



**Thiago de Andrade Machado**

**A High-Frequency Analysis of the Effects of Central Bank  
Communication on the Term-Structure of Interest Rates in  
Brazil**

**DISSERTAÇÃO DE MESTRADO**

Thesis presented to the Programa de Pós-Graduação em Economia, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Economia.

Advisor: Prof. Carlos Viana de Carvalho

Rio de Janeiro

April 2014



**Thiago de Andrade Machado**

**A High-Frequency Analysis of the Effects of Central Bank  
Communication on the Term-Structure of Interest Rates in  
Brazil**

Thesis presented to the Programa de Pós-Graduação em  
Economia, PUC-Rio as partial fulfillment of the  
requirements for the degree of Mestre em Economia.  
Approved by the following commission.

**Prof. Carlos Viana de Carvalho**

Advisor

Departamento de Economia, PUC-Rio

**Prof. Marcelo Cunha Medeiros**

Departamento de Economia, PUC-Rio

**Prof. Fernanda Feitosa Nechio**

Federal Reserve Bank of San Francisco

**Prof. Monica Herz**

Coordinator of the Centro de Ciências Sociais of PUC-Rio

Rio de Janeiro, April 11<sup>th</sup>, 2014

All rights reserved.

**Thiago de Andrade Machado**

B.A., Economics, PUC-Rio, 2006-2010 (Rio de Janeiro, Brazil). MsC. Candidate, Economics, PUC-Rio, 2012-2014.

Bibliographic data

Machado, Thiago de Andrade

A High-Frequency Analysis of the Effects of Central Bank Communication on the Term Structure of Interest Rates in Brazil / Thiago de Andrade Machado; advisor: Carlos Viana de Carvalho – 2014.

66 f. : il. (color.) ; 30 cm

Dissertação (Mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Economia, 2014.

Inclui Bibliografia

1. Economia – Teses. 2. Taxa de Juros. 3. Comunicado. 4. Banco Central. 5. Dados Intradiários. I. Carvalho, Carlos Viana de. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Economia. III. Título.

CDD: 330

## Acknowledgments

I would like to first thank my adviser Carlos Viana de Carvalho for all support and guidance during the whole process of the dissertation. I also want to thank Marcelo Medeiros and Fernanda Nechio for all suggestions and comments. I thank my friends for the uncountable study group hours and conversations about several issues. And last, but not least, I am very grateful to all support I received from my family, in particular to my parents, Camilo and Maria Cecília, my grandparents, Edivaldo, Haydée and Sílvia, and my girlfriend Cristiane.

## Abstract

Machado, Thiago de Andrade; Carvalho, Carlos Viana de (Advisor). **A High-Frequency Analysis of the Effects of Central Bank Communication on the Term-Structure of Interest Rates in Brazil.** Rio de Janeiro, 2014. 66p. MSc. Dissertation - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

This work builds semantic scores using the Google and the Factiva Dow Jones database, based on Lucca and Trebbi's (2011) methodology, in order to quantify the content of the COPOM's statements released by the Central Bank of Brazil shortly after the interest rate's decision and attributing to it a semantic orientation, "hawkish" or "dovish". Using daily and intraday data of swap contracts and DI1 futures contracts, respectively, we find that the content of the BCB's statement affects the yield curve only in the period prior to Tombini's tenure. In addition, we find that the yields respond one-to-one to the interest rate surprise, sometimes more, in the pre-Tombini period even for long term maturities, which we do not see in the period prior to Tombini, where the interest rate surprises affect only the short-to-medium rates. Furthermore, we see an intraday dynamic in the yield responses to the content of the statement in the Tombini period, which give evidence to a delay in its interpretation, differently from what we observe in the previous period. We also find that the interest rate surprises induce changes in the yield curve during the whole time that the market is open for both periods analyzed.

## Keywords

Interest Rates; Statement; Central Bank; Intraday data

## Resumo

Machado, Thiago de Andrade; Carvalho, Carlos Viana de. **Uma Análise em Alta Frequência dos Efeitos da Comunicação do Banco Central na Estrutura a Termo da Taxa de Juros no Brasil**. Rio de Janeiro, 2014. 66p. Dissertação de Mestrado - Departamento de Economia, Pontifícia Universidade Católica do Rio de Janeiro.

Este trabalho constrói índices de semântica, utilizando o Google e o Factiva Dow Jones, empregando a metodologia de Lucca e Trebbi (2011), na tentativa de quantificar o conteúdo do comunicado do Copom que o Banco Central do Brasil emite logo após a divulgação da decisão da taxa de juros e o designando uma orientação semântica restritiva ou expansionista. Utilizando dados diários e intradiários de contratos de swap e contratos futuros de DI, respectivamente, vemos que o comunicado afeta a curva de juros somente no período pré-Tombini. Ademais, vemos que a surpresa de política monetária tem um efeito de um para um, algumas vezes maior, nas taxas de juros para o período Tombini até mesmo para taxas longas, o que não ocorre no período pré-Tombini, com a surpresa afetando apenas as taxas entre o curto e médio prazo. Além disso, percebemos que há uma dinâmica intradiária no efeito do conteúdo do comunicado na curva de juros no período Tombini, o que evidencia uma demora na sua assimilação por parte dos agentes econômicos, diferentemente do que é observado no período anterior. Vemos que a surpresa de política monetária afeta a estrutura a termo da taxa de juros durante todo o tempo de funcionamento do mercado para os dois períodos analisados.

## Palavras-chave

Taxa de Juros; Comunicado; Banco Central; Dados Intradiários

# Summary

1 Introduction	11
1.1. Motivation	11
1.2. Literature	14
2 COPOM Statements Measurement	16
2.1. Factiva Semantic Score	17
2.2. Google Semantic Score	19
3 Data and Results	22
3.1. Daily Data	23
3.2. Intraday data	28
4 Conclusion	41
5 References	43
6 Tables and Figures	45
6.1. Daily Tables	45
6.2. Intraday Figures	50
6.2.1. $\Delta FSO\_GSO$	50
6.2.2. GSO Intraday Results (Alternative Interest Rate Surprise)	55
6.2.3. Intraday Dynamics Specification (without the residuals)	65

## List of Figures

Figure 1 FSO and the SELIC rate.....	18
Figure 2 - Discrepancies in BCB's series of swaps .....	24
Figure 3 – NSS fit (Yield Variation) .....	31
Figure 4 – NSS fit (Term Structure) .....	32
Figure 5 - Yield responses to the language for pre-Tombini period .....	34
Figure 6 - Yield responses to the language for Tombini period.....	34
Figure 7 - Yield responses to the interest rate surprise for pre-Tombini period ...	35
Figure 8 - Yield responses to the interest rate surprise for Tombini period.....	35
Figure 9 - Intraday dynamics of the language for pre-Tombini period .....	37
Figure 10 - Intraday dynamics of the language for Tombini period .....	38
Figure 11 - Intraday dynamics of the interest rate surprise for pre-Tombini period .....	38
Figure 12 - Intraday dynamics of the interest rate surprise for Tombini period ...	39
Figure 13 - Intraday dynamics of the residuals for pre-Tombini period .....	39
Figure 14 - Intraday dynamics of the residuals for Tombini period.....	40
Figure 15 - Yield responses to the language for pre-Tombini period .....	50
Figure 16 - Yield responses to the language for Tombini period.....	50
Figure 17 - Yield responses to the interest rate surprise for pre-Tombini period .	51
Figure 18 - Yield responses to the interest rate surprise for Tombini period.....	51
Figure 19 - Intraday dynamics of the language for pre-Tombini period .....	52
Figure 20 - Intraday dynamics of the language for Tombini period .....	52
Figure 21 - Intraday dynamics of the interest rate surprise for pre-Tombini period .....	53
Figure 22 - Intraday dynamics of the interest rate surprise for Tombini period ...	53
Figure 23 - Intraday dynamics of residuals for pre-Tombini period .....	54
Figure 24 - Intraday dynamics of the residuals for Tombini period.....	54
Figure 25 - Yield responses to the language for pre-Tombini period .....	55
Figure 26 - Yield responses to the language for Tombini period.....	55
Figure 27 - Yield responses to the interest rate surprise for pre-Tombini period .	56
Figure 28 - Yield responses to the interest rate surprise for Tombini period.....	56

Figure 29 - Intraday dynamics of the language for pre-Tombini period .....	57
Figure 30 - Intraday dynamics of the language for Tombini period .....	57
Figure 31 - Intraday dynamics of the interest rate surprise for pre-Tombini period .....	58
Figure 32 - Intraday dynamics of the interest rate surprise for Tombini period ...	58
Figure 33 - Intraday dynamics of the residuals for pre-Tombini period .....	59
Figure 34 - Intraday dynamics of the residuals for Tombini period.....	59
Figure 35 - Yield responses to the language for pre-Tombini period .....	60
Figure 36 - Yield responses to the language for Tombini period.....	60
Figure 37 - Yield responses to the interest rate surprise for pre-Tombini period .	61
Figure 38 - Yield responses to the interest rate surprise for Tombini period.....	61
Figure 39 - Intraday dynamics of the language for pre-Tombini period .....	62
Figure 40 - Intraday dynamics of the language for Tombini period .....	62
Figure 41 - Intraday dynamics of the interest rate surprise for pre-Tombini period.....	63
Figure 42 - Intraday dynamics of the interest rate surprise for Tombini period ...	63
Figure 43 - Intraday dynamics of the residuals for pre-Tombini period .....	64
Figure 44 - Intraday dynamics of the residuals for Tombini period.....	64
Figure 45 - Intraday dynamics of the language for pre-Tombini period .....	65
Figure 46 - Intraday dynamics of the language for Tombini period .....	65
Figure 47 - Intraday dynamics of the interest rate surprise for pre-Tombini period.....	66
Figure 48 - Intraday dynamics of the interest rate surprise for Tombini period ...	66

## List of Tables

Table 1 – Strings of the January 18, 2012 COPOM Statement.....	19
Table 2 - GSO results .....	25
Table 3 - Later GSO Results.....	26
Table 4 - Correlation between the Semantic Scores.....	27
Table 5 - Language Surprise (GSO) (since 2008) .....	28
Table 6 - Correlation between intraday and daily variation .....	31
Table 7 – GSO ifc (with initial and final cutting date - since 2008) .....	45
Table 8 – GSO2 (without strings - since 2008).....	46
Table 9 - FSO_GSO (since 2008) .....	46
Table 10 - FSO_GSO (since 2003) .....	47
Table 11 - FSO (since 2008).....	47
Table 12 - FSO (since 2003).....	48
Table 13 - Language Surprise (Later GSO) (since 2008).....	48
Table 14 - Language Surprise (GSO ifc) (since 2008).....	49
Table 15 - Language Surprise (GSO2) (since 2008) .....	49

# 1

## Introduction

### 1.1. Motivation

Over the past two decades, central bank communication has seen its importance increased in several countries. Inflation targeting regimes, adopted by some of them, intensified somehow this process. Even central banks that did not officially adopt this regime, such as the Federal Reserve (Fed) in the US, decided to increase their communication to the public.

Central banks seek to enhance their credibility and transparency towards the public in order to make their monetary policy actions more efficient. They do so by trying to develop ways to inform economic agents of how and when they will take action. That is the reason why communication is an important instrument for the economy as whole and why it became a relevant subject for economic research.

This work tries to quantify the statement of the Central Bank of Brazil (BCB), which is released shortly after the monetary policy committee (COPOM) meeting, after market close, and estimate its effect, along with a measure of interest rate surprise, on the term structure of the interest rates on the following business day, using daily and intraday data. We analyzed those effects for the periods prior and during Alexandre Tombini's mandate ahead of the Central Bank of Brazil. We choose to quantify the central bank statement, because it is the primary form of BCB's communication to the public after the interest rate (SELIC rate) decision.

