Department of Economics- FEA/USP

Central Bank Communication Affects Long-Term Interest Rates

FERNANDO D. CHAGUE

RODRIGO DE-LOSSO

BRUNO GIOVANNETTI

PAULO MANOEL

WORKING PAPER SERIES Nº 2013-07

DEPARTMENT OF ECONOMICS, FEA-USP
WORKING PAPER Nº 2013-07

Central Bank Communication Affects Long-Term Interest Rates

Fernando D. Chague (fchague@usp.br)

Rodrigo De-Losso (delosso@usp.br)

Bruno Giovannetti (bcg@usp.br)

Paulo Manoel (paulombfm@gmail.com)

Abstract:
We empirically study how the communication of the Central Bank of Brazil affects the term structure of interest rates. Using an algorithm that classifies the words from the Central Bank minutes into predetermined semantic themes, we estimate a time-series factor related to Central Bank optimism. Then, we show that the long-term interest rates are sensitive to the optimism factor: when Central Bank is more optimistic, long-term interest rates fall. The fact that minutes are released one week after the changes in the target rate allows us to identify the effect of the communication in isolation. Our result is in line with the idea that Central Bank communication can be an effective monetary policy instrument through its impact on market expectations, particularly at the longer maturities.

Keywords: Brazilian Central Bank; central bank communication; text mining

JEL Codes: G14, E58

Central Bank Communication Affects Long-Term Interest Rates*

Fernando Chague[†], Rodrigo De-Losso[‡], Bruno Giovannetti[§], Paulo Manoel[¶]

December 2012

Abstract

We empirically study how the communication of the Central Bank of Brazil affects the term structure of interest rates. Using an algorithm that classifies the words from the Central Bank minutes into predetermined semantic themes, we estimate a time-series factor related to Central Bank optimism. Then, we show that the long-term interest rates are sensitive to the optimism factor: when Central Bank is more optmistic, long-term interest rates fall. The fact that minutes are released one week after the changes in the target rate allows us to identify the effect of the communication in isolation. Our result is in line with the idea that Central Bank communication can be an effective monetary policy instrument through its impact on market expectations, particularly at the longer maturities.

*We would like to thank Gabriel Madeira, Ricardo Madeira, Carlos Eduardo Gonçalves, Fernando Botelho and Victor Westrupp for important comments and discussions.

[†]Department of Economics, University of São Paulo. Email for contact: fchague@usp.br

[‡]Department of Economics, University of São Paulo. Email for contact: delosso@usp.br

[§]Department of Economics, University of São Paulo. Email for contact: bcg@usp.br

[¶]UC Berkeley Haas School of Business. Email for contact: paulombfm@gmail.com

1 Introduction

In the recent scenario of low interest rates in the U.S. and in the Eurozone, the use of alternative monetary policy instruments has been an important research topic. As argued by Eggertsson et al. (2003) and Ben Bernanke et al. (2004), a key ingredient in this context is the skilful management of agents expectations through credible and clear central bank communication. In this work we provide empirical evidence that central bank communication can indeed help to shape expectations: we find that the communication of the Brazilian Central Bank (BCB) significantly affects long-term interest rates.

Our empirical strategy has two steps. In step one, we analyze the content of all minutes published by the BCB between 2000 and 2012. BCB minutes are the most important instrument of monetary policy communication in Brazil, as in many other countries. They are disclosed 8 days after the regular (every 40 days on average) decision of BCB about the target for the interest rate in Brazil (the SELIC rate). The minutes summarize the view of the BCB regarding the economic environment and provide justification for the monetary policy decision.

Our analysis of the content of the minutes follows an automated procedure. First, we classify the words from the minutes into predetermined semantic themes. For this purpose we use the Harvard IV dictionary, which maps words into semantic groups, such as “Positive”, “Negative”, “Strong”, among others. Then we compute for each minute the relative frequency of each semantic group and use Principal Component Analysis (PCA) to extract factors from the time-series of such frequencies. The results suggest that the informational content present in the minutes can be summarized in a single factor. The analysis of the factor loadings and the correlations of the factor with macroeconomic variables indicate that the factor can be interpreted as the level of optimism of the BCB regarding the economic environment. For this reason, we label it the “optimism factor” (OF) [1].

In the second step of our empirical strategy, we show that the market incorporates the information of the minutes in the pricing of interest rates future contracts, with stronger

[1]The OF series can be found in the FINBRAX website http://www.fipe.org.br/web/finbrax

influence in contracts with longer maturities. We run regressions which have the changes of the yield curve on the day of the release of a minute on the left-hand side and the change of the OF as the explanatory variable. Our estimates indicate that an increase of one standard deviation of the OF generates a decrease of 3 basis points (3 bps) per year in the 6-month future rate and a decrease of 5 bps per year in the 12- and 24-month future rates. These effects are economically important: the standard deviation of the daily change of the 12-month interest rate is 9 bps, while the average of the absolute value of this same variable is 5 bps.

A plausible concern is that our result may be contaminated by the changes in the SELIC rate. Indeed, a minute with a high OF should come along with a decrease in the SELIC rate, what would directly lower future interest rates . However, since the minutes are released 8 days after the changes in the SELIC rate, we should be capturing the effect of the communication in isolation, as the innovations in the SELIC rate should be already incorporated into prices by the time the minutes are made public.

