessary: alternative models different from the *benchmark*; a decision rule based on composed hypotheses; and restrictions on the distribution, according to the fundamental asymptotic properties for the specification of optimal tests, with the possibility of distributional variations. White (2000) specified a test for the superior predictive ability that would be more powerful and less sensitive to poor and irrelevant alternatives. In other words, additional hypothetical information would have little effect on the predictive capacity of the structure considered. In the structure *Reality Check*, White (2000),  $m$  alternative predictions would be compared to a basic model, and the predictive ability of each one of the  $m$  alternatives would be considered from an expected loss function, as specified by the equation 1:

$$\hat{f}_{k,t+1} = \log L_k(Y_{t+1}, X_{k,t+1}, \hat{\beta}_{k,t}) - \log L_0(Y_{t+1}, X_{0,t+1}, \hat{\beta}_{0,t}), k = 1, \dots, l \quad (1)$$

were:  $L_k(\cdot)$  = log-predictive likelihood of the  $k$  ( $k = 1, \dots, l$ ) model, alternatives to the *benchmark*.

The test, therefore, should impose the following joint decision rule among the alternatives to *benchmark* model and the respective model, namely:

$$H_0 : \max_{k=1, \dots, l} E(f_k^*) \leq 0 \quad (2)$$



In the hypothesis of nullity establishes that the alternatives would not be superior to the *benchmark* model. With this framework, the basic alternative would be a limit test like Bonferroni<sup>5</sup>, that could be specified as:

$$\text{Prob}\left\{\bigcup_{i=1}^l \left(P_i \leq \frac{\alpha}{l}\right)\right\} \leq \alpha, \quad (0 \leq \alpha \leq 1) \quad (3)$$

The equation (3) should represent the *inequality of Bonferroni*, built from the n-statistics of the p-values corresponding to  $P_1, \dots, P_l$ , that would represent the composition  $H_1, \dots, H_l$ , for which  $H_0 = \{H_1, \dots, H_l\}$  would be rejected if some of the p-values were lower than  $\alpha/l$ , in which  $\alpha$  represents the probability of the occurrence of the Type I Error to reject the nullity hypothesis when it is true, in independent tests. In general, the inequality of *Bonferroni* must be used in multiple tests, to determine a superior limit for the  $\alpha$  significance level, probability of making the Type I Error. In spite of the simplicity of the test, due to its tractability, it would be limited in power if highly correlated individual tests were considered, which seems to be the case of the data on predictions of some macroeconomic variables like inflation, exchange rate, growth, among others. If very closed models were considered in joint test, in other words. So, White (2000) was concerned about the settle of an optimal test based in the restrict maximization in which a smaller error of the Type II, could accept  $H_0$  when it is false, given the test power  $(1 - \beta)$ . In an unbiased test, according to Spanos (1989), in which  $H_0$  would be really true, its power should be equivalent to its size. Therefore, Sullivan, Timmermann, and White (2003) proposed two alternatives of simulation for the achievement of parameters for the test: Monte Carlo simulation and *bootstrap* procedure, but White (2000) proved that the distribution should be approximately asymptotic normal.

### 3.2 Superior Predictive Ability (SPA) – Hansen's Proposal

Hansen (2005) established an alternative for White (2000), employing a different statistical test, based on a dependent sample distribution, under a null hypothesis. Then, the contributions of Hansen (2005) would be: a theoretical analysis of the test problem, exposing some of their important aspects, in which the inclusion of poor and irrelevant alternatives could be manipulated, thus altering the results, for the benefit of one of the models; an detailed explanation of the bootstrapping procedure implementation; and, an application for the North-American inflation, using the *random walk* model as *benchmark*. The SPA test, proposed by Hansen (2005), is a test that may be used to compare with the performances of two or more forecasting models. The forecasts are evaluated using a pre-specified

5 Consult Simes (1986) for a modified version of the test in the ordering of the probabilistic values.

loss function, and the “best” model forecast indicates the smaller expected loss. Two loss functions were pre-suggested by the author: “the mean squared error (mse) and mean absolute deviation (mad)”. In analytical terms, the application of the SPA test, Hansen and Lunde (2007), would start with the specification of an expected loss function.

Let  $L(Y_t, \hat{Y}_t)$  denote the loss under the supposition that the institution/model has made the prediction,  $\hat{Y}_t$ , when, in fact, the realized value was  $Y_t$ ; then the performance of model  $k$ , relative to the *benchmark* model (at time  $t$ ), can be defined as:

$$X_k(t) = L(Y_t, \hat{Y}_{0t}) - L(Y_t, \hat{Y}_{kt}), \quad k = 1, \dots, l; \quad t = 1, \dots, n. \quad (4)$$

The issue of interest is to verify if any of the models are better than the *benchmark* model. For such, it is tested the hypothesis that the *benchmark* model is the best forecasting model among those analyzed in the test. This hypothesis can be expressed parametrically as follows:

$$H_0 : \mu_k = E[X_k(t)] \leq 0, \text{ for all } k = 1, \dots, l. \quad (5)$$

For notational convenience, the  $l$ -dimension vector  $\mu$  is defined by

$$\mu = \begin{pmatrix} \mu_1 \\ \vdots \\ \mu_l \end{pmatrix} = E \begin{pmatrix} X_1(t) \\ \vdots \\ X_l(t) \end{pmatrix}. \quad (6)$$

Provided that a positive value of  $\mu_k$ , corresponding to the model  $k$ , be better than the *benchmark*, it is possible to test the hypothesis  $H_0: \mu \leq 0$ , for  $k = 1, \dots, l$ . The equivalent vector formula would be  $H_0: \mu \leq 0$ .

A way of testing this hypothesis is to consider the test statistic,

$$T_n^{sm} = \max_k \frac{n^{1/2} \overline{X}_k}{\hat{\sigma}_k} \quad (7)$$

were:

$\overline{X}_k = \frac{1}{n} \sum_{t=1}^n X_k(t)$ ,  $\hat{\sigma}_k^2 = \hat{\text{var}}(n^{1/2} \overline{X}_k)$  is a consistent estimator of the asymptotic variance,  $\hat{\sigma}_k^2 = \lim_{n \rightarrow \infty} \text{var}(n^{1/2} \overline{X}_k)$ . The last one is calculated using the *bootstrap* method. The

term “sm” refers to the *standardized maximum*. Under the regularity conditions it holds that

$$n^{-\frac{1}{2}} T_n^{sm} = \max_k \frac{\bar{X}_k}{\hat{\sigma}_k} \xrightarrow{p} \max_k \frac{\mu_k}{\sigma_k} \quad (8)$$

which will be greater than zero if and only if  $\mu_k > 0$ , for some  $k$ . Thus, it is possible to test  $H_0$  using the test statistic  $T_n^{sm}$ . The problem is to derive the distribution of  $T_n^{sm}$ , under the presupposition of a true null hypothesis. To test multiple inequalities is more complicated than to test equalities (or a single inequality) because the distribution is not unique under the null hypothesis. However, a consistent estimate for the  $p$ -value can be achieved by the SPA program of Hansen (2005), as well as an upper and a lower bound.

