Leading Indicators of Inflation for Brazil*

MARCELLE CHAUVETϒ

Research and Studies Department
Central Bank of Brazil

June 2000

Abstract

The goal of this project is to construct leading indicators that anticipate inflation cycle turning points on a real time monitoring basis. As a first step, turning points of the IPCA inflation are determined using a periodic stochastic Markov switching model. These are the event timing that the leading indicators should anticipate. A dynamic factor model is then used to extract common cyclical movements in a set of variables that display predictive content for inflation. The leading indicators are designed to serve as practical tools to assist real-time monitoring of monetary policy on a month-to-month basis. Thus, the indicators are built and ranked according to their out-of-sample forecasting performance. The leading indicators are found to be an informative tool for signaling future phases of the inflation cycle out-of-sample, even in real time when only preliminary and unrevised data are available.

---

* The views expressed in this report are those of the author and do not necessarily reflect the views of the Central Bank of Brazil. This paper was written as part of consulting services in which the author developed leading indicators of inflation for the Central Bank of Brazil. The material in this paper draws heavily from the reports generated during the consulting: Activity Report Sept/1999; Final Report Dec/99: "Leading Indicators of Inflation for Brasil," Activity Report: Apr/2000a: "Turning Point Analysis of the Leading indicators of inflation."

ϒ Department of Economics, University of California, Riverside, CA 92521-0247; phone: (909) 787-5037 x1587; fax: (909) 787-5685; email: chauvet@mail.ucr.edu.

1. Introduction

Inflation targeting programs are made operational through frameworks that enable the Central Bank to compare inflation forecasts to the announced target. In fact, the forecast serves as an intermediate target for policy decisions. The implementation of the inflation monitoring process is based on an assessment whether price forecasts deviate from the target path. A policy action is then taken based on these forecasts. Thus, a proactive policy to control inflation depends crucially on the ability to estimate the future path of inflation trends and cycles.

The goal of this project is to build leading indicators that anticipate signals of changes in the level and variability of inflation as measured by the "Índice the Precos ao Consumidor Amplo" (IPCA) several months in advance. This instrument is one the forecasting tools used by the Central Bank for the inflation targeting program, in addition to the structural macroeconomic model of monetary transmission and linear vector autoregressive models. The indicators are designed to serve as practical tools to assist real-time monitoring of monetary policy in Brazil on a month-to-month basis. Thus, it is crucial that the leading indicators be constructed based on out-of-sample forecasting performance.

Leading indicators have been a successful forecasting tool with long tradition in the U.S, starting with the seminal work of Burns and Mitchell (1946) at the National Bureau of Economic Research (NBER). Recently, there has been a revival interest in this instrument, which is now widely used to predict economic turning points not only in the U.S., but also in the OECD countries. There has also been a renewal academic interest in this traditional method, as new econometric models and tools can be used to explore more formally the idea of differences in the dynamics of business cycle stages. Although originally the indicators were used mainly to anticipating business cycle turning points, they been also been used to anticipate regional growth cycles, international economic fluctuations, stock market changes, and inflation turning points signals, among others.[1]

[1] For a review of some of the related literature, see Lahiri and Moore (1991).

The method underlying the construction of economic indicators is distinct from econometric regression methods and can capture aspects of turning points in the inflation cycle that regression representations may miss. Unlike regression models, the focus is not in linear forecasting the level of inflation, but instead the leading indicator is designed to give early warning signals of imminent changes in inflation "trend-cycle" turning points. For example, based on an information set at t, linear regressions can yield t+h steps ahead forecasts whose accuracy decreases as the horizon h gets longer. On the other hand, the leading indicator is not a "forecast" of the inflation based on an information set, but it corresponds to set of variables that, under some economic theory, anticipate movements in inflation, without any loss of accuracy for longer horizons. The leading indicator is a tool to answer questions such as: "is inflation going to increase in the next couple of months?" "Is the economy in a high or low inflation phase?" As the leading indicator enters in a high (low) growth phase, this signals a high probability that inflation will also enter a high (low) growth phase a couple of months later. That is, the leading indicator is a combination of variables designed to signal cyclical changes in inflation, particularly the beginning and end of growth phases. Based on the answers to these questions, the indicator can serve as a real time tool for monitoring monetary policy.

In contrast with the goal of linear forecasting using regression methods, the leading indicators are built to form an 'event timing forecast.' The event is an inflation turning point, that is, the peaks or troughs of the inflation cycle phases. The event is certain and the outcome is known (e.g., if inflation is in a positive growth phase, the next event must be a peak, i.e., the end of this phase). However, the timing in which turning points occur is uncertain.

As the economy goes through growth phases, the index of leading indicators of inflation may give signals of future inflation fluctuations as a function of the stage of the economy. Thus, leading indicators may provide more insight into how the inflation process evolves than simply looking at economic time series over calendar periods. For example, changes in interest rates may have a stronger or weaker impact depending on whether the economy is close to an economic recession or in the beginning of an expansion. In addition, since the index of leading indicator is composed of several variables, it could be more informative than individual series by themselves in anticipating inflation fluctuations.

The leading indicator is constructed from a dynamic factor model, which is an unobserved variable that summarizes comovements of variables that lead the Brazilian inflation as measured by changes in the IPCA. The model is a signal-noise extractor that filters out idiosyncratic movements in the observable variables from common cyclical movements related to the inflation process.[2] The dynamic factor is composed of economic variables that display linear predictive performance in forecasting IPCA inflation and the ability to anticipate inflation turning points, such as price of inputs and energy, index of imported prices, price of sensitive materials, measures of demand pressure, prime movers such as fiscal or monetary policy changes, or forward-looking variables that reflect business expectations.[3] The resulting indicator (henceforth, leading indicator of inflation - LII) can be used to give early warning signals of the onset of inflation phases.

Stock and Watson (1989, 1991) use a dynamic factor model to construct a coincident indicator of business cycle, and then use this indicator in a VAR system to build a leading indicator as the six-month ahead forecast of the growth rate of the coincident indicator. Chauvet and Potter (2000) use a dynamic factor model to build a coincident indicator of the U.S. stock market, and leading indicators as one-step-ahead forecasts of this indicator. However, these authors do not use the dynamic factor to build leading indicators of a target variable. Since the dynamic factor model extracts common cyclical movements underlying the observable variables, this implies that these variables should exhibit a similar lead-lag relationship with inflation. Thus, an important criterion implied in this model is the historical conformity and the relationship of the leading variables with the reference inflation cycle as to the timing of changes.[4] A similar approach to the one developed here is found in Chauvet (2000b), in which a dynamic two-factor model is used to construct a leading indicator of business cycles using on promptly available financial variables.

[2] Notice that the dynamic factor constructed in this stage can also be used in the structural model and VAR models previously developed by the Brazilian Central Bank to improve their predictive performance.

[3] See the Activity Report I for a more detailed discussion.

[4] According to the NBER practices and an extensive number of other related studies (see for example Moore & Shiskin 1967, Beck, Bush & Hayes 1973 and Zarnowitz & Boschan 1975), historical conformity and the timing of changes with the reference cycle are regarded as the most important criteria to select economic time series to forecast turning points.

The primary goal of the leading indicator is in anticipating inflation turning points. Although it can not be used to give a linear forecast of inflation by itself, the leading indicator can be combined with inflation in vector autoregressions to provide linear forecasts of inflation. In fact, the leading indicator can be used in multivariate systems that also include other variables that have predictive power in forecasting inflation beyond just the leading indicator itself. Since the leading indicator is a scalar that summarizes information in a vector of variables, the system would be parsimonious, allowing the inclusion of more variables or lags. This is particularly important when the available sample is not very long, as in the case of Brazilian macroeconomic variables.

Leading indicators are studied at the monthly frequency for two sample data — one for the period post "Plano Real" (1994.08 – 1999:12) and the other for a longer sample (1980.01–1999:12), which includes the hyperinflationary process in the 1980s and several stabilization plans.

Since the goal is to use the leading indicators to forecasting turning points in real time, the model estimation and the variable selection process were based on out-of-sample forecasting performance. Out-of-sample estimation is crucial in order to avoid data mining and, consequently, poor forecasts in real time. Thus, the variable selection process and models were recursively re-estimated through the sample period, one-step-ahead forecast errors were computed, and the variables and models were then ranked according to their out-of-sample forecasting ability. This allows better understanding on how well the models would have performed if they had been applied month by month in real time.

A set of leading indicators of inflation was obtained for the shorter sample comprising the post-Real Plan period. These indicators were ranked according to their ability to forecast turning points and their performance in linear forecasting the IPCA inflation. Turning point analysis indicates that these leading indicators have been proving to be good real time forecasting tools for inflation in Brazil. All indicators predict all turning points of the inflation cycle. In addition, the indicators yield false signals only 15% of the time. The leading indicators have been proving to be informative tool for signaling futures phases of inflation cycles out-of-sample, even in real time, when only preliminary and unrevised data are available.

For the longer sample from 1980.01 to present, however, leading indicators of inflation exhibit a weaker ability to signal turning points. This result is not surprising and is a consequence of the unexpected changes in the economy introduced by the six major "pacotes econômicos," during this period, which most economic variables did not forewarned. These changes in policy regimes engendered structural breaks in the relation between nominal and real variables.

This paper is organized as follows. In the second section, the object of study – monthly inflation as measured by the IPCA growth since the Real Plan – is studied with respect to its long term trend, seasonal patterns, and short-term cyclical fluctuations. A turning point dating of the IPCA inflation is then established, which is the event timing the leading indicators of inflation should anticipate. In the third section, the process undertaken to select and rank the candidate leading variables is described. Sections 4 and 5 discuss the dynamic factor model and the estimation procedure. The sixth section presents the top 5 leading indicators of inflation and examines their performance in anticipating inflation turning points in an out-of-sample real time exercise. In the seventh section, the results of the leading indicators based on the longer sample are discussed. The eighth section concludes.

2. Analysis of Brazilian Inflation

The first step is to examine the object of study of the project – monthly inflation as measured by the log first difference of the IPCA seasonally unadjusted (heretofore, IPCA inflation) from 1994:08 to 2000:03.[5] The IPCA inflation was analyzed with respect to its trend, seasonal patterns, and short-term cyclical fluctuations. The idea is to establish a turning point dating of the cyclical growth phases of the IPCA inflation. The leading indicators are constructed to forecast the timing of these inflation turning points in real time.

[5] Since the inflation cycles and trend may be closely interwoven, important information to the understanding of cyclical changes may be lost by mechanically detrending inflation. In fact, removing the trend may lead to underestimation or overestimation of cyclical changes.

The seasonal inflation patterns were measured using two methods: a ratio-to-moving average and the X-11 additive technique[6]. The main difference between the X-11 and moving average methods is that the seasonal factors may change from year to year in the former while they are constant in the latter. From this analysis, there is evidence of a seasonal pattern from August to November in which IPCA inflation is substantially higher.