Our results are in line with the literature that shows that communication of the policy-maker plays an important role on the pricing of interest rate claims. Boukus et al. (2006) present evidence that the information content of the Federal Open Market Committee (FOMC) minutes can impact the Treasury yield around the time of the minutes release. Rosa et al. (2011) shows that the European Central Bank had been successful in moving their domestic asset prices using communication shocks. Lamlaa et al. (2011) show that long-horizon interest rates in the Euro zone are more sensitive to changes in the ECB communication. Janot et al. (2012) indicate that the BCB minutes are able to reduce uncertainties about the interest rates future path, diminishing the volatility of the future interest rates after minutes releases. There are no studies, however, about the impact of the *informational content* of the BCB minutes on the financial markets.

The rest of the paper is organized as follows. Section 2 describes the automated procedure to build the OF. Section 3 analyses the effects of the OF on the future interest rates. Section 4 presents robustness checks. Section 5 concludes.

2 The optimism factor

In June 1999 the Brazilian monetary authorities adopted an Inflation Targeting regime (IT) to set the basic interest rate (the SELIC rate) in the Brazilian economy that was largely inspired by the British IT model. In the Brazilian IT regime, the Brazilian National Monetary Council (CMN) deliberates the target for the SELIC rate on regular meetings (the COPOM meetings), which are held once every 6 weeks.

One important ingredient of every IT regime is the transparency. Eight days after each COPOM meeting, the committee releases the minutes summarizing the views of the Central Bank regarding the economic outlook and a justification for the policy decision. These minutes are closely monitored by market participants.

In order to assess econometrically how market participants react to the minutes, we first objectively measure the minutes' word content by giving a numeric value to it. We construct a time-series factor that summarizes the word content in the minutes. In this section we present our methodology to extract the time-series factors from minutes issued by the Brazilian Central Bank between 2000 and 2012.

2.1 Data

Our sample consists of 130 minutes issued between 2000 and 2012. Up to 2005, the COPOM meeting were held once every month, and so in this period we have one minute per month. After 2005, the meeting were held less frequently but regularly spaced, and so in this period we have 8 minutes per year.

We also use financial and macroeconomic variables to interpret our factors constructed from the minutes. To measure the economic activity we use the real GDP of the last twelve months, seasonally adjusted. The inflation is measured by the expected IPCA[2] consumer price index change in the next twelve months. To measure the level of prices of the stocks we use the Ibovespa stocks price index. The value of the local currency is measured by the BRL/USD exchange rate.

[2]IPCA consumer price index is the official measure of inflation used by the Brazilian Central Bank

We transform the macroeconomic and financial variables to have the same frequency of the COPOM minutes. For the variables that have lower frequencies, such as GDP growth and Consumer Prices (IPCA), we use the latest available observation. The variables that have higher frequencies, e.g. daily, such as the exchange rate and the stock market index (IBOVESPA), we compute accumulated changes from one minute disclosure to another.

2.2 Methodology

There are two main approaches to analyzing the information content of a document by identifying the semantic content of words. The first uses only the information contained in the sample of documents (e.g. Boukus et al. (2006)), in which the semantic content of words are derived from the correlations between their frequencies across documents. The second, seeking more efficiency in the use of the disposable data, uses external information in the semantic elucidation task (e.g. Lucca et al. (2009), in which the Google Search technology is used in the construction of a group of words correlated with monetary policy key terms). In this work we use the second approach and use Harvard IV dictionary to classify words. [1].

The Harvard IV dictionary is a system of classification of words (more then 7,000) according to semantic orientation. Each English word receives a semantic designations such as "Positive", "Negative" and "Strong", among others. It is possible for one word to have more than one semantic designation. For example, the word "Abandon" is classified as "Negative", "Week" and "Fail". Since december of 2012.

The dictionary classifies English words into a total 182 different semantic groups. In our analysis, we will restrict it to the 18 most relevant groups: "Positive", "Negative", "Pleasure", "Pain", "Feel", "Arousal", "Virtue", "Hostile", "Fail", "Strong", "Weak", "Power", "Active", "Passive", "Work", "Try", "Persist" and "Complete".

One problem that naturally arises in the classification of words in a document, and in information retrieval problems in general, is the identification of words with a common stem. Words with the same stem will usually have similar meanings, for example "con-

[1] For an example of the use of the Harvard IV psychological dictionary in a work in Finance, see Tetlock et al. (2008)

nect", "connected" and "connecting" , and thus should be treated as having the same semantic meaning. Fortunately, the family of algorithms in linguistic morphology called stemming (process used to trimm down words to their stem) is well developed and considerably accurate for the English language. In this paper we use Porter et al. (1980) stemming algorithm[2].

Table 1 contains the most frequent (stemmed) words in our sample in each of the 18 semantic groups. Inevitably, some words can have a misplaced semantic classification, such as "demand" in the group "Power", or "capit" in the group "Virtue". However, they are exceptions and will only contribute to the noise in our estimation of the content factor.