So, the structures of the hypothesis to be tested by the SPA model would be based on the statistics of the expressions (5) and (6), respectively:

$$T_n^{RC} = \max_k \sqrt{n} \bar{X}_k \quad (9)$$

$$T_n^{SPA} = \max_k \frac{\sqrt{n} \bar{X}_k}{\hat{\sigma}_k} \quad (10)$$

Were:  
 $\bar{X}_k = \frac{1}{n} \sum_{t=1}^n X_k(t)$ , and  $\hat{\sigma}_k^2 = \text{var} \left( n^{1/2} \bar{X}_k \right)$ .

The difference among the statistics (9), for the *Reality Check*, White (2000), and (10), for the SPA case, Hansen (2005), is that the statistics of the SPA would be less sensitive to the inclusion of the poor and irrelevant alternatives. Besides, the model is more powerful because it reduces the probability of the occurrence of the Type II Error. Even, the variance would be asymptotically consistent, obeying the conditions of regularity, as proposed by Sullivan, Timmermann, and White (2003).

## 4 Analysis and Discussion

In Brazil, the inflation target was defined considering the variation of the Broad Consumer Price Index (IPCA) calculated by the Brazilian Institute of Geography and Statistics (IBGE). This index aims to measure price variations for households with income among 1 and 40 minimum wages. The choice of price indexes to consumer is common in most inflation targeting regimes, because it is considered a correct mea-

sure to evaluate the evolution of the population's real income. The adoption of the full IPCA in the country was due to two reasons. The first is that, although in the long term the core and the inflation tend to converge, in the short term, they may diverge significantly. The second, and maybe the most important, is a question of transparency and credibility. Therefore, since the variation of the IPCA is adopted as an official measurement of the Brazilian inflation, the present work aimed to analyze the projections concerning this variable, not only for horizons considered of short term (one step ahead) but also for long term (twelve steps ahead), of the several institutions that took part of the research Focus-Market Report, from January 2002 to February 2007. It is useful to point out that, although the research Focus/BACEN had, in February 2007, 74 institutions participating of the research, according to the reference point adopted, dataset suggested the presence of 151 institutions. But not all of them responded from the beginning (Jan./2002) and, besides, many of them do not keep a constant participation. With that, the institutions went in and out, besides those that did not respond regularly, leading to a consequent lack of information in some points of time.

From this point on, it was chosen to analyze and test the superior predictive ability only of the institutions which presented inflation projections, month by month (one step ahead) in the mentioned period, that implied in a sample with 20 institutions, out of the 74 observed in February 2007, namely: institutions 04, 10, 18, 22, 31, 37, 40, 41, 45, 57, 59, 68, 72, 78, 86, 110, 112, 113, 114, and 115, identified this way in the data bank sourced by the Gerin. Döpke *et al.* (2006) also used a sample of around 20-30 projections. The same analysis could not be done for twelve steps ahead because of the constant lack of information. The alternative adopted was to compare with the projections of the following models: Focus/Bacen, Top5/Bacen, Bacen1 (central projection – reference scenery), Bacen2 (central projection – market scenery), AR (2), ARMA (3, 1), Random Walk, Top5\* and Focus20. The 20 institutions mentioned above are called “Focus20”, mentioned as part of a whole, named Focus/Bacen. It is important to point out that these are not the 20 institutions with the highest predictive power.

The Gerin recognizes the prediction excellence of the institutions which take part of the research on the market expectations by the *rankings Top 5* of short, medium and long terms<sup>6</sup>. Here, however, the five best predictive institutions, according to selection criteria, are used, out of the 20 institutions previously selected. However, for comparison purposes, the reference dates follow the last working day, previously to the date when the IPCA-15<sup>7</sup> is released, as the Gerin does to identify the TOP 5 institutions. Since the

6 The Gerin, in the short term *ranking*, evaluates the precision of the predictions with one month of difference in relation to the indicator's edition, for the previous 6 months. The mid-term ranking considers the average precision of the projections of three consecutive periods of four months in comparison to the effective results of three months – the reference month and the two months that precede it. The long term *ranking* considers the precision of the projections informed in 12 months for the annual indicator published in the subsequent January.

7 In May 2000, the IBGE also started to make the Broad Consumer Price Index-15 (IPCA-15) available on the internet, whose period of price collection is, approximately, from the 15th day of the previous month to the 15th day of the reference month.



adoption of the target regime, the effective inflation surpassed the tolerance interval in 2001, 2002 and 2003 (Table 1). In 2004 and 2005, the IPCA variation was above the center of the goal, but inside the superior limit of the tolerance interval, while in 2006 and 2007, the inflation was below the center of the goal, but inside the tolerance limit<sup>8</sup>. For 2008, the median of the market expectations for the inflation, collected by the Gerin and spread in the Focus, Market Report of May 9<sup>th</sup>, 2008, was 4,96%, a value above the inflation target of 4,5%. Furthermore, it is possible to see a rising tendency. For 2009, it is expected that the inflation remains next to the target, which was determined in 4,5% (4,47%).

It is possible to see, in Table 2, seemingly averages, medians and variances had very similar values, which demonstrates a certain predictive equivalence. Although some expectations seemed differences in absolute terms may be statistically equivalent.

To reinforce such evidence, tests of average and variance equivalence were applied among series<sup>9</sup> that demonstrated not to reject the equivalence null hypothesis of average and variance equivalence among the inflation expectations. It was also performed the calculation of the linear correlation coefficients among the expectations of the institutions which show significant positive correlations, indicating a high degree of direct linear association among them or, it can be said, co-movements among the expectations<sup>10</sup>.

---

8 For further details about the causes of the failure in meeting the inflation targets, as well as the alternatives to fix the situation, consult the site of the Central Bank of Brazil in System of Targets for Inflation (<http://www.bcb.gov.br/?SISMETAS>) which, seeking more transparency, started to inform society about the conduction of the target regime through many means of communication.