In order to compute the content of the statement, we build the Google Semantic Score (GSO) and the Factiva Semantic Score (FSO) based on Lucca and Trebbi (2011). These scores try to specify the semantic orientation of the central bank's statements. In other words, it shows whether the statement is "hawkish" or "dovish"<sup>1</sup>. The semantic orientation testifies as to how the public interprets the

---

<sup>1</sup> "Hawkish" or "dovish" are restrictive or expansionary monetary policy, respectively.

statement and, from its understanding of future interest rate movements, takes its decisions.

We build two types of GSO. In order to adapt the semantic scores to the Portuguese language, we follow Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013). For one of the GSOs, we separate the statements into strings and then analyze the co-occurrence with “hawkish” or “dovish” expressions. To build the other type, we do not separate the statement into strings, so the score is not subordinate to individual interpretations of how to separate the strings. Both are calculated directly from the Google search queries.

There are two main disadvantages with the GSO. One is due to noises that the Google search query carries. The reason of the existence of such noises is that we do not have full control over the search results. Some of the outcomes of the Google search queries are not related specifically to the statement. The other disadvantage comes from the fact that Google deletes old webpages and for this reason, we cannot repeat the same score in different periods of time.

An alternative method to the GSO is the FSO, for which there is no need to separate the statements into strings. The FSO is created from the Dow Jones Factiva database which includes business and financial news from newspapers, magazines, blogs and Internet pages. In this case, we consider only news that contain words related to the policy announcement, such as “COPOM” and “Central Bank”, and analyze its co-occurrence with our “hawkish” and “dovish” expressions.

Since we find relevant errors in the Central Bank of Brazil’s daily database of swaps, used by Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013), we update their GSO until the end of 2013 and rerun their regression specification, using data from Bloomberg. We notice that the semantic score affects the short-to-medium yields during the pre-Tombini period and, quantitatively, we find that one standard deviation increase in the semantic score, in other words, a change in the content of the statement towards the “hawkish” direction raises 6-month to 2-year yields by between 2 and 4 basis points. Those results are statistically significant at conventional levels.

Interest-rate surprises also affect the term structure of the interest rates. During pre-Tombini period, there is a one-to-one response until the 1-year maturity, decreasing for longer maturities. On the other hand, during the Tombini

period, we see a one-to-one response, sometimes more, of interest rate surprises for all maturities.

Since the GSO carries both expected and unexpected changes on the content of the central bank statement (Carvalho, Cordeiro and Vargas (2013)), we create an expectation semantic score in order to analyze just the unexpected change on the communication. We define the language surprise as the expectation semantic score subtracted from the GSO. Our results suggest that our language surprise's measure outcome changes dramatically the pattern from period prior to Tombini and during Tombini's tenure.

Furthermore, we create six months afterwards the GSO above, another GSO with the same strings and we find different results. The main changes occur in the period prior to Tombini, which is evidence towards the old webpages removed by the Google. The correlation between both semantic score in the period prior to Tombini is far lower than the correlation in the Tombini period. This confirms the difficulty of replication of the GSO, mentioned in Carvalho, Cordeiro and Vargas (2013).

Differently from Lucca and Trebbi (2011), we find most of the yield responses not statistically significant at conventional levels for our FSOs. This might happen due to the fact that the Factiva Dow Jones database has less news related to the COPOM's decision compared to the FOMC's decision.

Moreover, we use an intraday dataset to analyze the yield responses each moment of the market's operating time. It consists of all the interbank deposit futures contracts (DI1) traded and their respective times during all following business days after the COPOM's interest rate decision. We analyze the yield responses to the content of the central bank statement and to the interest rate surprises in different periods of the day. Alike the daily data, we study the effects of the period prior and during Tombini's tenure.

Since DI1 futures contracts do not have fixed maturities, such as swap contracts, we use the Nelson-Siegel-Svensson (NSS) model to estimate intraday term structure of interest rates. In order to achieve the best fit, we set some 'liquidity' parameters to drop illiquid DI1 futures contracts. Thereafter, we build the term structure of the DI1 futures contracts and the term structure of their respective price settles to find the yield variation for each time for each following business day of the COPOM meeting.

Using the GSO, we find that the yields appear to respond to changes in the COPOM language along the day at short to medium maturities, especially between 1 and 2 years maturities. This result only applies for period prior to Tombini's tenure. During Tombini period, the COPOM statements do not seem to have an effect on the yield changes. Furthermore, the interest rate surprises affect the yield curve. Prior to Tombini's tenure, short term maturities react more than long term maturities, while during Tombini period there is a one-to-one response, sometimes more, for all maturities. All these results confirm our daily data findings. This evidences loss of the communication's importance during Tombini period.

Moreover, we analyze the intraday dynamics along the market's operating time for both periods. We separate it into three parts and analyze the yield responses in each one. We find that the statements' content seem to be fully interpreted in the first part in the pre-Tombini period, while during Tombini's tenure, the language looks like to be assimilated later on in the second and in the third parts of the market's operating time.

Furthermore, the yields respond to the interest rate surprises in the first part for both periods. On one hand, in the pre-Tombini period, the short-to-medium rates react in the second and in the third parts of the market's operating time. On the other hand, in the Tombini period, we see the yields also responding in the second and in the third parts, notably in the long term maturities.

Likewise our FSO daily results, we find most of the yield responses not statistically significant at conventional levels for our FSO intraday results.

## 1.2. Literature

Earlier works tried to quantify the content of the statement of the central bank. Lucca and Trebbi (2011) built the GSO and the FSO for the FOMC statements, released by the Fed shortly after the interest rate decision. They analyzed the effect of the semantic orientation scores over Treasury yields for different maturities, using an intraday approach. They found that longer-dated maturities react more to changes in the semantic score, while shorter-dated maturities react more to interest rate surprise.

Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013) built the GSO for the statement of the BCB, using daily data. They analyzed the effect of the GSO over the Brazilian Treasury yields for different maturities prior and during Tombini's tenure. One main difference between the FOMC and the COPOM meeting is the fact that the first occurs when the market is open while the second takes place when the market is closed. Their results showed that the content of the statement has an effect on yields between 4 months and 1 year maturities only for the period prior to Tombini's tenure.

Rosa and Verga (2008) tried to quantify the ECB statements, assigning values to specific terms of the statements. This method is subject to personal judgments about the statement content. They find that changes in market interest rates can be explained by unexpected movements in the ECB announcements. Costa Filho and Rocha (2009, 2010) use their methodology on the COPOM minutes and find evidence that the COPOM actions go in the same direction as its own words.

Gürkaynak, Sack and Swanson (2005) also tried to quantify the content of the FOMC statement, using intraday data. They built a two-factor model in which one is related to the monetary policy surprise itself and the factor is associated to changes in the FOMC statement. They noticed that the latter has a more positive effect on longer maturities than the former.

This work proceeds as follows. In section 2, we show how we build our semantic scores, using the methodology proposed by Lucca and Trebbi (2011). Section 3 presents our daily and intraday data and shows their respective results. Section 4 concludes.

## 2 COPOM Statements Measurement

The challenge to measure the content of the central bank statements is how to convert an idea into an index which shows its meaning and at the same time its intensity. The measurement of the statements was based on the psycholinguistics literature. Church and Hanks (1990) introduced a measure of association ratio, in which the meaning of a word or a string can be noticed by its co-occurrence with another word or expression. This association ratio is called PMI (Pointwise Mutual Information) and is defined as below:

$$PMI(x, y) = \log \left( \frac{Pr(x, y)}{Pr(x)Pr(y)} \right)$$

As noticed above, the more “x” and “y” are associated, the higher the PMI will be. If “x” is in the text, PMI quantifies extra-information about a possible occurrence of “y”. Statistically, it compares the probability of noticing both “x” and “y” together in the same text with the probability of observing them independently. In other words, the ratio is a measure of the degree of statistical dependence between “x” and “y” (Turney and Littman (2002)).

We use the PMI to calculate the semantic orientation of words and strings. To do so, we choose two opposite concepts, “hawkish” and “dovish”, and evaluate the PMIs within both concepts. The semantic orientation is defined as follows:

$$\begin{aligned} SO(x) &= PMI(x, hawkish) - PMI(x, dovish) \\ &= \log \left( \frac{Pr(x, hawkish)Pr(dovish)}{Pr(x, dovish)Pr(hawkish)} \right) \end{aligned}$$

For example, if a string “x” appears more often with “hawkish” expressions relatively to “dovish” expressions, then this string has a “hawkish” connotation. In other words, its “hawkish” PMI is greater than its “dovish” PMI.

Based on the PMI and the semantic orientation concepts, we build the Factiva Semantic Score and the Google Semantic Score of the BCB’s statements.

## 2.1. Factiva Semantic Score

In order to build the Factiva Semantic Score, we use the Dow Jones Factiva database, which provides business and financial news, including newspapers, journals, magazines and Internet pages. We follow Lucca and Trebbi (2011) and apply the Factiva to focus only on news that is totally related to the monetary policy committee using a 4-day window around the moment of the interest rate release. We do so through the Factiva search mechanism, in which we look for news that contain the terms “COPOM”, “central bank” or “BC” and “rate” or “SELIC”.

In addition, we follow Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013), who adapted Lucca and Trebbi (2011) to the Portuguese language, and chose the most common Portuguese expressions used by the economic agents to indicate an expansionary or a restrictive monetary policy in order to build their GSO. They use “hike”, “elevation” and “increase” as “hawkish” terms and “cut”, “reduction” and “decrease” as “dovish” terms. Since the Factiva search queries are less subject to noises than the Google results, we add one more term to each group. We include “rise” and “diminish” to the “hawkish” and “dovish” groups, respectively<sup>2</sup>. Those terms are set to appear next to the words “rate” or “SELIC”. This adaptation to the Portuguese language is necessary, because there are not specific terms entirely related to the actions of the central bank, such as “hawkish” and “dovish”, used in English. Furthermore, we include other search specifications: the negations of the opposite groups<sup>3</sup> and the variations of all verbs selected above.

To build the Factiva Semantic Score, we based it on the semantic orientation and on the PMI definitions as follows:

$$FSO_t = \log \left( \frac{\sum_h hits(x, hawkish)}{\sum_d hits(x, dovish)} \right) + \delta$$

where  $hits(x, y)$  are the results of the Factiva search query using the keywords  $x$  within concept  $y$  for statement  $t$  and  $\delta = hits(dovish)/hits(hawkish)$ . As

<sup>2</sup> The Portuguese translation for those words are “alta”, “aumento”, “elevação” e “subida” for the “hawkish” terms and “corte”, “queda”, “redução” and “diminuição” for the “dovish” terms.

<sup>3</sup> We include the negations of the opposite groups by determining that the terms “not” and “interrupt” need to be before the terms listed above.

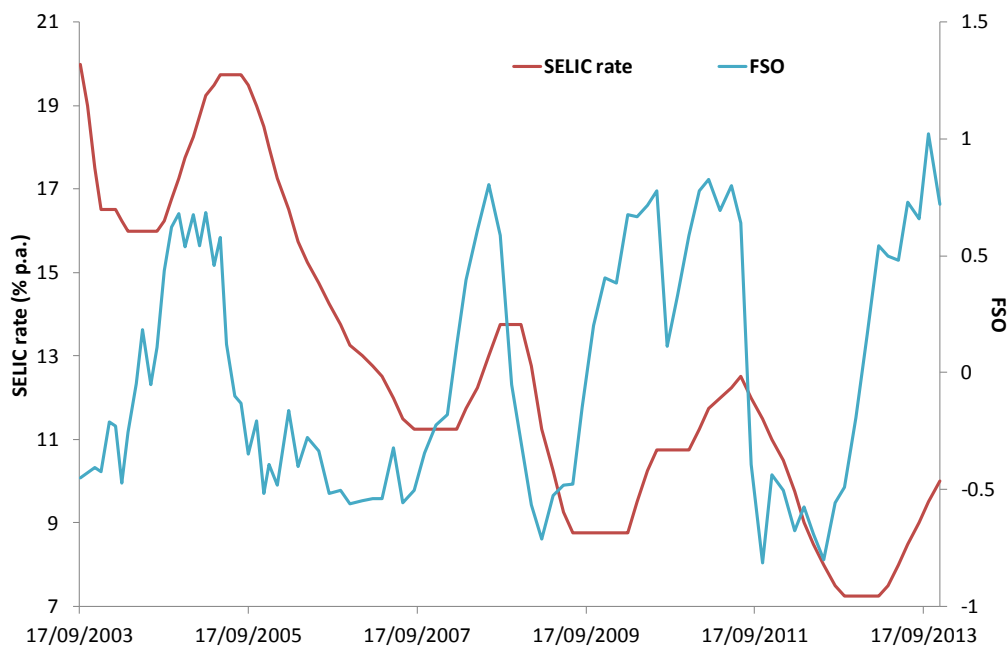
noted above, the constant  $\delta$  is inherent to any keyword, so it does not have any effect on our final results, thus becoming negligible.

