The role of expectations in modeling and forecasting industrial production

Resumo: Este estudo analisa a atual necessidade em se obter informações tempestivas sobre a trajetória de agregados macroeconômicos e propõe a modelagem e previsão da atividade industrial no Brasil, ao inovar uma função de produção que incorpora um insumo adicional para captar as expectativas empresariais acerca das perspectivas futuras de negócios. A modelagem segue um procedimento metodológico de adequada aplicação empírica para previsões baseado em modelos VAR. As estimativas baseadas no período Abr/1995-Nov/2013 confirmam a robustez das variáveis qualitativas e o sentimento dos empresários como um fator de inovação na decisão de produzir, além de mostrar que quanto maior a incerteza sobre o ambiente econômico, mais representativo será o papel das expectativas como determinante da produção industrial doméstica.

Palavras-chave: Produção industrial brasileira; Expectativas empresariais; Mudanças estruturais endógenas; Previsão.

Abstract: This study examines the current need to obtain timely information about the path of macroeconomic aggregates over time. It is proposed in the methodological framework to model and forecast the industrial activity in Brazil through a function that takes as arguments, besides proxies for capital and labor, a suitable variable that reflects subjective sentiment about the future business prospects. The modeling strategy for best forecasting relies on VAR models. Estimates from seasonally adjusted data in the period Apr/1995–Nov/2013 confirm the robustness of the qualitative variables and the entrepreneurs’ expectations, besides showing that the greater the uncertainty over the economic environment is, the more representative will be the role of expectations as a determinant of domestic industrial production.

Keywords: Brazilian industrial production; Subjective expectations; Endogenous structural changes; Forecasting.

JEL: C53; C82; E37

The authors are grateful to the Brazilian Institute of Economics at the Getulio Vargas Foundation (IBRE/FGV) for the dataset and the National Research Agency (CNPq) for the financial support.
1. Introduction

The dynamic of the business cycle requires obtaining increasingly timely information about the behavior of certain macroeconomic aggregates in order to make reliable predictions about future economic performance. However, for some reason, there is a considerable delay before official statistics are disclosed. In Brazil, for instance, it takes an average of six months for disclosure of prior quarterly GDP.

Aiming to overcome this problem, it is here proposed a model to forecast industrial output using a suitable set of quantitative and qualitative data, which include a variable that reflects entrepreneurs' subjective judgment. The advantage in so doing is that the output indexes can be obtained at the end of each month, in contrast to data from government agencies, which take a delay of at least two months to be disclosed. So, this procedure provides a more reliable way for investors, business executives and policymakers to anticipate trends in the economy.

The recent structural changes in so-called emerging economies pose an obstacle to the regularity of information, especially of a quantitative nature, over a time series. In the Brazilian case, for example, the impact of the interest rate or price indexes on GDP or industrial production in the 1980s, marked by hyperinflation interspersed with fleeting intervals of stability after a series of frustrated monetary stabilization plans, cannot be compared to the same effect from 1994 onward, as the implementation of the Real Plan succeeded in taming inflation. The sentiment of business executives, on the other hand, is non-dimensional since a situation of dissatisfaction with high inflation in the 1980s can be compared to dismay over the high interest rates in the 1990s. But it still cannot be securely inferred that confidence indicators of the Brazilian economy have not gone through structural changes.

Despite their relative advantages, studies involving qualitative information are still recent and very scarce in Brazil and many other developing countries, as are the sources of detailed and reliable information. The Brazilian Institute of Economics at the Getulio Vargas Foundation (IBRE/FGV) was the pioneer in this type of qualitative research on surveying industrial productions in the 1960s, whose objective is to characterize and anticipate trends in economic activity in Brazil, and more recently in Latin America.

Therefore, the paper aims to forecast industrial output based upon a model in reduced form for industrial performance, using as proxies for labor and capital, respectively, the level of employment predicted by business executives and the level of installed capacity utilization, one of the few quantitative variables in this study. Besides these, it is inserted a key-variable that reflects the expectation of executives on the future situation of their businesses. As predicted from macroeconomic theory, labor, capital and expectations are explanatory variables to the industrial production.

However, although the problem of structural changes is more common when working with quantitative data, the use of qualitative data is not free from the existence of structural breaks in the time series. A variety of econometric techniques have been developed to deal with the problem of structural breaks, largely prompted by Lucas critique\(^1\), as it has come to be known, which warns the risk of imprecision incurred by the application of time series econometric techniques to variables subject to structural changes. In other words, the problem of regime changes is particularly nettlesome, as pointed out by Lucas (1976), since the behavior of rational individuals should change whenever the policy environment within which they live changes. In the forecasting context this is particularly interesting critique because an economic forecast affects people’s expectations, which in turn affects the outcome of the variable being forecast, and so invalidates the forecast. This critique is largely ignored by econometricians, mainly because it does not appear to be of substantive magnitude, as argued by Stanley (2000), for instance, but a contrary standpoint is sustained by others as Linde (2001).

---

\(^1\) This refers to a critique formulated by Lucas (1976) regarding the indiscriminate use of econometric models to evaluate or predict the effects of proposed changes in economic policies based on relationships observed in historical time series of aggregated economic data. As a result, qualitative information is very important for economic policymaking.
On this point, important advances have been made in econometric theory involving structural changes, both to analyze and define economic cycles more precisely and to make inferences about regime changes in the main economic aggregates. Therefore, modeling industrial output through a combination of qualitative and quantitative variables is the core of this paper. This will be done by allowing for multiple structural breaks in the explanatory variables in agreement with Bai and Perron (1998), whose eventual occurrence will be incorporated to make further predictions. The results suggest a gradual deceleration of the Brazilian economy in 2011 and demonstrate that the greater the uncertainty is about the future economic scenario, the more representative is the role of executives’ expectations in determining industrial activity.

The article includes five additional sections. The second section contains a review of selected studies in the literature on the use of qualitative data, followed in the third section by presentation of empirical evidence and a brief discussion of economic cycles and the adjustment of the qualitative data to the indicator to be modeled. The methodology and the results are described in sections four and five, respectively, followed by the concluding remarks.