9 Consult EViews 6 User's Guide (2007, p. 315 - 321).

10 These results were omitted for reasons of space.

Table 1 – Report of the targets for inflation in

Year	Norm	Date	Target (%)	Band (p.p)	Inferior and Superior Limits (%)	Effective inflation (IPCA % a.a.)
<b>1999</b>	Resolução 2.615	30/6/1999	8	2	6-10	8,94
<b>2000</b>			6	2	4-8	5,97
<b>2001</b>			4	2	2-6	7,67
<b>2002</b>	Resolução 2.744	28/6/2000	3,5	2	1,5-5,5	12,53
<b>2003*</b>	Resolução 2.842	28/6/2001	3,25	2	1,25-5,25	9,30
	Resolução 2.972	27/6/2002	4	2,5	1,5-6,5	
<b>2004*</b>	Resolução 2.972	27/6/2002	3,75	2,5	1,25-6,25	7,60
	Resolução 3.108	25/6/2003	5,5	2,5	3-8	
<b>2005</b>	Resolução 3.108	25/6/2003	4,5	2,5	2-7	5,69
<b>2006</b>	Resolução 3.210	30/6/2004	4,5	2	2,5-6,5	3,14
<b>2007</b>	Resolução 3.291	23/6/2005	4,5	2	2,5-6,5	4,46
<b>2008</b>	Resolução 3.378	29/6/2006	4,5	2	2,5-6,5	
<b>2009</b>	Resolução 3.463	26/6/2007	4,5	2	2,5-6,5	

Source: BCB. \* The Open Letter, of 10/21/2003, established adjusted targets of 8,5% for 2003 and 5,5% for 2004.

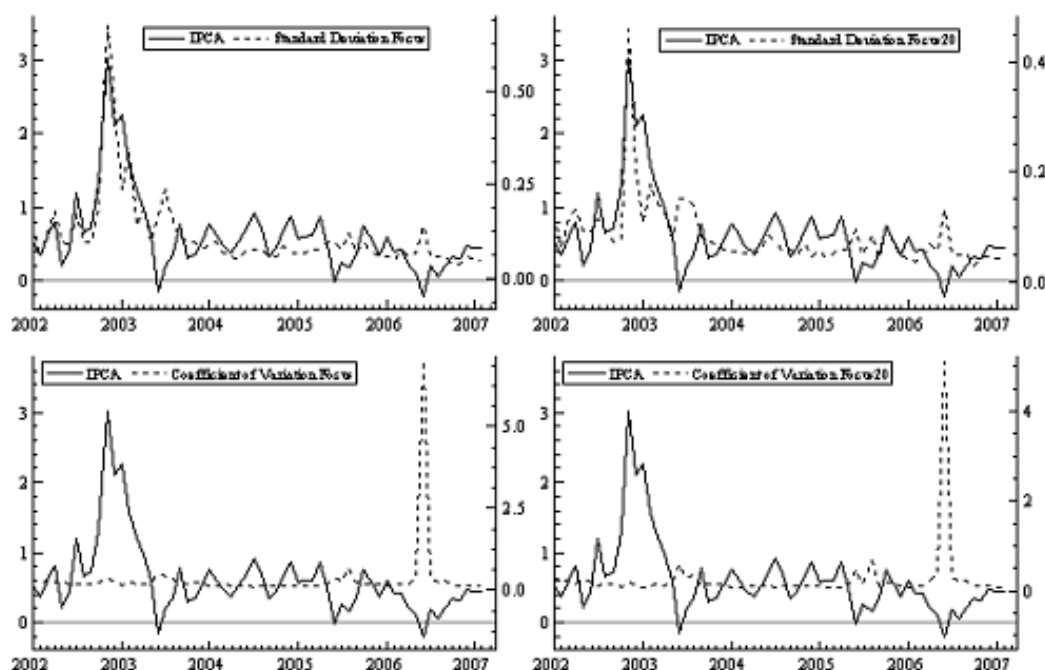
Table 2 – Descriptive statistics of the inflation expectations

Institutions	Average	Median	Max.	Min.	Std. Dev.	Skewness	Kurtosis	Jarque-Bera	P-Value
04	0,63	0,52	2,90	0,09	0,48	2,61	11,23	245,32	0,00
10	0,61	0,52	2,67	-0,16	0,47	2,31	9,59	167,12	0,00
18	0,59	0,47	2,90	-0,08	0,50	2,61	10,94	232,95	0,00
22	0,61	0,50	2,50	0,10	0,43	2,43	9,97	186,62	0,00
31	0,61	0,50	3,07	0,00	0,49	2,72	12,64	316,61	0,00
37	0,61	0,48	2,85	0,05	0,47	2,67	11,65	267,15	0,00
40	0,61	0,50	2,60	-0,09	0,43	2,28	9,83	174,32	0,00
41	0,57	0,50	1,80	0,20	0,33	2,15	7,66	104,09	0,00
45	0,58	0,50	1,90	0,05	0,36	1,93	7,31	86,53	0,00
47	0,58	0,48	2,10	-0,12	0,43	1,90	7,22	83,04	0,00
59	0,60	0,48	2,64	-0,15	0,48	2,35	9,65	171,28	0,00
68	0,60	0,50	2,50	0,00	0,44	2,45	9,62	175,46	0,00
72	0,55	0,50	2,00	0,05	0,38	2,12	8,46	123,33	0,00
78	0,60	0,50	2,80	-0,05	0,48	2,49	10,54	210,90	0,00
86	0,60	0,49	2,70	0,05	0,45	2,48	10,62	213,71	0,00
110	0,58	0,50	2,60	-0,05	0,43	2,41	10,50	205,35	0,00
112	0,59	0,50	2,40	-0,10	0,43	2,12	8,70	130,55	0,00
113	0,58	0,50	1,90	0,00	0,36	1,70	6,38	59,37	0,00
114	0,61	0,53	2,40	0,05	0,45	2,12	8,20	116,20	0,00
115	0,59	0,52	2,30	0,00	0,42	2,18	8,98	141,56	0,00

Source: GERIN/BACEN; Values calculated by the authors.

For one hand, these evidences of equivalence indicate similar behaviors of the institutions, which is something expected in a scenery with low inflation and a price stability, since this environment facilitates the inflation projection. On the other hand, such behavior is not expected in unfavorable sceneries, since a higher expectation divergence may exist simply because of the difficulty for prediction in unstable environments. However, the equivalence may remain even in environments of high uncertainty, with symmetric information, equivalent predictive capacity and an institution with good predictive capacity that displays its projection in the market without costs, among others. In this sense, it is relevant to go on exploring the predictive ability of the institutions, because they can be references for other agents in their decisions related to prices and salaries, for instance. Mankiw, Reis, and Wolfers (2003) are also an important reference for the investigation on the existence of divergence among the inflation expectations, as well as their relation with the effective inflation. Figure 1 presents the path of the pattern deviation and the variation coefficient<sup>11</sup> of the inflation expectations, either at complete (Focus) or partial (Focus20) sample levels, according to the behavior of the effective inflation, and the last one always presenting a scale in the first ordinate.

Figure 1 – Standard deviation, variation coefficient (FOCUS/FOCUS20) and observed inflation (IPCA)



Source: Development by the authors

11 The effect of variation or dispersion in relation to the average can be measured by the relative dispersion, defined by:  $\text{Relative Dispersion} = \text{Absolute Dispersion} / \text{Average}$ . If the absolute dispersion is the pattern deviation, the relative dispersion is called coefficient of variation C.V.