Based on Lucca and Trebbi (2011), we define the set of news released before and after the SELIC rate decision in a 4-day window. Thus, we establish the unexpected change on the content of the central bank statement as indicated below:

$$\Delta FSO_t = FSO_{t+} - FSO_{t-}$$

We define the  $FSO_{t-}$  as a measure of the expectation about the central bank's statement and its decision and the  $FSO_{t+}$  as a measure of what exactly happens to the interest rate and how the economic agents interpret the statement and the decision itself.

Figure 1 shows the time series of the FSO and the SELIC rate. It indicates that movements in the FSO anticipate movements in the interest rate. Following Carvalho, Cordeiro and Vargas (2013), we run a Granger-causality test, in which we reject the hypothesis that the FSO does not grangercause the SELIC rate, while we do not reject the other grangercause direction.



**Figure 1 FSO and the SELIC rate**

We also create an alternative measure of the unexpected change in the content of the central bank statement, using the Factiva. We build an analogous FSO related to the GSO created by Lucca and Trebbi (2011), but locking our Factiva search query to the day of the COPOM's decision. We define it as the

difference between the semantic scores of the statements released at meeting  $t$  and  $t - 1$  as shown below:

$$\Delta FSO\_GSO_t = FSO_t - FSO_{t-1}$$

## 2.2. Google Semantic Score

We build the Google Semantic Score straight from the Google search queries. Since it is not possible to get access to all webpages available in the past - over time they are deleted - and therefore we are not able to properly estimate the probabilities to find the semantic orientation, we follow Lucca and Trebbi (2011) and compute the Google search queries, which work as proxy for the joint and individual probabilities of the PMI. In order to build the Google Semantic Score, we choose the “hawkish” and “dovish” terms used by Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013).

We create two types of GSOs. For one of them, based on Lucca and Trebbi (2011), the COPOM statements are separated into segments that could somehow signalize to the public future movements of the monetary policy. In order to illustrate how the separation into strings occurs, we choose the statement released on January 2012 as an example:

*Dando seguimento ao processo de ajuste das condições monetárias, o Copom decidiu, por unanimidade, reduzir a taxa Selic para 10,50% a.a., sem viés. O Copom entende que, ao tempestivamente mitigar os efeitos vindos de um ambiente global mais restritivo, um ajuste moderado no nível da taxa básica é consistente com o cenário de convergência da inflação para a meta em 2012.*

The strings built from the statement above are:

**Table 1 – Strings of the January 18, 2012 COPOM Statement**

Strings "x"	
1	Dando seguimento ao processo de ajuste das condições
2	processo de ajuste das condições monetárias
3	tempestivamente mitigar os efeitos vindos de um ambiente global mais restritivo
4	ambiente global mais restritivo
5	um ajuste moderado no nível da taxa básica é consistente com o cenário de convergência da inflação para a meta em 2012
6	um ajuste moderado no nível da taxa básica
7	cenário de convergência da inflação para a meta
8	convergência da inflação

We use the Google search mechanism to establish a range for the quest so we can take into account only the Internet pages up to the day of the interest rate release. Furthermore, for comparison, we also create a GSO with an initial and final cut of dates for the quest. The GSO for a given string  $x$  is defined as below:

$$GSO(x) = \log \left( \frac{\sum_h hits(x, hawkish)}{\sum_d hits(x, dovish)} \right) + \delta$$

where  $hits(x, y)$  and  $\delta$  are defined in the same way as shown above for the FSO. However, instead of using the results of the Factiva query, we use the results of the Google query.

In order to build the GSO of COPOM statement  $t$ , we compute the average of the GSO's string as follows:

$$GSO_t = \sum_{x \in t} \frac{GSO(x)}{\#(x \in t)}$$

Given that the older is the statement, the fewer the webpages available online, we follow Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013) and establish the following rule: if there is no news with a particular string along with any “hawkish” term, its semantic score will be the highest value measured in the entire series. On the opposite hand, if the string has no co-occurrence with any “dovish” term, we choose the lowest semantic score calculated in the entire series. If a string has no co-occurrence with any term, we drop it from our list, since there is evidence that this string was not considered relevant to the economic agents.

The other GSO type is built in a way that enables us not to be dependent on personal judgments of how the statements' division into strings is made. In other words, we do not measure the semantic orientation of each segment, but instead of the entire statement. We do so by analyzing the co-occurrence of the word “COPOM” with the “hawkish” and “dovish” terms listed above, followed by “rate” or “SELIC”<sup>4</sup>. Since we are not separating the statements into strings and we are also using common words in Portuguese language, the search noises become more frequent so we lock the Google search queries to the day of the interest rate's release.

---

<sup>4</sup> In Portuguese language, since it is possible to use different prepositions between the “hawkish” and “dovish” terms and the words “rate” and “SELIC”, we need to employ in our Google search all those prepositions.

Following Lucca and Trebbi (2011), we define the unexpected change on the content of each statement  $t$  as the difference between the semantic score of date  $t$  and  $t - 1$  as showed below:

$$\Delta GSO_t = GSO_t - GSO_{t-1}$$

We also try to build a measure of the language's surprise from the statement's content. The GSOs calculated above take into account the language surprise and also what is expected to be written on the central bank statement. For this reason, we build an expectation statement score, based on the latter type of GSO, in which we just replace the term "COPOM" by the term "expectation" ("expectativa" in Portuguese language) to the Google query. Additionally, we define a range of one week for the Google search queries, ending on the day before the interest rate decision and then we subtract the expectation statement score from the GSO to create the language surprise score.

The main caveat of the expectation statement score is that the query results basically consider news about the expectation of the interest rate decision itself and not about the expectation of what will be written on the statement.

There are some important limitations in the construction of the GSO. We are not able to control the webpages included in the Google search queries. For example, since we are not able to control the distance between the statements strings and the meaning terms, the latter are likely to appear next to words or expressions not related to the interest rate decision, but next to words such as "GDP" or "unemployment". The webpages have the same weight in the GSO, regardless of the distance between the string and the meaning terms, even though the farther the string from the terms, the lower the connection between each other.

In addition, we are not capable of restricting the time at which the webpages are published. Thus, we cannot separate news released before and after the interest rate's decision, in order to build a less noisy expectation semantic score, as is done for the FSO.

### 3 Data and Results

In order to create our semantic scores, we use 92 BCB's statements, from September 2003, when the interest rate decision started to be released at night time, after market close, until November 2013, last meeting of the year. Given that the GSO has some limitations, since webpages are removed over time, it starts in the beginning of 2008. On the other hand, our FSOs encompass all COPOM statements.

We follow Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2013) and build a measure of interest rate surprise. We define it as the difference between the "Swap DI x pré" of 30-day maturity on the following business day after the interest rate decision and the swap rate on the COPOM's decision day. The "Swap DI x pré" of 30-day maturity is taken by the economic agents as an expectation about the interest rate decision. Consequently, the difference measures the unexpected change on the SELIC level.

Following Carvalho, Cordeiro and Vargas (2013), we assume rational expectations. Since the interest rate decision is revealed by the central bank to the public at night time, when the market is closed, we assume that what leads to changes on the yields on the following business day is the unexpected change on the SELIC rate level, represented by the interest rate surprise, the unexpected change on the content of central bank's statement or any relevant event<sup>5</sup> that could have happened in the meantime.

Our benchmark regression specification (daily data specification) is defined as follows (Carvalho, Cordeiro and Vargas (2013))<sup>6</sup>:

$$\Delta y_k^n = \beta_0^n + \beta_1^n MP_k \cdot (1 - Tomb) + \beta_2^n \Delta Score_k \cdot (1 - Tomb) + \beta_3^n MP_k \cdot Tomb + \beta_4^n \Delta Score_k \cdot Tomb + \varepsilon_k^n$$

<sup>5</sup> No relevant events regularly happened on the following business day of the interest rate decision. The relevant events considered were: Nonfarm payrolls, FOMC, CPI and GDP from the US, industrial production, CPI, GDP, unemployment rate from Brazil and CPI and GDP from the EU.

<sup>6</sup> We treat the COPOM meetings,  $k$ , as our cross section data, when we run the regression for each maturity,  $n$ .

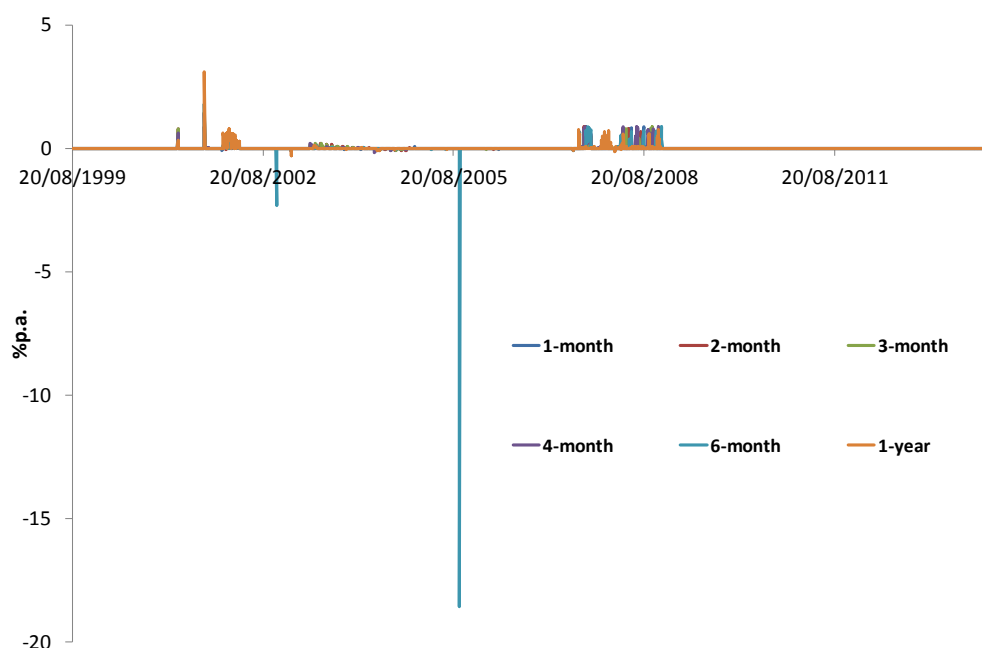
where  $\Delta y_k^n$  is the change in the yield of maturity  $n$  between the date  $k$  of the interest rate release and the next business day;  $MP_k$  is the interest rate surprise;  $\Delta Score_k$  is the difference between the semantic scores in the statements  $k$  and  $k - 1$ ;  $Tomb$  is a dummy variable, which takes value 1 for the period Tombini, that starts in January 2011, and zero for period prior Tombini. We include the dummy variable in order to analyze if the pattern of the yields response changes from period prior to, and during Tombini's tenure. Following Lucca and Trebbi (2011) and Carvalho, Cordeiro and Vargas (2013), we standardize changes in all semantic scores in the regressions, so their coefficients measure the effect of one standard deviation increase on the yield's variation (in basis point).