2. Qualitative data studies applied to economic models

The qualitative data gathered in business surveys have received growing attention from researchers, especially in the area of macroeconomics, which deals with production, investment and business cycles. These surveys are recognized for their utility and are widely used in academic studies, such as those reported in the publications of the Confederation of British Industry, KOF Swiss Economic Institute, National Institute of Economic Research in Sweden, The Conference Board (TCB) and Institute for Supply Management (ISM) in the United States, and the IFO Institute in Germany.

Abberger (2005) investigated the use of qualitative business tendency surveys to predict investment in Germany. He argued that official quarterly data regarding investment levels are disclosed some 16 months after data from qualitative surveys. To assess the fit of the data, he used a “spectral analysis” and autoregressive models to make predictions, and concluded that this type of study is a powerful tool to evaluate investment trends. He expanded on his previous study in 2007, incorporating other variables, such as business expectations, besides covering three time horizons – past, present and future – for certain variables. In a first approach to determine the forecasting power of the variables of the current business situation and for the next three months, he used simple regression analysis and concluded that the business expectation variable for three months beforehand is more important for predicting quarterly growth rates. In a second approach, he expanded the number of variables and applied them to four economic segments: industry, construction, retailing and wholesaling. Due to the large number of variables, he used a principal component regression. In this case the business sentiment variable was not as significant as in the previous one.

In a study of the contribution of diverse qualitative variables from business surveys on the forecast for growth of Swedish industrial production, Lindström (2000) concluded that this type of information is very useful for investigators, policymakers and financial market participants by reducing their uncertainties about the current state of the economy.

Also using qualitative data to make short-term predictions on industrial production, Kauppi, Lassila and Terasvirta (1996) conducted the analysis by sectors in a period before and during the Finnish recession at the start of the last decade. The results suggested that the use of data of that nature increased the precision of estimating some, but not all, segments studied, especially in the non-metallic materials and forest products industries.
Hansson, Jansson and Löf (2003) applied a dynamic factor model combined with the vector autoregressive (VAR) method with parsimonious parameters to verify the utility of business tendency surveys for short-term forecasting of some macroeconomic aggregates, such as real GPD (the main focus), unemployment, price and wage inflation. To test the model’s efficacy, they compared the results of the predictions with VARs that used the survey variables directly without filtering (the procedure proposed in the study), as well as economic activity indexes. The authors concluded that the method proposed surpassed the other alternatives in the majority of the cases, but the results were not conclusive in relation to the other aggregates. Also comparing different estimation methods, Siliverstovs (2009) employed data collected by business tendency surveys to forecast GPD and concluded that the use of such data for the predictions made six months before the first official estimate of GPD are more precise than those from application of univariate autoregressive models.

With the aim of investigating the Euro Zone, Bruno and Lupi (2003) considered a sample of its three largest economies: France, Germany and Italy. They used an approach based on dynamic factors and unobserved components models, mainly employing data from business surveys from the three countries. The results indicated high accuracy for a period of up to six months ahead. Similarly, in an attempt to relate the various surveys conducted in the Euro Zone countries, Lemmens, Croux and Dekimpe (2005) carried out a Granger causality test between the production expectation series and the observed data, individually for each country, and a simultaneous multivariate analysis among twelve countries sampled. There was statistical evidence toward Granger causality for only seven of the twelve countries, but a strong simultaneous effect of causality at the multinational level, indicating that some countries have a lower probability of being influenced by others.

From what is known to us, there is scarcity of papers in the literature focusing developing countries, particularly Brazil. Thus, it is worth mentioning Montes and Bastos (2013) that use qualitative and quantitative measured quarterly in the period 2000.01-2010.02 to capture the effect of the macroeconomic environment on the expectation of industrial entrepreneurs in Brazil, which is done through estimations of econometric equations by OLS and GMM methods. In connection to our paper, they run an equation relating effective industrial production index as a function of solely credit (banking operation to the industry private sector as a proportion of GDP) and industrial business confidence index (provided by 24 federations of industries in Brazil). They found positive and significant coefficients for both variables.

3. Empirical evidence

As presented in the preceding section, the literature on the application of qualitative survey data concentrates in developed countries. Scientific papers done for developing countries are still of lean contribution to the literature, although there is for Brazil a longstanding business survey focusing on the industrial sector published by the IBRE/FGV since the 1960s, which contains qualitative information on the manufacturing sector¹, and more recently the state federations of industries have been providing qualitative information concerning entrepreneurs’ expectations. A data series on industrial production dating back to the 1970s is disclosed by the Brazilian Federal Bureau of Geography and Statistics (IBGE). Descriptions of the series on industrial output and the fit of the qualitative sentiment data to the respective indicators are treated in the next section.

---

¹ Besides the public entity (IBGE) for data disclosure, The Brazilian Institute of Economics (IBRE) of Getulio Vargas Foundation (FGV) is a very important private entity that reports economic data and indexes in Brazil.

² More recently, the state industries federations in Brazil have been supplying information on qualitative data.
3.1 Industrial Production

The Industrial Production Index (IPI) used in this study comes from the Monthly Industrial Survey of Physical Production, carried out by the IBGE since 1971 for the transformation industry, and since 1975 for the extractive industry and for the industrial sector as a whole.

This survey involves various mutually independent geographic levels. In other words, there is a national selection and a series of other regional ones. The choice of “products and informants” is done using an intentional sampling selection technique in order to obtain “representative samples of at least 50% of the set of activities selected in each geographic detailing” 4, using as a basis the value of gross industrial output for 1998. Figure 1 depicts the evolution of the IPI, on a fixed basis, for the transformation industry in the monthly period 1995.04–2013.11.