Figure 1 reveals that the divergence among expectations decreases as the regime inflation targeting goes by and this result is confirmed by the correlation coefficient between the pattern deviation and the effective inflation (Table 3), when it shows that the divergence increases (reduces) with the increase (fall) of the effective inflation. It can be seen in some points that the divergence tends to increase when the inflation changes abruptly for any direction: cases occurred in 11/2002 (IPCA increased in 1,71 p.p.) and 06/2003 (IPCA registered downfall of 0,76 p.p.). As to the variation coefficient, which seeks to help in the inference on the variability of the inflation expectations, since the pattern deviation is influenced by the data magnitude, it can be seen that it tends to increase mainly when deflation occurs: cases of 06/2003 (-0,15%), 06/2005 (-0,02%) and 06/2006 (-0,21%). This inverse relation between the variation coefficient and the effective inflation is observed in the correlation coefficient between them, which indicates a negative linear association level.

The empirical results above corroborate those presented by Mankiw, Reis, and Wolfers (2003), in other words, the divergence increases (reduces) with the increase (downfall) of the effective inflation and this, many times, tends to increase when inflation changes abruptly for any direction. It can still be noticed some difficulty for the Brazilian institutions in previewing deflations. Before analyzing the problem of the superior predictive ability, Table 4 shows the institutions which present the highest predictive performance in an ascending order, according to the criteria of the Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and Mean Absolute Error (MAE), this way allowing the proposal of *benchmark* “models”, which is the case of the Top5\*, formed by the median of the projections of the five best institutions.

Table 3 – Correlation coefficient

	IPCA	
Pattern Deviation FOCUS20	0.66 **	(6,805)
C.V. FOCUS20	-0.26 *	(-2,086)

Source: Data calculated by the authors; ( )  $t_0$  calculated statistics; \*, \*\* It is rejected the null hypothesis of null correlation at 5% and 1% of significance, respectively.

Afterwards, the multiple comparison of projections is applied, according to Hansen (2005), in which the predictive power of the Top5\* is evaluated, inclusively in relation to the Top5/BACEN, which may have suffered changes in its composition in the period studied<sup>12</sup>. Table 4 reveals the good performance of the institutions 10, 78, 37,

12 There is not a definite imposition that the TOP5\* should be constant, an altered composition would also be possible along the period, since it is enough to make the introduction of new forecasts and then implement a new selection of the five best ones.

18 and 59, pointing out the institution 10. With that, there is only one Top 5\* model formed by these institutions. For purposes of comparison with the other models, the prediction median of these institutions is used, just like the Gerin does in the release of the median of the Top 5 – short term. The worst performance was presented by the institutions 45, 41 and 72. The *random walk* model, in which the best prediction for the inflation in the  $t+1$  period would be the inflation in the  $t$  instant, presented lower results, compared to the others, according to the RMSE, MSE and MAE criteria. In this case, there are signs that the institutions are at least developing “models” with a predictive power superior to the very observation of the inflation that is one period out of step. The Focus20 was also pointed out, demonstrating a better predictive performance than the Focus/Bacen itself<sup>13</sup>, maybe revealing the fact that these institutions which presented their expectations in a constant way had a better monitoring of the prices and/or used more sophisticated means (econometric software) to foresee inflation. Or, it is simply due to a higher dispersion when the sample is widened. Besides, the Focus20 showed higher predictive ability in comparison to other institutions.

Table 5 shows long term predictive ability, still maintaining the same evaluation criteria. Previously, there could be an explanation on the models used in the comparison.

As it can be seen, there was no comparison between the predictive capacity of the institutions individually, as it used to be, since the sample studied did not present enough information for such analysis. In this case, a choice was made for the median of the Top5\*, the Focus/Bacen, the Top5/Bacen, the Focus20 and Bacen1, which indicates a central-scenario projection of reference, Bacen2, suggesting a market central scenario projection, *Random Walk*, AR(2) and ARMA(3, 1), maintaining the same reference point defined in the short term reference. Although it is impossible to compare individually the predictive capacity of the institutions, this analysis is also relevant, since the information can be used as reference for other economic agents, except the models proposed here (Top5\* and Focus20). The AR (2) and ARMA (3, 1) models, on the other hand, are considered due to their large use and they can easily become reference for some economic agents.

Table 4 – REQM, EQM and EAM (in ascending order - *Ranking*)

<b>Model</b>	<b>EAM</b>	<b>Model</b>	<b>EQM</b>	<b>Model</b>	<b>REQM</b>
Top5*	0,091	Top5*	0,015	Top5*	0,123
10	0,096	10	0,017	10	0,132
78	0,100	78	0,019	78	0,137
37	0,101	Top5/BACEN	0,019	Top5/BACEN	0,137

(continues...)

13 Again, it is interesting to point out that the FOCUS20 refer to the median of the inflation expectations of the 20 institutions selected, while the FOCUS refers to the median of all the inflation expectations issued by the Gerin, evidently respecting the dates of reference for the purpose of comparison.

(conclusion)

<b>Model</b>	<b>EAM</b>	<b>Model</b>	<b>EQM</b>	<b>Model</b>	<b>REQM</b>
Top5/BACEN	0,104	37	0,019	37	0,139
Focus20	0,109	18	0,022	18	0,148
18	0,113	59	0,023	59	0,152
59	0,114	Focus20	0,024	Focus20	0,155
110	0,115	110	0,025	110	0,159
68	0,116	4	0,026	4	0,162
57	0,116	68	0,026	68	0,163
31	0,120	86	0,029	86	0,170
4	0,121	31	0,030	31	0,174
22	0,128	57	0,033	57	0,181
114	0,129	114	0,035	114	0,187
Focus/BACEN	0,129	22	0,036	22	0,189
86	0,129	Focus/BACEN	0,039	Focus/BACEN	0,197
115	0,132	112	0,039	112	0,198
40	0,144	40	0,040	40	0,200
112	0,147	115	0,040	115	0,200
113	0,148	113	0,067	113	0,258
45	0,162	72	0,069	72	0,263
41	0,177	41	0,072	41	0,267
72	0,194	45	0,083	45	0,288
SAR(1)	0,239	AR(1)	0,132	AR(1)	0,363
AR(1)	0,241	SAR(1)	0,137	SAR(1)	0,370
RW	0,282	RW	0,152	RW	0,390
SARMA(5,3)	0,342	ARMA(2,1)	0,215	ARMA(2,1)	0,464
ARMA(2,1)	0,348	SARMA(5,3)	0,218	SARMA(5,3)	0,467

Source: BACEN; IBGE. Values calculated by the authors. The ARMA(p, q) models were estimated using EViews 6, and chosen through the Schwarz selection criterion.