	Word Ranking (in frequency)		
Group	**1**	**2**	**3**
Positive	good (41.2 bps)	respect (30.6 bps)	consid (23.8 bps)
Negative	decreas (31.9 bps)	declin (20.4 bps)	lower (7.2 bps)
Pleasure	eas (3.9 bps)	optimist (0.4 bps)	resolv (0.2 bps)
Pain	shock (5.7 bps)	stress (1.6 bps)	tension (0.7 bps)
Feel	vigil (0.6 bps)	option (0.4 bps)	rigid (0.1 bps)
Arousal	determin (3.1 bps)	avers (2.0 bps)	anticip (1.9 bps)
Virtue	good (41.2 bps)	capit (19.0 bps)	real (14.2 bps)
Hostile	sever (4.5 bps)	exclus (3.4 bps)	exclud (3.0 bps)
Fail	delinqu (5.4 bps)	default (2.7 bps)	lag (2.6 bps)
Strong	increas (134.9 bps)	industri (45.1 bps)	reach (41.9 bps)
Weak	decreas (31.9 bps)	averag (27.1 bps)	declin (20.4 bps)
Power	mai (29.8 bps)	demand (20.4 bps)	employ (13.6 bps)
Active	increas (134.9 bps)	reach (41.9 bps)	oper (32.1 bps)
Passive	growth (36.4 bps)	expect (32.2 bps)	decreas (31.9 bps)
Work	adjust (38.9 bps)	oper (32.1 bps)	contribut (16.2 bps)
Try	avail (4.2 bps)	seek (0.2 bps)	redeem (0.1 bps)
Persist	continu (23.1 bps)	maintain (7.1 bps)	persist (6.1 bps)
Complete	reach (41.9 bps)	recoveri (12.8 bps)	establish (5.4 bps)

Table 1: **Most frequent words of each semantic group**

This table contains the most frequent (stemmed) words of each semantic group in all Central Bank minutes. Values in brackets represent relative frequencies, measured in basis points.

[2]The open code of Porter et al. (1980) algorithm can be found in several different computer languages in the website http://tartarus.org/martin/PorterStemmer

We now describe the automated algorithm. The first step is to group the documents (COPOM minutes) in a collection $\{d_1, \ldots d_T\}$ called *corpus*. For each $t = 1, \ldots, T$ let $\mathbf{v}_t$ denote 18 x 1 group-frequency vector of the document d_t, that is, the vector whose ith coordinate is the relative frequency of the stemmed words belonging to the semantic group i in the document d_t. Next we use the vectors $\{\mathbf{v}_1, \ldots \mathbf{v}_T\}$ to build the T x 18 matrix $\mathbf{X}$ composed by 18 group-frequency time-series and define $\mathbf{Z}$ as the matrix with the same series, but demeaned and normalized. The idea of using the matrix $\mathbf{Z}$ instead of the matrix $\mathbf{X}$ is that the reactions in the market are caused by *unexpected* changes in the communication. If we want to find which series had the most unexpected movement in a specific time period, we should compare their movements in terms of standard deviations.

The resulting data set, contained in the matrix $\mathbf{Z}$, has a large number of series (18), and a reduction in the number of series is desirable. An appropriate tool to deal with large number of series by reducing the dimension of a given sample is the mathematical procedure known as Principal Components Analysis (PCA).

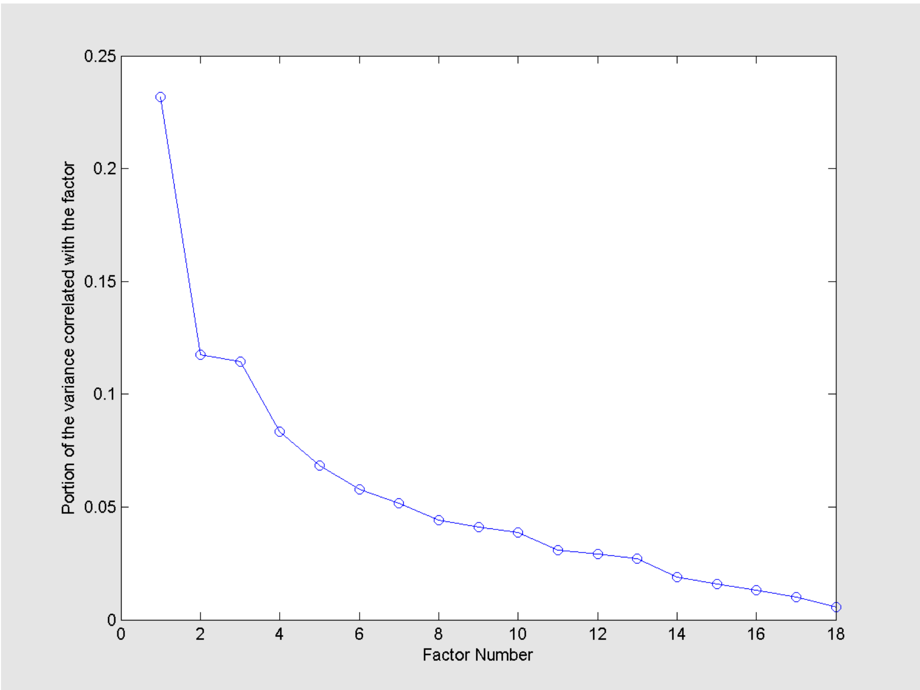

Figure 1: **Portion of the variance correlated with each factor**
This figure graphs, in descending order of magnitude, the portion of the variance correlated with each factor in the PCA based on the matrix of the 18 semantic indexes extracted from the Central Bank minutes.

PCA is defined as a orthogonal transformation of the dataset into a new coordinate system such that the projections of the time-series are ordered by their variances. It is widely used to summarize large amounts of information. For more information about PCA, see Jolliffe et al. (2002).