**Figure 1.** Brazilian industrial production index (IPI)

Several episodes in the Brazilian and global economy that influenced the path of industrial output can be noted in the Figure, reflected in different scenarios. There was modest growth from early 1995 to late 1997; a decline to the end of 1998, likely associated with the contagion from the Asian crisis; a new expansion cycle between 1999 and mid-2001, resulting largely from the change from a pegged to a floating exchange rate regime; a period of relative stagnation between 2001 and 2003, possibly associated with the American crisis of 2001; a period of high growth starting in 2003 with the change to a new government, which peaked in July 2008, one month before the global economic crisis hit with a vengeance, causing the index to plunge until December that year; and a period of economic recovery in 2009 and early 2010, driving output to the pre-crisis level in March, after which there were signals of a weak deceleration and then the stabilization started in mid-2012.

All these periods of economic crises and instability stimulate and justify the need for an approach that can deal with structural breaks, so as to avoid misspecification errors and to provide consistent results.

3.2 Explanatory power of the qualitative data

Three qualitative indicators have been chosen for the empirical exercises to be performed, which rely on the employment forecast, the installed industrial capacity utilization and the assessment of the business climate by company managers. Thus, there are two indicators closer connected to the quantity of the inputs and another that measures business sentiment and its respective effect on activity in the industrial sector.

---

4 Available (in Portuguese) at: www.ibge.gov.br.
Figures 2, 3 and 4 depict the evolution of the Industrial Production Index (IPI) for the transformation industry (manufacturing), seasonally adjusted, and computed by IBGE, together with each indicator used in the empirical exercise that will be presented next sections.

**Figure 2.** IPI of manufacturing and capacity utilization

![Image of Figure 2](image2.png)

*Source:* Brazilian Official Bureau of Statistics (IBGE - IPI) and IBRE/FGV (Capacity Utilization)
*Note:* Seasonally adjusted data – Apr/95 to Nov/13.

**Figure 3.** IPI of manufacturing and employment expectations

![Image of Figure 3](image3.png)

*Source:* Brazilian Official Bureau of Statistics (IBGE - IPI) and IBRE/FGV (Employment)
*Note:* Seasonally adjusted data – Apr/95 to Nov/13.

**Figure 4.** IPI of manufacturing and business climate expectations

![Image of Figure 4](image4.png)

*Source:* Brazilian Official Bureau of Statistics (IBGE - IPI) and IBRE/FGV (Busclim)
*Note:* Seasonally adjusted data – Apr/95 to Nov/13.
The correlations of the level of installed capacity utilization, entrepreneurs’ employment forecast and business situation for the coming six months in relation to the IPI are 66%, 79% and 50% respectively, evidencing that the indicators in the qualitative surveys moves closely to the IPI along the investigated time series.

Besides these variables obtained from the Industrial Sounding, the industrial production model adopted here uses a quantitative variable, the interest rate. A negative correlation can be expected between the interest rate and industrial output, because the higher the interest rate is, the more expensive financing for investments is. However, the effect of changes in the interest rate on production is not immediate. According to the Brazilian monetary authority institution (Banco Central do Brasil, 2007), the average lag between a change in the interest rate and the corresponding effect on industrial production is three months. Nevertheless, the correlations indicated that the optimal number of lags is four months. Figure 5 shows the paths of the Industrial Production Index and the SELIC interest rate (the benchmark rate), accumulated in each month, with four lags, for the same previous period. As expected, the correlation between these variables is negative (-77%).

Figure 5. IPI of manufacturing and official interest rate (SELIC)

Source: Brazilian Official Bureau of Statistics (IBGE - IPI) and Brazilian Central Bank (Selic)
Note: Seasonally adjusted data – Apr/1995 to Nov/2013 and Selic with four lags.

All these examples come to motivate the empirical study to be presented in the next section, where an econometric model to determine the level of industrial activity is built up by combining qualitative and quantitative variables.

4. Methodological framework

This section details the database and the methods for formulating a model to forecast industrial output and also propose a model in reduced form for this variable based on the combination of quantitative and qualitative data and a variable that measures business sentiment.

Vector Autoregression (VAR) technique is applied for the prediction model, plus the proposal of Diebold and Mariano (1995), who presented a robust criterion for choice of the best forecasting model. The process of modeling industrial production follows Bai (1997) and Bai and Perron (1998), whose methods have the advantage of allowing multiple endogenous structural breaks in the parameters of the proposed response equation, providing robustness to the Lucas critique.
4.1 Data

The variables have monthly frequency and cover the period from Apr/95 to Nov/2013, for a total of 224 observations, classified into five variables, according to their type, as described in Chart 1.

**Chart 1.** Classification and description of the variables included in the models

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Description</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPI</td>
<td>Quantitative</td>
<td>Index with base 100(^1)</td>
<td>IBGE</td>
</tr>
<tr>
<td>Employment</td>
<td>Qualitative</td>
<td>Indicator from survey responses on business executives’ hiring intentions(^2)</td>
<td>FGV</td>
</tr>
<tr>
<td>Capacity</td>
<td>Qualitative</td>
<td>Percentage of installed capacity utilization from entrepreneurs’ expectations(^3)</td>
<td>FGV</td>
</tr>
<tr>
<td>Busclim</td>
<td>Qualitative</td>
<td>Indicator from survey responses on business executives’ beliefs regarding the business climate(^4)</td>
<td>FGV</td>
</tr>
<tr>
<td>Interest</td>
<td>Quantitative</td>
<td>Deviation of the SELIC rate in relation to its natural level(^5)</td>
<td>BCB</td>
</tr>
</tbody>
</table>

Note: (1) monthly Industrial Production Index of the transformation industry; (2) indicator with values from 0 to 200, obtained from the difference between the optimistic and pessimistic intentions of executives in relation to the level of employment in the next period; (3) percentage of installed industrial capacity utilization reported by companies; (4) indicator with values from 0 to 200, obtained from the difference between the optimistic and pessimistic beliefs of executives in relation to the business climate; (5) deviation obtained in relation to the value calculated with a Hodrick-Prescott filter.

