A Monthly Indicator of Brazilian GDP

Marcelle Chauvet*

Abstract

This paper constructs an indicator of Brazilian GDP at the monthly frequency. The peculiar instability and abrupt changes of regimes in the dynamic behavior of the Brazilian business cycle are explicitly modeled within nonlinear frameworks. In particular, a Markov switching dynamic factor model is used to combine several macroeconomic variables that display simultaneous co-movements with aggregate economic activity. The model generates as output a monthly indicator of the Brazilian GDP and real time probabilities of the current phase of the Brazilian business cycle. The monthly indicator shows a remarkable historical conformity with cyclical movements of GDP. In addition, the estimated filtered probabilities predict all recessions in sample and out-of-sample. The ability of the indicator in linear forecasting growth rates of GDP is also examined within and out-of-sample. In both cases the estimated indicator displays a better predictive performance compared to a linear autoregressive model for GDP. In particular, the inclusion of lags of the indicator improves substantially forecasts of the severity of recessions and strength of expansions, as measured by the volatility of changes in GDP. These results suggest that the estimated monthly indicator can be used to forecast GDP and to monitor the state of the Brazilian economy in real time.

*Department of Economics, University of California, riverside, CA 92521-0247; phone (909)787-5037 x1587; fax (909)787-5685, chauvet@mail.ucr.edu. I am thankful for helpful comments and suggestions from an anonymous referee and from seminar participants at Getulio Vargas Foundation and at the Statistics Department, Federal University of Rio de Janeiro.
A Monthly Indicator of Brazilian GDP

Resumo

Esse artigo constrói um indicador do PIB brasileiro em frequência mensal. A instabilidade e mudanças bruscas de regime peculiares ao comportamento dinâmico do ciclo econômico brasileiro são explicitamente consideradas através de modelos não-lineares. Com efeito, um modelo de fator dinâmico com mudanças de Markov é utilizado para combinar várias séries macroeconômicas que apresentam movimentos simultâneos com a atividade econômica agregada. O modelo gera como produto um indicador mensal do PIB brasileiro e probabilidades da fase corrente do ciclo econômico em tempo real. O indicador mensal apresenta uma notável conformidade histórica com movimentos cíclicos do PIB. Além disso, as probabilidades filtradas estimadas preveem todas as recessões dentro e fora de amostra. O desempenho do indicador para previsões lineares das taxas de crescimento do PIB também é examinado dentro e fora de amostra. Em ambos os casos, o indicador estimado apresenta um melhor desempenho para previsão do PIB em comparação com modelos autoregressivos lineares. Com efeito, a inclusão de defasagens do indicador melhora substancialmente previsões da intensidade de recessões e expansões, medida pela volatilidade das mudanças do PIB. Esses resultados sugerem que o indicador mensal estimado pode ser utilizado para prever o PIB e monitorar as fases da economia brasileira em tempo real.

Key Words: Business Cycle, Dynamic Factor, Markov Switching, Composite Indicators, Kalman Filter, Filtered Probabilities, Forecast.

JEL Code: C32, C50, E32.

1. Introduction.

In recent years, reduction in capital restrictions in emerging market countries substantially increased their international financial inflows. In Brazil, the stability of the currency post-Real plan set the basis for reforms, which turned Brazilian markets more competitive. Trade liberalization and a stronger currency led to cheaper imports and propelled national firms to modernize their technology. These changes combined with privatization and deregulation of state monopolies attracted record high foreign investments.
The economic reforms, however, have not decreased the entrenched volatility of the Brazilian economy. Abrupt downturns and upturns have been a regularity affecting both financial markets and the goods producing sector. As international markets become more integrated, there has been a surging interest in the overall economic performance of Brazil. In particular, there has been an increasing international demand for timely and reliable data that would allow real time assessment of the dynamics of the Brazilian economy. The availability of more accurate real time information is not only important for policy or investment decisions, as it can also contribute to make it less likely the occurrence of self-fulfilling financial crisis driven by misperceptions regarding fundamentals of the Brazilian economy and the country risk¹.

A traditional method of monitoring the economic activity is through the use of composite leading and coincident indicators. In the U.S., composite indicators have been used since their first construction by the Department of Commerce in the 1960s², based on the work of Burns and Mitchell (1946). As a result of a thorough statistical analysis of macroeconomic variables, Burns and Mitchell classified over 500 series into lagging, coincident, or leading according to the timing of their cyclical movements with the U.S. economic activity, in a research sponsored by the National Bureau of Economic Research (NBER). The combination of some of these vari-

¹ There is a vast literature on self-fulfilling financial crisis. Some of the classical theoretical papers are Obstfeld (1996) and Krugman (1996), and an empirical approach is provided, for example, by Jeanne (1997). For a partial survey of self-fulfilling models of currency crisis see Rajan (1999).

² Since 1995 the Conference Board, a private non-profit organization, has been responsible for compiling the U.S. composite indicators.
ables resulted in the composite indexes, which are used to signal turning points of business cycles\textsuperscript{3}. These indicators are one of the most watched series by the press, businesses, policymakers, and stock market participants.

The popularity of the U.S. economic indicators led to the construction of these indexes for other countries, such as Australia, Japan, Canada, and other members of the OECD in the 1970s and 1980s. More recently, progressive market integration has induced a worldwide interest in the analysis of cyclical fluctuations through the use of economic indicators. This has placed the traditional method to construct the indicators in the spotlight, which has been under increased scrutiny and criticism in recent years\textsuperscript{4}. One of the main criticisms is as old as the method itself. Koopmans (1947) criticized Burns and Mitchell's analysis for being centered on measurement without any explicit model attached to it. This problem is also extended to composite indicators, which are constructed as simple weighted averages of statistical transformations of their components. This implies that the dynamic relationships among the variables are not taken into account. For example, the method ignores their cross-correlations or potential long-term relationships. Another important criticism is that the method does not include a probabilistic model to determine turning points in the indicators, and the existing ad-hoc rules require the use of substantial ex-post data. That is, turning points can only be identified and predicted a couple of months af-

\textsuperscript{3}In this context, turning points refer to peaks and troughs of business cycle phases. Peaks mark the beginning of recession phases (end of expansions) and troughs define the end of recessions (onset of expansions).

\textsuperscript{4}Boldin (1998/1997) summarizes the main criticisms to the composite indicators.
Marcelle Chauvet

ter their occurrence, which undermines the usefulness of indicators for real time forecasting. Finally, large revisions have been implemented in the indicators over the years to improve their ability in predicting turning points ex-post. These revisions, which include changes in the components, in their weights, and in the construction procedures, have substantially changed the historical record of the indicators’ performance. In fact, analysis of their real time forecasting ability can only be evaluated using the unrevised versions of the indicators, as examined in Diebold and Rudebusch (1991)\textsuperscript{5}.

The disseminated use of economic indicators and awareness of their shortcomings had a corresponding resurgent academic interest in this instrument. Frontier econometric models have been used to formally explore potential dynamic differences across business cycle phases, to generate indicators, and to determine turning points.

The goal of this paper is to obtain an indicator of Brazilian GDP at the monthly frequency that can be used to forecast business cycles and to monitor the current state of the economy in real time. In particular, a Markov switching dynamic factor model is used to combine several macroeconomic variables that display simultaneous comovements with aggregate economic activity\textsuperscript{6}. The model generates as output a monthly indicator of the Brazilian GDP and real time probabilities of the current phase of the Brazilian business cycle.