In order to investigate short-term cyclical movements in the IPCA inflation, Hamilton's (1989) Markov Switching model (MS) was used to determine phases of high and low inflation growth. However, in order to capture seasonal changes in the inflation process, the model was extended to a periodic stochastic regime switching model, as suggested by Ghysels (1993). An AR(1) two-state periodic Markov model was fitted to the seasonally unadjusted IPCA inflation, π_t:

$$\pi_t - \mu_{st} = \phi(\pi_{t-1} - \mu_{st-1}) + \varepsilon_t \qquad \varepsilon_t \sim \text{i.i.d. } N(0, \sigma^2), \text{ and } |\phi| < 1 \tag{1}$$

where $\mathrm{s}_t \equiv (k_t, s_t)$, that is, the state of inflation growth is described by a stochastic switching regime process k_t and a deterministic seasonal process s_t = tmod(12), where 12 corresponds to the monthly frequency sampling throughout the year.

In this model, the intercept μ_{st} can take the value of μ_0 representing a low inflation state (s_t=0), or $\mu_0 + \mu_1 k_t$ representing a high inflation state (s_t=1). The switches between the first order Markov chain s_t and the relation between $\{k_t\}$ and $\{s_t\}$ processes are ruled by the transition probabilities:

$$p_{ij} = \mathrm{Prob}[\mathrm{s}_t=j|\mathrm{s}_{t-1}=i] = \sum_{s=1}^{12} 1_{st}\, p_{ij}^{s},$$

where $\sum_{j=0}^{1} p_{ij}^{s} = 1,\ \forall i, \forall s$ for all t, and 1_{st} is the indicator function:

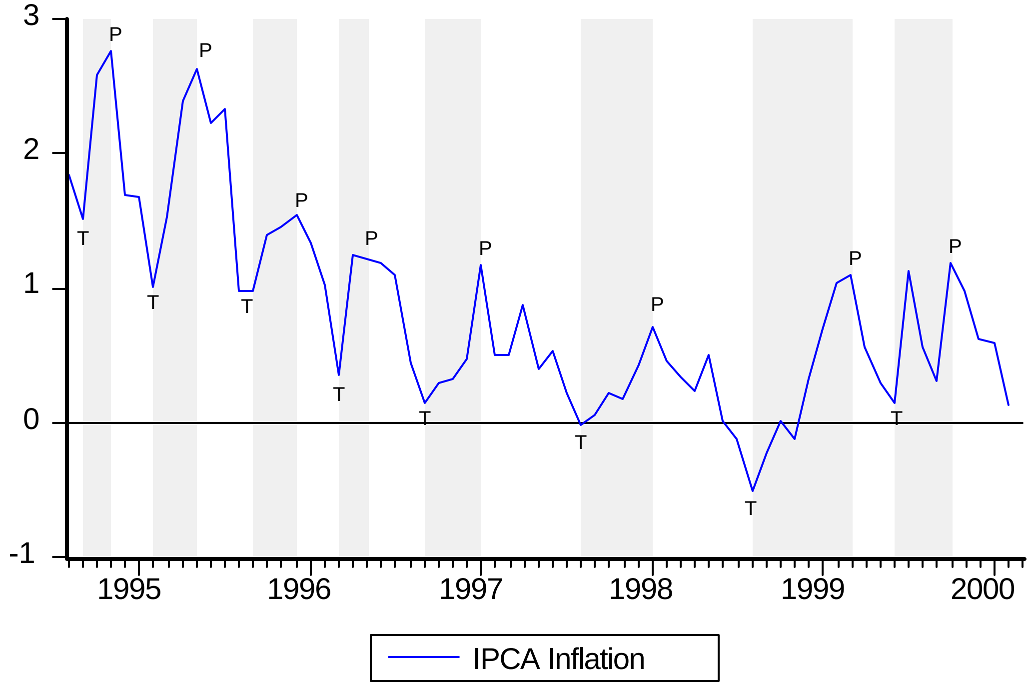
$$1_{st} \begin{cases} = 1 & \text{if } \quad s_t = s \\ = 0 & \text{otherwise} \end{cases}$$

Thus, the transition probabilities are allowed to vary stochastically and periodically according to monthly seasons.[7] The model yields inferences of the probabilities of high or low inflation

[6] The X-11 method is the standard U.S. Bureau of the Census adjustment method.
[7] For a discussion of the estimation procedure for this model, see Ghysels (1993) and Hamilton (1994).

phases, which are used to identify cyclical and seasonal changes in the IPCA inflation. Figure 1 plots the IPCA inflation its high growth phases since the Real Plan.

Figure 1 – IPCA Inflation and its Turning Points (Shaded Area) – Cyclical and Seasonal Changes (P for Peaks and T for Troughs)



Combining the Markov switching and seasonal adjustment techniques, eight half-cycles of high inflation were found, representing seasonal fluctuations and cyclical changes due to internal and external shocks.[8] In particular, the following dating of the IPCA inflation was established, based on the filtered probabilities that the economy is in a high inflation state:

[8] See the Activity Report I for a more detailed discussion.

Table 1 - Dating of High Inflation Phases
Seasonal and Cyclical Changes - Trough to Peak

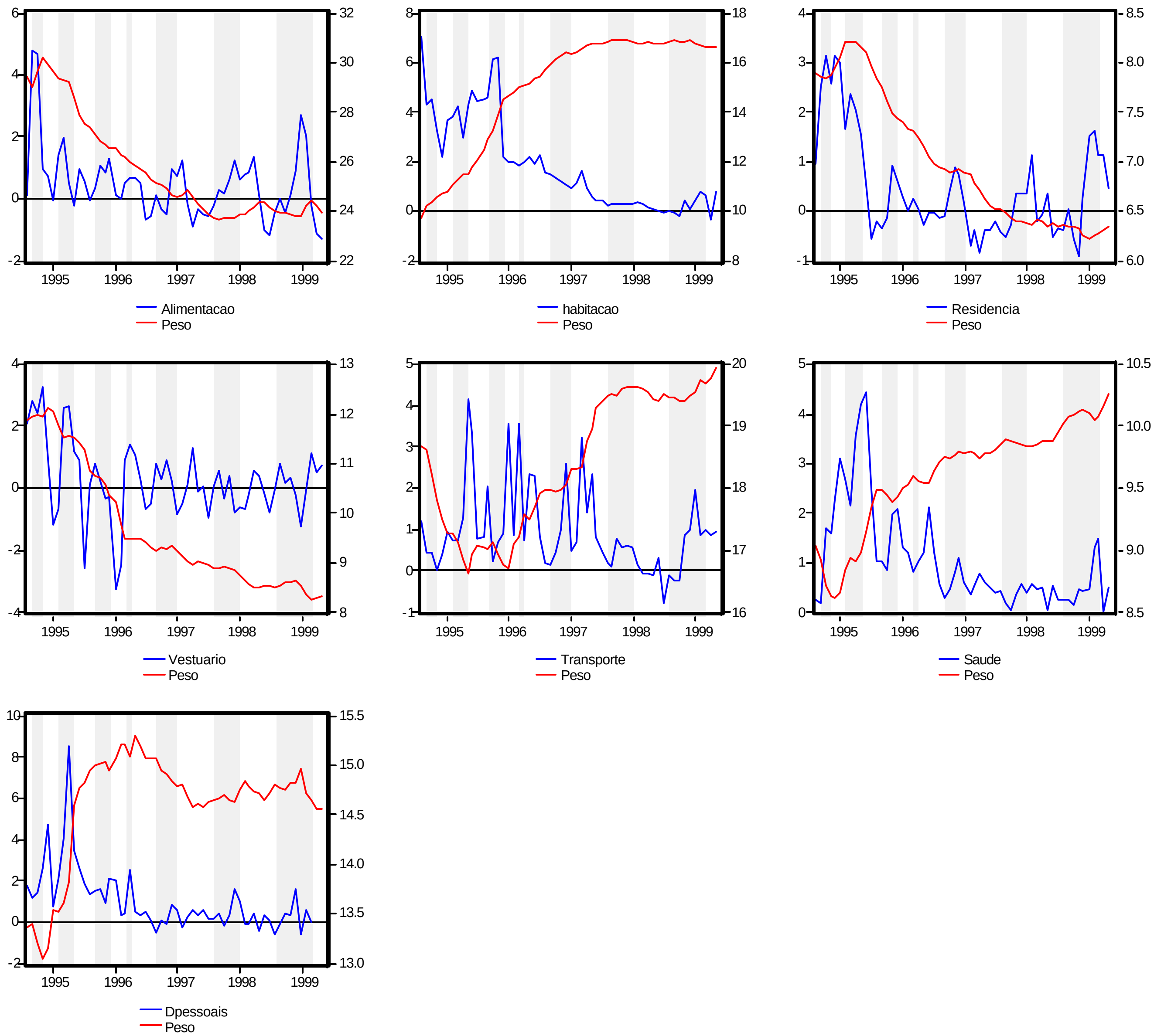
Trough-Peak	**Dating**	**Causes**
Phase 1	1994:9 – 1994:11	Seasonal + others
Phase 2	1995:2 – 1995:5	External Shock (Mexico)
Phase 3	1995:9 – 1995:12	Seasonal + others
Phase 4	1996:3 – 1996:5	Energy
Phase 5	1996:9 – 1997:1	Seasonal + others
Phase 6	1997:8 – 1998:1	Seasonal + others
Phase 7	1998:8 – 1999:3	Seasonal+External Shock
Phase 8	1999:6 – 1999:10	Energy + Seasonal

The dynamic behavior of the components of the IPCA was then examined to determine the main factors driving increases in inflation. Figure 2 plots monthly variations in each of the seven components of the IPCA as well as the corresponding changes in their underlying weight over time. The IPCA components are: food and drinks, housing, 'housing articles', clothing, transport and communication, health and personal care, and personal expenditures.[9]

Accordingly, phases 1, 3, 5, 6, and 7 correspond mainly to seasonal changes in the underlying components and internal market pressures, although other factors have contributed to both trigger and intensify the high inflation phases. In particular, food and clothing display a strong seasonal pattern, generally reflecting in the IPCA inflation from approximately August/September until the end of the year. The other inflation phases were mainly caused by either increases in energy prices and/or external shocks. Accordingly, phases 4 and 8 (and to a lessen degree phase 3) were triggered by energy shocks, while phases 2 and 7 were driven by external shocks — the financial crisis in Mexico and Russia, respectively, which led to exchange rate crises in Brazil. In particular, phase 7 combined both an external shock and a seasonal increase in inflation, which resulted in a longer high growth inflation state, from August 1998 to March 1999.

[9] The components are: *alimentação e bebidas, habitação, artigos de residência, vestuário, transporte e comunicação, saúde e cuidados pessoais e despesas pessoais.*

Figure 2 – Changes in the Components and Weights of the IPCA



3. Selection of Candidate Leading Inflation Variables

3.1 Data

The second step of the project was to gather the extensive available Brazilian data in the Central Bank as well as data sets from other Brazilian institutions, such as the Instituto Brasileiro de Geografia Econômica (IBGE), Fundação Getúlio Vargas, and the private sector. Around

200 economic variables were found as potential candidates to predict IPCA inflation. Then, a throughout research of the quality and reliability of the data was implemented.[10] The selection and treatment of the data are a crucial basic step to a robust empirical exercise, as sudden changes or different patterns in a series may arise from data handling and not from economic dynamics. Particular attention was given to changes in methodology or collection procedure that could cause spurious inferences in the series. Variables from unreliable sources or that presented changes in calculation and collection methods were excluded from the analysis. The variables were also selected according to: a) their availability at higher economic frequency available (monthly); b) their sample size (the ones that present longest history); and c) their timeliness to enable real time analysis—that is, how fast new releases of the series are available.