The data on industrial production and the SELIC rate were obtained from the IBGE and Brazilian Central Bank (BCB), respectively, while the values of the qualitative indicators were obtained from the IBRE/FGV. These two indicators come from its “Transformation Industry Sounding” survey in which business executives provide information on their firms (headcount and production capacity and limitations), products (production, demand, inventory, order book, prices, etc.) and also on the envisioned business climate (current and for the next six months), whose indicators are employment, level of installed capacity utilization (Capacity) and current business climate (Busclim) have been chosen.

In relation to employment, a question is asked to the interviewers with three responses regarding the headcount in the next three months (“higher”, “equal” or “lower”), to reflect the intention of firms to hire or lay off workers. Analogously, the question for calculating the business sentiment variable can be answered with “better”, “equal” or “worse”. In both cases, the qualitative variables are quantified in the traditional way as done by international organizations such as TCB or ISM, by the balance of responses, given by the difference between the favorable and unfavorable responses to each question\(^7\). For the quantitative variable Capacity, the response is given in relation to a range of 10% in amplitude. All the explanatory variables used in the model are stationary: Employment and Busclim can vary from 0 to 200, Capacity is within the interval [0; 100] and Interest rate is measured as a deviation in relation to its natural level. So, the task now is to analyze the stationarity of the Industrial Production Index (IPI)\(^6\).

4.2 Predicting industrial production

In the first investigation proposed, it is aimed to contribute to the existing debate in the international literature on business cycles by making predictions for the IPI based upon the proposal of Diebold and Mariano (1995) for the choice of the best forecasting model to the traditional vector autoregressive (VAR) method. This is done through the inclusion of sentiment variables in the models, which follow a general specification according to the equation:

\[^1\] One hundred is added to the balance of extreme values so as not to have negative values, so that the maximum value is equal to 200.

\[^2\] Besides the stability of each VAR model, even though the controls varied in a limited range, ADF tests are performed for all of the series as a way to attest the absence of a short-term explosive trend of the controls utilized.
\[ Z_{nt} = \sum_{k=1}^{m} \Phi_k Z_{n-t-k} + \varepsilon_t, \]

where \( Z_{n-t} \) for \( n = 1, 2, \ldots, 6 \), is a vector of seasonally adjusted endogenous variables for each group of two to four variables tested below, and \( k \) is the optimal number of lags, chosen according to the Schwarz criterion, or Bayesian information criterion (BIC). Therefore, we consider six models, according to equation (1)\(^7\).

\[
\begin{align*}
Z_{1t} &= \begin{bmatrix} IP1 \\ Employment \\ Capacity \end{bmatrix}, \\
Z_{2t} &= \begin{bmatrix} IP1 \\ Employment \\ Capacity \\ Busclim \end{bmatrix}, \\
Z_{3t} &= \begin{bmatrix} IP1 \\ Employment \end{bmatrix}, \\
Z_{4t} &= \begin{bmatrix} IP1 \\ Busclim \end{bmatrix}, \\
Z_{5t} &= \begin{bmatrix} IP1 \\ Busclim \end{bmatrix}, \\
Z_{6t} &= \begin{bmatrix} IP1 \\ Busclim \end{bmatrix}.
\end{align*}
\]

For each model, it is made in-sample predictions with the choice criterion being the minimization of the mean squared error (MSE) of these predictions. Considering the variability of each set of predictions, it is desirable to be able to accurately rank the models according to the magnitude of the MSE for the predictions made. More specifically, small differences between the MSEs of the in-sample predictions can be statistically nil.

### 4.2.1 Forecast testing

As already mentioned, the forecast relies on Diebold and Mariano (1995, p.2-4) who proposed a method to ascertain whether there is a difference between two MSEs. According to this approach, given a determined time series \( \{y_t\}_t=1 \) for which there are two good predictive models \( \{\hat{y}_t\}_t=1 \) and \( \{\hat{y}_t\}_t=1 \), with the respective forecast errors \( \{e_t\}_t=1 \) and \( \{e_t\}_t=1 \), each \( y_t \) realized and a predicted value of \( \hat{y}_t \) (or \( \hat{y}_t \)) are associated to a function \( g(y_t, \hat{y}_t) \) such that \( g(y_t, \hat{y}_t) = g(e_t) \), which is the MSE. Defining the loss function as:

\[
d(t) \equiv [g(e_t) - g(e_t)]
\]

(2)

The structure of the hypothesis testing concerning the power of the prediction is based upon the expectation of that function, that is: \( H_0: E[d(t)] = 0 \) and \( H_1: E[d(t)] \neq 0 \).

The test statistic is given by:

\[
S = \frac{d}{\sqrt{\hat{\sigma}_d(0)}}
\]

where \( \bar{d} = \frac{1}{T} \sum_{t=1}^{T} [g(e_t) - g(e_t)] \) is the sample mean of the loss function, \( \hat{\sigma}_d(0) = \frac{1}{2T} \sum_{t=1}^{T} \gamma_d(t) \) is the spectral density of the loss function at frequency zero, \( \gamma_d(t) = E[(d_t - \mu)(d_{t-t} - \mu)] \) is the autocovariance of the loss differential at displacement \( t \) and \( \mu \) is the population mean loss function.

### 4.3 Modeling industrial production

The next application is to model industrial output by inserting a measurable qualitative variable given by the sentiment of business executives as argument, which turns the model into an innovative production function with the following specification:

\[
Y_t = f(K_t, L_t, \tilde{X}_t)(3)
\]

where \( K_t \) and \( L_t \) represent fixed capital and labor, respectively, and \( \tilde{X}_t \) denotes a vector of other variables that affects output, in which includes the expectations of industrial firms’ entrepreneurs. Since a production function measures potential rather

\footnote{Nevertheless, only the estimates of the two best models are presented in the appendix.}
The role of expectations in modeling and forecasting industrial production

than actual output, which is the real objective of the modeling and prediction of this study, our formulation consists of an industrial production model that combines qualitative and quantitative variables, whose econometric specification is as follows:

\[ IPI_t = \mu + \alpha \text{Employment}_{t-1} + \beta \text{Capacity}_t + \lambda \text{Busclim}_t + \theta \text{Interest}_{t-4} + \nu_t \quad (4) \]

As defined in the first application, \( IPI \) is the Industrial Price Index at time \( t \); Employment represents the hiring/firing intentions of firms in comparison with the immediately preceding period; Capacity is the level of installed capacity utilization; Busclim represents the business climate as envisioned for the next six months by the executives surveyed. We also include a quantitative variable, Interest, which is measured by the deviations of the basic interest rate in relation to its natural value via application of a Hodrick-Prescott filter. With respect to the causality of Interest on the IPI, it was estimated an optimal level of four lags.