Let $\mathbf{Y}$ denote the PCA projection of the data set $\mathbf{Z}$. As our objective is to reduce the dimension of the matrix $\mathbf{Z}$, we need a criterion to choose the number of series with higher variances in $\mathbf{Y}$. Figure 1, which graphs the portion of the variance correlated with each factor in the projection $\mathbf{Y}$ in a descending order, give us a hint: the first component has, by far, the highest variance, while the other factors variances fall very slowly, suggesting the use of a single We use this single factor as the measure of COPOM minutes' content.

2.3 Interpreting the factor

Table 2 contains the weights by which each standardized original semantic-frequency time-series should be multiplied to get the first factor. The groups with higher weights are "Power" (with a negative signal) and "Positive" (with a positive signal). The weights on the other revelant groups show that "Virtue", "Pleasur", "Arousal" and "Feel" have positive signs, while "Pain", "Fail" and "Hostile" have negative signs. Overall, the signs of the resulting weigths indicate a consistent semantic pattern. For instance, a higher weight on words with positive, virtue, pleasure, arousal semantic content is given but lower weight on words with negative, pain, fail semantic content. Based on that, we label it the "optimism factor" (OF). However, we should emphasize that this factor is a result of a PCA analysis, that is, is the single factor that explains most of the variance. The fact that it may have some semantic interpretation is (in principle) a coincidence.

Semantic Group	Weight
Negative	-0.357
Positive	0.352
Strong	0.307
Complet	0.306
Virtue	0.297
Work	0.280
Weak	-0.267
Pleasur	0.251
Persist	0.249
Arousal	0.223
Feel	0.212
Pain	-0.210
Passive	-0.138
Power	-0.112
Try	-0.096
Active	0.088
Fail	-0.074
Hostile	-0.065

Table 2: **First Factor Loadings**

This table contains the weights by which each standardized original index should be multiplied to get the first factor. Weights are ordered by their absolute values.

Figure 2 shows some interesting features about the OF. First, the OF has a clear change of level between 2005 and 2007. The OF reaches its first peak on December of 2012, in the first period of full employment of the Brazilian economy after the stabilization of the inflation. Then, the OF reaches its two smallest values on 23/05/2001 and on 18/07/2001, both the dates of the COPOM meetings preceded by the highest log-variation of the IPCA[1] in the year of 2001, in a period of crisis in the energy supply. This is reasonable because 2001 is the first year under the inflation targeting system in which the inflation exceeded the tolerated upper band. The OF reaches its next peak in December of 2002, with the announcement of Henrique Meirelles for the management of the BCB, a name with high credibility in the Brazilian financial market. After 2007 the OF reaches its two smallest values in 01/21/2009 and in 03/07/2012, recessionary periods in Brazil. We can see that the OF reacts to shocks in the inflation and in the economic activity.

The analysis of the contemporaneous correlations of the factors with macroeconomic and financial variables in the table 3 follows the same script. As expected, OF is positively correlated with the GDP growth (+39%) and negatively correlated with the ex-

[1]IPCA is the Consumer Price Index used to define the inflation targets in Brazil.

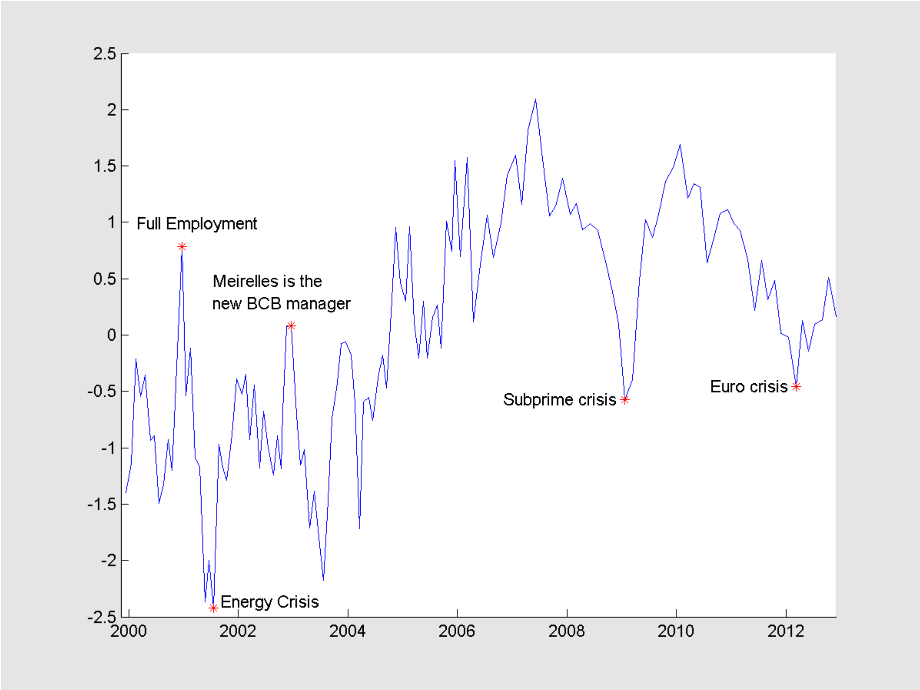

Figure 2: **Time-series of the first factor**
This figure graphs the time-series of the factor based on Central Bank minutes over the period from 2000 to 2012.