The following sections will show our daily and intraday data and their respective results.

### 3.1. Daily Data

Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2012) build their yield curves using "Swap DI x pré" fixed-rate swap contracts. They use the 1, 2, 3, 4, 6 month and 1, 2, 3, 5 and 10 year maturities of those contracts. For maturities up to 1 year values were taken from the Central Bank of Brazil database and beyond it, values were taken from Bloomberg.

Unfortunately, BCB's dataset displayed some serious errors. Some of the swap rates from BCB's website were completely different from the primary source, BMF&Bovespa. To illustrate this, let's consider an example, the "Swap DI x pré" of 1-year maturity on 01/23/2008 was 11.041% taken from the primary source, as well as on Bloomberg. However, the same swap rate was 11.41% on BCB's website. For some reason, this mistake appears to happen regularly. Figure 2 shows the difference between the correct swap rates and the ones from BCB's website from the beginning of the series in mid-1999 until February 2014.



**Figure 2 - Discrepancies in BCB's series of swaps**

As noticed above, all swap rate maturities up to 1 year were incorrect on BCB's website for a certain period, mainly between mid-2007 and the end of 2008. Since the GSO, built by Vargas (2012), Cordeiro (2013) and Carvalho, Cordeiro and Vargas (2012) started in October 2007, their results were compromised. In order to rectify their results, we run the same regression specification, using the same GSO, updated until the end of 2013, and the correct swap rates, taken from Bloomberg. Table 2 shows the results with the correct data.

We find that the semantic score has an effect on the short-to-medium yields during the pre-Tombini period. Quantitatively, we find that a unit standard deviation increase in the semantic score, in other words, a change in the content of the statement towards the "hawkish" direction leads to a response in the 6-month to 2-year yields between 2 and 4 basis points, which is significant at conventional levels. On the other hand, during Tombini's mandate, the semantic score has no statistically significant effect on the term structure of interest rates for any maturity.

The interest rate surprises have an effect on the yields up to 1 year maturity at conventional levels, during the pre-Tombini period. However, for long term maturities, the coefficients become negative and not statistically significant. Quantitatively, there is a one-to-one response in the short term, but it decreases for

longer maturities. On the other hand, during Tombini's period, there is a one-to-one response, sometimes more, for all maturities. These responses are statistically significant at conventional levels.

**Table 2 - GSO results**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔGSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔGSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.985*** (0)	0.00196 (0.585)	1.034*** (0)	-0.000944 (0.876)	48	0.958
Δ 3-mo. Yield	0.949*** (0)	0.00566 (0.439)	1.277*** (0)	0.00599 (0.593)	48	0.863
Δ 4-mo. Yield	0.954*** (0)	0.0127 (0.158)	1.521*** (0)	0.00602 (0.664)	48	0.836
Δ 6-mo. Yield	0.926*** (2.69e-10)	0.0213* (0.0661)	1.673*** (0)	0.00883 (0.647)	48	0.78
Δ 1-yr. Yield	0.684*** (0.000364)	0.0375** (0.0431)	1.704*** (8.57e-09)	0.0124 (0.647)	48	0.647
Δ 2-yr. Yield	0.203 (0.512)	0.0434* (0.0955)	1.535*** (1.09e-09)	0.00303 (0.900)	48	0.512
Δ 3-yr. Yield	-0.0496 (0.897)	0.0366 (0.232)	1.215*** (0)	-0.0122 (0.489)	48	0.328
Δ 5-yr. Yield	-0.243 (0.552)	0.0397 (0.233)	1.110*** (0)	-0.0155 (0.308)	48	0.289
Δ 10-yr. Yield	-0.405 (0.353)	0.0191 (0.665)	0.759*** (3.84e-07)	-0.0124 (0.512)	48	0.153

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

Carvalho, Cordeiro and Vargas (2013) left an important remark: the difficulty of the GSO's replication, due to the elimination of old webpages. In order to address this problem with the GSO, we create another GSO using the same strings six months after the one built by Carvalho, Cordeiro and Vargas (2013).

Table 3 shows the results of the GSO, built six months later. The coefficients of the semantic score change dramatically, especially in the pre-Tombini period, as expected, because it is the period, during which the most webpages have been removed. The GSO's coefficients change sign, making it difficult to derive an economic interpretation. The coefficients from the Tombini

period remain basically the same as well as the yield responses to the interest rate surprises. This highlights an important instability of the GSO.

We also create some GSO variations. We fix a range of two weeks for the Google search queries, defined as GSOifc (GSO with initial and final cut), as well as another GSO construction, not depending on personal judgments about the strings' division, defined as GSO2. Unfortunately, we create those GSO variations in the same period as the score created six months later. Consequently, we find results with no clear economic interpretation in the pre-Tombini period, due to the removed webpages.

**Table 3 - Later GSO Results**

Dep. Variable	$MP_k(1 - Tomb)$	$\Delta GSO_k(1 - Tomb)$	$MP_kTomb$	$\Delta GSO_kTomb$	Obs.	R-squared
$\Delta$ 2-mo. Yield	0.983*** (0)	0.00127 (0.708)	1.034*** (0)	-5.51e-05 (0.994)	48	0.957
$\Delta$ 3-mo. Yield	0.950*** (0)	-0.000313 (0.960)	1.251*** (0)	0.00862 (0.510)	48	0.862
$\Delta$ 4-mo. Yield	0.958*** (0)	-0.00152 (0.859)	1.495*** (0)	0.00851 (0.617)	48	0.831
$\Delta$ 6-mo. Yield	0.934*** (1.97e-10)	-0.00319 (0.767)	1.645*** (0)	0.00974 (0.658)	48	0.764
$\Delta$ 1-yr. Yield	0.717*** (0.000417)	-0.0143 (0.377)	1.680*** (5.16e-08)	0.00996 (0.722)	48	0.618
$\Delta$ 2-yr. Yield	0.241 (0.455)	-0.0162 (0.504)	1.545*** (5.61e-09)	-0.000873 (0.971)	48	0.472
$\Delta$ 3-yr. Yield	-0.00456 (0.991)	-0.0189 (0.514)	1.236*** (8.08e-09)	-0.00627 (0.756)	48	0.301
$\Delta$ 5-yr. Yield	-0.173 (0.683)	-0.0299 (0.284)	1.147*** (1.61e-09)	-0.0110 (0.513)	48	0.269
$\Delta$ 10-yr. Yield	-0.291 (0.518)	-0.0505 (0.112)	0.766*** (4.00e-05)	-0.00125 (0.950)	48	0.200

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

Table 4 shows the correlation between the semantic scores for the periods prior to and during Tombini's tenure. In the shaded area, we notice that the correlation between the GSO built later and the other semantic scores are lower in the period prior to Tombini. This is evidence that the information stock in the

GSO depreciates over time. Thus, the GSO built later is less informative about the period prior to Tombini, because using the same strings and building the semantic score six months later, we get less correlation for this period.

**Table 4 - Correlation between the Semantic Scores**

Pre-Tombini								
	GSO	Later GSO	GSOifc	GSO2	FSO_GSO	FSO	FSO-	FSO+
GSO	1.00							
Later GSO	0.75	1.00						
GSOifc	0.63	0.60	1.00					
GSO2	0.85	0.73	0.63	1.00				
FSO_GSO	0.83	0.65	0.69	0.90	1.00			
FSO	0.86	0.67	0.73	0.92	0.96	1.00		
FSO-	0.84	0.63	0.68	0.92	0.95	0.97	1.00	
FSO+	0.86	0.70	0.76	0.91	0.94	0.99	0.92	1.00
Tombini								
	GSO	Later GSO	GSOifc	GSO2	FSO_GSO	FSO	FSO-	FSO+
GSO	1.00							
Later GSO	0.87	1.00						
GSOifc	0.79	0.73	1.00					
GSO2	0.81	0.89	0.82	1.00				
FSO_GSO	0.81	0.80	0.86	0.91	1.00			
FSO	0.86	0.81	0.88	0.91	0.98	1.00		
FSO-	0.79	0.77	0.90	0.90	0.97	0.98	1.00	
FSO+	0.88	0.82	0.84	0.90	0.97	0.99	0.95	1.00

In addition, we estimate the yield responses to our language surprise measure, defined as the difference between the GSO, which also includes the expected changes in the communication, and our expectation semantic score. Table 5 shows the language surprise results. We notice that there is a significant change on the pattern of the yield responses to the communication surprises from pre-Tombini period to Tombini's tenure. Furthermore, the interest rate surprise's effects remain the same as showed above in the GSO results.

In order to avoid this GSO problem with the removed webpages, we follow Lucca and Trebbi (2011) and create the Factiva Semantic Score, using the Factiva Dow Jones database. Differently from Lucca and Trebbi (2011), we find most of the yield responses not statistically significant at conventional levels, for both  $\Delta FSO\_GSO$  and  $\Delta FSO$ <sup>7</sup>. In the next subsection, we will analyze the intraday results and then with this in hands, we will discuss the FSO problems<sup>8</sup>.

<sup>7</sup> Although, the FSO starts on 2003, we also analyze the results with the FSO starting in 2008 in order to compare with the GSO outcomes.

<sup>8</sup> All results are in the appendix.

### 3.2. Intraday data

In this section, we analyze the yield responses to the content of the central bank's statement and to the interest rate surprise, using an intraday approach.

**Table 5 - Language Surprise (GSO) (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	Surp_GSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	Surp_GSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.981*** (0)	0.00145 (0.766)	1.025*** (0)	-0.00554 (0.286)	46	0.959
Δ 3-mo. Yield	0.911*** (0)	0.0111 (0.258)	1.255*** (0)	-0.0180* (0.0580)	46	0.874
Δ 4-mo. Yield	0.879*** (2.90e-09)	0.0216* (0.0581)	1.484*** (0)	-0.0270** (0.0284)	46	0.861
Δ 6-mo. Yield	0.823*** (7.49e-07)	0.0299** (0.0334)	1.627*** (0)	-0.0330* (0.0634)	46	0.808
Δ 1-yr. Yield	0.577*** (0.00802)	0.0315 (0.130)	1.655*** (0)	-0.0360 (0.145)	46	0.676
Δ 2-yr. Yield	0.147 (0.616)	0.0183 (0.327)	1.494*** (0)	-0.0283 (0.231)	46	0.570
Δ 3-yr. Yield	-0.0770 (0.832)	0.00957 (0.677)	1.194*** (2.46e-09)	-0.0114 (0.639)	46	0.368
Δ 5-yr. Yield	-0.274 (0.499)	0.00974 (0.699)	1.113*** (8.88e-09)	0.00655 (0.760)	46	0.319
Δ 10-yr. Yield	-0.447 (0.295)	0.0113 (0.649)	0.789*** (2.91e-06)	0.0235 (0.305)	46	0.197

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

We use tick-by-tick data of the interbank deposit futures contracts (DI1)<sup>9</sup> traded on the following business day after all COPOM meetings in our sample. Our data encompasses all DI1 contracts traded and the times at which they were negotiated. These contract rates incorporate the economic agents' expectation about the interest rate level from the date that they are traded until the contract's maturity and also a risk premia.

The DI1's maturity rule works as follows: it always matures on the first business day of the contract's month. Moreover, the contract months are the four months subsequent to the month in which a trade has been made and, after that, the months that initiate a quarter. For example, if the current month is January,

<sup>9</sup> The interbank deposit is the rate at which banks lend to each other.

there will be contracts maturing on the first business days of February, March, April, May, July, October and so on.

Unlike swap contracts, the DI1's maturities are not fixed. Every day, depending on the liquidity, there are fixed maturities for the swap contracts, such as 2, 3, 4, 6 months and 1, 2, 3, 5 and 10 years. On the other hand, since COPOM meetings take place at any part time of the month and the DI1 contracts always mature in the beginning of the contract months, their time to maturity are not regular. Rarely, we have DI1 contracts maturing in the exact fixed maturities mentioned above for the swap contracts.