To overcome the obstacles exposed by the Lucas critique, as stated by Stock (2004), it is estimated a model that does not neglect the possible existence of structural breaks in the parameters of equation (4). The literature contains two main approaches for this type of estimation. The first, proposed by Andrews (1993), is applied to nonlinear models and permits selecting only one break. This is ruled out because industrial production is a series that \textit{a priori} has a deterministic trend, besides the restriction that the asymptotic distributions for the F-tests must be constructed for non-trended regressors.

Therefore, the estimation model used here is that proposed by Bai and Perron (1998), which possesses the advantage of permitting multiple endogenous structural breaks and makes no requirements about the existence or not of trends in the regressors. Moreover, it is necessary to reduce the computational effort to perform the estimations, which is based on the algorithm described by Perron (1997).

Following this approach, we estimate a model that permits estimation of equation (4) for sample sub-periods according to the successive verification of structural breaks in the parameters \( \alpha, \beta \) and \( \lambda \), as shown in equation (5), with \( m (m=0, 1, 2...) \) structural breaks:

\[ IPI_t = \mu + \sum_{i=0}^{m+1} \iota_{i} \{ (\alpha, \text{Employment}_{t-1} + \beta \text{Capacity}_t + \lambda \text{Busclim}_t) + \theta_i \text{Interest} \} \quad (5) \]

Note that, as \( m = 0 \), expression (4) reduces to (5). \( I_i \) is the sub-period determined by the structural breaks \( t_i \), and \( \iota_i \) and \( \iota_{i-1} \) corresponds to the indicator function, which assumes value 1 if \( t_{i-1} \leq t \leq t_i \) and 0 otherwise. Moreover, for each sub-period \( i \) there is a vector \( (\alpha; \beta; \lambda) \), so that \( \beta \) and the intercept are defined from the complete sample.

There are two possible methods of estimation regarding whether the number of breaks is known or not. If there is evidence on a priori number of breaks, the procedure used is to estimate the first break point \( t_1 \), such that \( t_1 = \arg \min_{t} S_T(t) \), with \( S_T(t) \) defined as the sum of the squared residuals (SSR) resulting from the estimation of model (5) in the complete period. Upon identification of the first break, there are then two subsamples. The previous procedure is repeated for each of them, to obtain two more possible break points (one for each subsample), from which the one that causes the greatest reduction in the SSR of the full sample is chosen as the second break point \( t_2 \). This process is repeated until all the breaks are found.

If the number of breaks is unknown, as is the case here, since graphical analysis of the IPI series does not show an exact number of breaks, the procedure consists of testing the null hypothesis of \( m \) structural breaks against the alternative of \( m+1 \) breaks \( (m=0, 1, 2...) \) until the null hypothesis is rejected. In other words, the number of break points corresponds to the largest value of \( m \) for which that hypothesis is not rejected.
The critical values for such sequential tests, $F_T (m+1/m)$, are provided by Bai and Perron (1998, p. 61).

Sequential least squares method estimates (5) with $m=1$ for the full period and identify that $\hat{t}_1 = \arg \min S_T (t_1)$ with $S_T (t_1)$ defined as the SSR of the model with one break, where $\hat{t}_1$ is a candidate at $t_1$. The sample is then divided into two and for each sub-period, $[1,t_1]$ and $[t_1,T]$, a model is estimated with on break, providing two new potential break points, $\hat{t}_1$ and $\hat{t}_2$, respectively. The value of the second break, $\hat{t}_2$, is identified as follows: If $S_T (\hat{t}_1, \hat{t}_2) < S_T (\hat{t}_1, \hat{t}_1)$, then $\hat{t}_2 = \hat{t}_1$, otherwise $\hat{t}_2 = \hat{t}_1$. Note that $S_T (\hat{t}_1, \hat{t}_1)$ represents the SSR for model (2) with $m=2 (\hat{t}_1, \hat{t}_1)$. Bai and Perron (1998) showed that if $\hat{t}_1$ and $\hat{t}_2$ are true break points, then $\hat{t}_1$ is consistent for $\hat{t}_1$. The sample is then subdivided into three: $[1,\hat{t}_1], [\hat{t}_1,\hat{t}_2], [\hat{t}_2,T]$, and so on.

An alternative way to test that procedure is to combine the above sequential method with the Bayesian information criterion $BIC (m)$ of Yao (1998) and the $LWZ(m)$ of Liu et al. (1997). The optimal number of breaks is attained when the minimum values of these information criteria are reached. These criteria are appropriate for models with multiple breaks because they introduce a penalty factor for additional breaks, which necessarily reduces the value of the SSR.

5. Results

This section is divided into two parts, according to the two prediction and modeling exercises proposed, whose methods were described in the preceding section. Initially we present the forecasts of two models with a sample time horizon redefined consistent with the structural changes explained in the second subsection and the differential for choice of the models is the business sentiment variable. The second exercise consists of only one model to quantify the impact of business expectations and the adherence of the qualitative variables.

5.1 In-sample predictions

To investigate the possible existence of an explosive trend in the variables, even of short duration, we performed ADF unit-root tests in order to specify the order of integration of the IPI series for the transformation industry, which result a first-order integration of the IPI, but the VAR models remained stable.