Table 5 – REQM, EQM and EAM (in ascending order - *Ranking*)

<b>Model</b>	<b>EAM</b>	<b>Model</b>	<b>EQM</b>	<b>Model</b>	<b>REQM</b>
AR(2)	0,381	AR(2)	0,320	AR(2)	0,566
ARMA(3,1)	0,409	ARMA(3,1)	0,350	ARMA(3,1)	0,592
RW	0,556	RW	0,679	RW	0,824
TopFive/BACEN	2,224	TopFive/BACEN	10,219	TopFive/BACEN	3,197
Focus20	2,419	Focus20	11,982	Focus20	3,462
Focus/BACEN	2,463	Focus/BACEN	12,673	Focus/BACEN	3,560

(continues...)

(conclusion)

Focus5*	2,589	Focus5*	14,999	Focus5*	3,873
BACEN 2	3,367	BACEN 2	22,784	BACEN 2	4,773
BACEN 1	3,880	BACEN 1	24,881	BACEN 1	4,988

Source: BACEN; IBGE. Values calculated by the authors. The AR(p) and ARMA(p, q) models were estimated using Eviews 6, and chosen through the Schwarz selection criterion.

P.S.: BACEN1 implies in the central scenario projection of reference and BACEN2 in the central scenario projection of market, both presented in the “Report of Inflation”, of the BACEN.

The empirical results now show that the largest predictive capacities are in the AR (2) and ARMA (3, 1) models selected. The highest predictive power of the *random walk* is remarkable, in comparison to the projections of the market and Bacen here considered.

So, when following the criteria of RMSE, MSE and MAE, which are very much used, the empiric results show the existence of heterogeneity in the predictive capacity of the Brazilian inflation, diverging from the results of the simple equivalence tests. Based on this information, it is possible to imagine that the agents would form new prices from projections different from the inflation. If these projections are significant for the determination of prices, then such divergence may produce changes in the intergroup and intergroup relative prices. In the next subsection, a more powerful statistical test is applied, the SPA test of Hansen (2005), to verify the superior predictive ability, allowing a stronger corroboration of the problem of disagreement in the inflation expectations, because, if institutions with superior predictive power exist, then the divergence is statistically reaffirmed in the presence of co-movement among the projections. For such, it was used the MULCOM (*Multiple Comparison*) package version 1.00, which can be used interactively together with the OxPack for OxMetrics<sup>14</sup>.

#### 4.1 Superior Predictive Ability Test (SPA)

The analysis of the results of the SPA test can be divided into three parts: The first one contains the descriptive statistics, such as the number of competing models, the size of the sample, bootstrap parameters, among others. The second one informs about model performance, and six pair-wise comparisons. The six models compared to the *benchmark* model are: (1) The “Most Significant” model, whose performance is the most “significant” relative to the *benchmark* model in the sample analyzed; and (2)-(6) are those models with a performance that corresponded to the 75%, 50% (median), 25% and 0%. The performance is measured in terms of specified loss function. In the present case, the MSE is the measure considered. The *t-stat* indicates the significance of the model in comparison to its performances

14 Hansen and Lunde (2007) and Doornik and Hendry (2006).



compared with the *benchmark* model. But it is important to notice that this statistic is not t-distributed, and the “p-value” reported next to it cannot be interpreted as a p-value, rather it is a number that is calculated like a p-value. The third part offers information to verify if the hypothesis that the *benchmark* model is the best forecasting model, in a certain level of significance (HANSEN, 2005; HANSEN; LUNDE, 2007).

For the short term analysis, the investigation starts with the use of *benchmark* models considered weak for inflation prediction. It would be the case of the *random walk* and the institutions that had the worst predictive performance, according to Table 4, in other words, the institutions 45, 41 and 72. With these models, it is sought to corroborate their predictive inferiority, which would already confirm the existence of disagreement among the inflation expectations. Other four models served as *benchmark*: Top 5/Bacen, Top5\*, Focus/Bacen and Focus20. In these cases, it is investigated if it is reasonable to consider that the Focus/Top 5 surveys, for example, carried out by the Bacen and that may be adopted in the market, are good predictors of the inflation rate in a short term prediction horizon. Another objective was to verify if these inflation projections have a greater predictive ability than some univariate linear models AR (1), SAR (1), ARMA (2, 1) and SARMA (5, 3). For the long term analysis, the rationale is the same, but, due to the reasons already commented, only four models are considered as *benchmark*: *random walk*, AR (2), Top5/Bacen and Focus/Bacen.

The empiric results for the short term cases are presented in Tables A1 and A2, divided in panels, while the long term cases may be seen in Table A3.

The panels A-D, Table A1, show *SPA-Consistent p-values*, suggesting that the *benchmark* models, in this case, the *random walk* and the institutions 41, 45 and 72, respectively, are surpassed at least for one of the models of the other institutions, thus confirming the existence of a disagreement in the Brazilian inflation expectations. On the other hand, the results presented in the panels E-H showed: one *SPA-Consistent p-value* of 0.8619, demonstrating that the Top5/Bacen is not inferior to the other models mentioned in Table 3, except for the Top5\*, which is not included in this test; one *SPA-Consistent p-value* of 0.9996, revealing that the Top5\* would not also be inferior to the others, including the Top5/BACEN itself; and, finally, the inferiority, not only of the Focus/Bacen but also of the Focus20, since in a reasonable level of significance, they would be surpassed by at least one of the alternatives.

The panels I-L, Table A3, showed a *SPA-Consistent p-value* of 0.02470, suggesting that the *random walk* model, in the long term analysis, is surpassed by at least one of the alternative models, but that the AR(2) would not be surpassed by the others (*SPA-Consistent p-value* de 0.70). It is still possible to observe that the

Top5/Bacen and Focus/Bacen would be inferior to the other models in the prediction of the twelve steps ahead.

In short, both short and long term analyses offer empiric results favorable to the existence of disagreement in the inflation expectations, in which some models/institutions exert superior predictive ability. They are in accordance with the results of Lima and Cespedes (2003), when these authors say that simple autoregressive models are able to generate more exact long term predictions, but do not corroborate their results that the market expectations are predictions for inflation statistically similar to those of the Bacen.

It can also be observed that, if the institutions present not only divergences in the projections of the Brazilian inflation, but also some degree of inertia in the dynamics of their expectations, then such persistence would be different. For such, the variance ratio is used, a measure systematically used for the measurement of the persistence of long term shock<sup>15</sup>. Inertia is a persistence of long duration. In the analysis of time series, the completely inertial case corresponds to the random walk process, in which a shock of one per cent becomes completely persistent. The presence of certain degree of inertia in the expectations would indicate that the adjustment of these expectations would gradually occur, according to the shocks in the determinants, such as Selic interest rate, effective inflation, among others. Table A4 shows the estimate of the persistence measures,  $\hat{\nu}_{1_k}$ ,  $\hat{\nu}_{2_k}$ ,  $\hat{s}_{1_k}$  e  $\hat{s}_{2_k}$  for the inflation expectations, one step ahead, of the institutions that presented the best and the worst predictive performances, according to Table 4. In all the case is observed that the  $\hat{\nu}_{1_k}$  and  $\hat{\nu}_{2_k}$  decreases as the value of k increases and presents similar values. For k = 2 and k = 3, these measures suggest sharp shock effects, but the estimates decrease, so that k = 24, namely, 24 months or 2 years after the initial shock,  $\hat{\nu}_{1_k}$  and  $\hat{\nu}_{2_k}$  remain next to 0.20 and 0.10, with some cases a little above or a little below these values. Supported by these results, it can be said that the inertia in the dynamics of the expectations of these institutions are of second order. In this sense, shock effects must be quickly dissipated. Such results become stronger with the estimates, but they are not clear as to the institutions 41 and 72, because they present contradictory estimates  $\hat{s}_{1_k}$  and  $\hat{s}_{2_k}$ . The last institutions mentioned presented weak predictive performances. In general, it is possible to say that the measures of estimated persistence suggest sharp shock effects, but that such effects decrease as time passes, for most of the cases investigated.