\textsuperscript{5}Following Diebold and Rudebusch (1991), several other authors have examined the reliability of the unrevised leading indicator in predicting turning points in real time, such as Emerson and Hendry (1996), Lahiri and Wang (1994), Perez-Quiros and Hamilton (1996), and Chauvet (1998/1999), among others.\textsuperscript{6}Chauvet (1995, 1998) uses a similar approach to build a coincident indicator of the U.S. business cycle, which is extended in Kim and Nelson (1998) to allow for time-varying probabilities.
The formal representation of business cycle phases through the dynamic factor model with Markov switching overcomes the drawbacks of the traditional method while maintaining the main points of Burns and Mitchell (1946) and the NBER's definition of business cycles, as suggested by Diebold and Rudebusch (1996). First, since the model includes several variables, it captures pervasive cyclical fluctuations in various sectors of the economic activity. Since recessions and expansions are caused by different shocks over time, the inclusion of different variables increases the ability of the model in representing and signaling phases of the business cycles. In addition, the combination of variables reduces measurement errors in the individual series and, consequently, the likelihood of false signaling turning points, which is a common problem in monthly data. For example, one of the findings is that the individual variables that compose the indicator, especially industrial production, do not represent broad changes in the economy and, by themselves, give several false turning point signals. Finally, the model allows potential asymmetries across expansions and recessions and analysis of turning points, albeit here their evaluation can be followed in real time.

In order to compile a composite indicator as accurately as possible, a large number of economic time series were evaluated in-depth, with special attention to the peculiar dynamical behavior of the Brazilian business cycle. In particular, the instability and abrupt changes of regimes in the history of Brazilian economic activity are explicitly modeled within nonlinear frameworks.

The primary step in developing composite indicators is the definition of a business cycle chronology to be used as a benchmark. That is, it is necessary to first determine peaks and troughs of the
Brazilian business cycle, which are then used to evaluate the predictive performance of the indicators. The NBER has been dating the U.S. business cycle for the last fifty years. Peaks and troughs are determined from a subjective consensus among the members of the Business Cycle Dating Committee. Decisions about business cycle turning points are based on the cyclical movements of several coincident variables, such as manufacturing and trade sales, personal income, industrial production, and non-agricultural employment, among others. However, the NBER dating can not be used to monitor the economy in real time, since the Business Cycle Committee meets only months after a turning point has occurred.

In this paper, the results of the monthly indicator of GDP are compared to the business cycle chronology obtained in Chauvet (2002), who uses several methods to date the Brazilian business cycle and growth cycle in the last 100 years. The results from Markov switching models are compared with several non-parametric rules, and the dating obtained from all methods is very similar. However, the ad-hoc rules require that a couple of months have passed for the turning point dating to be determined. On the other hand, the results of the model-approach method enable analysis of business cycles on a current basis.

The monthly indicator of Brazilian GDP includes variables measuring aggregate economic activity such as employment, industrial production, capacity utilization, compensated hours, and real wages. The estimated indicator shows a remarkable historical conformity

---

7 More specifically, Markov switching models are fitted to quarterly and annual real GDP, and the resulting smoothed probabilities are used as filtering rules to determine turning points. The results are compared to the dating obtained from applying Bry-Boschan’s routine and the rule of thumb of two quarters decline in GDP.
with cyclical movements of GDP with respect to volatility, timing of peaks and troughs, and duration of the phases. In addition, the estimated filtered probabilities predict all recessions and expansions in the sample, as dated in Chauvet (2002).

The performance of the indicator is also examined out-of-sample. In order to emulate forecasting in real time, the model is recursively re-estimated out-of-sample. This allows us to reproduce the information content that was available to forecasters at any point in time. The indicator summarizes information in several macroeconomic variables, which is found to be relevant in forecasting economic activity. In particular, the resulting filtered probabilities predict all recession turning points out-of-sample.

The ability of the indicator in linear forecasting growth rates of GDP is also examined within and out-of-sample. In both cases the estimated indicator displays a better predictive performance compared to a linear autoregressive model for GDP. In particular, the inclusion of lags of the indicator improves substantially forecasts of the severity of recessions and strength of expansions, as measured by the volatility of changes in GDP.

The results suggest that the indicator and the associated probabilities of business cycle phases could be useful tools to forecast GDP and to monitor the state of the Brazilian business cycle in real time.

The paper is organized as follows. The next section describes the procedure used to analyze and select the variables that compose the model. Section 3 presents the model. In the fourth section, the empirical results are discussed. In particular, the probabilities of business cycle phases and the indicator of the Brazilian GDP are examined within and out-of-sample. The fifth section concludes.
2. Data Analysis.

In order to compile the composite indicator, a large number of economic time series was considered. The data were obtained from Getulio Vargas Foundation (FGV) and the Central Bank of Brazil databases. In order to obtain a historically representative composite indicator of the Brazilian GDP, an important criterion for selecting the series was their sample size. Although there are many candidate variables whose cyclical variation could potentially coincide with the real Brazilian GDP, only a small subset of monthly series have a longer sample going back to the 1970s and 1980s\(^8\).

As a starting point for the analysis, the variables were seasonally adjusted using the X-11 additive method from the U.S. Bureau of the Census. Variables expressed in nominal terms were deflated using the Global Price Index, IGP-DI, or the Broad Consumer Price Index, IPCA\(^9\). Finally, all variables were transformed to achieve stationarity\(^{10}\).

Several selection criteria were then implemented to find the series that display simultaneous movements with the Brazilian business cycles. The underlying guidelines were the economic significance of the variables, their statistical adequacy, their timing at turning points, and their overall conformity to the business cycle. First, the series

---

\(^8\) In fact, most series start in the 1990s. Those were not selected, since a decade of observations would not be representative of the historical behavior of recessions and expansions in Brazil.

\(^9\) Since the Broad Consumer Price Index is only available from 1979 on, the IGP-DI was utilized for data with longer sample.

\(^{10}\) During the period analyzed there were several stabilization plans that engendered structural breaks in the Brazilian macroeconomic variables. Thus, both the augmented Dickey-Fuller (1979) and Perron’s (1989) tests were used to verify the hypothesis of unit roots versus linear deterministic trends with breaks.
were classified as coincident, leading or lagging according to their ability to Granger-cause changes in GDP\textsuperscript{11}, their cross-correlation with changes in GDP in time domain, and coherence in frequency domain. Second, they were ranked according to the timing of their turning points with cyclical variation in GDP, which is taken as the reference cycle. Peaks and troughs of GDP mark recession and expansion phases in Brazil, and were obtained from Chauvet (2002). In particular, two-state Markov switching models were fitted to each of the candidate variables and to GDP, with the states representing economic expansions or contractions. The estimated probabilities of contractions were then used to determine peaks and troughs in the series, and the timing of change of each series was then compared to GDP turning points.

From these procedures, fifteen variables were found to display coincident movements with changes in GDP, which are described in Table 1. These series represent different measurements of industrial production, capacity utilization, real wages, compensated hours, retail sales, employment, unemployment rate, fuel consumption, and production of cement.