After assortment of the reliable data, a database was set up containing over 100 candidate leading, coincident, and lagging inflation variables. No reliable variable was overlooked in the scrutiny to determine their economic patterns and their cyclical relationship with the Brazilian inflation.[11]

3.2 Selection of the Variables

Analysis of the inflation dynamics developed in section 2 suggests four types of sources causing changes in inflation phases: exchange rate shocks, energy shocks, internal markets factors (supply and demand pressures, etc.), and seasonal changes. Accordingly, 68 variables[12] can be classified into 4 major categories:

1) Changes in input prices and variables sensitive to market conditions. These variables reflect exchange rate and energy shocks, internal supply shocks, and seasonal factors. E.g.: prices of industrial materials, energy prices, import prices, commodity prices, etc.
2) Measures of inflationary pressures from labor markets, capital markets, and commodity markets. These variables capture internal market pressures and seasonal

[10] I am thankful to the invaluable assistance of Jose Ricardo Costa e Silva in this part of the project.
[11] For more details see Chauvet (1999a, 1999b).
[12] Some of the series measure closely related definitions.

factors. E.g.: rate of capacity utilization, measurements of employment, measurements of growth in debt outstanding, etc.

3) Variables that may themselves contribute to generate economic and inflation fluctuations, such as monetary and fiscal policies. E.g.: M1, M2, monetary base, changes in taxes, etc.
4) Proxies for business expectations. E.g.: stock market prices, prices of future contracts, etc. These variables may reflect overall inflation expectations based on all available information to the market participants.

3.2.1 The Problem of Overfitting

A critical issue is how to select variables and build the leading indicators from this list of 68 variables. Thousands and even millions of combinations of the variables are possible. To illustrate the dimension of the exercise, if the variables were combined in a group of 4, this would result in 814385 possible models. If instead groups of 5, 6 or higher number of variables were used, this would yield millions of combinations of the leading variables.

On the other hand, a selection of a shorter list containing the best variables among the 68 variables runs the risk of overfitting. In fact, a search of this dimension, with the specific goal of finding the best indicators for predicting inflation can be expected to do one thing: to find a good fit to the sample period used.

Overfitting refers to the procedure of adapting a model to maximize its fit to historical data. A consequence of overfitting is that although the model may fit historical data well, it performs poorly in out-of-sample forecasting. This is because the model fits not just the signal it intends to extract, but also idiosyncrasies of historical data that are not necessarily observed in future sample data.[13] In fact, overfitting can easily lead to wild unreasonable predictions and large variances of the forecast error.

---

[13] The problem of over-fitting can be illustrated as follows: we are given a set of data points that we want to fit with a function. Now, from numerical analysis we know that we can fit the data exactly with a polynomial of high degree. However, this does not tell anything about its behavior outside of data sample used. Generally, the polynomial may behave wildly between successive points on the grid. This phenomenon is exactly what makes it difficult for the model to perform well outside of the sample. In fact, over-specialized functions merely memorize the sample data used, and thus does not generalize well for new observations.

The idea here is to obtain a general model of the behavior of the variables that will do well also on unseen data. In fact, the goal is to use the leading indicators of inflation for forecasting in real time, so that they can be used as an informative tool for monitoring monetary policy in a month-to-month basis. Thus, the critical issue in searching for candidate leading variables is to understand how well they would make predictions for cases that are not in the sample used. The best way to minimize overfitting, and hence get more realistic estimates, is to select the variables (and the implicit models to select the variables) depending on their out-of-sample forecasting performance. To ensure that, the variable selection process was recursively re-estimated through the sample period. That is, the procedures were estimated repeatedly, using larger and larger subsets of the sample data. The first estimation was obtained for the first n observations, where n is equal to the number of parameters in the model. For each subsequent month, the models were recursively re-estimated, and the process was repeated until the end of the sample. For each re-estimation of the model, the estimates of the parameters were used to predict the one-step-ahead forecast value of the dependent variables, and to compute the one-step ahead forecast error.

Then, for each of the procedures described below, root mean squared error, Theil inequality coefficients, and mean absolute percentage error were used as criteria to classify the variables according to their incremental predictive power out-of-sample. This procedure allows better understanding on how well the models would have performed if they had been applied month by month in real time.

3.2.2 Linear Procedures

Several econometric procedures were implemented to select and rank the variables that lead inflation. First, all series were transformed to achieve stationarity and were normalized to have mean zero and unity variance.[14]

[14] The Augmented Dickey-Fuller (1979) and Phillips-Perron (1988) tests were used to test for unit roots. In addition, Perron's (1989) test was also used for the null of integration against the alternative of deterministic trend in the presence of a structural break. In the case of deterministic trends, the best specifications were selected using Akaike Information Criteria and BIC criteria. One of the problems of this analysis is that the sample is small. Variables that

The variables were then classified according to: a) their ability to Granger-cause inflation; b) their marginal predictive content for inflation; c) their bivariate relation with inflation—cross-correlation in time domain, and coherence and phase lead in frequency domain;[15] and d) their ability to anticipate the peaks and troughs of the inflation process.

In addition to Granger causality tests and cross-correlations, some autoregressive systems were examined to assess the marginal predictive content of the variables for inflation. Although it is desirable to use VARs with a large number of lags and variables to forecast macroeconomic variables, in practice the number of observations available does not allow much flexibility in this exercise. Thus, VAR systems were used with inflation and a small number of variables and lags, as well as univariate models with inflation and a larger number of lags and variables.[16]

This yields several leading indexes. Using these basic frameworks, alternative additional variables were included one at a time in the autoregressive systems.[17] Then, it was verified whether lags of the additional variables help predict inflation beyond what other variables and lags of inflation itself already predict, using Schwarz Information Criteria (SIC), and Akaike information Criteria (AIC). The exercise was repeated recursively, and the forecasting performance of the models were evaluated out-of-sample using the root mean squared error, Theil inequality coefficients, and mean absolute percentage error, as described above.

A throughout examination of the data according to this procedure yielded a ranking of the 68 leading variables of inflation for Brazil, based on the optimality of linear one-step-ahead least squared predictors in the out-of-sample exercise. However, some variables that that did not perform well according to the linear criteria were not eliminated if economic theory

---

have stochastic trends may appear to have deterministic trends in a sub-period, and this information will only be revealed as the sample size increases. Thus, these tests should be revised as more observations are available.

[15] Spectral analysis requires a sample size four times larger than the available. This technique was mainly applied for the longer sample from 1980:1 to 1999:12 (using subsamples to avoid nonstationarities arising from the several structural breaks during this period).

[16] Since the number of parameters increases rapidly with the number of lags, even systems of moderate size become overparameterized relatively to the total number of observations. This leads to poor and inefficient estimates of the short-run cyclical features of the data. However, if the lags are too small, the residuals may contain important relevant information for the variables and only part of the available information is used to characterize the data. As a consequence, this leads to spurious significance in the coefficients.

[17] Details regarding the basic models and variables included in the analysis can be found in Chauvet (1999a, 1999b).

suggested that they should have some predictive content for inflation. It could be the case that in this small sample they did not do so well due to some major changes in the inflation dynamics, such as the currency crisis in 1995 and in 1998-1999. In fact, next section shows evidence that the correlation between inflation and these variables display a structural break in 1998-1999. Thus, as more observations are collected they may prove to be good candidate leading variables for inflation.

The procedure undertaken here is similar to the NBER approach and the one pursued by Stock and Watson (1989, 1991), which list a large number of variables and reach a shorter list of variables that enter their leading indicators as weighted averages. A critical difference is that the procedure in this project is based on out-of-sample forecasting performance, rather than the predictive content in-sample. The idea is to avoid overfitting and produce reasonable forecasts in real time. The main criticism of the leading indicators proposed by Stock and Watson (1989, 1991) was their reliance on variable selection s and, ultimately, the selection of the leading indicator, based on in-sample performance.[18]

3.2.3 Non-Linear Procedures

One important drawback of the linear approach to causality testing and marginal predictive content is that such models can have low power detecting certain kinds of nonlinear causal relations. The main goal of the leading indicators of inflation is to give early out-of-sample signals of peaks and troughs of inflation, which Granger causality tests and linear autoregressive systems can fail to uncover.

Thus, the nonlinear relationship of each of the series was studied using probability methods to determine if they anticipate the peaks and troughs of the inflation process. In particular, different specifications of two-state first-order Markov switching models were fitted for each of the candidate leading variables.[19] The estimated probabilities of high or low states

[18] In fact, Stock and Watson's leading economic indicator was first released in 1988, and failed to forecast the subsequent U.S. recession in 1990.

[19] Different specifications includes or not seasonal factors, switching intercepts, switching volatility, switching autoregressive parameters, and different autoregressive processes.

for each series were used in an analysis of the nonlinear lead-lag relationship with inflation IPCA. In particular, the growth phases of the leading variables were compared to the growth phases of inflation using the Quadratic Probability Score, which is a nonlinear counterpart for the mean squared error. Again, the models were re-estimated recursively and the filtered probabilities of high growth phase were computed for each date in an out-of-sample exercise.

Turning points analysis indicates that some of the variables that were ranked low in the previous linear exercise actually display good nonlinear predictive power in terms of forecasting inflation turning points, rather than the level of inflation. This result will be further discussed in the next section.

3.2.4 Structural Change

One problem of using linear models such as Granger causality, VAR models, and linear regressions is that they can be sensitive to nonstationarities associated with structural breaks. In fact, it is important that in the periods studied the variables can be considered stationary, otherwise the correlation between inflation and the leading variables may display structural breaks around times in which monetary policy procedures changed, such as in 1999. This could be one of the reasons why there were discrepancies in the findings using linear versus nonlinear models to classify the leading variables.

In this section, structural stability tests are used to estimate a break in the inflation process around the period of the currency crisis in late 1998 early 1999. From previous results in section 2, inflation π_t is represented by an AR(1) process, and the inflation process is tested for structural breaks in its mean and autoregressive parameters:

$$\pi_t = \mu_1 D_{1t} + \mu_2 D_{2t} + \phi_1 \pi_{t-1} D_{1t} + \phi_2 \pi_{t-1} D_{2t} + \varepsilon_t \qquad (2)$$

where ε_t is distributed normal and

$$D_1 = \begin{cases} 0 & \text{if } t \leq T \\ 1 & \text{if } t > T \end{cases}, \qquad D_2 = \begin{cases} 1 & \text{if } t \leq T \\ 0 & \text{if } t > T \end{cases}.$$

where t = 1998:05, 1998:06,…, 1999:05.

First, a jointly test of a break in both the mean and the coefficient on lagged inflation is performed. Then tests for breaks in the mean and the lag coefficient separately are implemented. The null of no break cannot be rejected for the autoregressive parameter, but it is rejected for the mean, using LM test. Using a Chow test with the estimated break date imposed on 1998:11, the null of no break for the mean is also rejected.