For this reason, in order to estimate fixed maturities to the DI1 futures contracts, we use the Nelson-Siegel-Svensson (NSS) model, which consists of the Nelson Siegel (1987) model with its extension by Svensson (1994)<sup>10</sup>. The NSS model is defined as follows:

$$y(m) = \beta_0 + \beta_1 \frac{[1 - \exp(-m/\lambda_1)]}{m/\lambda_1} + \beta_2 \left\{ \frac{[1 - \exp(-m/\lambda_1)]}{m/\lambda_1} - \exp(-m/\lambda_1) \right\} \\ + \beta_3 \left\{ \frac{[1 - \exp(-m/\lambda_2)]}{m/\lambda_2} - \exp(-m/\lambda_2) \right\}$$

where  $m$  is the maturity,  $\beta_0, \beta_1, \beta_2, \beta_3, \lambda_1$  and  $\lambda_2$  are the parameters and  $y(m)$  is the yield of the curve at maturity  $m$ .

All these parameters are related to the format of the term structure of interest rates. The constant,  $\beta_0$ , represents the long term interest rate level,  $\beta_1$  represents the slope of the curve and  $\beta_2$  and  $\beta_3$  represent the curvature of the term structure.  $\lambda_1$  and  $\lambda_2$  are the shape parameters.  $\beta_1$  and  $\beta_2$  are also interpreted as the short term and medium term components. Both  $\lambda_2$  and  $\beta_3$  are additional parameters included by Svensson (1994) to the original Nelson-Siegel (1987) model.

As noticed above, the NSS model is highly nonlinear, due to the two parameters,  $\lambda_1$  and  $\lambda_2$ . First, we try to estimate all parameters simultaneously. However, we do not find any global minimum, as well as any local minimum. To circumvent this problem, we follow Nelson and Siegel (1987) and use ordinary least square to estimate four parameters conditional to a grid of the two shape

<sup>10</sup> To the best of my knowledge is the first intraday NSS estimated for Brazil.

parameters. The chosen grid value was 0.1 ranging from 0.1 to 20 for both shape parameters. We select the optimal parameters with the highest  $R^2$ .

We estimate the NSS model for different periods of time on the following business day of the COPOM's decision. In order to have a minimum amount of contracts traded to be able to build the term structure of interest rates, we begin to analyze the intraday yield curve starting from the 30<sup>th</sup> minute after the financial market opens'. Furthermore, we study the term structure behavior in the 60<sup>th</sup>, 120<sup>th</sup>, 180<sup>th</sup>, 240<sup>th</sup>, 360<sup>th</sup>, 420<sup>th</sup>, 480<sup>th</sup> and 540<sup>th</sup> minutes<sup>11</sup>.

Moreover, we create a 'liquidity control' parameter,  $\theta$ . Some contracts are illiquid and for this reason, could bias the estimation. We set  $\theta = 60$ , so we only consider contracts traded within the last 60 minutes of each interval mentioned above, regardless of their maturities<sup>12</sup>. From this set of contracts, we pick up the last contract traded of each maturity and then we estimate our NSS model and build our intraday yield curve. Furthermore, we define a minimum of eight contracts to be considered in order to build the term structure. If there are less than eight contracts traded in the 'liquidity control' set, we do not build the term structure for this period of time.

We also estimate a yield curve of the price settle<sup>13</sup> of each intraday contract inside our liquidity control set. If a contract has a price settle, but it was not traded on the following business day of the COPOM's decision, we drop it from our term structure estimation.

After building the term structures of the DI1 futures contracts and the ones of their respective price settles each minute mentioned above for each COPOM meeting, we build the 3-, 4-, 6-, 9-, 12-, 15-, 18-, 21-, 24-, 30- and 36-month maturities. In order to compute the yield variation for each maturity, we calculate the difference between the term structures of the DI1 and the price settles.

Since the DI1 and the swap contracts are closely related, the intraday yield variation of the market's closing time, built by the NSS model, should be similar to the yield variation of our daily data. In order to verify this similarity, Table 6 shows the correlation between the yield variations of our daily and intraday data.

<sup>11</sup> The financial market opens for 9 hours.

<sup>12</sup> An alternative 'liquidity parameter' is the contracts' volume. It can be established a minimum contract volume to be considered in the liquidity control set.

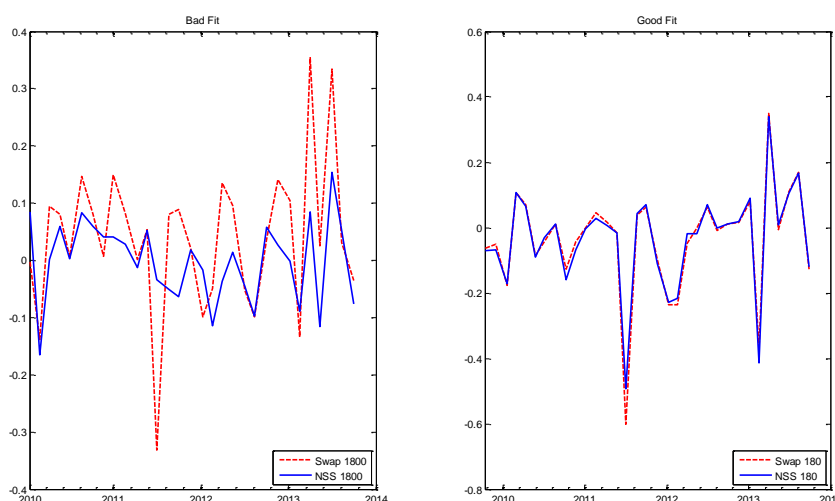
<sup>13</sup> Price settle is typically used as the closing price for future contracts. It is calculated as the weighted average of transactions' prices during the market's closing period.

As noted, the correlations of the yield variations are considerably high. Since there are some difficulties to obtain a good fit in the short term part of the curve, it can be noticed that the lowest correlation is for the 3-month yield across all maturities.

**Table 6 - Correlation between intraday and daily variation**

3-month	4-month	6-month	9-month	12-month
0.75	0.86	0.95	0.98	0.98
15-month	18-month	21-month	24-month	36-month
0.98	0.98	0.98	0.98	0.98

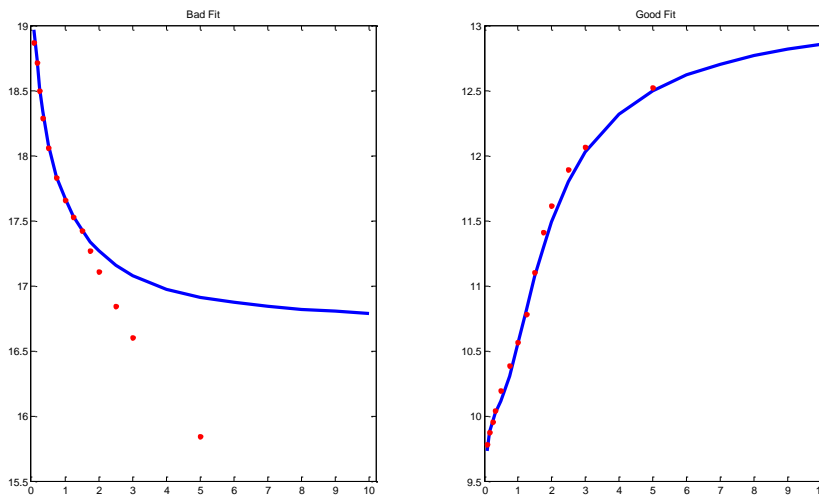
In order to analyze how bad and how good the NSS's fit can be, we show in Figure 3 one example of each type. On the left, we plot the yield variation of the 5-year (1800 days) swap contract together with the yield variation estimated by the NSS for the same maturity in the market's closing time. Since the DI1 futures contracts of longer maturities are in general more illiquid, the NSS model does not fit well the yield variation<sup>14</sup>. On the right, we plot the yield variation of the 6-month (180 days) swap contract together with the yield variation estimated by the NSS. We notice that the NSS fits well the yield variation of the swap contract.



**Figure 3 – NSS fit (Yield Variation)**

<sup>14</sup> The NSS does not estimate properly maturities far from the range used to fit the term structure. For example, if the NSS estimates the yield curve using maturities ranging from the 3-month to the 3-year maturities, the 5-year maturity will not be correctly estimated.

In addition, we show another example of how bad the NSS fit can be due to illiquid contracts. Figure 4 shows the term structure of the DI1 futures contracts estimated by the NSS for the last hour of the market for two different COPOM meetings together with the swap contract rates of the same day, represented by red bullet points. On the left, one can notice that the term structure poorly fits the swap contract rates beyond the 1-year maturity, since only contracts up to around the 1-year maturity were traded in this period of time. On the right, we show a term structure estimated from contracts traded up to the 10-year maturity and we note that it fits well the swap contract rates<sup>15</sup>.



**Figure 4 – NSS fit (Term Structure)**

Thus, the intraday regression specification is defined as follows:

$$\begin{aligned} \Delta y_k^{n,t} = & \beta_0^{n,t} + \beta_1^{n,t} MP_k \cdot (1 - Tomb) + \beta_2^{n,t} \Delta Score_k \cdot (1 - Tomb) \\ & + \beta_3^{n,t} MP_k \cdot Tomb + \beta_4^{n,t} \Delta Score_k \cdot Tomb + \varepsilon_k^{n,t} \end{aligned}$$

There are three important differences from the daily specification. First, the dependent variables are different. While on the daily specification, our dependent variable is defined as the difference between the swap rates on the following business day after the interest rate decision and the swap rates on the COPOM's decision day, our the dependent variable on the intraday specification is defined as the difference between the DI1 futures contracts traded on the following business day of the interest rate decision at a particular time and its own price settle on the

<sup>15</sup> We choose a term structure that is outside our 'liquidity control' set, in order to expose a bad NSS fit. Only 6 different maturities were traded in this period of time, while on the good fit term structure 18 different maturities were traded.

COPOM's decision day. Second, we include the superscript  $t$ , which represents the period of time that we run the regression. Third, we also change our interest rate surprise measure in the intraday specification. Now, instead of using the variation of the swap contracts, the interest rate surprise is the difference between the rate of the last DI1 future contract, traded in the first hour in the following business day after the COPOM's decision, which maturity is the closest to the 30-day maturity and its respective price settle on the day before. Since we lose nine data points in our sample, we complete the missing values by the DI1 future contract - if it exists - with shorter maturity than the one mentioned above and the remaining missing values, we just complete by our previous interest rate surprise measure. By doing this, we could better estimate the intraday dynamics effect, because using only the daily data could give us an artificial intraday dynamics<sup>16</sup>.

All the upcoming figures follow the same pattern. On the primary axis, we display the yield responses to the content of the COPOM statement or to the interest rate surprises, in other words, it shows their regression coefficients in each maturity and each period of time. On the secondary axis, we set the p-values up to 10% of significance level, which are represented by red bullet points. We also include the 95% confidence interval.

Figures 5 and 6 show the yield responses to the GSO for each maturity at each period of the day for the periods prior to and during Tombini's tenure, respectively. We find that the yield responds to the content of the COPOM statements only in the period prior to Tombini in most part of the day, which confirms our findings using daily data. There is evidence of the 1- and 2-year maturities responding almost all day long.

Figures 7 and 8 show the yield responses to the interest rate surprise. There is a one-to-one response for the pre-Tombini period up to the 6-month maturity and it is statistically significant at conventional levels up to the 1-year maturity. On the other hand, there is a one-to-one yield response, sometimes more, for the interest rate surprise in all maturities for the Tombini period. This also confirms our daily data findings.

---

<sup>16</sup> In the appendix, we expose both intraday results using our daily measure of the interest rate surprise and using our intraday interest rate surprise measure without replacing the missing values.