\textsuperscript{11} For comparison with growth rates of GDP, two procedures were undertaken. First, all monthly series were converted to quarterly frequencies as simple averages. Second, changes in GDP were converted into monthly frequency using quadratic-match average for comparison with the monthly series.
Marcelle Chauvet

Table 1
Original Variables at Monthly Frequency

<table>
<thead>
<tr>
<th>Variables</th>
<th>Source unit</th>
<th>Measuring</th>
<th>Sample Mnemonic</th>
<th>Transform.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Physical Production-General Industry Indicator</td>
<td>IBGE</td>
<td>Index</td>
<td>1975:01 PRODBR</td>
<td>△log</td>
</tr>
<tr>
<td>Unemployment Rate-Average of 6 Major Metropolitan Regions</td>
<td>IBGE</td>
<td>Percentage</td>
<td>1980:01 URBR</td>
<td>△</td>
</tr>
<tr>
<td>Total Employment-Brazil</td>
<td>MT</td>
<td>Index</td>
<td>1985:01 EMPLBR</td>
<td>△log</td>
</tr>
<tr>
<td>Employees in the Industrial Production Sector-Brazil</td>
<td>IBGE</td>
<td>Index</td>
<td>1985:01 EMPLIP</td>
<td>△log</td>
</tr>
<tr>
<td>Value Added Tax Revenue, ICMS - Brazil</td>
<td>CONFAZ</td>
<td>R$million</td>
<td>1980:01 ICMS</td>
<td>△log</td>
</tr>
<tr>
<td>Retail Sales</td>
<td>FCESP</td>
<td>Index</td>
<td>1979:01 SALES</td>
<td>△log</td>
</tr>
<tr>
<td>Manufacturing Industry-Capacity Utilization(*)</td>
<td>BACEN</td>
<td>Percentage</td>
<td>1986:01 CAPBR</td>
<td>△</td>
</tr>
<tr>
<td>Manufacturing Industry-Real Wages(*)</td>
<td>BACEN</td>
<td>Index</td>
<td>1986:01 WAGEMI</td>
<td>△log</td>
</tr>
<tr>
<td>Fuel Consumption-Brazil</td>
<td>PETROBRAS</td>
<td>ThousandsM3</td>
<td>1980:01 COMB</td>
<td>△</td>
</tr>
<tr>
<td>Electricity Consumption-Brazil</td>
<td>ELETROBRAS</td>
<td>GWH</td>
<td>1976:01 ELETBR</td>
<td>△log</td>
</tr>
<tr>
<td>Total Unemployment, São Paulo</td>
<td>SEADE</td>
<td>Percentage</td>
<td>1984:01 UNEMSP</td>
<td>△</td>
</tr>
<tr>
<td>Manufacturing Industry-Total Number of Employees São Paulo</td>
<td>FIESP</td>
<td>Index</td>
<td>1975:01 EMPLOY</td>
<td>△log</td>
</tr>
<tr>
<td>Manufacturing Industry-Capacity utilization, São Paulo</td>
<td>FIESP</td>
<td>Percentage</td>
<td>1975:01 CAPSP</td>
<td>△</td>
</tr>
<tr>
<td>Manufacturing Industry-Total of Compensated Hours São Paulo</td>
<td>FIESP</td>
<td>Index</td>
<td>1975:01 HOURS</td>
<td>△log</td>
</tr>
<tr>
<td>Manufacturing Industry-Total Wages, São Paulo</td>
<td>FIESP</td>
<td>Index</td>
<td>1975:01 WAGESP</td>
<td>△log</td>
</tr>
<tr>
<td>Real GDP(quarterly)</td>
<td>IBGE</td>
<td>Index</td>
<td>1975:01 GDP</td>
<td>△log</td>
</tr>
</tbody>
</table>

△ stands for first difference. IBGE is the Brazilian Institute of Economic Geography. FIESP is the State of São Paulo’s Industry Federation. FCESP is the State of São Paulo’s Commerce Federation. MT is the Labor Ministry. CONFAZ is the National Treasury Council. BACEN is the Central Bank of Brazil. SEADE is State of São Paulo System of Data Analysis. ELETROBRAS is the Brazilian Electricity Holding Company.

(*) These data were generated from surveys of the Industry Federation in the following Brazilian States: Amazonas, Ceará, Pernambuco, Bahia, Espírito Santo, Minas Gerais, Rio de Janeiro, São Paulo, Paraná, Santa Catarina, Rio Grande do Sul e Goiás.

The indicator of the Brazilian business cycle is constructed using a Markov switching dynamic factor model, as in Chauvet (1995, 1998). Let $Y_t$ be a vector of $n \times 1$ observable macroeconomic variables that move simultaneously with the Brazilian Gross Domestic Product:

\[ \Delta Y_t = \lambda \Delta F_t + \Delta v_t, \]

where $\Delta = 1 - L$ and $L$ is the lag operator. Changes in these $n$ macroeconomic variables, $\Delta Y_t$, are modeled as a scalar common unobserved factor, $\Delta F_t$, and $n$ individual idiosyncratic terms, $\Delta v_t$. The factor loadings, $\lambda$, measure the sensitivity of the series to the dynamic factor, $\Delta F_t$. Both the factor and the idiosyncratic terms follow autoregressive processes:

\[ \Delta F_t = \mu S_t + \theta \Delta F_{t-1} + \eta_t \sigma_{\eta S_t}, \quad \eta_t \sim N\left(0, \sigma_{\eta S_t}^2\right), \]

\[ \Delta v_t = d(L) \Delta v_{t-1} + \epsilon_t, \quad \epsilon_t i.i.d. \sim N\left(0, \sum\right). \]

where $\eta_t$ is the common shock that engenders changes in the phases of the business cycle and $\epsilon_t$ are the measurement errors.

In order to capture potential asymmetries across different states of the business cycle, the intercept and variance of the factor switch

\textsuperscript{12}The factor loading for the production series is set equal to one to provide a scale for the latent dynamic factor. This normalization is a necessary condition for identification of the factor and the choice of parameter scale does not affect any of the time series properties of the dynamic factor or the correlation with its components.
regimes according to a Markov variable, $S_t$, where $\mu S_t = \alpha_1 + \alpha_0 S_t$, and $S_t = 0, 1$. That is, the economy can be either in an expansion state ($S_t = 1$), where the mean growth rate is positive; or in a contraction phase ($S_t = 0$), with a negative mean growth rate. The volatility of the factor as measured by $\sigma_{\eta S_t}^2$ can also assume different values across states. The switches from one state to another is determined by the transition probabilities of the first-order two-state Markov process, $p_{ij} = \text{Prob} [S_t = j | S_{t-1} = i]$, where $\sum_{j=0}^1 p_{ij} = 1, i, j = 0, 1$.

The model separates out common signal underlying the observed variables from individual variations in each sector of the economic activity. The dynamic factor captures widespread simultaneous downturns and upturns movements of several sectors of the Brazilian economy, corresponding to Burns and Mitchell’s (1946) definition of business cycles. On the other hand, if only one of the variables declines (e.g. industrial production), this would not characterize a recession in the model, and it would be captured by the industrial production idiosyncratic term. A recession (expansion) will occur when all $n$ variables decrease (increase) at about the same time. That is, $\eta_t$ and $\Delta v_t$ are assumed to be mutually independent at all leads and lags, for all $n$ variables, and $d(L)$ is diagonal.

The dynamic factor is the outcome of averaging out the discrete states. Although the $n$ variables measure economic activity in specific sectors, the estimated dynamic factor is a nonlinear combination of them, representing broader movements in the Brazilian economy.