The possibility that there is a break in the residual variance is also examined using procedure suggested in McConnell and Perez-Quiros (1998). The following model is jointly estimated using GMM:

$$\pi_t = \mu_1 + \phi_1\pi_{t-1} + \varepsilon_t \tag{3}$$

$$\sqrt{\frac{\pi}{2}}\left|\hat{\varepsilon}_t\right| = \alpha_1 D_{1t} + \alpha_2 D_{2t} + \mu_t$$

where t = 1998:05, 1998:06,…, 1999:05, ε_t is distributed normal and $\sqrt{\frac{\pi}{2}}\left|\hat{\varepsilon}_t\right|$ is an unbiased estimator of the standard deviation of ε_t. The null of no break is rejected for the variance using LM test. Again, using a Chow test with the estimated break date imposed on 1998:11, the null of no break for the variance is also rejected. Thus, there is evidence of a break in both the mean and variance of inflation around 1998:11.[20] The apparent break in the series is relatively recent, and the tests should be implemented again as more observations become available.

However, given the evidence of structural break in the inflation dynamics around 1999, the results of selecting variables and specifications based on linear models of Granger causality, VARs, and regressions should be interpreted with caution. The procedure used here partially overcome this problem, since the variables were selected based on their out-of-sample recursive forecasting ability.

4. Models for the Leading Indicators of Inflation

4.1 The Dynamic Factor Model

The leading indicators of inflation are constructed from a dynamic factor model, using an approach similar to the ones developed in Chauvet (2000b). The dynamic factor is a latent

[20] The same tests were applied to detrended inflation, and the evidence in this case is for a structural break only in its variance.

variable that summarizes comovements of some variables that lead the Brazilian inflation as measured by changes in the IPCA. It is a signal-noise extractor that filters out idiosyncratic sectoral movements in the observable variables from common cyclical movements related to the inflation process. The dynamic factor model is:

$$y_t = \delta + \Lambda(L) F_t + \varpi_t \tag{4}$$

$$\Phi(L)F_t = \gamma + \upsilon_t \tag{5}$$

where y_t is the nx1 vector of observable economic variables that exhibit predictive power in forecasting inflation, δ and γ are constant terms, Λ is the vector of factor loadings, and F_t is the scalar dynamic factor. $\Lambda(L)$ and $\Phi(L)$ are finite lag polynomials and L is the lag operator and Δ=1-L. Anticipating the empirical results in section 6, for most of the variables considered unit roots tests cannot reject the null hypothesis of integration. Further, a stochastic trend is not included in the dynamic factor based on evidence that the series studied are integrated but not cointegrated. Thus, the model is transformed using the first difference of the observable variables, Δy_t:

$$\Delta y_t = \beta + \Lambda(L)\, lii_t + \varepsilon_t \qquad \varepsilon_t \sim \text{ i.i.d. } N(0, \Sigma) \tag{6}$$

$$\Phi(L)\, lii_t = \alpha + \eta_t \qquad \eta_t \sim \text{i.i.d. } N(0, s_h^2) \tag{7}$$

where $\varepsilon_t = \Delta\varpi_t$ are the nx1 measurement errors, $\eta_t = \Delta\upsilon_t$ is the scalar transition shock, and $lii_t = \Delta F_t$ is the scalar dynamic factor, that is, the Leading Indicator of Inflation. Notice that in this specification, the sample mean of y_t does not separately identifies β and α. A simple way to solve this problem is to write the model in deviations from means, thus, "concentrating out" of the likelihood function the constant parameters in equations (6) and (7).[21] The model used in the empirical analysis is:

$$\Delta Y_t = \Lambda\, LII_t + \varepsilon_t \qquad \varepsilon_t \sim \text{ i.i.d. } N(0, \Sigma) \tag{8}$$

$$LII_t = \Phi\, LII_{t-1} + \eta_t \qquad \eta_t \sim \text{i.i.d. } N(0, s_h^2) \tag{9}$$

where $\Delta Y_t = \Delta y_t - \Delta\overline{y}$, and $LII_t = \Delta lii_t - \alpha/(1-\phi)$. For identification of the dynamic factor, a scale has to be assigned to it. This can be achieved by normalizing the factor variance or one of

---

[21] An alternative way to identify the parameters is by imposing restrictions on their relationship.

the factor loadings to one. In the estimation exercise the factor variance, s_h^2, is set to one, and the variables are transformed as deviation from their means divided by their standard deviation.

The model assumes that $H_t \sim \begin{bmatrix} \sigma_\eta^2 & 0 \\ 0 & \Sigma \end{bmatrix}$ and Σ are diagonal, which implies that the leading inflation indicator (LII_t) and the nx1 vector ε_t are mutually uncorrelated at all leads and lags. Thus, the dynamic factor is driven by, η_t, the shocks common to all observed variables, ΔY_t. Sector-specific shocks, ε_t, are idiosyncratic movements inherent to the observable variables, and they do not affect the dynamic factor. The output of the model is the Leading Inflation Indicator, LII_t, constructed as a combination of the underlying observable variables ΔY_t, using the Kalman filter. The elements of the vector Λ correspond to the factor loadings, which measure the sensibility of each of the ΔY_t series to the leading inflation indicator LII_t.

Given the above assumptions, all the observational information for identification of the model is subsumed in the covariance matrix of the observable variables, and necessary and sufficient conditions for identification of all the model parameters are met.[22]

4.2 The VAR Model

Although the primary goal of the leading indicator is in anticipating turning points, it can also be used to form linear forecasts of inflation. The leading indicator of inflation is composed of variables that anticipate the inflation process such as price of inputs and energy, index of imported prices, price of sensitive materials, measures of demand pressure, prime movers such as fiscal or monetary policy changes, or forward-looking variables that reflect business expectations.[23] By itself, it can not be used to give a linear forecast of inflation. However, the leading indicator can be combined with inflation in vector autoregressions to provide a linear forecast of inflation:

[22] See Anderson and Rubin (1956), Bollen, and Joreskog (1985), Bollen (1989), Deistler (1993), Dunn (1973), Fisher (1966), Rothenberg, (1971), Geweke (1977), and Sargents and Sims (1977).

[23] See the Activity Report I for a more detailed discussion.

$$\Delta IPCA_t = a_1 \Delta IPCA_{t-1} + \ldots + a_p \Delta IPCA_{t-p} + b_1 LII_{t-1} + \ldots + b_p LII_{t-p} + \zeta_t \qquad (10)$$

$$LII_t = c_1 \Delta IPCA_{t-1} + \ldots + c_p \Delta IPCA_{t-p} + d_1 LII_{t-1} + \ldots + d_p LII_{t-p} + \nu_t \qquad (11)$$

$$\zeta_t \sim \text{i.i.d. } N(0, \sigma_\zeta^2) \qquad\qquad \nu_t \sim \text{i.i.d. } N(0, \sigma_\nu^2)$$

where ζ_t and ν_t are serially uncorrelated error terms.

The IPCA is projected forward h-step ahead to obtain linear forecasts using the history of inflation and the dynamic factor to predict its future values in this VAR system.

5. Estimation Procedure

The estimation is implemented using the Kalman filter. The model is first cast in state space as:

$$\Delta Y_t = \Lambda LII_t + \varepsilon_t \qquad \text{Measurement Equations}$$

$$LII_t = \Phi \, LII_t + h_t \qquad \text{Transition Equations}$$

The objective of the Kalman filter is to form forecasts of the unobserved state vector and the associated mean squared error matrices (MSE) at t based on information available up to time t-1, $I_{t-1} \equiv [\Delta Y'_{t-1}, \Delta Y'_{t-2}, \ldots, \Delta Y'_1]'$:

$$LII_{t|t-1} = E(LII_t \mid I_{t-1})$$

$$P_{t|t-1} = E[(LII_t - LII_{t|t-1})(LII_t - LII_{t|t-1})' \mid I_{t-1}].$$

The Kalman filter is a set of recursions that, given an initial state estimate LII_0 with MSE P_0, it provides linear least square predictions $LII_{t|t-1}$ and updates LII_t, along with the corresponding MSE matrices $P_{t|t-1}$ and P_t. That is, given the parameters in Λ, Φ and H, the filter uses as inputs an inference about the state vector using information up to t-1, $\{LII_{t-1|t-1}\}$; and the mean squared error matrices, $\{P_{t-1|t-1}\}$. The outputs are their one-step updated values. The algorithm is:

Step 1: Initial state estimate and MSE

$$LII_0 = E(LII_0)$$

$$P_0 = E(LII_0 - \hat{L}II_0)(LII_0 - \hat{L}II_0)'$$

Step 2: one-step-ahead state prediction and MSE (prediction recursions):

$$LII_{t|t-1} = \alpha + \Phi LII_{t-1|t-1}$$
$$P_{t|t-1} = \Phi P_{t-1|t-1}\Phi' + \sigma^2$$

Step 3: extraction and MSE (updating recursions):

$$LII_{t|t} = LII_{t|t-1} + K_t N_{t|t-1}$$
$$P_{t|t} = (I - K_t Z)P_{t|t-1}$$

where:

$K_t = P_{t|t-1}\Lambda'[Q_t]^{-1}$, $N_{t|t-1} = \Delta Y_t - \Lambda LII_{t|t-1}$ is the conditional forecast error of ΔY_t, and $Q_t = \Lambda P_{t|t-1}\Lambda' + \Sigma$ is the conditional variance of $N_{t|t-1}$.

Step 4: maximize the likelihood function:

$\text{Log } f(\Delta Y_T, \Delta Y_{T-1}, \ldots \mid I_0) =$

$$\sum_{t=1}^{T} \log f(\Delta Y_t \mid I_{t-1}) = \sum_{t=1}^{T} \log\{(2\pi^{-n/2} \mid Q_t \mid^{-1/2} \exp(-\frac{1}{2} N'_{t|t-1} Q_t^{-1} N_{t|t-1})\}$$

The filter evaluates this likelihood function, which can be maximized with respect to the model parameters using a linear optimization algorithm. The parameters estimated and the sample data are then used in a last application of the filter to draw inferences about the dynamic factor based on information available at time t.

The parameters of the model are estimated as follows: the model is cast in state-space form, where equations (5) and (6) are, respectively, the measurement and transition equations. Then, the Kalman algorithm is applied to construct an optimal linear prediction of the latent dynamic factor. The filter tracks the course of the dynamic factor, which is calculated using only observations on ΔY_t. It computes recursively one-step-ahead predictions and updating equations of the dynamic factor and the associated mean squared error matrices. The output is the leading inflation indicator, $LII_{t|t}$, which is an optimal estimator of the state vector constructed as a linear combination of the variables ΔY_t, using information available through time t. As new information becomes available, the Kalman filter can be applied to update the leading indicator on a real time basis.

6. Empirical Results

6.1 Model Selection and Specification

After consideration of structural breaks, the linear and nonlinear out-of-sample procedures were used to rank the 68 leading variables, with particular attention to changes around the structural break in the inflation series around the end of 1998.

As mentioned above, if the variables were combined in a group of 4, this would result in 814385 possible models. However, some of the series measure closely related definitions. For example, there are 4 variables measuring capacity utilization, 15 measurements of employment, etc. Of course, combinations of variables that reflect only one type of measurement should be excluded. For example, an indicator composed only of employment variables would miss large part of the dynamics of inflation. In fact, many combinations simply lack economic content as leading indicators of inflation. Thus, a first guideline in the combination of the variables was not to include more than one variable measuring closely related concepts. This reduces the search to 10626 possible models.