The prediction models are estimated for the period Jan.2009-Nov.2013 as a result of the most recently of those three break dates estimated according methodology presented in subsection 4.3 and showed next subsection. It is also worth noting the suitability of such period for forecasting since after the last break it encompasses a period of post worldwide financial crises. Each group of two equations is differentiated by the sentiment variable, and the MSE associated with the in-sample forecasts for two time horizons – six and twelve months – of the VAR models specified in equation (1) are shown in Table 1. The $S(1)$ statistic, according to the proposal of Diebold and Mariano (1995), tests the null hypothesis of equality between the MSEs of the forecasts in each group of models.

The results of the three groups of prediction models are presented in Table 1, where the first is composed with capital and labor (employment) proxies, the second with capital only and the third with labor only. To each of these three models is added a sentiment variable, and the predictive power of each of them is compared with its respective version without it through the test of Diebold and Mariano (1995), for the in-sample periods of twelve and six months.

---

10 For $m$ structural breaks, these criteria are given by:

\[ BIC(m) = n \log \text{RSS} - (p^* + m + p) \log n \],

\[ LWZ(m) = n \log \text{RSS} - (p^* + m + p) \log n + \frac{mp^*}{2} \log n \],

where $p^* = (m + 1)q + m + p + 1$ when represents a penalty factor that offsets the reduction of the SSR for each additional break, with $p$ and $q$ defined as the number of coefficients that change and remain the same, respectively, between the regimes (Perron, 1998).

11 According to Bai (1997), in the case of regressors with trend, a 95% confidence interval (CI) for the estimated break point, $\hat{t}_1$, can be computed as $[\hat{t}_1 - c_{\alpha/2}, \hat{t}_1 + c_{\alpha/2}]$, where $c_{\alpha/2}$ is the nearest whole number to $\alpha/2$, $\alpha$ is the 97.5% quantile and $\hat{t}_1$ and $\hat{t}_2$ represent the estimates of the government response before and after the break, $\hat{t}_1$ and the estimated variance of $\hat{v}^2$, respectively.

12 For more details, see Appendix A

13 Unit roots for the best models are shown in Appendix C.

14 In this case, it is considered the possibility that the expectation variable could contribute more in the short run because of the short memory of economic agents.
Table 1. Mean squared error of forecasts and Diebold-Mariano tests

<table>
<thead>
<tr>
<th>Models</th>
<th>Variables</th>
<th>12 months</th>
<th>6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Z_{i}$</td>
<td>IPI Employment, Capacity</td>
<td>31.951</td>
<td>16.152</td>
</tr>
<tr>
<td>$Z_{s}$</td>
<td>IPI Employment, Capacity, Busclim</td>
<td>3.466</td>
<td>5.021</td>
</tr>
</tbody>
</table>

Diebold Mariano Statistic: $S(1)$  
comparing $Z_{i}$ and $Z_{s}$  
$p$-value*  
0.000  
0.000

---

Diebold Mariano Statistic: $S(1)$  
comparing $Z_{s}$ and $Z_{o}$  
$p$-value*  
0.000  
0.001

Diebold Mariano Statistic: $S(1)$  
comparing $Z_{o}$ and $Z_{m}$  
$p$-value*  
0.000  
0.000

Source: Own calculation.  
Note: (*) Under the null hypothesis, the MSEs are equals.

For the period of twelve months, the inclusion of the sentiment variable increases the predictive power of the model with the labor proxy only, it is indifferent with installed capacity only, and weakens power as both proxies are put together. Out of the three models for the period of six months, all of them with the sentiment variable perform better, but the one with employment only. This indicates that the inclusion of business sentiment variables particularly improves the performance of the prediction for shorter term, but loses power for longer period of time. This may be due to a natural instability, subjectivity and short memory that encompass such variables.

5.2 Out-of-sample predictions

For the out-of-sample predictions, the two best models\textsuperscript{b} – for the groups with and without the inclusion of the sentiment variable – were the VARs estimated according to $Z_{i}$ and $Z_{s}$.

According to the extreme low $p$-values of the $S(1)$ statistic, the null hypotheses that the models’ predictions have the same accuracy are rejected for both six and twelve months time spans. Yet, it is also established that the sentiment variable improves the accuracy of the forecasts for longer time horizon, as shown by shaded figures in Table 1.

Additionally, Figure 6 depicts the predictive power of the two models with the out-of-sample forecasts from Dec/13 to Nov/14. The forecasts trend show there is stability and low growth rates for the Brazilian IPI in 2014. This may be due to the reduction of government incentives to avoid contraction of industrial activity.

\textsuperscript{b}They were ranked according to the minimization of the MSE of the forecasts in-sample.
5.3 Econometric modeling

As described in subsection 4.3, the estimation based upon the procedure proposed by Bai and Perron (1998) is carried out, and the results are reported in Table 2.

Table 2. Estimates of number of structural changes and dates

<table>
<thead>
<tr>
<th>Breaks: m=3</th>
<th>Dates</th>
<th>C.I. (90%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Dec./1997</td>
<td>[after Oct./1995]</td>
</tr>
<tr>
<td></td>
<td>Dec./2008</td>
<td>[After Nov./2008]</td>
</tr>
</tbody>
</table>

Source: Own estimates according to LWZ Criteria

According to the LWZ criteria, three structural changes for the parameters are identified in the estimates, which occurred in Dec/97, Jul/05 and Dec/08. In relation to the confidence intervals, the estimates indicate four different economic scenarios relating the controls and the dependent variable: the first in the beginning of the Real Plan; the second after December of 1997, a period of high exchange rate volatility together with the alteration of the exchange rate regime in Brazil, and the Asian and Russian crises; the third is characterized by a period of high growth of Brazilian economy until the international financial crisis; the fourth period encompasses the 2008 crisis.

These results become even more robust and plausible because the date of the last break is closely linked to the variability of the São Paulo Stock Exchange Index, which in May/08 had already started the downward trend, and to the sharp international crisis that got incited in Sep/08 with the bankruptcy of Lehman Brothers Bank.

The increase in the Brazil risk indicator, the sharp valuation of national currency and the increasing inflation expectations reflected the uncertainty about the economic outlook in the last structural change period. The initial skepticism of economic agents, regarding the political actions that were impending from the government, was largely dampened in the next few months, with the decision to place priority on avoidance the international crisis’ effects, even at the cost of higher inflation.