The model is estimated using a combination of Hamilton’s (1989) Markov switching algorithm and the Kalman filter, as suggested by Kim (1994) based on the work of Harrison and Stevens (1976). The model is cast in state space form, where (4) and (5) are the measurement and transition equations, respectively:
A Monthly Indicator of Brazilian GDP

\[(4) \quad \Delta Y_t = Z\xi_t \]

\[(5) \quad \xi_t = \alpha\xi_{st} + T\xi_{t-1} + u_t, \]

A particular state space representation for the estimated indicator using five variables is:

\[
\Delta Y_t = \begin{bmatrix}
\Delta Y_{1t} \\
\Delta Y_{2t} \\
\Delta Y_{3t} \\
\Delta Y_{4t} \\
\Delta Y_{5t}
\end{bmatrix}, \quad Z = \begin{bmatrix}
\lambda_1 & 1 & 0 & 0 & 0 & 0 \\
\lambda_2 & 0 & 1 & 0 & 0 & 0 \\
\lambda_3 & 0 & 0 & 1 & 0 & 0 \\
\lambda_4 & 0 & 0 & 0 & 1 & 0 \\
\lambda_5 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad \alpha_{\xi_{st}} = \begin{bmatrix}
\mu_{st} \\
0 \\
0 \\
0 \\
0 \\
0 \\
0
\end{bmatrix},
\]

\[
T = \begin{bmatrix}
\theta_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & d_1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & d_2 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & d_3 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & d_4 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & d_5 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}, \quad \text{and} \quad u_t = \begin{bmatrix}
\eta_t \sigma_{\eta_{st}} \\
\epsilon_{1t} \\
\epsilon_{2t} \\
\epsilon_{3t} \\
\epsilon_{4t} \\
\epsilon_{5t} \\
0
\end{bmatrix}.
\]

The term $F_{t-1}$ is included in the state vector to allow estimation of the dynamic factor in levels from the identity $\Delta F_{t-1} = F_{t-1} - F_{t-2}$. 

Brazilian Review of Econometrics 21 (1) May 2001
The nonlinear filter forms forecasts of the unobserved state vector, $\xi^{(i,j)}_{t|t-1}$, and the associated mean squared error matrices, $\theta^{(i,j)}_{t|t-1}$, based on information available up to time $t - 1$, $I_{t-1} \equiv [\Delta Y'_{t-1}, \Delta Y'_{t-2}, \ldots, \Delta Y'_1]'$, on the Markov state $S_t$, taking on the value $j$, and on $S_{t-1}$ taking on the value $i$, for $i, j = 0, 1$:

\[
\xi^{(i,j)}_{t|t-1} = E(\xi_t|I_{t-1}, S_t = j, S_{t-1} = i) \tag{6}
\]

\[
\theta^{(i,j)}_{t|t-1} = E\left[ (\xi_t - \xi_{t|t-1}) (\xi_t - \xi_{t|t-1})' | I_{t-1}, S_t = j, S_{t-1} = i \right]. \tag{7}
\]

The filter uses as inputs the joint probability of the Markov-switching states at time $t - 1$ and $t$ conditional on information up to $t - 1$, \{Pr ($S_{t-1} = i, S_t = j | I_{t-1}$)\}; an inference about the state vector using information up to $t - 1$, given $S_{t-1} = i$ and $S_t = j$, that is, \{ $\xi^{(i,j)}_{t-1|t-1}$ \}; and the mean squared error matrices, \{ $\theta^{(i,j)}_{t-1|t-1}$ \}. The outputs are their one-step updated values. The nonlinear Kalman filter is:

\[
\xi^{(i,j)}_{t|t} = \alpha \xi_{st} + T \xi^{i}_{t-1|t-1} \tag{8}
\]

(prediction equations)

\[
\theta^{(i,j)}_{t|t-1} = T \theta^{i}_{t-1|t-1} T' + H \tag{9}
\]

\[
\xi^{(i,j)}_{t|t} = \xi^{(i,j)}_{t|t-1} + K^{(i,j)}_{t} N^{(i,j)}_{t|t-1} \tag{10}
\]

(updating equations)

\[
\theta^{(i,j)}_{t|t} = \left( I - K^{(i,j)}_{t} Z \right) \theta^{(i,j)}_{t|t-1} \tag{11}
\]
where $H$ is the variance-covariance matrix of the vector of disturbances $u_t$, $I$ is the identity matrix, $K_t^{(i,j)} = \theta_{t|t-1}^{(i,j)}Z'[Q_t^{(i,j)}]^{-1}$, $N_t^{(i,j)} = \Delta Y_t - Z\xi_{t|t-1}^{(i,j)}$ is the conditional forecast error of $\Delta Y_t$, and $Q_t^{(i,j)} = Z\sigma_{t|t-1}^{(i,j)}Z'$ is the conditional variance.

The probability terms are computed using Hamilton’s filter:

(12) $\text{Prob} (S_{t-1} = i, S_t = j | I_{t-1}) = 
\rho^{ij} \sum_{h=0}^{1} \text{Prob} (S_{t-2} = h, S_{t-1} = i | I_{t-1})$

From these joint conditional probabilities, the density of $\Delta Y_t$ conditional on $S_{t-1}$, $S_t$, and $I_{t-1}$ is:

(13) $f (\Delta Y_t | S_{t-1} = i, S_t = j, I_{t-1}) = 
(2\pi)^{-n/2} |Q_t^{(i,j)}|^{-1/2}\exp \left( -\frac{1}{2}N_t^{(i,j)}Q_t^{(i,j)^{-1}}N_t^{(i,j)} \right)$.

The joint probability density of states and observations is then calculated by multiplying each element of (12) by the corresponding element of (13):

(14) $f (\Delta Y_t, S_{t-1} = i, S_t = j | I_{t-1}) = 
\sum_{j=0}^{1} \sum_{i=0}^{1} f (\Delta Y_t, S_{t-1} = i, S_t = j | I_{t-1}) \text{Prob} (S_{t-1} = i, S_t = j | I_{t-1})$.

The probability density of $\Delta Y_t$ given $I_{t-1}$ is:

(15) $f (\Delta Y_t | I_{t-1}) = \sum_{j=0}^{1} \sum_{i=0}^{1} f (\Delta Y_t, S_{t-1} = i, S_t = j | I_{t-1})$.

The joint probability density of states is calculated by dividing each element of (14) by the corresponding element of (15):
Finally, summing over the states in (16), we obtain the filtered probabilities of expansions or recessions:

\[ \text{Prob} (S_{t-1} = i, S_t = j | I_t) = \frac{f (\Delta Y_t, S_{t-1} = i, S_t = j | I_{t-1})}{f (\Delta Y_t | I_{t-1})} \]

As in the linear Kalman filter, the algorithm calculates recursively one-step-ahead predictions and updating equations of the dynamic factor and the mean squared error matrices, given the parameters of the model and starting values for \( \xi^j_{t|t} \), \( \theta^j_{t|t} \), and the probabilities of the Markov states. However, for each date \( t \) the non-linear filter computes \( 2^k \) forecasts, where \( k \) is the number of states, and at each iteration the number of cases is multiplied by \( k \). This implies that the algorithm would be computationally unfeasible even for the simplest cases. Kim (1994) based on Harrison and Stevens (1976), proposes an approximation introduced through \( \xi^j_{t|t} \) and \( \theta^j_{t|t} \) for \( t > 1 \). This approximation consists of truncating the updating equations into averages weighted by the probabilities of the Markov states.

The conditional likelihood of the observable variables is obtained as a by-product of the algorithm at each \( t \), from equation (13), which is used to estimate the unknown model parameters. The filter evaluates this likelihood function, which is then maximized with respect to the model parameters using a nonlinear optimization algorithm\(^{13}\).