An important criterion in the combination of the variables is the historical conformity and the relationship of the leading variables with the reference inflation cycle as to the timing of changes. The dynamic factor model is designed to extract common cyclical movements underlying the observable variables. This implies that the variables composing each indicator should exhibit a similar lead-lag relationship with inflation. That is, cyclical movements in each of the four variables composing the leading indicator should coincide. For example, a variable that anticipates inflation movements with a lead of 4 to 7 months should be combined with others with approximately the same forecasting lead. If this criterion is not met, the upturn in one variable may offset a lagged upturn in the other variables and there is no common cyclical movement to be summarized by the dynamic factor. In this case, the Kalman filter will either not converge or will converge to one of the candidate variables. In fact, a random search of different combinations of the variables will lead exactly to that in the majority of the cases.

Thus, implicit in the dynamic factor model is the conformity of the lead-lag relationship of the variables with inflation. This reduces the number of possible combinations to just a few hundreds.

6.2 Classification of the Best Leading Indicators of Inflation

As in the procedure to select the variables, a major concern in all steps of this project was to avoid overfitting the data. Thus, in the construction of the leading indicators of inflation, the models were also recursively re-estimated out-of-sample and, for each re-estimation of the model, the estimates of the parameters were used to compute one-step-ahead forecast values and forecast errors of the dynamic factors. Then, scale invariant statistics, such as Theil inequality coefficients and the mean absolute percentage error were used as criteria to rank the leading indicators according to their forecasting ability out-of-sample.

An overriding criterion is the ability of the indicators to anticipate inflation turning points. This requires analysis of the lead-lag relationship of the indicators with inflation. The procedure utilized was as follows. First, Markov switching models were fitted to the indicators. Different specifications were estimated allowing the switching mean, switching variance, or both. Then, specification tests were applied to study whether the switches reflect changes from low/high phase or whether permanent changes (structural breaks) characterize the indicators. If the switching reflects short-run changes in regimes, the filtered probabilities were then used to determine turning points.[24] However, if the filtered probabilities reflect instead a major switch in the mean or variance of the indicators around some specific dates (possibly related to currency crises), then the filtered probabilities from the Markov switching model were used to segment the different volatility periods. Upper and lower bounds thresholds were established as the mean minus/plus half the standard deviation of the series, where the standard deviation assumes pre and post break values. These thresholds were used to date turning points of low/high growth phases of the leading indicators. Then, 0/1 dummy variables were constructed, where the value of 1 indicates high growth phases. Finally, after the turning points were determined, the

Quadratic Probability Score at different leads was used to compare the filtered probabilities of high inflation growth obtained from equation (1) with the 0/1dummies.[25] The QPS is a nonlinear counterpart to the Mean Squared Error, and corresponds to a loss function in which the turning points of the leading indicators of inflation are compared to the IPCA inflation turning points at different leads. Again, this analysis of turning points was implemented out-of-sample.

6.3 Analysis of the Top Five Leading Indicators

The linear forecasting ability of the leading indicators and the analysis of turning points based on out-of-sample exercises were used to classify the top 20 leading inflation indicators. In this section, we explore the ability of the indicators in predicting inflation turning points using information available in real time.

First, the ability of the leading indicators in anticipating inflation turning points is examined using full sample information. For historical analysis, the models were estimated using data from 1994.08 to 1999.07. The adequacy of the model specification was verified through analysis of the whiteness of the one-step-ahead forecast errors and dynamic multipliers behavior. The diagnostic tests indicated that the specification selected were adequate for all equations.[26] In addition, the autocorrelation functions for the disturbances are within the limit of two times their asymptotic standard deviation.[27]

Second, the parameters were estimated up to 1998.07, and the estimates obtained recursively from 1998:08 to 2000:03 were used to generate forecasts and examine out-of-sample performance of the leading indicators in predicting inflation turns. This tests the ability of the models in predicting out-of-sample even when major events such as the Brazilian currency crisis in January 1999 are excluded from the sample.

---

[24] For example, a peak occurs if the probabilities of high growth phase fall above their mean plus one-half their standard deviation.

[25] A Bayesian procedure was also implemented, as described in Chauvet (1999b).

[26] The hypothesis of cointegration was tested using Stock and Watson's (1988) test and Engle and Granger (1987) pairwise test.

[27] However, this holds marginally for some indicators, as discussed in Chauvet (1999b). An interesting extension would be to model AR(1) processes for the idiosyncratic terms of some of the observable variables.**

Finally, the performance of the leading indicators in predicting cyclical turns of inflation is examined, using only real time data available at the date of each forecast. An important aspect of the criteria adopted to build the leading inflation indicator is the possibility of real-time prediction and monitoring of the inflation cycle. The idea here is to reproduce the forecasting problem the Central Bank faces when only preliminary data is available. In this part, the parameters of the dynamic factor model were estimated using data up to 1999.05. For each subsequent month, the model was re-estimated and only real time unrevised data were used to generate out-of-sample forecasts of the filtered dynamic factors from 1999.06 through 2000:03. Real time data for the economic variables correspond to preliminary and unrevised data.[28]

6.3.1 Turning Point Analysis

The top 20 leading indicators display similar cyclical movements. In fact, they can be classified into 5 major groups, according to their common cyclical dynamics. In order to represent a broad spectrum of cyclical movements, the analysis below reports the top 5 leading indicators of inflation from each of the five groups. Thus, analysis of turning point prediction as well as linear forecasting performance will be examined for the leading indicators of inflation F2, F6, F8, F18, and F23.[29]

Figures 3 to 5 plot the five leading indicators of inflation against the IPCA inflation and its turning points. A visual inspection reveals that the indicators anticipate all inflation turns. This will be carefully examined in the analysis of turning points below.

---

[28] These data were obtained from several issues of the publication 'Indicadores Econômicos" published by the DEPEC/BACEN, and from continuously collection by the DEPEP. One of the important criteria for selecting variables to compose the leading indicator is their prompt availability (timeliness), if real time analysis is to be performed. Many variables that are good candidates are released with a long delay and, therefore, were not included in the analysis.

[29] The top 5 leading indicators can actually be further divided into three groups according to the similarities in their cyclical movements: leading inflation indicators LIIF18 and LIIF23 as one group; LIIF2 and LIFF6 as another group, and LIIF8.

Figure 3 – Leading indicators of inflation LIIF18, LIIF23, and High Growth Phases of the

IPCA Inflation (Shaded Area)

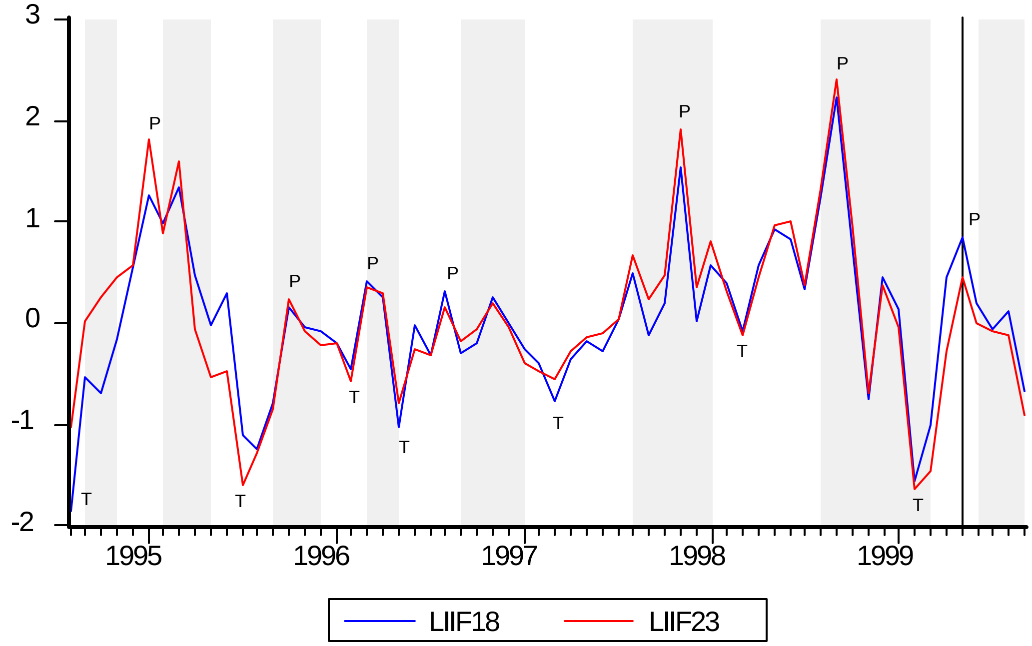


Figure 4 – Leading indicators of inflation LIIF2, LIIF6, and High Growth Phases of the IPCA Inflation (Shaded Area)

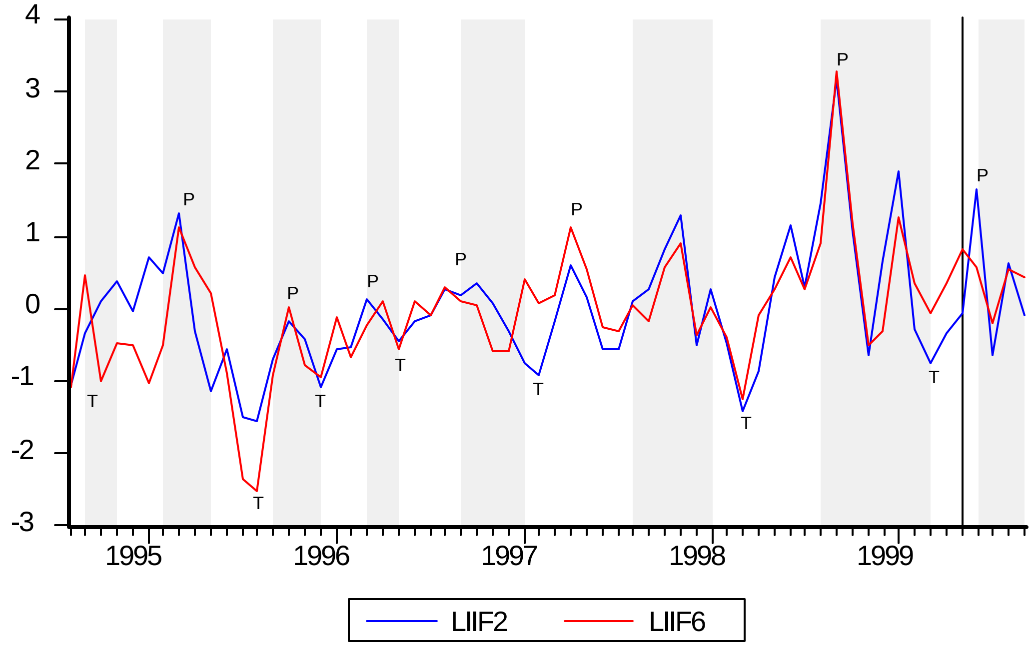


Figure 5 – Leading indicators of inflation LIIF8, and High Growth Phases of the IPCA Inflation (Shaded Area)