	SELIC	INFL	GDP	FX	IBOV	OF
SELIC	1.00					
INFL	0.64	1.00				
GDP	-0.38	-0.33	1.00			
FX	-0.05	-0.06	0.25	1.00		
IBOV	0.07	0.04	-0.19	-0.62	1.00	
OF	-0.64	-0.51	0.39	-0.16	0.09	1.00

Table 3: **Sample time-series correlations**

This table contains the contemporaneous correlations between the optimism factor (OF) and four macroeconomic variables over the period from 2000 to 2012, with 129 observations. Macroeconomic variables are (1) the SELIC target rate; (2) expected inflation, measured by BCB inflation survey for the next twelve months; (3) GDP growth, measured by the intermeeting log-variation of the twelve months accumulated real GDP; (4) foreign exchange rate, measured by the intermeeting log-variation of the USD/BRL exchange rate; and (5) stocks returns, measured by the intermeeting log-variation Ibovespa stocks index.

pected inflation (-51%), in line with the view that the Central Bank uses the minutes to communicate their views about the economic environment, with more optimist views in situations with high rates of GDP growth and in situations with small rates of inflation. The correlations with the financial variables also corroborate the interpretation of the OF as a measure of optimism: a negative correlation with the USD/BRL exchange rate log-variation (-16%) means that the minutes are less optimist when the national currency loses value when compared to the USD, a situation that generates inflationary pressures. The positive correlation with the Ibovespa stocks Index (+9%) means that drop-downs in the stocks market walk together with more pessimistic evaluations of the economy by the central bank in the COPOM minutes.

3 OF and future interest rates

In this section we test if the communication of the BCB, measured by the optimism factor (OF), affects the pricing of the future interest rates contracts traded in the BM&F Bovespa Stock Exchange.

We use a database of time-series of future rates disclosed by the BM&F Bovespa. We use six different maturities (one month, three months, six months, one year and two years) to measure their sensitivity to shocks in the BCB communication. Table 4 contains basic descriptive statistics.

Data from 03/01/2005 to 28/12/2012, 1,967 observations					
	1 month	3 months	6 months	1 year	2 years
Mean	12.29	12.25	12.24	12.34	12.63
Median	11.62	11.55	11.56	11.88	12.22
Stdev	3.37	3.34	3.24	3.03	2.62
Min	6.97	7.03	6.92	6.87	7.31
Max	19.82	19.88	19.74	19.42	18.78

Table 4: **Descriptive statistics of the future interest rates**

This table contains descriptive statistics for the sample of future rates with constant maturities. The data is from BM&F Bovespa.

Let $\{1, 2, \ldots, T\}$ denote the set of COPOM meetings. Our variable of interest is the variation of the future interest rate with m months to maturity during the time window around the disclosure of the minutes of the t meeting (closing price on the day that the minute is issued against the closing price one day before), denoted by Δy_t^m. We measure the shock in the communication from meeting t using the variation of the optimism factor relatively to the minutes issued in the $t-1$ COPOM meeting, that is, $\Delta\mathrm{OF} = \mathrm{OF}\ _t - \mathrm{OF}_{t-1}$.

An important fact regarding the monetary policy information flow is that the minutes are issued eight days after the disclosure of the SELIC interest rates. Therefore, when measuring the sensitivity of the future rates to shocks in the communication we do not need to control for the monetary policy decision, as this event is already incorporated in the future rates .

Let $\mathbf{m} = (m_1, m_2, \ldots, m_n)^\top$ denote the vector of (constant) maturities of the sample of future rates. We carry out the empirical analysis by running a linear regression model for $y_t^{m_i}$ against the optimism factor, with sensibility β^i. Thus the system of regressions that describes the evolution of the future rates can be written as

$$\Delta y_t^{m_i} = c_i + \beta_i \Delta\mathrm{OF}_t,\ i \in \{1 \ldots n\},\ t \in \{1 \ldots T\}. \tag{3.1}$$

Table 5 contains the regression results for maturities $\mathbf{m}$ = (1-month, 3-months, 6-months, 12-months, 24-months)$^\top$.

The estimates of table 5 show that a positive shock of one standard deviation of the optimism factor has the following monotonic effects: a decrease of one basis point on the three month future rate, a decrease of three basis point on the six month future rate, and a decrease of five basis point on the one and two years future rate. These results enable us to formulate some interesting interpretations regarding the communication of the Brazilian Central Bank.

Coefficient	Estimate	Std. Error	
Dependent Variable: Δ1-month yield			
c	0.09	[0.16]	
ΔOF	0.02	[0.38]	
R^2	0.00		
Dependent Variable: Δ3-months yield			
c	-0.20	[0.41]	
ΔOF	-0.75	[0.82]	
R^2	0.01		
Dependent Variable: Δ6-months yield			
c	-0.03	[0.76]	
ΔOF	-2.75	[1.64]	*
R^2	0.03		
Dependent Variable: Δ1-year yield			
c	-0.39	[1.27]	
ΔOF	-4.80	[2.14]	**
R^2	0.04		
Dependent Variable: Δ2-years yield			
c	0.55	[1.50]	
ΔOF	-4.81	[2.18]	**
R^2	0.03		

Table 5: **Regression Results for Future Interest Rates**

This table contains the estimates of the coefficients of the regression of the future interest rates changes (measured in basis points) against the changes in the optimism factor (OF) over the period from 2005 to 2012, with 68 observations. Changes in the future interest rate are the closing price on the day of disclosure of the minutes less the closing price one day before. Changes of the optimism factor are calculated as the difference of the OF between the contemporaneous minutes and the last minutes. Statistical significance is indicated by *, **, *** which correspond to significance levels of 10%, 5%, and 1%.