\(^{13}\)In the out-of-sample analysis the maximum likelihood estimates and the associated dynamic factor index and filtered probabilities were recursively estimated at each time \( t \).
The maximum likelihood estimators and the sample data are then used in a final application of the filter to draw inferences about the dynamic factor and probabilities, based on information available at time $t$. The final estimated state vector is calculated as:

$$\xi_{t|t} = \sum_{i=0}^{1} \Pr [S_t = j|I_t] \xi_{t|t}^{ij}.$$  

4. Empirical Results.

Given that the sample availability differs across variables, the model was estimated using three periods: a) from 1975:01 to 2000:06, b) from 1980:01 to 2000:06, and c) from 1986:01 to 2000:06. Six variables are available from 1975:01 on, which are used to estimate the indicator with the longest period. For the second sample, these variables are combined with the other four that start in 1980:01. Finally, these ten variables are combined with the remaining five to find the best composite indicator from 1986:01 on.

The dynamic factor summarizes the comovements underlying the macroeconomic series used in the model. Thus, it is crucial that the variables composing the indicator display coincident cyclical movements with each other$^{14}$. Factor analysis and principal components were used to test the variables to be included in the dynamic factor model$^{15}$. The magnitude of the eigenvalues of the common factor

---

$^{14}$The inclusion of variables with low correlation with each other will generally result in the convergence of the factor to one of the component variables.

$^{15}$Factor analysis corresponds to the study of the covariance or correlation structure of observed variables through a small number of hypothetical unobservable common factors that account for their correlations. In exploratory factor analysis, a model relating $n$ observable variables
correlation matrix indicates whether the factor structure is a reasonable representation of the data. In fact, this method can also be used to test for single or multi-factor dynamic specifications, as discussed in the next section.

Finally, the selection among those that represent closely related definitions was based on two guidelines. First, it was taken into consideration the economic significance of the specific activity or process as more representative of the Brazilian GDP. For example, total industrial production is chosen instead of production in the manufacturing industry, and so on. Second, the evaluation process described in section 2 was applied to rank and select between closely related variables.

From this analysis, three coincident indicators are found from the combination of a small set of variables. For the sample from 1975:01 on, five variables are combined to yield a monthly indicator of GDP (group 1): total industrial production (PRODBR), total number of employees (EMPLOY), capacity utilization (CAPSP), compensated hours (HOURS), and total wages (WAGESP)\(^{16}\). For the sam-

---

with \( k \) latent factors is first analyzed with one factor only, and then it is verified whether most of the partial correlation coefficients among the observed variables vanish after accounting for the factor effect. If not, two underlying factors are introduced in order to explain the partial correlations among each pair of the observable variables, and the procedure is repeated until there remain almost no correlations between the \( n \) variables. A similar approach is principal components analysis, which extracts latent components from the observable variables representing their maximum variance. The main difference between these two approaches is that factor analysis does not extract all variance from the variables, but only that proportion that is due to the common factors and shared by all the observable variables (i.e. their common variance). In addition, the set of hypotheses that form the conceptual basis of confirmatory factor analysis enables interpretation of the factor.

\(^{16}\) Variables measuring electricity usage were found to lag the economic activity. This result might reflect in part the way some of these series are calculated. For example, observation for
ple from 1980:01, six variables are combined to obtain the indicator (group 2): total industrial production (PRODBR), unemployment rate (URBR), compensated hours (HOURS), retail sales (SALES), capacity utilization (CAPSP), and total wages (WAGESP). Finally, for the sample from 1985:01 on, six variables enter into the composition of the coincident indicator (group 3): total industrial production (PRODBR), total employment (EMPLBR), compensated hours (HOURS), retail sales (SALES), capacity utilization (CAPBR), and total wages (WAGEMI)17.

The models using the three sets of variables were estimated by maximizing the likelihood function through a numerical procedure. The nonlinear discrete filter produces two outputs: the dynamic factor and the associated probabilities of the Markov state. The filtered probabilities give at time $t$ the probability of the Markov state using only information available at $t$, $\Pr (S_t = 0, 1 | I_t)$. On the other hand, the smoothing probabilities are obtained through backward recursion using the information in the full sample, $\Pr (S_t = 0, 1 | I_T)$.

Figure 1 shows the smoothed probabilities of recession for the three sets of variables. The results are striking similar. The probabilities match in terms of timing of changes, duration, and amplitude. This also holds for the dynamic factor from the three models, where the correlation for each pair is around 0.999. This finding is due to

time $t$ corresponds to electricity usage from mid-month $t-1$ to mid-month $t$. The series on fuel consumption exhibits low correlation with most of the other components and was, therefore, not included in the final composite indicator. The value-added tax revenue series (ICMS) did not score very high in the evaluation process - in addition to being a very noisy variable, it also displays low correlation with current economic conditions at the monthly frequency, which may be a consequence of measurement error.

17The hypothesis of cointegration for these three groups of series was rejected at the 5% level using Johansen's (1991, 1995) test.
the fact that the series used in each subset represent the same sectors. Basically, the variables are broader or narrower measurements of employment, wages, industrial production, capacity utilization, and sales\textsuperscript{18}. Thus, the empirical analysis is shown for the monthly indicator composed of the variables with a longer historical sample, group 1.

\textbf{Figure 1 - Smoothed Probabilities of Recessions for Group 1 (---), Group 2 (- - -), and Group 3 ( - - - )}:

\footnotesize
\textsuperscript{18}Notice that these are roughly the same variables used by the NBER, the Conference Board, and the OECD to construct their coincident indicators of the business cycle. The similarity of the results for the three groups also indicate that the series representing the production process in São Paulo are good proxies for the broader production process in the whole country. A further examination of the data explains this finding - São Paulo's GDP has corresponded to around 60\% of the total Brazilian GDP in the last twenty years (data for GDP per State is obtained from IBGE).
4.1 Specification Tests.

The dynamic factor structure captures cyclical comovements underlying the observable variables. The resulting dynamic factor is highly correlated with all the series used in its construction, indicating that the structure was not simply imposed on the data by assuming large idiosyncratic errors. In addition, tests for the number of states support the single factor specification. This is tested in different ways. First, the eigenvalues of the correlation matrix of the common factor indicate adequacy of the single factor specification\textsuperscript{19}. Second, the model assumes that the factor summarizes the common dynamic correlation underlying the observable variables, which implies that the $n$ residuals $\Delta v_{k,t}$ for $k = 1, \ldots, n$, are uncorrelated with each other and with the observed variables $\Delta Y_{h,t}$, $h = 1, \ldots, n$, for $k \neq h$\textsuperscript{20}. In order to test this assumption, the disturbances $\Delta v_{k,t}$ are regressed on six lags of the observable variables $\Delta Y_{h \neq k,t}$, and the parameters of the equations are found to be insignificantly different from zero. In addition, the one-step-ahead conditional forecast errors, $N_{t+1|t-1}$, obtained from the filter described in section 3, are not predictable by lags of the observable variables. These results support the single factor specification, since these error terms are not capturing common information underlying the observable variables.

With respect to the residuals $E_t$ the i.i.d. assumption is tested using Ljung-Box statistics on their sample autocorrelation, and Brock, Dechert, and Scheinkman's (1996) diagnostic test\textsuperscript{21}. Both

\textsuperscript{19}The magnitude of the $n$ eigenvalues for each factor reflects how much of the correlation among the observable variables is explained by $k \leq n$ potential factors. For each of the three composite indicators, there is only one eigenvalue greater than one, while the others are much smaller and close to zero.

\textsuperscript{20}The model was estimated allowing either AR(1) or AR(0) processes for the disturbances $\Delta v_t$. The likelihood ratio test favors the AR(1) specification at the 1% level.

\textsuperscript{21}Leads of 2, 3, 4, 5, and 6 months are used for the residuals and the distance between the two
tests fail to reject the i.i.d. assumption.