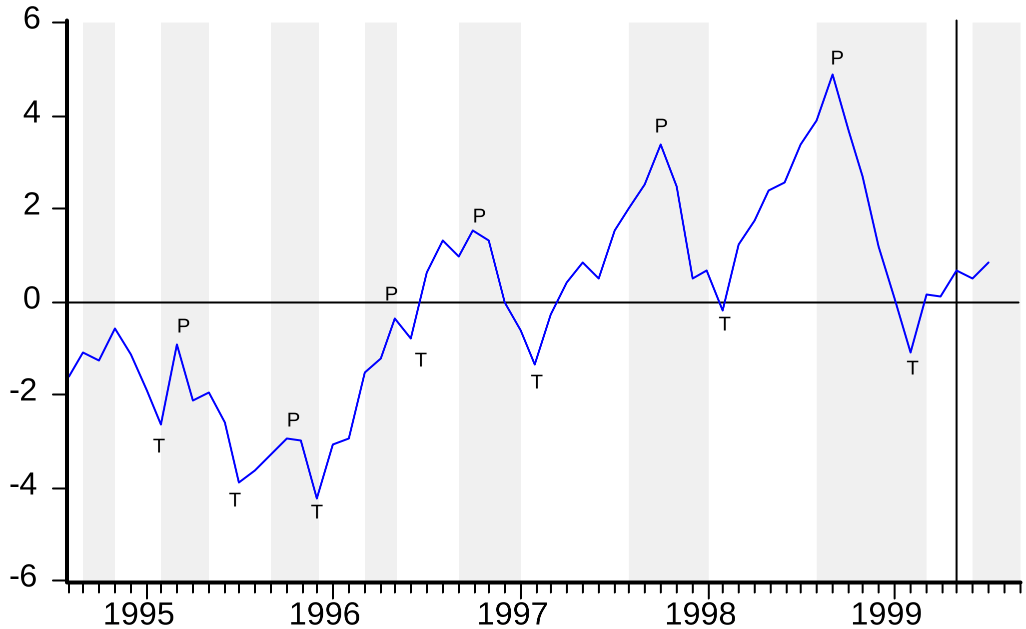


Table 2 shows how consistently the indicators turn before inflation turning points. All five indicators signal all 14 inflation turns with different leads. In addition, the leading indicators do not exhibit multiple spikes around turning points. The median lead of the turns is around 5 months, while the average lead of the indicators is around 4 months with a standard deviation of 1.9 month.

There are two types of turning point errors: predicting a turning point when one does not occur, and predicting no turning point when one does occur. A perfect forecast is obtained when these two errors are zero. Table 3 summarizes evaluation of the turning point signals of the leading indicators. The leading indicators LII F2 and LII F6 signal one false peak and one false trough, while the LII F18 and LII F23 gives two false peaks and two false troughs. The performance of these indicators is based on out-of-sample selection of the variables and models. Thus, the results indicate a very good performance of the leading indicators in predicting turning points – each of them would have signaled correctly all the peaks and troughs of the inflation phase in an out-of-sample exercise. Cautious should be exercised for false signals from the indicators. However, there were only 2 to 3 false signals out of 14 peaks and troughs – an occurrence in only 15% of the turning point events.

Table 2- Signals of IPCA Inflation Turning Points from the Leading indicators of inflation

Troughs and Peaks of the IPCA Inflation	LII F2	LII F6	LII F18	LII F23	LII F8
In-sample					
1995:2 (T)	-6	-6	-6	-6	-1
1995:5 (P)	-2	-2	-4	-4	-2
1995:9 (T)	-3	-3	-1	-2	-2
1995:12 (P)	-2	-2	-2	-2	-2
1996:3 (T)	-3	-3	-1	-1	-8
1996:5 (P)	-2	-1	-2	-2	-1
1996:9 (T)	-4	-4	-4	-4	-4
1997:1 (P)	-5	-5	-5	-5	-3
1997:8 (T)	-6	-1	-5	-5	-6
1998:1 (P)	-7	-7	-5	-5	-3
Out-of-sample					
1998:8 (T)	-5	-5	-5	-5	-6
1999:3 (P)	-6	-6	-6	-6	-6
Real time:					
1999:6 (T)*	-3	-3	-6	-6	-6
1999:10 (P)*	-4	-5	-5	-5	NA
Average Lead	-4.1	-3.8	-4.1	-4.1	-4.1
Standard Deviation	1.7	1.9	1.8	1.7	2.2
Median	-4.0	-3.5	-5.0	-5.0	-3.5

The (-) sign indicates leads, that is, how many months ahead the indicator signal an inflation peak or trough.
The criterion adopted to determine turning points is whether the series display growth plus or minus one half their standard deviation.
(*) Results from real time analysis using only unrevised data from 1997.06 to 2000.03.

Table 3- Evaluation of Turning Point Signals

Turning Point Evaluation	LII F2	LII F23	LII F18	LII F8	LII F6
Correct TP	14	14	14	14	14
Correct TP with lead	14	14	14	14	14
Missed TP	0	0	0	0	0
False Peaks	1	2	2	0	1
False Troughs	1	1	1	0	1

A missed TP occurs when the indicator does not signal inflation turns at any lead

Table 4 compares the accuracy of the indicators in predicting inflation turning points, using the Quadratic Probability Score (QPS):

$$QPS = \frac{2}{T}\sum_{t=1}^{T} [\hat{N}_t - N_t]^2$$

where $\hat{N}_t$ is a 0/1 dummy variable that takes the value of one if the series are above a threshold determined by plus half the standard deviation of the growth of each series. N_t is the probabilities of high inflation phase obtained from equation (1). The QPS ranges between 0 and 2, with the maximum accuracy corresponding to zero. The leading indicators LII F23 and LII F18 display the lowest QPS for most horizons. The smallest QPS for all indicators is found around the 6-month horizon, although it is also small at the 12-month horizon. Recall that the QPS is a counterpart for the mean squared error. Thus, this result is equivalent to say that the loss function associated with event timing forecast is minimized at 6-step ahead. Based on this result, the best use of the leading indicators is to forecast inflation turning points 6 months ahead.

The performance of the leading indicators is compared with a benchmark model. Take $\hat{N}_t$ to be a constant equal to the historical fraction of quarters for which the economy was in a high growth inflation phase. The QPS in this case is equal 0.61. Thus, the leading indicators display better out-of-sample performance compared to naïve forecasting model.

Table 4 - Evaluation of In-Sample Peak Forecasts of the IPCA Inflation Using the QPS

Forecast Horizon	LII F2	LII F23	LII F18	LII F8	LII F6
0-month	0.43	0.44	0.44	0.42	0.43
1-month	0.41	0.42	0.42	0.45	0.41
2-month	0.44	0.47	0.47	0.41	0.43
3-month	0.46	0.41	0.41	0.39	0.45
4-month	0.32	0.31	0.31	0.42	0.31
5-month	0.30	0.29	0.29	0.45	0.29
6-month	0.29	0.26	0.26	0.31	0.27
7-month	0.32	0.34	0.34	0.39	0.32
8-month	0.28	0.35	0.35	0.48	0.28
9-month	0.35	0.34	0.34	0.51	0.34
10-month	0.31	0.39	0.39	0.53	0.31
11-month	0.29	0.36	0.36	0.44	0.30
12-month	0.28	0.30	0.30	0.38	0.27

6.3.2 Recent Performance of the Leading indicators of inflation – Turning Point Analysis in Real Time

Revisions of the series that compose the leading indicators are sometimes substantial. Large revisions are made in subsequent releases of the series that compose the leading indicators. This suggests that a reliable prediction of turning points in real time is more difficult due to the availability of only preliminary and unrevised data.

Nonetheless, all indicators predict all inflation turning points in a real time exercise. The last two rows of Table 2 reports the result for out-of-sample real time analysis. On average, the leading indicators signal inflation turns with a shorter lead in the out-of-sample exercise.

Thus, turning point analysis indicates that leading indicators have been proving to be informative about futures phases of inflation cycles in real time out-of-sample, and it can be a useful monitoring tool for monetary policy in a current basis.

Figure 6 – Real Time: Leading indicators of inflation LIIF2, LIIF6, and High Growth Phases of the IPCA Inflation (Shaded Area)

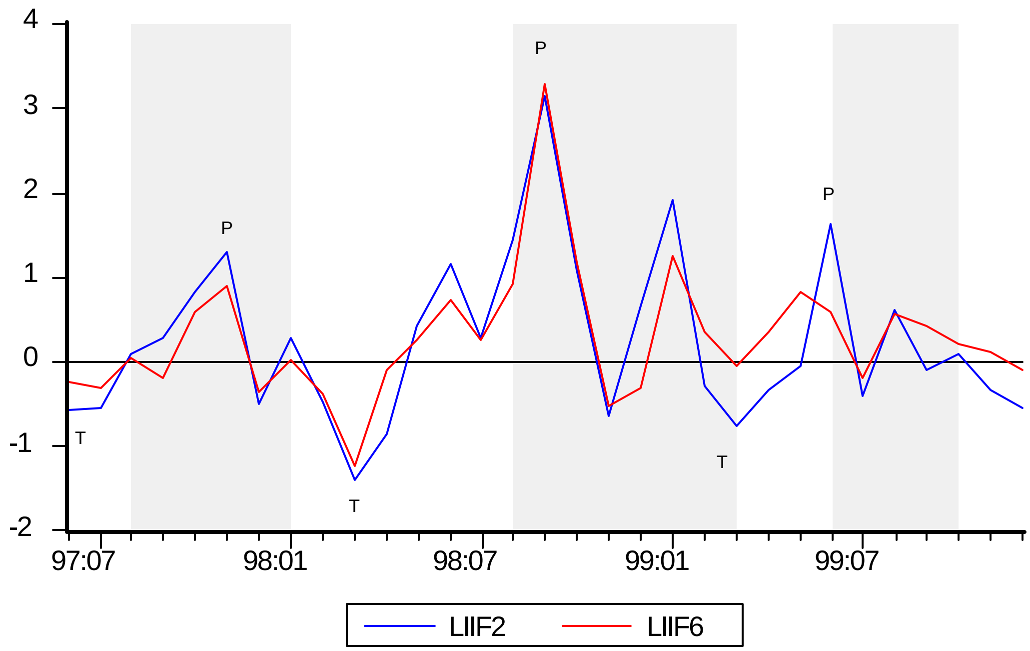


Figure 7 – Real Time: Leading indicators of inflation LIIF2, LIIF6, and High Growth Phases of the IPCA Inflation (Shaded Area)