First, increases in the interest rates are associated with pessimist minutes. This is coherent with a Central Bank signalling concern with inflation in its communication, driving the market to raise their forecasts about the future values of the monetary policy interest rates (SELIC). Note that this view is also coherent with results from Table 3, which shows that the macroeconomic variable with the highest correlation with the optimism factor is the expected inflation.

Another relevant result, in line with previous studies about impact of the monetary authorities communication (see Lucca et al. (2009) for the American case), is that contracts with higher maturities are more sensitive to shocks in the communication of the Central Bank. This is coherent with the view that the minutes contain information about future policy rate and that market incorporate this information when pricing longer contacts.

The general picture that can be drawn from results of Table 5, backing the original motivation of this paper, is that the Brazilian Central bank has *some* degree of influence in the formation of the market expectations (and in the future interest rates) through communication. We are not able, however, to say if this influence is strong or marginal. A precise answer of this question would require (i) the decomposition of the shocks in the future interest rates in a framework of a model with both latent and macroeconomic variables (such as in Moralesa et al. (2008)), with the inclusion of a communication factor, or (ii) a comparative analysis of our findings with other inflation targeting regime. We leave this analysis for future work.

4 Robustness Analysis

In this Section we perform a robustness analysis of our the main results. We first analyze the behaviour of a fake factor based on a randomization of the Harvard IV dictionary. We expect that the fake factor should not affect the term structure of interest rates.

Then, in a second robustness check, we estimate regressions with the changes in the future interest rates one day *before* the disclose of the minute as dependent variables. We expect the OF to be not significant in this regression.

The random dictionary is constructed by a random swap in the meaning of all the words covered by the dictionary. For example, in our randomization the meaning of the word "abolish" is substituted by the meaning of the word "wild"; the meaning of the word "accomplish" is substituted by the meaning of the word "equate", and so on.

Table 6 contains the loadings of the factor constructed with the random dictionary (using exactly the same methodology used in the building of the optimism factor). Evidently, the optimism interpretation is no longer valid for this factor. In fact, the loading doesn't seen to have a clear interpretation as the optimism factor because semantic groups with opposite meanings have the same signal.

Semantic Group	**Weight**
Hostile	0.407
Strong	0.352
Try	-0.337
Negatie	0.313
Work	0.272
Arousal	0.266
Complet	0.264
Fail	0.251
Pain	0.194
Feel	0.189
Weak	-0.189
Passive	0.178
Persist	-0.165
Positive	0.151
Virtue	0.130
Active	-0.088
Pleasur	0.071
Power	0.031

Table 6: **First Factor Loadings (Random Dictionary)**

This table contains the weights by which each standardized original index should be multiplied to get the factors (based on the random dictionary) values. Weights are ordered by their absolute values.

Coefficient	Estimate	Std. Error
Dependent Variable: Δ1-month yield		
c	0.34	[0.59]
ΔFakeFactor	0.17	[0.35]
R^2	0.00	
Dependent Variable: Δ3-months yield		
c	-0.82	[1.15]
ΔFakeFactor	-0.43	[0.67]
R^2	0.01	
Dependent Variable: Δ6-months yield		
c	-1.37	[1.64]
ΔFakeFactor	-0.93	[1.10]
R^2	0.01	
Dependent Variable: Δ1-year yield		
c	-2.45	[2.01]
ΔFakeFactor	-1.43	[1.56]
R^2	0.01	
Dependent Variable: Δ2-years yield		
c	-1.41	[2.26]
ΔFakeFactor	-1.36	[1.66]
R^2	0.00	

Table 7: **Regression Results for the Factor Based on the Random Dictionary**

This table contains the estimates of the coefficients of the regression of the future interest rates changes (measured in basis points) against the changes in the optimism factor based on the random dictionary over the period from 2005 to 2012, with 68 observations. Changes in the future interest rate are the closing price on the day of disclosure of the minutes less the closing price one day before. Changes of the optimism factor are calculated as the difference of the OF (based on the random dictionary) between the contemporaneous minutes and the last minutes. Statistical significance is indicated by *, **, *** which correspond to significance levels of 10%, 5%, and 1%.

Table 7 contains the estimates of the regression of the interest rates against the factor based on the random dictionary. As expected, the coefficients are statistically insignificant at the level 10%. This shows that our factor is not related to a spurious dynamic on the number of words, but rather to the meaning attached to them.

In our second robustness analysis, we estimate a similar version of the original regression but instead of using contemporaneous shocks in the interest rates, we use the shocks in the interest rates one day before the disclosure of the COPOM minutes.

Table 8 contains the regression results. None of the coefficients of the optimism factor are statistical significant at the level 10%.