15 Consult Cribari-Neto and Cassiano (2005) for see some measures of the variance ratio. One of them, specifically, was achieved from the alterations in test statistics proposed by Wright (2000), and it is more adequate in the presence of discordant observations, either outliers or inliers.

## 5 Final Considerations

Inflation expectations are considered a key-variable in the specification of macroeconomic models, in which such relevance comes from Friedman (1968) and Phelps (1969). Recently, the expectations, especially those of inflation, have drawn great interest in the discussion about the regime inflation targeting and the informational rigidity that would lead to price rigidity. Therefore, the present work aimed to offer literature the results of the Brazilian data, exploring the problem of disagreement in the inflation expectations. To achieve this, it evaluates the predictive ability of the institutions that took part in the research “Focus-Market Report” of the Gerin/Bacen in the period from January 2002 to February 2007, starting from a multiple comparison of predictions of a certain *benchmark* model via test for superior predictive ability proposed by Hansen (2005) – test SPA.

The empiric results reinforce the statement that the divergence in the expectations increases (reduces) with the increase (fall) of the effective inflation and that it sometimes tends to increase when inflation changes abruptly for any direction, according to the result presented in Mankiw, Reis, and Wolfers (2003). But the existence of disagreement in the expectations becomes stronger with its corroboration via RMSE, MSE e MAE criteria and, mainly, with the application of the SPA statistical test in short and long term horizons.

Here, the disagreement is set over time, not among economic agents on a specific date, but the disagreement in terms of ability to predict inflation over the studied period.

Some of the results achieved take the same direction of Lima and Cespedes (2003) when they show that simple autoregressive models are able to generate more exact long term predictions, but other authors do not corroborate their result that market expectations are predictions for the inflation statistically similar to those of the Bacen. It can also be said, based on the estimated persistence measures, that there is a certain degree of inertia in the expectation dynamics, in which the sharp shock effects occur, but that such effects decrease as time goes by, for most part of the cases analyzed.

So, there would be, besides disagreement in the expectations, a co-movement among the series and a certain degree of inertia, indicating that their adjustment would gradually occur, due to shocks in their determinants, such as the Selic interest rate, effective inflation, among others, depending on the speed of the shock effect fall, disorganized and in the same direction. Consequently, the presence of rigidity in the costs would be inevitable, because do not adjust coordinately and may not automatically adjust when shocks occur. The reason for that might be informational rigidity.

Depending on the level of disagreement among economic agents, at time  $t$ , it may affect relative prices causing problem on the conduct of monetary policy in coordinating expectations to maintain price stability. Expectations are important because they are a component in the formation of prices, wages and economy contracts, besides determining the term structure of ex-ante real interest rates associated to a determined nominal interest curve, which, by its turn, affects the private decisions of consumption and investments.

## References

- BANK OF ENGLAND. *Economic Models at the Bank of England*: Park Communications Ltd. London, 1999.
- BRANCH, W. A. The theory of rationally heterogeneous expectations: evidence from survey data on inflation expectations. *Economic Journal*, Cambridge, US, v. 114, p. 592-621, Jul. 2004.
- CENTRAL BANK OF BRAZIL – BACEN. Evaluation of the Central Bank of Brazil's Inflation Forecasts. *Inflation Report*, Brasília, v. 5, n. 1, p. 107-111, Mar. 2003.
- \_\_\_\_\_. Monetary Policy Decision. *Inflation Report*, Brasília, v. 6, n. 4, p. 114-117, Dec. 2004.
- COWLES, A. Can Stock Market Forecasters Forecast? *Econometrica*, Chicago, US, v. 1, n. 3, p. 309-324, Jul. 1993.
- CRIBARI-NETO F.; CASSIANO, K. Uma Análise da Dinâmica Inflacionária Brasileira. *Revista Brasileira de Economia*, Rio de Janeiro, v. 59, n. 4, p. 535-566, out./dez. 2005.
- DOORNIK J. A.; HENDRY, D. F. *OxMetrics: An Interface to Empirical Modelling*. 5th ed. London: Timberlake, 2006.
- DÖPKE J. et al. Stick information Phillips Curve: European evidence. *Discussion Papers of DIW*, Berlin, AL, v. 615, 2006.
- FRIEDMAN, M. The Role of Monetary Policy. *American Economic Review*, Nashville, US, v. 58, n. 1, p. 1-17, Mar. 1968.
- HANSEN, P. R. A Test for Superior Predictive Ability. *Journal of Business and Economic Statistics*, Washington, US, v. 23, p. 365-380, Oct. 2005.
- HANSEN, P. R.; LUNDE, A. A forecast comparison of volatility models: does anything beat a GARCH (1,1)? *Journal of Applied Econometrics*, Chichester, GB, v. 20, n. 7, p. 873-889, 2005.
- HANSEN, P. R.; LUNDE, A. MULCOM 1.0: Econometric Toolkit for Multiple Comparisons, 2007.
- KOOPMAN, S.; JUNGBACKER, B.; HOL, E. Forecasting daily variability of the S&P 100 stock index using historical, realized and implied volatility measurements. *Journal of Empirical Finance*, Amsterdam, NL, v. 12, n. 3, p. 445-475, Jun. 2005.
- KORENOK, O. Empirical comparison of sticky price and sticky information models. *Journal of Macroeconomics*, Detroit, US, v. 30, n. 3, p. 906-927, Sept. 2008.