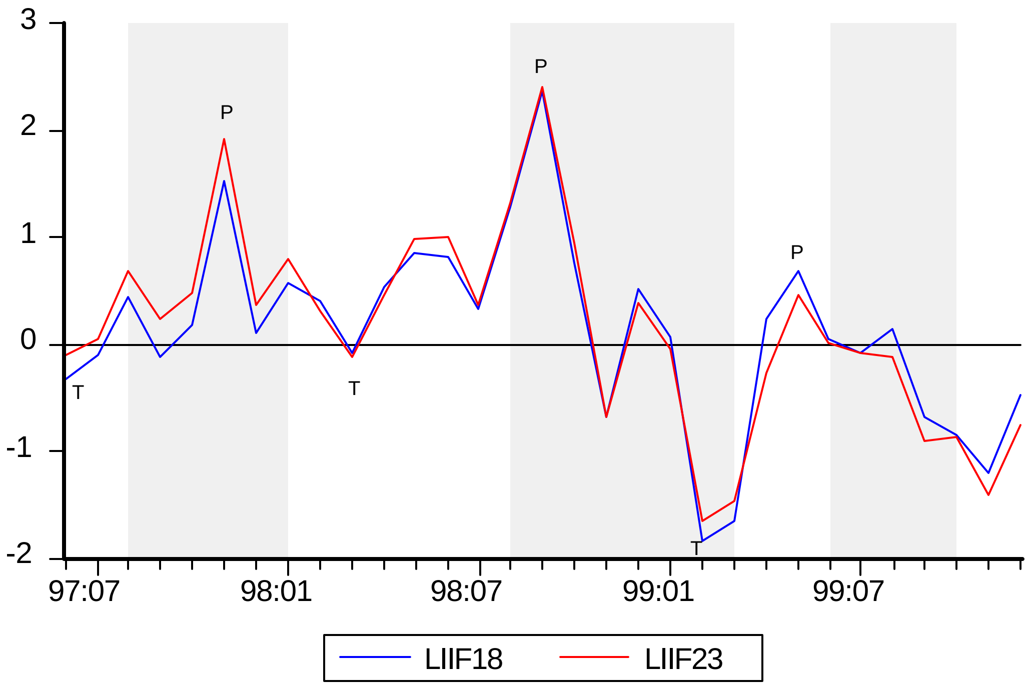
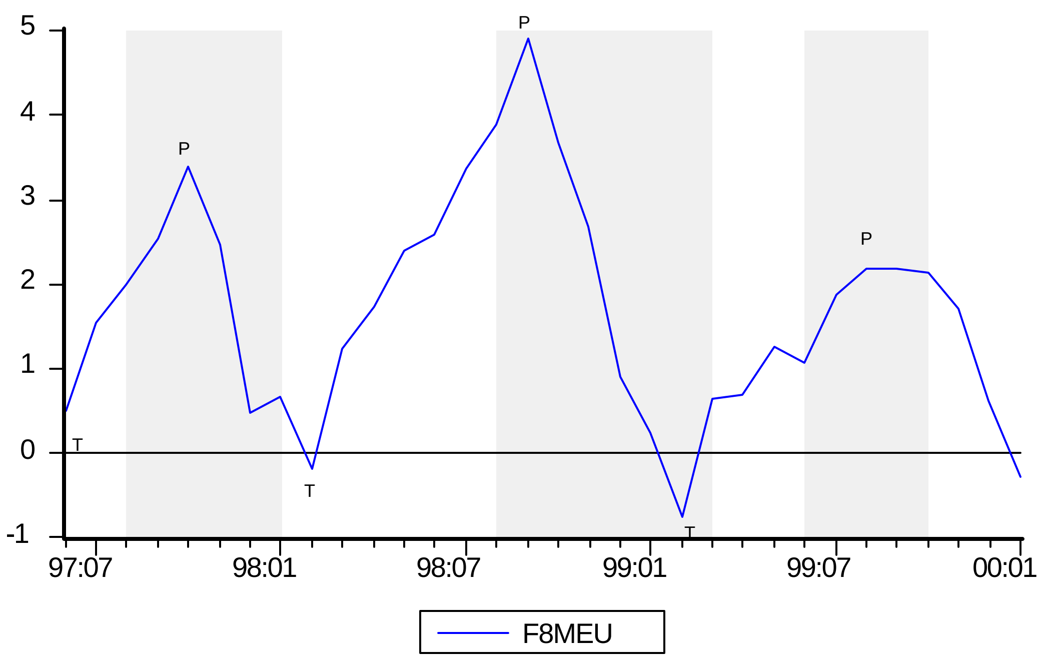


Figure 8 – Real Time: Leading indicator of inflation LIIF8 and High Growth Phases of the IPCA Inflation (Shaded Area)



6.3.3 Linear Forecasts of Inflation – The VAR Model

The indicators can be combined with inflation in bivariate vector autoregressive processes to yield linear forecast of inflation. As an illustration, the indicator LII F23 is combined with IPCA inflation in a VAR. The last observations available for the components of LII F23 are for September 1999. Thus, dynamic forecast can be implemented projecting forward the series IPCA inflation based on its own history and past values of the leading inflation indicator. Figures 9 and 10 plot the dynamic forecast of the Brazilian inflation six-months ahead, from October 1999 to March 2000, given information up to September 1999. The leading indicator correctly signaled a decrease in inflation in the next couple of months, which is consistent with the seasonal pattern of low inflation growth in the beginning of the year. In addition, turning points analysis also indicates that inflation enter a low growth phase in the beginning of the year.

Figure 9 – Linear Forecast of Inflation from VAR (6) between IPCA Inflation and the Leading Inflation Indicator LII F23 – 1999:10 to 2000:03

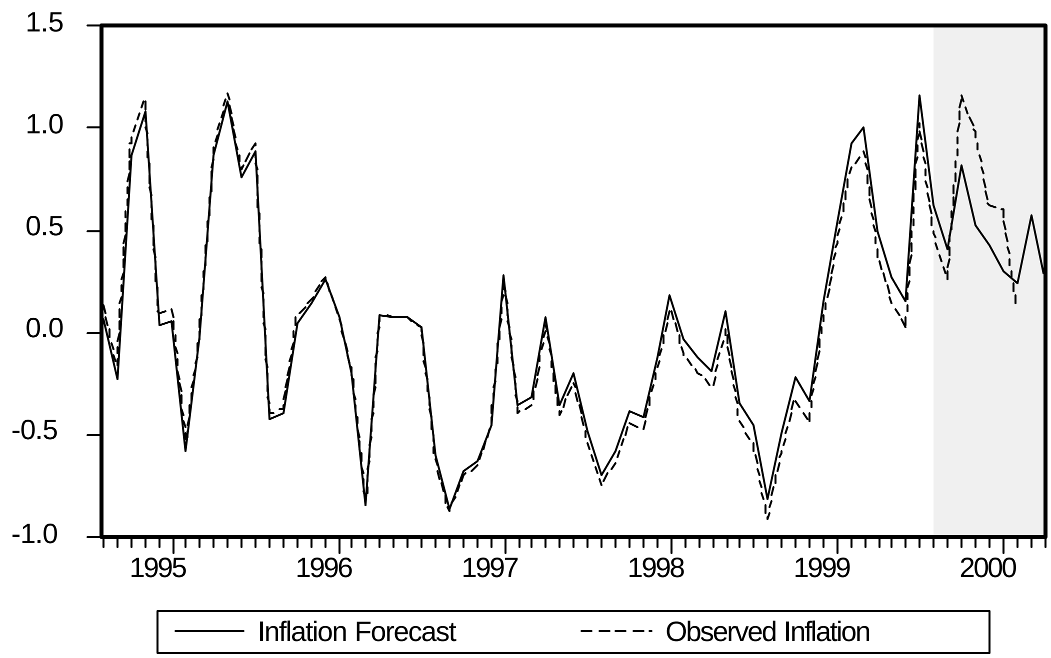
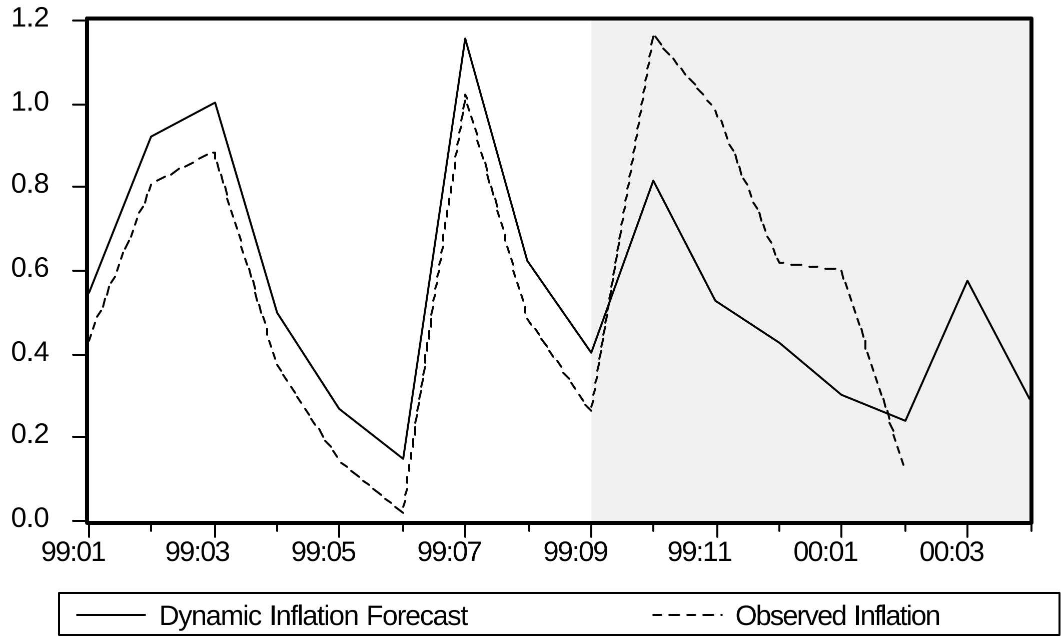


Figure 10 – Linear Forecast of Inflation from VAR (6) between IPCA Inflation and the Leading Inflation Indicator LIIF23 –1999:10 to 2000:03

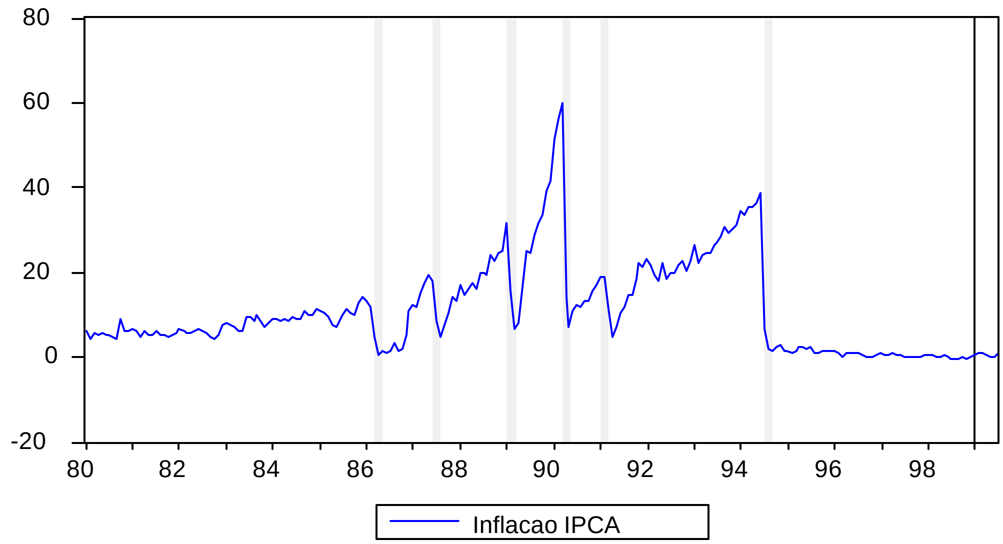


7. Leading Indicators of IPCA Inflation: 1980:01 on

In this part, analysis of IPCA inflation is extended for the period from 1980:01 to 1999:12, which includes the hyperinflationary process in the 1980s and the several Brazilian stabilization plans. Figure 11 plots the behavior of inflation in Brazil during this period. The series displays several structural breaks corresponding to the stabilization plans.