Finally, the number of states in the Markov switching process is tested using the approach proposed by Garcia (1998), based on Hansen (1993, 1996). The test provides strong evidence for the two-state specification against the assumption of single state. Even though the likelihood ratio test for comparing the Markov switching model with a non-switching model has an unknown sampling distribution (see Hansen 1996), one can evaluate different two-state specifications using standard chi-squared sampling distributions. A likelihood ratio test comparing a model with constant variance and switching mean versus the switching mean and variance model rejects at the 1% level the hypothesis that the additional parameter for the state dependent variance is zero.

4.2 Results.

Table 2 shows the maximum likelihood estimates of the Markov switching dynamic factor model using the five variables from group 1 for the period from 1975:01-2000:06. The Markov states for the factor are statistically significant. State 1 has a positive mean growth rate, a smaller volatility, and a higher transition probability, while in state 0 the factor has a negative mean growth rate, high variance, and a smaller transition probability. That is, the positive state is associated with the longer and calmer expansion phases of the business cycle and the negative state reflects the more volatile and shorter economic contractions in Brazil. These asymmetries in the phases of business cycles are also found in the U.S., Australia, and several other OECD countries²², and corroborate the findings in Chauvet (2002) for the Brazilian business cycle.

---

²²See Chauvet and Yu 2002.
The factor loadings measure how changes in the dynamic factor affect changes in the observable variables. Industrial production, real wages, and compensated hours are the most sensitive variables to business cycles as measured by the GDP indicator, while employment is relatively less so. That is, some variables are more flexible in adjusting to cyclical economic variation than others. For example, firms facing the prospective of an economic contraction may opt first in reducing the number of compensated hours rather than actually firing workers.

Table 2

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Parameters</th>
<th>Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$</td>
<td>1.23</td>
<td>$\lambda_{\text{capac}}$</td>
</tr>
<tr>
<td></td>
<td>(0.51)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\mu_0$</td>
<td>-0.79</td>
<td>$\lambda_{\text{wages}}$</td>
</tr>
<tr>
<td></td>
<td>(0.44)</td>
<td>(0.10)</td>
</tr>
<tr>
<td>$\sigma_1^2$</td>
<td>1.73</td>
<td>$\lambda_{\text{hours}}$</td>
</tr>
<tr>
<td></td>
<td>(0.39)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\sigma_0^2$</td>
<td>9.74</td>
<td>$\lambda_{\text{employ}}$</td>
</tr>
<tr>
<td></td>
<td>(2.70)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$p_{11}$</td>
<td>0.93</td>
<td>$d_{\text{capac}}$</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$p_{00}$</td>
<td>0.83</td>
<td>$d_{\text{wages}}$</td>
</tr>
<tr>
<td></td>
<td>(0.10)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>$\varnothing$</td>
<td>-0.20</td>
<td>$d_{\text{hours}}$</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

$\text{LogL}(\theta) = -2526.07$

Asymptotic standard errors in parentheses. The factor loading for production is set to one to normalize the factor.
Figure 2 plots the smoothed probabilities of recessions from the monthly dynamic factor, and the smoothed probabilities of recessions obtained from fitting an AR(0) Markov switching model to quarterly GDP, as in Chauvet (2002). The probabilities from both models are very similar and capture all recessions. In addition, the timing and duration of the business cycle phases are very close. One difference is the first recession of the quarterly sample, in 1980-82. Since the sample for quarterly GDP starts only in 1980:01, the model using only GDP captures less information about this recession period than the dynamic factor model, whose sample begins in 1975:01.

The smoothed probabilities can be used as a metric to identify historical business cycle turning points in Brazil. One rule is to consider that the economy is in a recession if the smoothed probabilities of recession are above 50%. Alternatively, the frequency distribution of the probabilities can also be used to define a turning point - a rule of thumb to call a peak is when the probabilities of recessions are greater than their mean plus one-half their standard deviation. Figure 3 shows turning points of the monthly dynamic factor and business cycle dating from GDP, obtained from Chauvet (2002). Both these rules yield very similar dating of expansions and recessions for both models.

Figure 4 shows the smoothed probabilities of recessions and

---

23 The quarterly smoothed probabilities were converted to monthly frequency using the quadratic-match average method.

24 Notice that some of the other minor differences in the figure are due to the transformation from quarterly to monthly frequency.

25 In order to rule out short-term events, such as strikes, tax law changes, etc. from a broad recession in the economy, one of the NBER rules is that a recession corresponds to a general downturn in the economy of at least six months. Here we also consider the minimum duration of business cycle phases to be six months.
the dating of the Brazilian business cycle at the monthly frequency (shaded area). In the last 25 years Brazil experienced ten recessions and ten expansions. Recessions are generally shorter than expansions. In the sample studied, there were five recessions lasting only six months while the longest recession occurred between 1980 and 1982, lasting 20 months. The longest expansion took place between 1983 and 1987, while the shorter one (6 months) occurred in between the recession induced by the Collor’s Plan in 1990-1991 and the subsequent recession in the end of 1991 and beginning of 1992.

The variables that enter the dynamic factor model are available in a timely basis. However, they are very volatile, which makes it difficult to use them to discern between broad cyclical movements in the economic activity from individual noise in the different sectors these variables represent. That is, individually, these series give very mixed signals about the state of the economy. This can be seen from Figure 5, which shows the filtered probabilities of recession obtained from fitting a Markov switching model to each individual component of the dynamic factor model. The probabilities display several extra peaks and troughs that do not correspond to an overall contraction of the Brazilian economic activity. It is particularly interesting to notice that, in the case of Brazil, monthly industrial production by itself is a poor measurement of business cycles – in several periods industrial production decreases or increases substantially while the other sectors of the economy (as well as total GDP) do not show a corresponding movement.

The dynamic factor model combines these variables and extracts the cyclical variation that is common to all of them. In the process of extracting the signal, the noise is isolated into the idiosyncratic terms, $\xi_t$, and the resulting index is a smoother variable that summarizes the common correlation underlying the factor components.
Figure 2: Smoothed Probabilities of Recessions from a MS-Model for GDP (- - -), and from the Dynamic Factor Model (---), at the Monthly Frequency.
Figure 3 - Dating of the Brazilian Business Cycle for Quarterly (---) and Monthly Frequencies (Shaded Area)
Figure 4 - Smoothed Probabilities of Recessions and Dating of the Brazilian Business Cycle (Shaded Area)
Figure 5 - Filtered Probabilities of Recessions Obtained from Fitting a Univariate AR(1) Markov Switching Model to the Factor Components, and Dating of the Brazilian Business Cycle (Shaded Area)
Figure 6 - Smoothed (---) and Filtered (---) Probabilities of Recessions and Dating of the Brazilian Business Cycle (Shaded Area)
In contrast to the probabilities of the individual series, the probabilities of recession obtained from the dynamic factor capture closely business cycle expansions and contractions, as measured by GDP. Figure 6 plots both the smoothed and filtered probabilities of recessions. The filtered probabilities give at time $t$ the probability of a recession using only data available at $t$. Thus, these probabilities can be used as a monitoring tool to determine the state of the economy in real time. By construction, the filtered probabilities display more spikes than the corresponding smoothed probabilities. However, as a comparison of Figures 5 and 6 illustrates, using the filtered probabilities from the dynamic factor to monitor the economy would not lead to calling many false peaks and no recessions would be missed (Figure 6). This is contrary to the result that would be obtained by using the filtered probabilities of the individual series that compose the factor (Figure 5).