In relation to the estimation technique proposed, the structural changes, captured by the statistical tests, imply that any structural or reduced-form model that ignores these changes will incur a serious specification error. To overcome this problem, the procedure proposed by Bai and Perron (1998) proved to be the most suitable for this analysis of the upcoming dates.
The role of expectations in modeling and forecasting industrial production

determinants of the IPI. Estimates are displayed in Table 3, which testify that all control variables, specially the qualitative ones, have statistically significant impact on the IPI, but the deviations of the interest rate in relation to its natural level. The policies adopted by the government in each period can be identified by the changes in the estimates of the model (4) presented in Table 3.

Table 3. Estimates of model (5) in four regimes

<table>
<thead>
<tr>
<th></th>
<th>Apr/95-Nov/97</th>
<th>Dec/97-Apr/05</th>
<th>May/05-Apr/08</th>
<th>May/08-Nov/13</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPI</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Capacity</td>
<td>1.7346</td>
<td>-0.1983</td>
<td>0.2101</td>
<td>1.5565</td>
</tr>
<tr>
<td></td>
<td>(0.19)</td>
<td>(0.11)</td>
<td>(0.03)</td>
<td>(0.15)</td>
</tr>
<tr>
<td>Employment_{i,t-1}</td>
<td>0.4633</td>
<td>-0.1058</td>
<td>1.4339</td>
<td>0.5933</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.02)</td>
<td>(0.18)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Busclim</td>
<td>-0.0639</td>
<td>2.0238</td>
<td>0.1208</td>
<td>0.0111*</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.16)</td>
<td>(0.06)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Interest_{i,t-4}</td>
<td>1.2447</td>
<td>1.2447</td>
<td>1.2447</td>
<td>1.2447</td>
</tr>
<tr>
<td></td>
<td>(0.60)</td>
<td>(0.60)</td>
<td>(0.60)</td>
<td>(0.60)</td>
</tr>
<tr>
<td>C</td>
<td>-59.5457</td>
<td>-59.5457</td>
<td>-59.5457</td>
<td>-59.5457</td>
</tr>
<tr>
<td></td>
<td>(10.89)</td>
<td>(10.89)</td>
<td>(10.89)</td>
<td>(10.89)</td>
</tr>
<tr>
<td>N</td>
<td>32</td>
<td>89</td>
<td>36</td>
<td>67</td>
</tr>
<tr>
<td>R²</td>
<td>0.976</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Own estimations with data seasonally adjusted.
Notes: Standard errors in parenthesis. (*) Non-significant at 5%.

The coefficients in the four sub-periods indicate that the sentiment variable of Business Climate (Busclim) started to have a higher impact on the IPI after 1997 until mid-2005. This suggests that in periods of higher economic growth the importance of business expectations as an indicator of industrial production is enhanced. Despite the ideological change in the federal public administration in 2003, the economic growth after that was higher and the entrepreneurs were more optimistic than in the previous decade.

On the other hand, the installed capacity utilization (Capacity) had greater impact in the period associated with saturation of the industrial sector, as a consequence of high level of installed capacity utilization itself. Data from the National Confederation of Industry reveal that before the first quarter of 2002 until the global crisis hit in mid-2008, the percentage of installed capacity utilization maintained a predominantly increasing path, reaching 87% in the last quarter of 2007, and after two quarters of recession (fourth quarter of 2009 and first of 2009) it steadily recovered, reaching 85.4% in the third quarter of 2010. Indeed, after 2008 the variable Capacity recovered its power over IPI. Concerning employment, it was more representative in the period of highest Brazilian growth rates, until preceding the 2008 crisis.

The upward relationship between the deviation of interest rates over its natural level and the IPI might apparently be an awkward result. Economic scenarios that abnormal interest rates arise may be conducive to capitalization by the domestic enterprises, which would tend to increase production in the next periods.

As long as the model estimation relies on unofficial qualitative variables, which provided reliable high explanatory power, it can be considered an acceptable indicator for economic policy. We stress the advantage of considering the existence of structural changes in the model parameters, in so agreeing with Stock (2004) who states that neglecting this aspect implies specification errors, besides weakening reliability in the results and the inference process.
Concluding remarks

This study addresses the need of economic agents to obtain more timely information that uses data from traditional qualitative surveys, as those conducted by The Conference Board and Institute for Supply Management in the United States and the IFO Institute in Germany, to model and make predictions about the level of industrial, as we have done here for a developing country – Brazil. Additionally, we incorporate a sentiment variable in the statistical models related to business executives’ expectations for their companies that permits more precise measurement of the impact of uncertainty about the economic environment on the level of industrial activity.

We conducted two empirical exercises, the first using VAR models combined with the proposal of Diebold and Mariano (1995) for choice of the model with the best predictive capacity and the second considering the possibility of multiple endogenous structural changes in the reduced-form model for industrial production proposed, following the method proposed by Bai and Perron (1998).

The prediction models demonstrated high explanatory power for the qualitative variables, and the business expectations played a very relevant role to anticipate trends especially over a 12-monthes horizon. In some cases the models that incorporate business sentiment were equal to the best models composed only by variables considered to be inputs of industrial output. The out-of-sample forecasts also suggested stability for industrial activity in Brazil in 2014, a result that was borne out.

The results of the model for the determinants of the Industrial Production Index suggest the occurrence of three structural changes: Dec/97, Jul/05 and Dec/08. These findings suggest four different regimes to the relation between Installed Capacity, Employment and Business Climate with Brazilian Industrial Production: the first in the beginning of the Real Plan, the second after Dec/97 in a period of high exchange rate volatility including the alteration of the exchange rate regime in Brazil and the Asian and Russian crises; the third in a period of high growth of Brazilian economy until the international financial crisis and the fourth after the 2008 crisis.