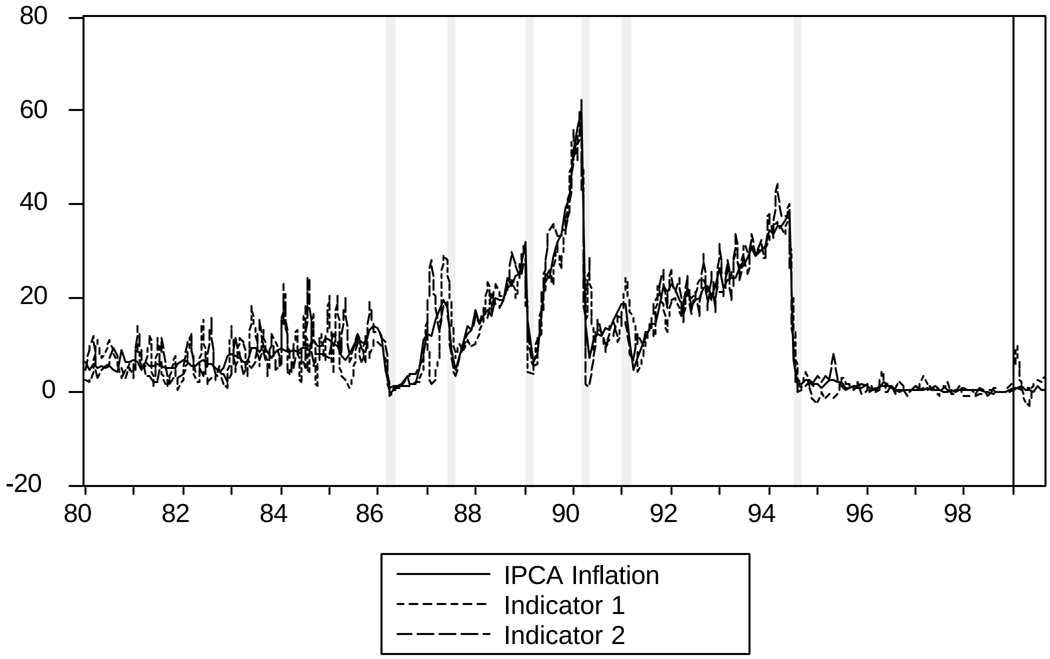
The first issue in searching for a leading inflation indicator for this period is the availability of data. Unfortunately, data-generating Institutions have changed methods of calculating and collecting Brazilian economic data, creating new series that were not extended very far back. In fact, most of the Brazilian series go as far as 1990. This is a serious drawback as historical analysis of the Brazilian economy is compromised by the lack of further information. From hundreds of variables studied for the smaller sample (1994:08-1999:08), there were only 50 candidate variables, whose sample period starts in the beginning of the 80s. From those, only 14 variables can be classified as leading inflation variable according to the rationales in described in section 3.2. The lead-lag analysis as described in section 3 was implemented between each of these candidates and the IPCA inflation for sub-samples in between the structural breaks. However, the results for this sample were not as favorable – all the variables display cyclical movements that coincide with the IPCA inflation. That is, none of the variables analyzed are able to anticipate IPCA inflation turning points. In addition, using a time-varying version of the Granger causality test, the candidate variables do not show significant predictive power in linear forecasting inflation.

Figure 11 – Inflation IPCA and Brazilian Economic Plans: 1980:01 to 1999:07



In order to investigate this further, a time-varying dynamic factor model was used to obtain potential leading indicators of inflation. This method allows investigation of the time-varying relationship between inflation and the indicator across different policy regimes. Figure 12 shows some of the resulting indicators of inflation. All indicators exhibit weak ability to signal turning points. These results are consequence of the unexpected changes in the economy introduced by the six major "pacotes econômicos," which most economic variables did not forewarned – changes in policy regimes engendered breaks in the relationships of the inflation and candidate variables.

Figure 12 – Inflation IPCA and Indicators of Inflation



8. Conclusions and Remarks

This project had as a goal the construction of leading indicators that anticipate inflation cycle turning points on a real time monitoring basis. As a first step, turning points of the IPCA inflation were dated using a periodic stochastic Markov switching model. A dynamic factor model was then used to extract common cyclical movements in a set of variables that display predictive content for inflation.

Since the idea is to use the leading indicators as practical tools to assist real-time monitoring of monetary policy, the econometric procedure to rank variables and select models were based on recursively out-of-sample forecasting performance. Out-of-sample estimation is crucial in order to avoid data mining and, consequently, poor forecasts in real time. This allows better understanding on how well the models would have performed if they had been applied month by month in real time.

The leading indicators are found to be an informative tool for signaling future phases of the inflation cycle out-of-sample, even in real time when only preliminary and unrevised data are available.

For the first set of leading indicators, the sample period used is very small (65 observations), which makes it difficult to infer aspects that can be predicted in the future.

Thus, the leading indicator should be revised as more observations become available, since the relationship between the variables may change over time. In the absence of major shocks, however, the frequency of revisions can be low as long as the procedure used is based on out-of-sample performance.

There is evidence of a structural break in the relation between inflation dynamics and candidate variables in the end of 1998. Thus, the results of selecting variables and specifications based on linear models of Granger causality, VARs, and regressions should be interpreted with caution. The procedure used here partially overcomes this problem, since the variables were selected based on their out-of-sample recursive forecasting ability.

The resulting indicators for the longer period do not anticipate inflation turning point. In fact, they display coincident movements with inflation. However, this does not preclude forecasting analysis of inflation using other approaches than turning point evaluation for this sample. For example, an interesting extension of the analysis would be the application of Pesaran and Timmermann (1999) recursive forecasting method in the presence of structural breaks.

REFERENCES

Anderson, T.W. and H. Rubin, 1956, "Statistical Inference in Factor Analysis," *Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability*. vol. 5, Berkeley: University of California Press, 111-150.

Bollen, K. and K. Joreskog, 1985, "Uniqueness does not Imply Identification," *Sociological Methods and Research*, vol. 14, nº 2, 155-163.

Bollen, K.A., 1989, *Structural Equations with Latent Variables*, Willey Series in Probability and Mathematical Statistics.

Burns, A. and W. Mitchell, 1946, *Measuring Business Cycles*, (New York: National Bureau of Economic Research).

Chauvet, M., 1998, "An Empirical Characterization of Business Cycle Dynamics with Factor Structure and Regime Switching," *International Economic Review*, vol. 39, No.4, 969-996.

Chauvet, Marcelle, *Turning Point Analysis of the Leading Inflation Indicators*, *Activity Report*. Central Bank of Brasil. April/May, 2000a.

Chauvet, M., 2000b, "Stock Market Fluctuations and the Business Cycle," forthcoming, *Journal of Economic and Social Measurement*.

Chauvet, Marcelle. *Activity Report*. Central Bank of Brazil. September 1999a.

Chauvet, Marcelle. *Leading Inflation Indicators for Brazil*. *Final Report.* Central Bank of Brazil. December 1999b.

Chauvet, M. and S. Potter, 2000, "Coincident and Leading Indicators of the Stock Market," forthcoming, *Journal of Empirical Finance*.

Dìckey, D. A., and W. A. Fuller, "Distribution of the Estimators for Autoregressive Time Series with a Unit Root," Journal of the American Statistical Society 74 (1979), 427-31.

Diebold, F. X. and G. D. Rudebusch, 1989, "Scoring the Leading Indicators," *Journal of Business* 62, 369-391.

Engle, Robert F. and C.W.J. Granger (1987) "Co-integration and Error Correction: Representation, Estimation, and Testing," *Econometrica* 55, 251–276.

Geweke,J., 1977, "The Dynamic Factor Analysis of Economic Time-Series Models," in *Latent Variables in Socio-Economic Models* by Aigner, D.J. and A.S. Goldberger; Contributions to Economic Analysis 103, North-Holland, 365-383.

Ghysels, E., 1993, "A Time Series Model with Periodic Stochastic Regime Switching," Federal Reserve Bank of Minneapolis, Discussion Paper 84.

Granger, Clive W. J., 1996 **"C**an we improve the perceived quality of economic forecasts?" Journal of Applied Econometrics, 11(5):455-473.

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

Hamilton, J.,1989 "A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle," *Econometrica* 57, 357-384.

Joreskog, K.G. and D. Sorbom, 1977, "Statistical Models and Methods for Analysis of Longitudinal Data," in *Latent Variables in Socio-Economic Models* by Aigner, D.J. and A.S. Goldberger; Contributions to Economic Analysis 103, North-Holland, 285-325.

Kling, J. L., 1987, "Predicting the Turning Points of Business and Economic Time Series," *Journal of Business*, 60, 201-38.

Neftci, S., 1982, "Optimal Prediction of Cyclical Downturns," *Journal of Economic Dynamics and Control*, 4, 225-41.

McConnell, M. and G. Perez-Quiros, 1998, "Output Fluctuations in the United States: What Has Changed Since the Early 1980s? Federal Reserve Bank of New York, Staff Reports, 41.

Roth, Howard L. "Leading Indicators of Inflation", in Kajal Lahiri and Geoffrey H. Moore, eds., *Leading Economic Indicators: New Approaches and Forecasting Records.* Cambridge: Cambridge University Press (1991), pp 275-302.

Phillips, P. and P. Perron, 1988, "Testing for a Unit Root in Time Series Regression," *Biometrika*, 75:335-46.

Perron

Sargent, T.J. and C.A. Sims (1977), "Business Cycles Modeling Without Pretending to Have Too Much *a Priori* Economic Theory," in Sims (ed.), *New Methods in Business Cycle Research*, Minneapolis: Federal Reserve Bank of Minneapolis.

Stock, J. H. and M. W. Watson, 1988, "Testing for Common Trends," Journal of the American Statistical Association 83, 1097-1107.

Stock, J. and Watson, M., 1989, "New Indexes of Coincident and Leading Economic Indicators," in O. Blanchard and S. Fischer (eds.), *NBER Macroeconomics Annual*, (Cambridge: MIT Press).

________, 1991, "A Probability Model of the Coincident Economic Indicators," in K. Lahiri and G.H. Moore (eds.), *Leading Economic Indicators: New Approaches and Forecasting Records*, (Cambridge: Cambridge University Press).

Stock, J. H. and M. W. Watson, "A Procedure for Predicting Recessions with Leading Indicators: Econometric Issues and Recent Experience," in J.H. Stock and M.W. Watson (eds.), Business Cycles, Indicators and Forecasting, (Chicago: University of Chicago Press for NBER, 1993), 255-284.

Wecker, W. E., 1979, "Predicting Turning Points of a Time Series," *Journal of Business*, 52, 35-50.

Zellner, A. and C. Hong, 1991, "Bayesian Methods for Forecasting Turning Points in Economic Time Series: Sensitivity of Forecasts to Asymmetry of Loss Structures," in Lahiri K. and G.H. Moore (eds.), Leading Economic Indicators: New Approaches and Forecasting Records, (Cambridge: Cambridge University Press, 1991).