**Monthly Indicator of GDP.**

The resulting dynamic factor extracted from the model, $F_{t|t}$, is the monthly indicator of Brazilian GDP. Table 3 presents some statistics for the indicator and its components. All components are pro-cyclical, hence displaying a positive correlation with the indicator. The growth rate of the coincident indicator, $\Delta F_{t|t}$, is highly correlated with $\Delta$Hours (0.87), followed by $\Delta$Production (0.61). The other components $\Delta$Capacity and $\Delta$Employment have a balanced contribution to the indicator, with a correlation around 52%. The least correlated is $\Delta$Wages (0.47).
Table 3 - Statistics for the Monthly Indicator and its Components

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Correlation with $\Delta F_{t \mid t}$</th>
<th>Sample Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$F_{t \mid t}$ in level</td>
<td>-</td>
<td>102.062</td>
<td>14.088</td>
</tr>
<tr>
<td>$\Delta F_{t \mid t}$quarterly (*)</td>
<td>1</td>
<td>0.104</td>
<td>2.041</td>
</tr>
<tr>
<td>$\Delta F_{t \mid t}$monthly</td>
<td>-</td>
<td>0.027</td>
<td>1.070</td>
</tr>
<tr>
<td>$\Delta$Capacity</td>
<td>0.522</td>
<td>-0.003</td>
<td>2.243</td>
</tr>
<tr>
<td>$\Delta$Hours</td>
<td>0.871</td>
<td>-0.127</td>
<td>1.588</td>
</tr>
<tr>
<td>$\Delta$Wages</td>
<td>0.474</td>
<td>-0.077</td>
<td>2.495</td>
</tr>
<tr>
<td>$\Delta$Employment</td>
<td>0.512</td>
<td>-0.099</td>
<td>0.676</td>
</tr>
<tr>
<td>$\Delta$Production</td>
<td>0.611</td>
<td>0.176</td>
<td>3.998</td>
</tr>
<tr>
<td>$\Delta GDP_{quarterly}$</td>
<td>0.705</td>
<td>0.493</td>
<td>2.279</td>
</tr>
<tr>
<td>GDP</td>
<td>0.973(**)</td>
<td>101.81</td>
<td>14.221</td>
</tr>
</tbody>
</table>

(*) Here $\Delta F_{t \mid t}$ was converted to quarterly frequency for comparison with $\Delta GDP$.
(**) This is the correlation of GDP with the coincident indicator in level, $F_{t \mid t}$.

In order to compare the results with GDP, two alternative procedures were undertaken. First, monthly variables were converted to quarterly frequency using simple average. Second, GDP was converted to monthly frequency using the local quadratic interpolation method with average matched to the observed data. The growth rate of the indicator displays a 70% correlation with $\Delta GDP$ at the quarterly frequency.

Figure 7 plots the growth rate of the dynamic factor indicator, the growth rate of GDP at the monthly frequency, and the recession phases (shaded areas). The estimated indicator is strongly related to movements in GDP. In particular, the volatility of these two series is very close, as well as the timing of changes and amplitude of the oscillations. From Table 3, the standard deviation of the growth rate of GDP is 2.3 while the growth rate of the factor is 2.0. One exception was the abrupt change in the economy during the Collor Plan. The
monthly indicator matches the steep drop in the economic activity upon the introduction of the Plan in the second quarter of 1990, when GDP decreased at a quarterly average rate of -6.7%. The indicator, however, did not go up as much as GDP did subsequently in the third quarter (6.8%). An analysis of the components of the dynamic factor explains this. The series employment, wages and hours did not display this strong upswing after the first impact of the Collor plan. In fact, only industrial production rebounded back. Thus, the indicator reflects more accurately the negative impact of the plan in the Brazilian labor market. The series on GDP per sector, obtained from IBGE, corroborates this finding – the responses of GDP from the service and agricultural sectors to the Collor Plan were much smoother than the one from industrial GDP.

Figures 8 and 9 show the estimated indicator in level at the monthly and quarterly frequency, respectively. The indicator in level was obtained directly from the Kalman filter using the identity \( F_{t-1} = \Delta F_{t-1} + F_{t-2} \). The indicator follows closely cyclical movements in GDP. In particular, the timing of business cycle peaks and troughs coincides for most recessions in the sample, which can also be seen in Figure 3. Conformity with the reference cycle in terms of the turning points is one of the most important contributions of the indicator, since this implies that it can be a useful monitoring tool to assess the state of the business cycle. Notice that some of the differences between the two series arise from the conversion to different frequencies that was used to compare them. However, GDP decreased a bit earlier than the factor in the beginning of the 1988 and 1991 recessions.

\[ F_{t-1} = \Delta F_{t-1} + F_{t-2} \]

---

\( F_{t-1} = \Delta F_{t-1} + F_{t-2} \)

26For graphical comparison, the factor is adjusted to have the same trend as GDP.
Figure 7 - Monthly Indicator of Brazilian GDP (—), Growth Rate of Real GDP (- - -), and Recession Phases (Shaded Area)

Figure 8 - Monthly Indicator of Brazilian GDP in Level (—), Real GDP (- - -), and Recession Phases (Shaded Area) at Monthly Frequency
A Monthly Indicator of Brazilian GDP

Figure 9 - Indicator of Brazilian GDP (—), Real GDP (- - -), and Recession Phases (Shaded Area) at Quarterly Frequency

Figure 10 - Out-of-Sample (—) and In-Sample (- - -) Filtered Probabilities of Recession, and Recession Phases (Shaded Area)
4.3 Out-of-Sample Analysis.

The in-sample analysis shows a remarkable historical conformity of the indicator with cyclical movements of GDP. In this section, turning point conformity of the indicator as well as its linear forecasting ability are verified out-of-sample.

In this exercise, the model is estimated from 1975:02 to 1993:12. The model is then recursively re-estimated for each month from 1994:01 to 2000:06 and out-of-sample recursive forecasts of the filtered coincident indicator and probabilities are computed. This tests the ability of the indicator to predict GDP business cycle phases out-of-sample in real time, and allows us to reproduce the information content that was available to forecasters at any point in time.

Probabilities.

Figure 10 compares the in-sample recession probabilities obtained from estimating the model for the entire sample with the out-of-sample filtered probabilities of recession. Both probabilities are very similar. In particular, the out-of-sample probabilities predict the two recessions in the period, in 1995 and in 1997-98. The out-of-sample probabilities, however, exhibit more pronounced spikes during the short-lived contraction that occurred in mid-1996. Although this event is not considered a recession given its mildness and short length (only three months), it illustrates the harder task of discerning in real time false peaks from signals of more severe upcoming recessions.

Linear Forecasts.

Composite indicators are generally used to predict expansion and recessions phases. The indicators, however, have also been shown to
display good linear predictive power to forecast output as well. This section studies the linear forecasting ability of the monthly indicator. Two linear models are used to evaluate the performance of the indicator \( \Delta F_{t\mid t} \) in predicting growth rates of GDP \( \Delta GDP \). Model A is a regression of \( \Delta GDP \) on five lags of itself, and Model B is a regression of \( \Delta GDP \) on five lags of \( \Delta GDP \) and one lag of \( \Delta F_{t\mid t} \). The comparison of the two models is a test of the marginal predictive content of the indicator in predicting \( \Delta GDP \) beyond lags of \( \Delta GDP \). Notice that the idea here is not to select the best forecasting model for \( \Delta GDP \). Instead, this exercise is performed to verify the ability of the indicator in forecasting \( \Delta GDP \) beyond lags of \( \Delta GDP \) itself.