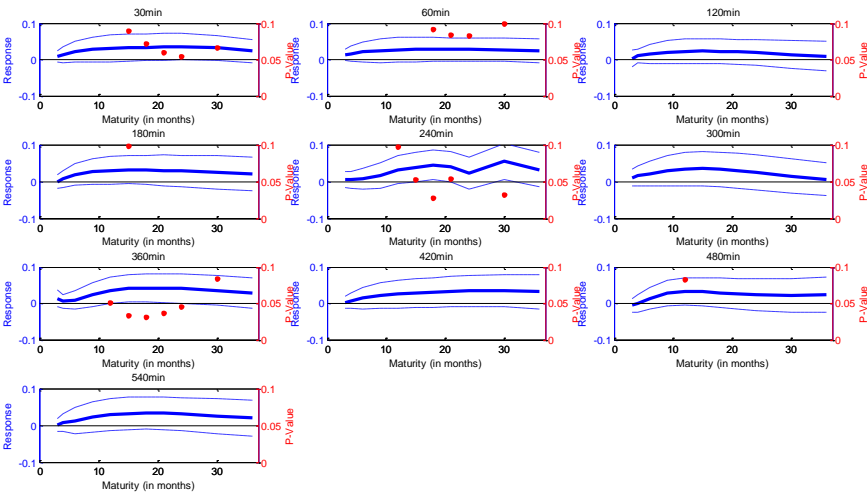


Figure 5 - Yield responses to the language for pre-Tombini period

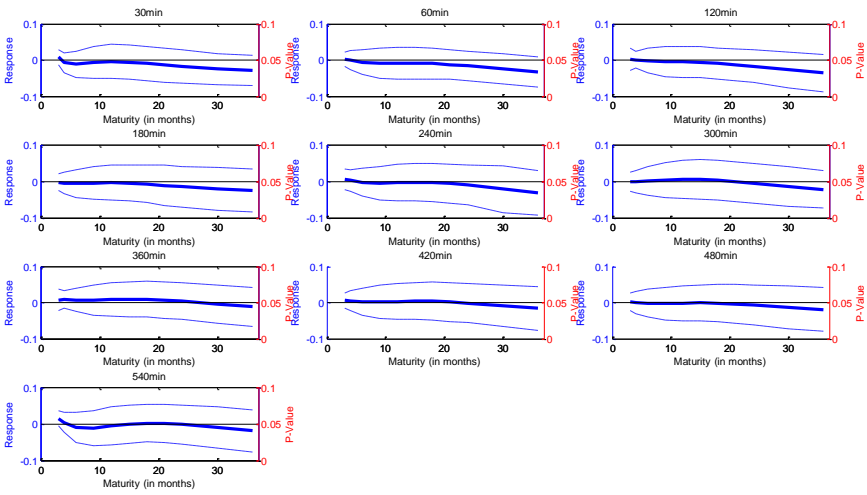
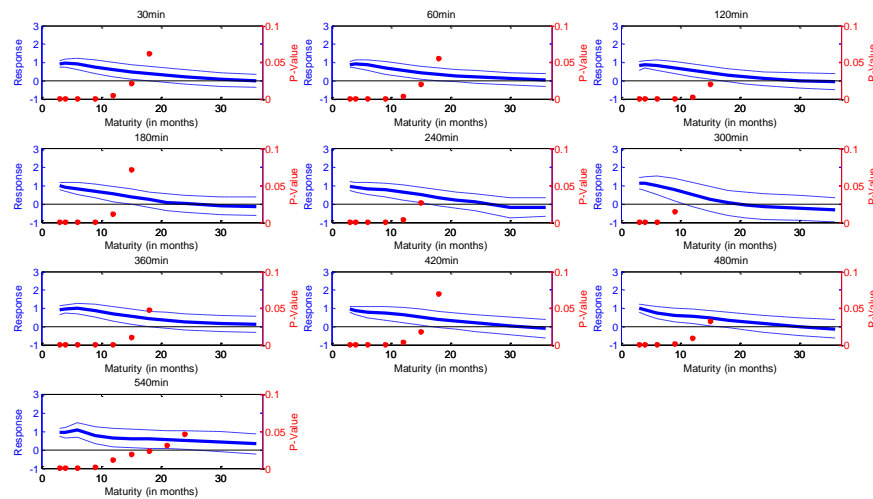
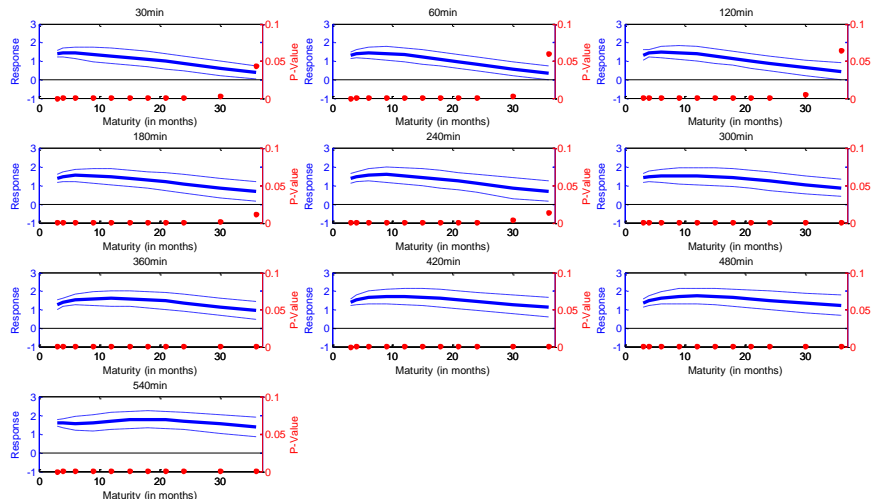


Figure 6 - Yield responses to the language for Tombini period



**Figure 7 - Yield responses to the interest rate surprise for pre-Tombini period**



**Figure 8 - Yield responses to the interest rate surprise for Tombini period**

In addition, we analyze the existence of intraday dynamics of the DI1 contracts. If there is not intraday dynamics, we expect the contracts to react instantaneously to the content of the statement and to the interest rate surprise shortly after the financial market opens. However, it may be the case that the information is not fully interpreted as soon as the financial market opens and that message is assimilated over time.

In order to capture the intraday dynamics effect, we separate the operating time of the financial market into three parts. The first part corresponds to the first hour, the second corresponds to the interval from the second to the fifth hour and the third comprises the last four hours of the market. The regression specification of the market's first part is the intraday one. Thus, we define our intraday dynamics regression specification of the second and third parts as follows<sup>17</sup>:

$$\begin{aligned}\Delta y_k^{n,p} - \Delta y_k^{n,p-1} &= \beta_0^{n,p} + \beta_1^{n,p} MP_k \cdot (1 - Tomb) + \beta_2^{n,p} \Delta Score_k \cdot (1 - Tomb) \\ &+ \beta_3^{n,p} MP_k \cdot Tomb + \beta_4^{n,p} \Delta Score_k \cdot Tomb \\ &+ \beta_5^{n,p} \hat{\varepsilon}_k^{n,p-1} \cdot (1 - Tomb) + \beta_6^{n,p} \hat{\varepsilon}_k^{n,p-1} \cdot Tomb + \varepsilon_k^{n,p}\end{aligned}$$

where  $\Delta y_k^{n,p} - \Delta y_k^{n,p-1}$  is defined as the difference between the yield variation of part  $p$  and part  $p - 1$ ;  $\hat{\varepsilon}_k^{n,p-1}$  is the residual of part  $p - 1$  regression also iterated with the Tombini dummy variable<sup>18</sup>. The other independent variables are defined in the same way as the intraday specification. In this case, the interest rate surprises and the semantic score coefficients are interpreted as the incremental effect on the yield variation only in part  $p$ .

Including the dependent variable defined above, we try to explain the yield variation that happened between  $p$  and  $p - 1$ . We add the estimated residuals of the previous part in order to estimate the effect of any other variable that could impact the yield curve, excluding the interest rate surprises and the semantic scores.

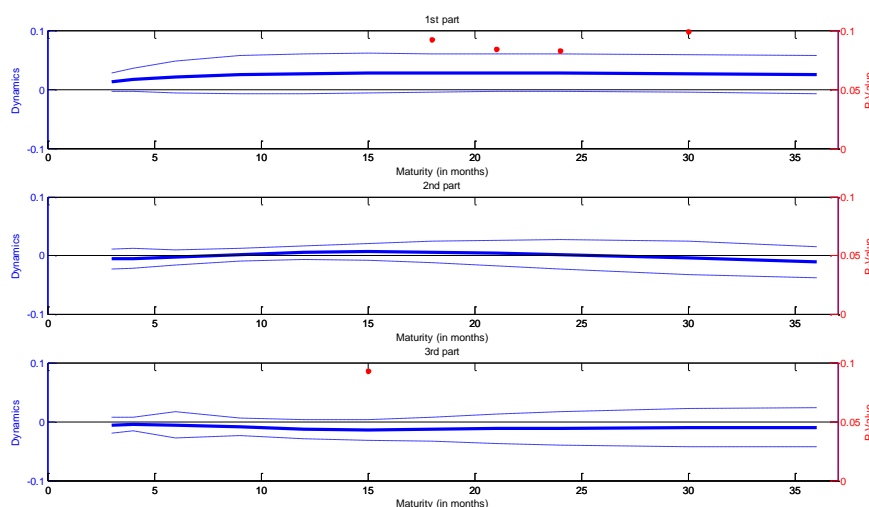
<sup>17</sup> In this case,  $p = 2, 3$ .

<sup>18</sup>  $\hat{\varepsilon}_k^{n,p-1}$  is a generated regressor. The main problem when we include it is the fact that we are using a proxy instead of the actual variable, which we do not observe, and we are measuring it with error. By using generated regressors, we get biased standard errors, which impede to make a proper inference. We run the intraday dynamics specification without the generated regressors and we still find similar results (in the appendix).

Figures 9 and 10 show the intraday dynamics of the content of the statement in each market's part. As noted, the yield responds only in the first part in the pre-Tombini period. On the other hand, the COPOM language has an effect in the second and third parts in the Tombini period. This evidences that, prior to Tombini's tenure, the message on the statement is quickly assimilated, while during the latter period, the economic agents delay to fully understand the message of the COPOM's statement.

Figures 11 and 12 show the intraday dynamics of the interest rate surprises in each market's part. During the pre-Tombini period, there is a one-to-one response in the short term portion of the yield curve, but it decays with longer maturities. In the second and third parts, the interest rate surprises still affect the yield variation in the short-to-medium maturities. On the other hand, during Tombini's tenure, there is a one-to-one response, sometimes more, in the first part for almost all maturities. Likewise the pre-Tombini period, the interest rate surprise still affects the yield changes in the second and third parts during Tombini's tenure, but unlike the previous period, the long term end of the yield curve responds to the interest rate surprises.

Figures 13 and 14 show the yield responses to the residuals of the respective previous market parts for both periods. We find that there is a sharp response on the short term end in the last four hours of the market during Tombini period, maybe due to an event that could somehow strongly impact the yield variation.