- KORENOK, O.; SWANSON, N. R. How sticky is sticky enough? A distributional and impulse response analysis of New Keynesian DSGE models. *Journal of Money, Credit and Banking*, Columbus, US, v. 39, n. 6, p. 1481-1508, Sep. 2007.
- LIMA, E.; CÉSPEDES, B. O desempenho do mercado (focus) e do Bacen na previsão da inflação: comparações com modelos lineares univariados. *Boletim de Conjuntura*, Rio de Janeiro, v. 60, p. 75-83, mar. 2003.
- LO, A.; MACKINLAY, C. Data-Snooping Biases in Testes of Financial Asset Pricing Models. *Review of Financial Studies*, Cary, US, v. 3, n. 3, p. 431-467, 1990.
- LUCAS, R. E. Some International Evidence on Output-Inflation Tradeoffs. *American Economic Review*, Nashville, US, v. 63, n. 3, p. 326-334, Jun. 1973.
- MANKIW, N. G. A Letter to Ben Bernanke. *American Economic Review*, Nashville, US, v. 96, n. 2, p. 182-184, May 2006.
- MANKIW, N. G.; REIS, R. Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve. *Quarterly Journal of Economics*, Cambridge, US, v. 117, p. 1295-1328, Nov. 2002.
- MANKIW, N. G.; REIS, R.; WOLFERS, J. Disagreement About Inflation Expectations, *NBER – Working Papers*, n. 9796, Jun. 2003.
- MARQUES, A.; FACHADA, P.; CAVALCANTI, D. Sistema Banco Central de Expectativas de Mercado. *Notas Técnicas do Banco Central do Brasil*, Brasília, n. 36, p. 1-27, jun. 2003.
- PHELPS, E. S. The new microeconomics in inflation and employment theory. *American Economic Review*, Nashville, US, v. 59, n. 2, p. 147-160, May 1969.
- ROBERTS, J. M. New Keynesian Economics and the Phillips Curve. *Journal of Money, Credit and Banking*, Columbus, US, v. 27, n. 4, p. 975-984, Nov. 1995.
- SAFFI, P. A. C. Análise técnica: sorte ou realidade? *Revista Brasileira de Economia*, Rio de Janeiro, v. 57, n. 4, p. 953-974, out./dez. 2003.
- SIMES, R. J. An Improved Bonferroni Procedure for Multiple Tests of Significance. *Biometrika*, London, UK, v. 73, n. 3, p. 751-754, 1986.
- SPANOS, A. *Statistics Foundations of Econometric Modelling*. Cambridge, US: Cambridge University Press, 1989.
- SULLIVAN, R.; TIMMERMAN, A.; WHITE, H. Forecast evaluation with shared data sets. *International Journal of Forecasting*, Amsterdam, NL, v. 19, n. 2, p. 217-227, Apr./Jun. 2003.
- SVENSSON, L. Inflation forecast targeting: implementing and monitoring inflation targets. *European Economic Review*, Amsterdam, NL, v. 41, n. 6, p. 1111-1146, Jun. 1997.
- TAYLOR, J. B. Aggregate Dynamics and staggered contracts. *Journal of Political Economy*, Chicago, US, v. 88, n. 1, p. 1-23, Feb. 1980.
- WEST, K. D. Asymptotic Inference about Predictive Ability. *Econometrica*, Chicago, US, v. 64, n. 5, p. 1067-1084, Sep. 1996.
- WHITE, H. A Reality Check for Data Snooping. *Econometrica*, Chicago, US, v. 68, n. 5, p. 1097-1126, Sep. 2000.
- WOODFORD, M. *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press: New Jersey, 2003.
- WRIGHT, J. H. Alternative variance-ratio tests using ranks and signs. *Journal of Business and Economic Statistics*, Washington, US, v. 18, n. 1, p. 1-9, Jan. 2000.

# Attachment A

Table A1 – Tests for superior predictive ability (short term)

Panel A: Results for the Benchmark <i>Random Walk</i>					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	RW	0.15175	–	–
m = 20 (number of models)	Most Significant	113	0.06650	3.79073	0.00060
n = 62 (sample size)	Best	10	0.01734	2.48782	0.01580
B = 10,000 (resamples)	Model_25%	110	0.02516	2.48531	0.01570
q = 0.5 (dependence)	Median	57	0.03277	2.74673	0.00940
	Model_75%	40	0.03987	2.20649	0.02660
	Worst	45	0.08306	3.21059	0.00210
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.00200	0.00200	0.00200
Panel B: Results for the Benchmark 41					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	41	0.07153	–	–
m = 19 (number of models)	Most Significant	57	0.03277	2.45947	0.01420
n = 62 (sample size)	Best	10	0.01734	2.07410	0.03460
B = 10,000 (resamples)	Model_25%	110	0.02516	1.99207	0.03810
q = 0.5 (dependence)	Median	31	0.03034	1.54431	0.07510
	Model_75%	40	0.03987	1.38385	0.08900
	Worst	45	0.08306	-0.64552	0.74910
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.03240	0.03250	0.03250

Panel C: Results for the Benchmark 45					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	45	0.08306	–	–
m = 19 (number of models)	Most Significant	113	0.06650	1.79606	0.05110
n = 62 (sample size)	Best	10	0.01734	1.56794	0.07690
B = 10, 000 (resamples)	Model_25%	110	0.02516	1.47945	0.08450
q = 0.5 (dependence)	Median	31	0.03034	1.25139	0.09340
	Model_75%	40	0.03987	1.12294	0.12020
	Worst	41	0.07153	0.64552	0.25090
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.10830	0.10830	0.10830
Panel D: Results for the Benchmark 72					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	72	0.06917	–	–
m = 19 (number of models)	Most Significant	57	0.03277	3.98665	0.00010
n = 62 (sample size)	Best	10	0.01734	2.93896	0.00570
B = 10, 000 (resamples)	Model_25%	110	0.02516	2.91920	0.00580
q = 0.5 (dependence)	Median	31	0.03034	2.09497	0.03020
	Model_75%	40	0.03987	1.95368	0.03860
	Worst	45	0.08306	-0.51415	0.71910
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.00140	0.00140	0.00140

Table A2 – Tests for superior predictive ability (short term)

Panel E: Results for the Benchmark Top 5					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	Top 5	0.01881	–	–
m = 27 (number of models)	Most Significant	10	0.01734	0.86760	0.18500
n = 62 (sample size)	Best	10	0.01734	0.86760	0.18500
B = 10, 000 (resamples)	Model_25%	04	0.02619	-1.70405	0.95080
q = 0.5 (dependence)	Median	22	0.03555	-2.50679	0.99010
	Model_75%	41	0.07153	-2.03848	0.96250
	Worst	SARMA	0.21820	-3.00996	0.99440
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.44500	0.86190	0.91010
Panel F: Results for the Benchmark Top5*					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	Top5*	0.01521	–	–
m = 28 (number of models)	Most Significant	10	0.01734	-1.01752	0.84450
n = 62 (sample size)	Best	10	0.01734	-1.01752	0.84450
B = 10, 000 (resamples)	Model_25%	110	0.02516	-1.92955	0.95730
q = 0.5 (dependence)	Median	22	0.03555	-2.76910	0.99370
	Model_75%	72	0.06917	-2.92735	0.99360
	Worst	SARMA	0.21820	-3.00281	0.99440
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.92870	0.99960	1.00000