First, both models were estimated for the full sample available for \( \Delta GDP \), from 1980:2 to 2000:1, and one-step-ahead forecasts were computed in-sample. Second, both models were estimated from 1980:2 up to 1993:4, and the estimates were then used to calculate recursive one-step-ahead forecasts out-of-sample, from 1994:1 to 2000:1. Out-of-sample results are compared to forecasts using the full sample.

Table 4 reports the predictive performance of the models. The model including a lag of the composite indicator (Model B) displays a better forecasting performance than using only lags of GDP. The in-sample adjusted coefficient of determination (\( R^2 \)) for Model B is 64%, while for Model A it is only 18%. In addition, Model B has a smaller Root Mean Squared Error (RMS), Mean Absolute Error (MAE) and Theil Inequality Coefficient (Theil IC) compared

---

28 Since the two series have different frequencies, the indicator was converted to quarterly using simple average.
29 These models were selected using Box-Jenkins procedure, Akaike Information Criteria (AIC) and Schwarz Information Criteria (SIC).
Marcelle Chauvet

to Model A in-sample. In particular, the Theil IC is substantially smaller for Model B (0.30) compared to Model A (0.55)\(^{30}\). Although both models do well in forecasting the mean of \(\triangle GDP\) (bias proportion close to zero), Model B has a superior ability in forecasting the variance of \(\triangle GDP\). More specifically, the variance proportion of the Theil coefficient – which measures how far the forecast is from the variance of the actual series – is much smaller for the model including a lag of the factor (Model B). This supports previous findings in the literature, in which the nonlinearity of the Markov regime switching generates additional cyclical movements that are useful in replicating the variability of the business cycle\(^{31}\).

<table>
<thead>
<tr>
<th>Table 4 - Sample and Out-of-Sample Forecasting Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Forecasting Performance</strong></td>
</tr>
<tr>
<td>In-Sample: 1980.2-2000.1</td>
</tr>
<tr>
<td>In-Sample: 1980.2-1993.4</td>
</tr>
<tr>
<td>Model A</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>(\bar{R}^2)</td>
</tr>
<tr>
<td>RMS</td>
</tr>
<tr>
<td>MAE</td>
</tr>
<tr>
<td>Theil IC</td>
</tr>
<tr>
<td>Bias</td>
</tr>
<tr>
<td>Variance</td>
</tr>
<tr>
<td>Covariance</td>
</tr>
</tbody>
</table>

Model A: Regression of \(\triangle GDP\) on 5 lags of \(\triangle GDP\).
Model A: Regression of \(\triangle GDP\) on 5 lags of \(\triangle GDP\) and 1 lag of \(\triangle F_{t|t}\).

\(^{30}\)The Theil IC is divided into three components: bias proportion, variance proportion, and covariance proportion. The bias and variance proportions measure, respectively, how far the mean and the variance of the forecast is from the mean and the variance of the actual series. The covariance proportion is obtained by residual as the three components add up to one. Thus, the smaller the bias and variance proportions the better the forecasts are, that is, most of the bias should be on the covariance proportion.

Figures 11 and 12 plot in-sample one-step-ahead forecasts, forecast intervals\(^{32}\), and actual changes in GDP from Models A and B, respectively. A comparison of the figures illustrates how close forecasts from Model B track variation in actual $\Delta$GDP. The one-step-ahead forecast from Model A is much smoother than actual $\Delta$GDP and does not mimic as well its oscillations. In fact, focusing on the recession phases as represented by the shaded areas, it is noticeable how forecasts from Model B turn down together with $\Delta$GDP around peaks and troughs.

These findings also hold for out-of-sample forecasting. The inclusion of a lag of the indicator in the GDP equation improves somewhat the model forecasting performance as measured by the Theil IC and the adjusted coefficient of determination, which are smaller for Model B compared to Model A (Table 4). Figures 13 and 14 show out-of-sample one-step-ahead forecasts, forecast intervals, and actual changes in GDP from Models A and B, respectively. As within sample, the most striking difference between the two models is how forecasts from Model B (Figure 14) match closely the volatility of changes in GDP (smaller variance proportion in the Theil coefficient). This feature is also observed around $\Delta$GDP turning points (shaded area).

Adding up, overall Model B displays a better forecasting performance for $\Delta$GDP than Model A, both in-sample and out-of-sample, which indicates that a lag of the dynamic factor contain useful information for forecasting changes in GDP. The factor, which is a parsimonious representation of information in several variables, is useful for both assessing the state of the business cycle as well as to form linear forecast of $\Delta$GDP.

\(^{32}\)Forecast intervals are plus or minus twice the forecast standard errors. The standard error bands give an approximate 95% forecast interval.
Marcelle Chauvet

Figure 11 - In Sample One-Step-Ahead Forecasts of ΔGDP from Model A, Actual ΔGDP, and Recession Phases (Shaded Area)

Figure 12 - In Sample One-Step-Ahead Forecasts of ΔGDP from Model B, Actual ΔGDP, and Recession Phases (Shaded Area)
Figure 13 - Out-of-Sample One-Step-Ahead Forecasts of $\Delta$GDP from Model A, Actual $\Delta$GDP, and Recession Phases (Shaded Area)

Figure 14 - Out-of-Sample One-Step-Ahead Forecasts of $\Delta$GDP from Model B, Actual $\Delta$GDP, and Recession Phases (Shaded Area)
5. Conclusions.

This paper constructs an indicator of Brazilian GDP at the monthly frequency that can be used as a monitoring tool for real time assessment of the current state of the economy. Special attention was given to the peculiar dynamical behavior of the Brazilian business cycle. In particular, the instability and abrupt changes of regimes in the history of Brazilian economic activity were explicitly modeled within nonlinear frameworks. The indicator is the output of a Markov switching dynamic factor model, which combines several macroeconomic variables that display simultaneous comovements with aggregate economic activity. The model generates as outputs a monthly indicator of Brazilian GDP and real time probabilities of the current phase of the Brazilian business cycle.

The formal representation of the indicator through the Markov switching dynamic factor model overcomes the drawbacks of the NBER traditional method. First, the model takes into account the dynamic relationship among the variables and generates an indicator that summarizes common cyclical movements underlying several sectors of the economy. Second, the model simultaneously yields probabilities that can be used to determine turning points in the indicators in a current basis, and the model can be evaluated in real time.

The monthly indicator of Brazilian GDP shows a remarkable historical conformity with cyclical movements of GDP with respect to volatility, timing of peaks and troughs, and duration of the phases. In addition, the estimated filtered probabilities predict all recessions in-sample and out-of-sample. The indicator, which is a parsimonious representation of information in several variables, can also be used to form linear forecasts of GDP. A linear autoregressive model including a lag of the indicator in addition to lags of the growth rate of GDP displays a better in-sample and out-of-sample forecasting
A Monthly Indicator of Brazilian GDP

performance than a univariate autoregressive model of GDP. In particular, the inclusion of a lag of the indicator improves substantially forecasts of the severity of recessions and strength of expansions, as measured by the volatility of changes in GDP. This feature is particularly accentuated around GDP turning points.

The results suggest that the estimated indicator could be useful in assessing the state of the Brazilian business cycle, through the filtered probabilities, and in forecasting GDP, through the filtered indicator, in real time on a monthly basis.


References


and Econometrics, (John Wiley & Sons), 53–73.

Hansen, B. E. 1996. ”Inference When a Nuisance Parameter is not Identified Under the Null Hypothesis”. *Econometrica* 64:413–430.


Marcelle Chauvet