**Figure 9 - Intraday dynamics of the language for pre-Tombini period**

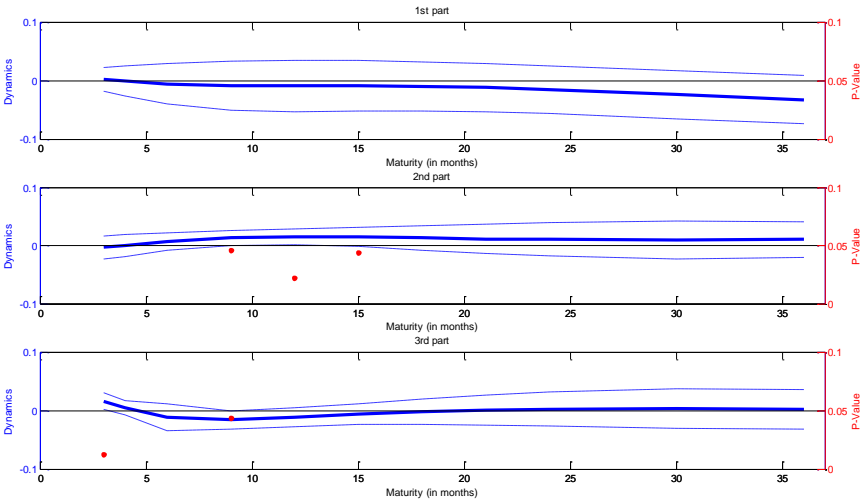


Figure 10 - Intraday dynamics of the language for Tombini period

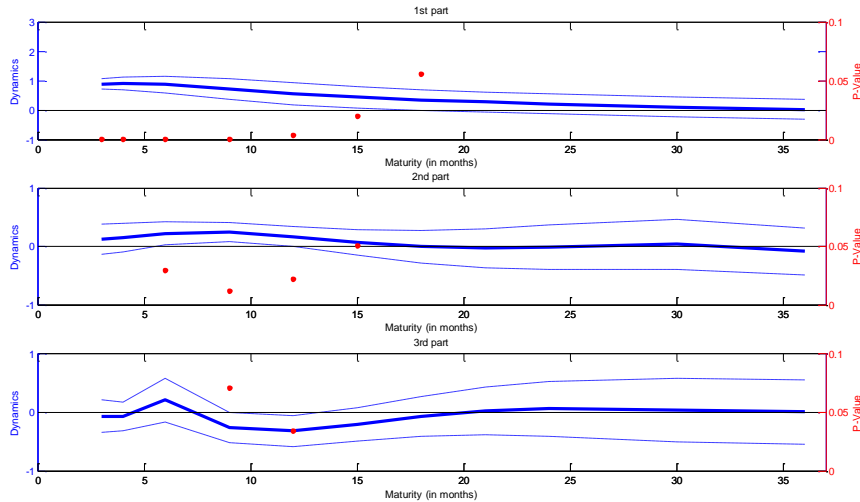
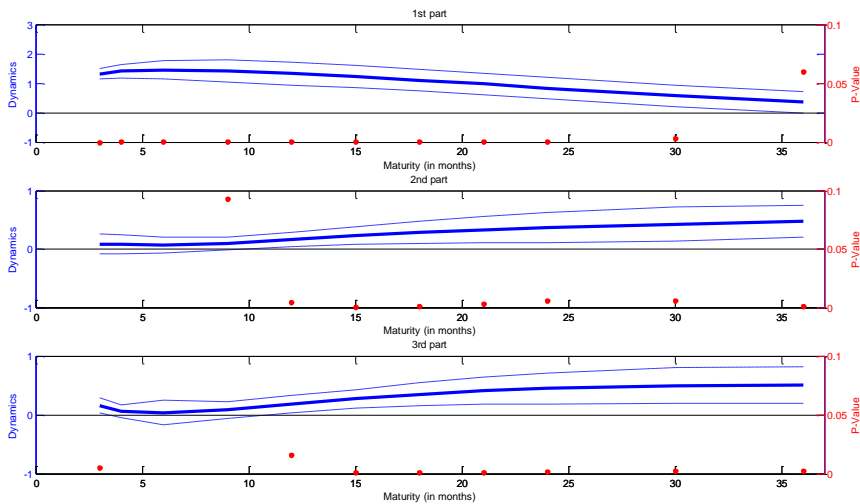
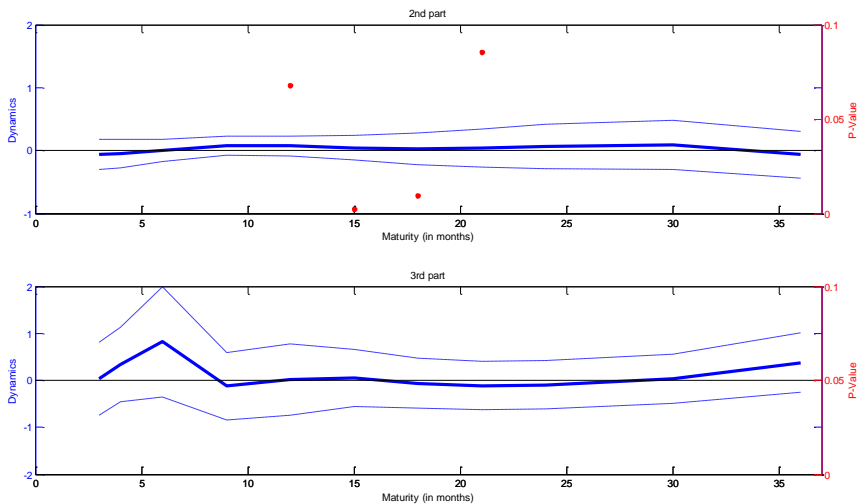


Figure 11 - Intraday dynamics of the interest rate surprise for pre-Tombini period



**Figure 12 - Intraday dynamics of the interest rate surprise for Tombini period**



**Figure 13 - Intraday dynamics of the residuals for pre-Tombini period**

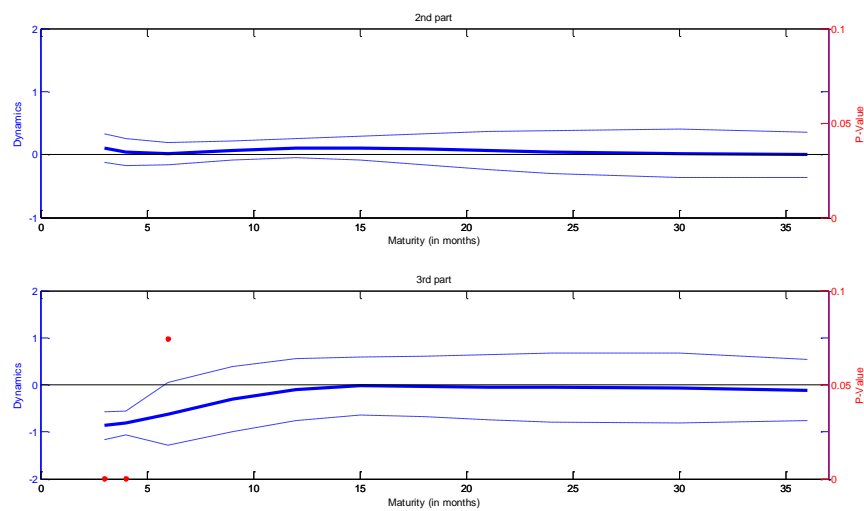


Figure 14 - Intraday dynamics of the residuals for Tombini period

## 4 Conclusion

Even though, we have full control over the construction of the FSO, there are some critical problems in building it for the COPOM statements. We can list three main problems with the FSO. First, the Factiva Dow Jones database has less Brazilian media sources than Americans. There are only 96 source searches for Brazil in the Factiva, compared to more than 3000 for the United States. Second, there is much less media coverage over the COPOM meeting compared to the FOMC. The FOMC's decision is reported worldwide, since any economic decision in the USA affects in some way all other economies. Only Brazilian domestic media describes the COPOM statements and its impact on the economy, while international media in general only reports the interest rate decision, not properly reporting the content of the COPOM statements. The third reason is related to each country's culture. While in the USA, people are used to discuss the economic environment, in Brazil the same does not happen frequently. In other words, the public is more aware about economic discussions, including the FOMC's decision about the interest rate level and its statement.

It is clear that there is a trade-off between the GSO and the FSO. While on the one hand, in spite of having full control over the query, our news sample is considerably low in the construction of the FSO. On the other hand, even though we have a lack of control over the Google search queries, we have a high amount of webpages in order to build the GSO. Since the low amount of our news universe on the FSO is rather costly, we consider the GSO, the semantic score which best represents the content of the BCB statement.

Our GSO results, both using daily and intraday data, suggest that communication lost somehow its importance among the economic agents during Tombini's tenure compared to the period prior to Tombini, considering that the COPOM statement stopped to affect the yield curve during Tombini period. Additionally, it seems that the content of the BCB's statements are fully interpreted in the beginning of the market's operating time during the pre-Tombini

period, while during Tombini's tenure, it appears to be assimilated along the day, affecting the yield curve in the middle and at the end of the market's operating time. This might happen due to the loss of credibility and confidence on the BCB's communication.

In addition, the interest rate surprise effects on the yield curve differ in both periods. In the Tombini period, these effects seem to influence the entire yield curve, what we do not observe in the pre-Tombini period, in which only short-to-medium maturities respond. This is evidence that economic agents expect that interest rate surprises will last long. This might be the case, considering the unexpected interest rate cut in August 2011 COPOM meeting and the unexpected increase in April and May 2013 meetings.

Moreover, we find evidence that the interest rate surprises impact the yield curve during all day for both periods analyzed. Thus, it shows that the financial market does not react to the interest rate surprises in just one moment, but rather it reacts in several moments, evidenced by the fact that we see their incremental effects on the yield changes during the entire market's operating time.

## 5 References

CARVALHO, C.; CORDEIRO, F.; VARGAS, J. Just Words? A Quantitative Analysis of the Communication of the Central Bank of Brazil. **Revista Brasileira de Economia**, 67, 443-455, 2013.

CHURCH, K.; HANKS, P. Word Association Norms, Mutual Information, and Lexicography. **Computational Linguistics**, 16 (1), 22-29, 1990.

CORDEIRO, F. Quantificando os Efeitos da Comunicação do BC. **Monografia do Programa de Educação Tutorial** – Pontifícia Universidade Católica do Rio, 2013.

COSTA FILHO, A.; ROCHA, F. Comunicação e Política Monetária no Brasil. **Revista Brasileira de Economia** 63, 405-422, 2009.

COSTA FILHO, A.; ROCHA, F. Como o Mercado de Juros Futuros Reage à Comunicação do Banco Central?. **Revista Brasileira de Economia** 14, 265-292, 2010.

GURKAYNAK, R.; SACK, B; SWANSON, E. Do Actions Speak Louder Than Words? The Response of asset Prices to Monetary Policy Actions and Statements. **International Journal of Central Banking**, 1, 55-93, 2005.

LUCCA, D.; TREBBI, F. Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. **National Bureau of Economic Research, Working Paper** nº 15367, 2011.

NELSON, C.; SIEGEL, A. Parsimonious Modeling of Yield Curves. **Journal of Business** 60, 473-489, 1987.

ROSA, C.; VERGA, G. The Impact of Central Bank Announcements on Asset Prices in Real Time: Testing the Efficiency of the Euribor Futures Market. **International Journal of Central Banking** 4, 175-217, 2008.

SVENSSON, L.; Estimating and Interpreting Forward Interest Rates: Sweden 1992-1994. **International Monetary Fund, Working Paper** n° 114, 1994.

TERMINAL CARTEZYAN, **Intraday Data Source**.

TURNEY, P.; LITTMAN, M. Unsupervised Learning of Semantic Orientation from a Hundred-Billion-Word Corpus. **National Research Council**, Insitute for Information Technology, Technical Report ERB-1094, 2002.

VARGAS, J.; Mensurando a Comunicação do BACEN: Análise Quantitativa dos Comunicados do Copom. **Monografia de Graduação**, Departamento de Economia da PUC-Rio, 2012.