Coefficient	Estimate	Std. Error
Dependent Variable: Δ1-month yield		
c	-0.53	[0.25] **
ΔOF	-0.34	[0.29]
R^2	0.01	
Dependent Variable: Δ3-months yield		
c	-0.85	[0.48] *
ΔOF	-0.76	[0.46]
R^2	0.01	
Dependent Variable: Δ6-months yield		
c	-1.48	[0.60] **
ΔOF	-1.11	[0.87]
R^2	0.01	
Dependent Variable: Δ1-year yield		
c	-2.17	[0.99] **
ΔOF	-1.16	[1.13]
R^2	0.00	
Dependent Variable: Δ2-years yield		
c	-2.67	[1.19] **
ΔOF	-1.80	[1.78]
R^2	0.01	

Table 8: **Regression Results for Past Interest Rates**

This table contains the estimates of the coefficients of the regression of the lagged (past) future interest rates changes (measured in basis points) against the changes in the optimism factor (OF) over the period from 2005 to 2012, with 68 observations. Changes in the future interest rate are the closing price one day before the disclosure of the minutes less the closing price two days before. Changes of the optimism factor are calculated as the difference of the OF between the contemporaneous minutes and the last minutes. Statistical significance is indicated by *, **, *** which correspond to significance levels of 10%, 5%, and 1%.

5 Optimism and Volatility

In this section we analyse how the information contents of the BCB minutes affects the conditional volatility of the interest rates market. Janot et al. (2012) shows that the volatility of the Brazilian future rates are lower immediately after the disclose of the BCB minutes. We expect that the meaning of such minutes, represented by the optimism factor, has some degree of influence in the volatility dynamics after the release of the minutes.

Let $Vol_{5,t}^m = \sqrt{5^{-1}\sum_{i=0}^{4}\Delta y_{t+i}^2}$ denote the five days ahead volatility of the future rates with maturity m in the date t, where y_t is the m-months future rate measured in basis points. We use $Vol_{5,t}^m$ as a proxy of the conditional volatility of the maturity m. We study the influence of the BCB communication in the volatility of the yield market by estimating the following four regressions:

$$Vol_{5,t}^m = \beta_0 + \beta_1 \cdot \text{minute}_t + \varepsilon_t \tag{5.2}$$

$$Vol_{5,t}^m = \beta_0 + \beta_1 \cdot \text{minute}_t + \beta_2 \cdot \text{pessimism}_t + \varepsilon_t \tag{5.3}$$

$$log(Vol_{5,t}^m) = \beta_0 + \beta_1 \cdot \text{OF}_t + \varepsilon_t \tag{5.4}$$

$$log(Vol_{5,t}^m) = \beta_0 + \beta_1 \cdot \text{OF}_t + \beta_2 \cdot \text{pessimism}_t \cdot \text{OF}_t + \varepsilon_t \tag{5.5}$$

where minute_t is a dummy variable for the day of disclose of the minutes, pessimism_t is a dummy with value 1 if t is a day of issue of minutes and if the optimism factor in t is less then the optimism factor of the last minutes and OF_t is the value of the optimism factor if there are minutes in t and 0 otherwise.

Regression 5.2 studies the volatility of the future rates conditionally to the existence of minutes in a given date, ignoring the meaning of the minutes. We expect a negative value for the angular coefficient β_1. Such regression is generalized in the regression 5.3 by the inclusion of a dummy for the pessimism of the minutes. We expect a positive coefficient for this dummy because negative evaluations of the economic outlook are usually associated with situations of high volatility.

Regression 5.4 studies the relation between the volatility and the magnitude of the shocks in the communication. We use the logarithm of the volatility to avoid the problem of negative forecasts for $Vol^{m}_{5,t}$. Regression 5.5 look for an eventual asymmetric effect of the optimism factor in the volatility. If $\beta_2 \neq 0$ we can conclude that pessimist and optimist minutes affects the volatility with different strengths.

Table 9 contains the estimates for the volatility regressions of the maturities 3-, 6-months and 1-, 2-years.

Results of the regression of the volatility against the dummy for minutes are in line with the findings of Janot et al. (2012): future rates volatility are lower after the discolse of the BCB minutes. This is coherent with the view that the minutes are able to reduce uncertainties about the interest rates future path. However, this effect is not statistically significant at the level 10% for longer (higher then 6-months) yields. In other words, if we don't take into account the meaning of the minutes, effects on volatility are significant only in the short rates.

The second regression, including a dummy for pessimist minutes, tell us something more. First, more pessimist minutes are associated with higher volatilities, and this effect is significant at the level 10% for the 6-months maturity. Second, the estimates of the minute dummy coefficient are always significant when we take the pessimism minute dummy into account. Thus, if we ignore the meaning of the words when analysing the conditional volatility of the interest rates, we can incorrectly conclude that the communication affects only the short rates.

The magnitude of the impact of the communication in the volatility can be found in the estimates of the third regression: a negative shock of one standard deviation of the optimism factor (a pessimist minute) if followed by an increase of 30%, 18%, 10% and 8% of the 3-, 6-months, 1-, 2-years future rates, respectively. Thus, the size of the shocks in the optimism factor affects the short rates volatility more intensively. The last regression contains insignificant estimates for the cross term, an evidence that such communication effect is symmetric in the volatility.