The qualitative variables possess high adherence and statistical relevance for predicting production, and the randomized coefficients through regime structural breaks allowed inferring that uncertainty over the economic environment associated to the ideological change in the government impacts significantly and is reflected in business expectations. In sum, this means that the greater the uncertainty, the lower the marginal response of the qualitative variable.

It is believed that this paper has shed contributing remarks to the debate over the insertion of new variables and the use of techniques that allow considering subjective parameters in quantitative models, by attesting to the important role of qualitative surveys., Although there is already a consolidated practice for this type of modeling in developed economies, but very scarce in developing ones, it is suggested here an innovative way for applying Brazilian qualitative data by taking into account an estimation process that does not neglect the presence of structural changes in the parameters of economic models.

References

The role of expectations in modeling and forecasting industrial production


BRASIL. Banco Central do Brasil. site: www.bcb.gov.br. nov/2012.


___. IPEADATA - Instituto de Pesquisa Econômica Aplicada. site: www.ipeadata.gov.br. nov/2012.


STANLEY, T.D. An Empirical Critique of the Lucas Critique. Journal of Socio-


APENDIX A – UNIT ROOT TEST

Table 5. Unit root ADF tests

<table>
<thead>
<tr>
<th></th>
<th>Level</th>
<th>1st Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Statistic</td>
<td>p-value</td>
</tr>
<tr>
<td>IPI</td>
<td>-0.923</td>
<td>0.7796</td>
</tr>
<tr>
<td>Employment$_{t-1}$</td>
<td>-2.842</td>
<td>0.0541</td>
</tr>
<tr>
<td>Capacity$_{t-1}$</td>
<td>-3.066</td>
<td>0.0307</td>
</tr>
<tr>
<td>Busclim$_{t-4}$</td>
<td>-4.131</td>
<td>0.0011</td>
</tr>
<tr>
<td>Interest$_{t-4}$</td>
<td>-3.248</td>
<td>0.0186</td>
</tr>
</tbody>
</table>

Source: Own calculation

APENDIX B – Best VAR Models’ Estimates

Table 6a. Best VAR model without sentiment

<table>
<thead>
<tr>
<th></th>
<th>IPI</th>
<th>Capacity</th>
<th>IPI</th>
<th>Capacity</th>
<th>Busclim</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPI$_{t-1}$</td>
<td>0.7549</td>
<td>0.1078</td>
<td>IPI$_{t-1}$</td>
<td>0.5105</td>
<td>0.0641</td>
</tr>
<tr>
<td>[ 4.777]</td>
<td>[ 3.653]</td>
<td></td>
<td>[ 3.736]</td>
<td>[ 2.446]</td>
<td>[-0.140]</td>
</tr>
<tr>
<td>IPI$_{t-2}$</td>
<td>0.0568</td>
<td>-0.0533</td>
<td>IPI$_{t-2}$</td>
<td>0.1446</td>
<td>-0.0378</td>
</tr>
<tr>
<td>[ 0.343]</td>
<td>[-1.73]</td>
<td></td>
<td>[ 1.074]</td>
<td>[-1.46]</td>
<td>[-0.830]</td>
</tr>
<tr>
<td>Capacity$_{t-1}$</td>
<td>1.7594</td>
<td>1.1735</td>
<td>Capacity$_{t-1}$</td>
<td>0.4064</td>
<td>0.9321</td>
</tr>
<tr>
<td>[ 2.301]</td>
<td>[ 8.221]</td>
<td></td>
<td>[ 0.612]</td>
<td>[ 7.315]</td>
<td>[ 1.396]</td>
</tr>
<tr>
<td>Capacity$_{t-2}$</td>
<td>-1.6363</td>
<td>-0.3867</td>
<td>Capacity$_{t-2}$</td>
<td>-0.3611</td>
<td>-0.1598</td>
</tr>
<tr>
<td>[-2.780]</td>
<td>[-3.518]</td>
<td></td>
<td>[-0.698]</td>
<td>[-1.611]</td>
<td>[-1.971]</td>
</tr>
<tr>
<td>C</td>
<td>13.1484</td>
<td>11.0456</td>
<td>Busclim$_{t-1}$</td>
<td>0.2701</td>
<td>0.0480</td>
</tr>
<tr>
<td>[ 0.642]</td>
<td>[ 2.890]</td>
<td></td>
<td>[ 5.660]</td>
<td>[ 5.244]</td>
<td>[ 10.877]</td>
</tr>
<tr>
<td>R$^2$</td>
<td>0.83</td>
<td>0.95</td>
<td>Busclim$_{t-2}$</td>
<td>-0.1886</td>
<td>-0.0334</td>
</tr>
<tr>
<td>N</td>
<td>67</td>
<td></td>
<td></td>
<td>[-3.177]</td>
<td>[-2.925]</td>
</tr>
</tbody>
</table>

Source: Own calculation

APENDIX C – VAR Stability

Table 7a. Stability of model Z5

<table>
<thead>
<tr>
<th>Characteristic Roots Polynomial</th>
<th>Modulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.802748 - 0.161523i</td>
<td>0.819</td>
</tr>
<tr>
<td>0.802748 + 0.161523i</td>
<td>0.819</td>
</tr>
<tr>
<td>0.596107</td>
<td>0.596</td>
</tr>
<tr>
<td>-0.273162</td>
<td>0.273</td>
</tr>
</tbody>
</table>

Source: Own calculations

Table 7a. Stability of model Z6

<table>
<thead>
<tr>
<th>Characteristic Roots Polynomial</th>
<th>Modulus</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.873401 - 0.213892i</td>
<td>0.899</td>
</tr>
<tr>
<td>0.873401 + 0.213892i</td>
<td>0.899</td>
</tr>
<tr>
<td>0.729013</td>
<td>0.729</td>
</tr>
<tr>
<td>0.658532</td>
<td>0.659</td>
</tr>
<tr>
<td>-0.182900 - 0.051645i</td>
<td>0.190</td>
</tr>
<tr>
<td>-0.182900 + 0.051645i</td>
<td>0.190</td>
</tr>
</tbody>
</table>

Source: Own calculations