Panel G: Results for the Benchmark Focus					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	Focus	0.03862	–	–
m = 28 (number of models)	Most Significant	Top5*	0.01521	1.98268	0.04030
n = 62 (sample size)	Best	Top5*	0.01521	1.98268	0.04030
B = 10, 000 (resamples)	Model_25%	Focus20	0.02397	1.75596	0.06130
q = 0.5 (dependence)	Median	114	0.03487	0.53356	0.28250
	Model_75%	72	0.06917	-3.47075	0.99870
	Worst	SARMA	0.21820	-3.14314	0.99620
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.12120	0.12620	0.14230
Panel H: Results for the Benchmark Focus20					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	Focus20	0.02397	–	–
m = 28 (number of models)	Most Significant	Top5*	0.01521	2.43416	0.01570
n = 62 (sample size)	Best	Top5*	0.01521	2.43416	0.01570
B = 10, 000 (resamples)	Model_25%	110	0.02516	-0.36423	0.65930
q = 0.5 (dependence)	Median	22	0.03555	-2.20945	0.98020
	Model_75%	72	0.06917	-2.95053	0.99420
	Worst	SARMA	0.21820	-3.01386	0.99430
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.04580	0.06730	0.07610



Table A3 – Tests for superior predictive ability (long term)

Panel I: Results for the Benchmark *Random Walk*

		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	RW	0.67871	–	–
m = 8 (number of models)	Most Significant	AR	0.31984	2.37776	0.01860
n = 60 (sample size)	Best	AR	0.31984	2.37776	0.01860
B = 10, 000 (resamples)	Model_25%	TopFive	10.21928	-2.38843	0.97910
q = 0.5 (dependence)	Median	Focus	12.67272	-2.47603	0.98310
	Model_75%	Top5*	14.99866	-2.41508	0.98130
	Worst	BACEN1	24.88102	-3.06999	0.99500
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.02470	0.02470	0.02470

## Panel J: Results for the Benchmark AR(2)

		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	AR	0.31984	–	–
m = 8 (number of models)	Most Significant	ARMA	0.35043	-0.50590	0.69590
n = 60 (sample size)	Best	ARMA	0.35043	-0.50590	0.69590
B = 10, 000 (resamples)	Model_25%	TopFive	10.21928	-2.46717	0.98190
q = 0.5 (dependence)	Median	Focus	12.67272	-2.53833	0.98470
	Model_75%	Top5*	14.99866	-2.46490	0.98280
	Worst	BACEN1	24.88102	-3.09904	0.99520
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.50340	0.70000	0.96740

Panel K: Results for the Benchmark TopFive					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	TopFive	10.21928	–	–
m = 8 (number of models)	Most Significant	ARMA	0.35043	2.47804	0.01790
n = 60 (sample size)	Best	AR	0.31984	2.46717	0.01810
B = 10, 000 (resamples)	Model_25%	RW	0.67871	2.38843	0.02090
q = 0.5 (dependence)	Median	Focus	12.67272	-2.26885	0.97540
	Model_75%	Top5*	14.99866	-2.03633	0.96280
	Worst	BACEN1	24.88102	-3.36574	0.99780
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.01860	0.01860	0.01890
Panel L: Results for the Benchmark Focus					
		<b>“Model”</b>	<b>Sample Loss</b>	<b>t-statistic</b>	<b>“p-value”</b>
Evaluated by MSE	Benchmark:	Focus	12.67272	–	–
m = 8 (number of models)	Most Significant	ARMA	0.35043	2.54877	0.01510
n = 60 (sample size)	Best	AR	0.31984	2.53833	0.01530
B = 10, 000 (resamples)	Model_25%	RW	0.67871	2.47603	0.01690
q = 0.5 (dependence)	Median	Focus20	11.98228	1.80776	0.05240
	Model_75%	Top5*	14.99866	-1.76487	0.94420
	Worst	BACEN1	24.88102	-3.59028	0.99880
	SPA p-values:		SPA <sub>Lower</sub>	SPA <sub>Consistent</sub>	SPA <sub>Upper</sub>
			0.03050	0.03050	0.03070

Table A4 – Persistence measures

	K	2	4	6	8	10	12	14	16	18	20	22	23	24
Focus	Vr	0.91	0.74	0.64	0.52	0.44	0.41	0.36	0.35	0.30	0.28	0.28	0.28	0.26
	vr2	0.91	0.72	0.63	0.50	0.41	0.27	0.19	0.18	0.14	0.13	0.14	0.14	0.12
	S1	0.88	0.59	0.59	0.59	0.52	0.58	0.60	0.65	0.69	0.71	0.78	0.79	0.77
	S2	0.89	0.58	0.54	0.50	0.39	0.41	0.40	0.42	0.44	0.42	0.46	0.46	0.42
	Vr	0.89	0.73	0.51	0.44	0.39	0.34	0.32	0.31	0.26	0.25	0.24	0.24	0.22
	vr2	0.89	0.72	0.50	0.42	0.36	0.25	0.18	0.17	0.14	0.12	0.12	0.12	0.10
	S1	0.74	0.45	0.34	0.28	0.23	0.22	0.19	0.18	0.17	0.16	0.19	0.19	0.18
	S2	0.78	0.58	0.51	0.48	0.43	0.45	0.45	0.42	0.36	0.30	0.27	0.24	0.21
45	Vr	0.92	0.71	0.69	0.56	0.44	0.41	0.35	0.32	0.28	0.26	0.26	0.26	0.24
	vr2	0.92	0.70	0.68	0.54	0.41	0.28	0.19	0.17	0.14	0.12	0.13	0.13	0.11
	S1	1.01	0.66	0.66	0.62	0.47	0.50	0.41	0.36	0.32	0.25	0.27	0.27	0.24
	S2	1.08	0.82	0.84	0.77	0.58	0.60	0.51	0.44	0.36	0.27	0.27	0.25	0.21
10	Vr	0.84	0.70	0.63	0.51	0.38	0.34	0.28	0.27	0.24	0.22	0.22	0.21	0.19
	vr2	0.84	0.70	0.62	0.50	0.37	0.24	0.15	0.15	0.13	0.11	0.11	0.11	0.08
	S1	0.81	0.49	0.47	0.44	0.33	0.35	0.33	0.35	0.39	0.42	0.48	0.51	0.50
	S2	0.82	0.49	0.45	0.41	0.28	0.28	0.25	0.25	0.28	0.29	0.34	0.36	0.34
18	Vr	0.90	0.66	0.60	0.48	0.38	0.35	0.29	0.27	0.24	0.21	0.22	0.21	0.19
	vr2	0.89	0.66	0.60	0.47	0.36	0.24	0.16	0.15	0.12	0.10	0.12	0.11	0.09
	S1	0.94	0.56	0.53	0.47	0.38	0.40	0.32	0.31	0.26	0.20	0.21	0.17	0.13
	S2	0.95	0.69	0.61	0.54	0.43	0.47	0.39	0.34	0.27	0.20	0.19	0.16	0.11

(continues...)