Regression	(1)		(2)		(3)		(4)	
Dep. Variable	$Vol_{5,t}$		$Vol_{5,t}$		$log(Vol_{5,t})$		$log(Vol_{5,t})$	
Maturity: 3-month								
constant	3.12		3.12		0.87		0.87	
	[0.15]	***	[0.15]	***	[0.04]	***	[0.04]	***
minute	-0.84		-1.00					
	[0.21]	***	[0.26]	***				
pessimist			0.33					
			[0.36]					
OF					-0.30		-0.31	
					[0.07]	***	[0.10]	***
pessimist · OF							0.03	
							[0.14]	
R^2	0.0030		0.0031		0.0050		0.0051	
Maturity: 6-months								
constant	4.51		4.51		1.27		1.27	
	[0.21]	***	[0.21]	***	[0.04]	***	[0.04]	***
minute	-0.54		-1.13					
	[0.33]		[0.37]	***				
pessimism			1.21					
			[0.60]	**				
OF					-0.18		-0.20	
					[0.06]	***	[0.07]	***
pessimism · OF							0.09	
							[0.14]	
R^2	0.0006		0.0015		0.0020		0.0021	
Maturity: 1-year								
constant	6.76		6.76		1.71		1.71	
	[0.32]	***	[0.32]	***	[0.03]	***	[0.03]	***
minute	-0.39		-1.02					
	[0.48]		[0.53]	*				
pessimism			1.31					
			[0.87]					
OF					-0.10		-0.11	
					[0.06]	*	[0.06]	*
pessimism · OF							0.03	
							[0.16]	
R^2	0.0001		0.0006		0.0008		0.0008	
Maturity: 2-years								
constant	9.25		9.25		2.04		2.04	
	[0.43]	***	[0.43]	***	[0.03]	***	[0.03]	***
minute	-0.70		-1.41					
	[0.62]		[0.63]	**				
pessimism			1.46					
			[1.13]					
OF					-0.08		-0.09	
					[0.06]		[0.06]	
pessimism · OF							0.03	
							[0.15]	
R^2	0.0003		0.0006		0.0006		0.0006	

Table 9: **Regression Results for the Future Rates Volatility**

This table contains the estimates of the coefficients of the future interest rates volatility in a time window with five network days against some selected variables. The dependent variable is defined by $Vol_{5,t} = \sqrt{5^{-1}\sum_{i=0}^{4}\Delta y_{t+i}^2}$, where y_t is the future interest rate measured in basis points; "minute" is a dummy with value 1 when the minute is issued and 0 otherwise; "pessimism" is a dummy with value 1 when the variation of the optimism factor is negative in comparison with optimism factor of the last minutes and 0 otherwise. OF is the optimism factor when the minute is issued and 0 otherwise. We use data over the period from 2005 to 2012, with 2,001 observations. The dummy "minute" and the variable OF have nonzero values in 68 observations, and the dummy "pessimism" have nonzero values in 33 observations. Statistical significance is indicated by *, **, *** which correspond to significance levels of 10%, 5%, and 1%.

6 Conclusion

In this paper study the ability of the Brazilian Central Bank (BCB) to influence the interest rates market through communication. For this purpose, we analyze the COPOM minutes (the Brazilian equivalent to the FOMC minutes) issued between 2000 and 2012 and propose a new method to measure the information content of documents. Our method relies on the classification of words according to theirs semantic meaning and in the statistics of large data sets. The advantage of our methodology is that it is direct, without subjectivity in the interpretation of contents, and operationally simple.

We find evidence that the information of the COPOM minutes can be summarized in a single factor. The analysis of factor loadings, as well as the correlations with economic and financial time-series, indicates that the factor is primarily related to the optimism of the COPOM minutes about the economic outlook. We call our factor the optimism factor (OF).

We show that the variation of the future rates in a time window around the COPOM meeting respond to variations in the OF. Our results indicates that longer contracts reacts more intensely to changes in the communication, in line with the view that the minutes contain information about the future monetary policy actions and that the market incorporate this information when pricing longer contracts.

We also give evidence that the minutes are able to reduce the uncertainness about the future path of the short-run interest, diminishing, in the average, the volatility of the interest rates market. The reduction of the volatility, however, is less intensive when minutes are pessimist.

Our result is in line with the idea that Central Bank communication can effectively change market expectations, as measured by the impact on longer maturity interest rates.

References

Martin Porter (1980). An algorithm for suffix stripping. *Program 40*(3), 211–218.

I T Jolliffe (2002). *Principal Component Analysis* (2 ed.). Springer Series in Statistics.

Gauti Eggertsson and Michael Woodford (2003, May). Zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity 1*(1).

Ben Bernanke, Vincent Reinhart and Brian Sack (2004). *Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment* (2 ed.). Brookings Papers on Economic Activity.

Ellyn Boukus and Joshua Rosenberg (2006, June). The information content of fomc minutes. *Working Papers Series* (265).

Paul C. Tetlock and Maytal Saar-Tsechansky and Sofus Macskassy (2008, 06). More than words: Quantifying language to measure firms' fundamentals. *Journal of Finance 63*(3), 1437–1467.

Marco Moralesa (2008). The real yield curve and macroeconomic factors in the chilean economy. *Applied Economics 42*(27), 3533–3545.

David O. Lucca and Francesco Trebbi (2009, September). Measuring central bank communication: An automated approach with application to fomc statements. *NBER Working Papers* (15367).

Carlo Rosa (2011). Talking less and moving the market more: Evidence from the ecb and the fed. *Economic Notes 58*(1), 51–81.

Michael Lamlaa and Sarah Leinb (2011). What matters when? the impact of ecb communication on financial market expectations. *Applied Economics 43*(28), 4289–4309.

Marcio Janot and Daniel El-Jaick de Souza Mota (2012, January). O impacto da comunicacao do banco central do brasil sobre o mercado financeiro. *Working Papers Series* (265).