## 6 Tables and Figures

### 6.1. Daily Tables

**Table 7 – GSO ifc (with initial and final cutting date - since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔGSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔGSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.980*** (0)	-0.00537 (0.162)	1.058*** (0)	-0.00499 (0.220)	47	0.960
Δ 3-mo. Yield	0.932*** (0)	-0.0154 (0.298)	1.312*** (0)	-0.00651 (0.446)	47	0.869
Δ 4-mo. Yield	0.934*** (0)	-0.0193 (0.251)	1.595*** (0)	-0.0144 (0.201)	47	0.843
Δ 6-mo. Yield	0.903*** (5.73e-10)	-0.0238 (0.150)	1.810*** (0)	-0.0268* (0.0983)	47	0.786
Δ 1-yr. Yield	0.690*** (0.00163)	-8.27e-05 (0.997)	1.899*** (0)	-0.0368* (0.0913)	47	0.637
Δ 2-yr. Yield	0.245 (0.456)	0.0349 (0.235)	1.716*** (0)	-0.0322* (0.0720)	47	0.536
Δ 3-yr. Yield	0.00706 (0.986)	0.0494 (0.229)	1.369*** (0)	-0.0272* (0.0895)	47	0.387
Δ 5-yr. Yield	-0.200 (0.644)	0.0378 (0.424)	1.249*** (0)	-0.0245* (0.0752)	47	0.315
Δ 10-yr. Yield	-0.373 (0.412)	0.0314 (0.518)	0.863*** (7.03e-06)	-0.0174 (0.287)	47	0.186

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

**Table 8 – GSO2 (without strings - since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔGSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔGSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.986*** (0)	0.00340 (0.414)	1.051*** (0)	-0.00260 (0.567)	47	0.958
Δ 3-mo. Yield	0.948*** (0)	-0.000686 (0.946)	1.231*** (0)	0.00766 (0.399)	47	0.862
Δ 4-mo. Yield	0.953*** (0)	0.00685 (0.609)	1.480*** (0)	0.00682 (0.562)	47	0.831
Δ 6-mo. Yield	0.926*** (8.78e-11)	0.0146 (0.385)	1.630*** (3.13e-10)	0.00740 (0.663)	47	0.766
Δ 1-yr. Yield	0.682*** (0.000843)	-0.00298 (0.917)	1.651*** (8.13e-07)	0.0101 (0.656)	47	0.618
Δ 2-yr. Yield	0.196 (0.531)	-0.0118 (0.763)	1.496*** (1.01e-06)	0.00901 (0.643)	47	0.499
Δ 3-yr. Yield	-0.0572 (0.878)	-0.0228 (0.617)	1.205*** (2.59e-06)	0.00413 (0.812)	47	0.334
Δ 5-yr. Yield	-0.248 (0.533)	-0.0250 (0.608)	1.159*** (1.79e-07)	-0.00514 (0.729)	47	0.286
Δ 10-yr. Yield	-0.411 (0.332)	-0.0234 (0.660)	0.837*** (5.72e-05)	-0.00952 (0.570)	47	0.171

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 9 - FSO\_GSO (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔFSO_GSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔFSO_GSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.979*** (0)	-0.00607 (0.300)	1.050*** (0)	-0.00302 (0.384)	48	0.959
Δ 3-mo. Yield	0.927*** (0)	-0.0198 (0.308)	1.239*** (0)	0.00816 (0.287)	48	0.875
Δ 4-mo. Yield	0.931*** (5.90e-11)	-0.0204 (0.361)	1.472*** (0)	0.0103 (0.354)	48	0.842
Δ 6-mo. Yield	0.902*** (5.56e-09)	-0.0217 (0.358)	1.606*** (1.09e-09)	0.0141 (0.480)	48	0.775
Δ 1-yr. Yield	0.657*** (0.000385)	-0.0251 (0.310)	1.608*** (3.53e-06)	0.0201 (0.476)	48	0.629
Δ 2-yr. Yield	0.176 (0.560)	-0.0263 (0.381)	1.490*** (1.52e-06)	0.00985 (0.687)	48	0.479
Δ 3-yr. Yield	-0.0921 (0.802)	-0.0417 (0.284)	1.234*** (6.57e-07)	-0.00392 (0.852)	48	0.323
Δ 5-yr. Yield	-0.298 (0.438)	-0.0540 (0.265)	1.175*** (1.35e-07)	-0.0128 (0.517)	48	0.300
Δ 10-yr. Yield	-0.462 (0.249)	-0.0568 (0.207)	0.921*** (8.23e-07)	-0.0314* (0.0659)	48	0.213

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 10 - FSO\_GSO (since 2003)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔFSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	Δ_FSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	1.076*** (0)	-0.00270 (0.412)	1.075*** (0)	-0.00373 (0.264)	91	0.965
Δ 3-mo. Yield	1.164*** (0)	-0.00537 (-0.549)	1.261*** (0)	0.00708 (0.313)	91	0.899
Δ 4-mo. Yield	1.210*** (0)	-0.00595 (0.578)	1.498*** (0)	0.00894 (0.380)	91	0.851
Δ 6-mo. Yield	1.233*** (0)	-0.0104 (0.367)	1.641*** (0)	0.0124 (0.497)	91	0.793
Δ 1-yr. Yield	1.066*** (2.28e-09)	-0.0154 (0.272)	1.645*** (9.02e-08)	0.0183 (0.473)	91	0.602
Δ 2-yr. Yield	0.782*** (0.000197)	-0.0147 (0.431)	1.476*** (1.07e-07)	0.0102 (0.648)	91	0.382
Δ 3-yr. Yield	0.493*** (0.00137)	-0.0248 (0.217)	1.163*** (2.86e-07)	-0.00114 (0.952)	90	0.264
Δ 5-yr. Yield	0.238* (0.0970)	-0.0305 (0.211)	1.089*** (1.86e-07)	-0.00919 (0.605)	90	0.152
Δ 10-yr. Yield	-0.434 (0.197)	-0.0620* (0.0736)	0.893*** (1.18e-06)	-0.0287* (0.0803)	55	0.219

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 11 - FSO (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔFSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔFSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	1.002*** (0)	-0.00973*** (0.00478)	1.033*** (0)	-0.000122 (0.976)	48	0.960
Δ 3-mo. Yield	0.972*** (0)	-0.0127 (0.116)	1.284*** (0)	-0.00277 (0.738)	48	0.864
Δ 4-mo. Yield	0.963*** (0)	-0.00450 (0.561)	1.526*** (0)	-0.00168 (0.884)	48	0.830
Δ 6-mo. Yield	0.931*** (1.42e-10)	-0.00207 (0.817)	1.678*** (0)	-6.16e-05 (0.997)	48	0.763
Δ 1-yr. Yield	0.658*** (0.000703)	0.0164 (0.349)	1.714*** (9.67e-11)	-0.00104 (0.958)	48	0.615
Δ 2-yr. Yield	0.148 (0.614)	0.0337 (0.191)	1.547*** (1.13e-10)	-0.00491 (0.766)	48	0.479
Δ 3-yr. Yield	-0.108 (0.767)	0.0356 (0.272)	1.214*** (1.65e-09)	-0.000606 (0.971)	48	0.307
Δ 5-yr. Yield	-0.300 (0.445)	0.0350 (0.274)	1.101*** (6.41e-10)	0.00411 (0.764)	48	0.262
Δ 10-yr. Yield	-0.432 (0.315)	0.0164 (0.611)	0.742*** (1.69e-06)	0.00972 (0.519)	48	0.148

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 12 - FSO (since 2003)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	ΔFSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	ΔFSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	1.080*** (0)	-0.00114 (0.675)	1.054*** (0)	-0.000804 (0.856)	92	0.961
Δ 3-mo. Yield	1.178*** (0)	-0.00405 (0.426)	1.298*** (0)	-0.00330 (0.697)	92	0.891
Δ 4-mo. Yield	1.208*** (0)	0.00122 (0.847)	1.542*** (0)	-0.00225 (0.848)	92	0.840
Δ 6-mo. Yield	1.229*** (0)	0.00122 (0.873)	1.702*** (0)	-0.000772 (0.961)	92	0.781
Δ 1-yr. Yield	1.029*** (4.17e-09)	0.0108 (0.338)	1.740*** (0)	-0.00156 (0.937)	92	0.594
Δ 2-yr. Yield	0.723*** (0.000603)	0.0177 (0.286)	1.531*** (0)	-0.00401 (0.811)	92	0.381
Δ 3-yr. Yield	0.430** (0.0145)	0.0188 (0.295)	1.152*** (1.53e-10)	0.00183 (0.917)	91	0.254
Δ 5-yr. Yield	0.194 (0.234)	0.0117 (0.551)	1.031*** (2.32e-10)	0.00682 (0.638)	91	0.131
Δ 10-yr. Yield	-0.355 (0.374)	-0.00698 (0.813)	0.719*** (2.65e-06)	0.0112 (0.473)	55	0.128

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 13 - Language Surprise (Later GSO) (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	Surp_GSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	Surp_GSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.981*** (0)	0.00123 (0.838)	1.032*** (0)	-0.00265 (0.512)	46	0.958
Δ 3-mo. Yield	0.905*** (8.85e-10)	0.0113 (0.360)	1.274*** (0)	-0.00827 (0.183)	46	0.866
Δ 4-mo. Yield	0.854*** (4.82e-08)	0.0259* (0.0597)	1.509*** (0)	-0.0147* (0.0708)	46	0.851
Δ 6-mo. Yield	0.819*** (5.98e-07)	0.0283* (0.0754)	1.656*** (0)	-0.0189* (0.0894)	46	0.791
Δ 1-yr. Yield	0.613** (0.0114)	0.0201 (0.366)	1.683*** (0)	-0.0269* (0.0604)	46	0.662
Δ 2-yr. Yield	0.159 (0.669)	0.0133 (0.617)	1.517*** (0)	-0.0225* (0.0644)	46	0.565
Δ 3-yr. Yield	-0.0777 (0.864)	0.00889 (0.800)	1.206*** (7.81e-10)	-0.00437 (0.662)	46	0.365
Δ 5-yr. Yield	-0.240 (0.632)	0.00153 (0.966)	1.105*** (2.66e-09)	0.00656 (0.450)	46	0.318
Δ 10-yr. Yield	-0.315 (0.552)	-0.0205 (0.651)	0.759*** (2.15e-06)	0.0143 (0.257)	46	0.193

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 14 - Language Surprise (GSO ifc) (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	Surp_GSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	Surp_GSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.987*** (0)	0.00154 (0.865)	1.037*** (0)	0.00106 (0.782)	46	0.958
Δ 3-mo. Yield	0.939*** (0)	-0.0127 (0.592)	1.264*** (0)	-0.00495 (0.389)	46	0.864
Δ 4-mo. Yield	0.948*** (0)	-0.0114 (0.646)	1.494*** (0)	-0.00892 (0.269)	46	0.840
Δ 6-mo. Yield	0.924*** (4.68e-09)	-0.00921 (0.742)	1.632*** (0)	-0.0135 (0.223)	46	0.780
Δ 1-yr. Yield	0.689*** (0.00104)	-0.00502 (0.877)	1.644*** (3.64e-10)	-0.0212 (0.119)	46	0.654
Δ 2-yr. Yield	0.195 (0.525)	-0.0182 (0.571)	1.490*** (7.08e-10)	-0.0150 (0.215)	46	0.558
Δ 3-yr. Yield	-0.0722 (0.848)	-0.0331 (0.410)	1.191*** (5.70e-09)	-0.00468 (0.720)	46	0.373
Δ 5-yr. Yield	-0.266 (0.521)	-0.0356 (0.468)	1.101*** (9.84e-10)	0.00253 (0.821)	46	0.326
Δ 10-yr. Yield	-0.430 (0.327)	-0.0370 (0.477)	0.757*** (2.37e-06)	0.00541 (0.732)	46	0.193

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

**Table 15 - Language Surprise (GSO2) (since 2008)**

Dep. Variable	MP <sub>k</sub> .(1 - Tomb)	Surp_GSO <sub>k</sub> .(1 - Tomb)	MP <sub>k</sub> .Tomb	Surp_GSO <sub>k</sub> .Tomb	Obs.	R-squared
Δ 2-mo. Yield	0.987*** (0)	-0.000562 (0.895)	1.032*** (0)	-0.00344 (0.534)	46	0.958
Δ 3-mo. Yield	0.942*** (0)	0.00269 (0.714)	1.278*** (0)	-0.00294 (0.732)	46	0.861
Δ 4-mo. Yield	0.931*** (0)	0.00937 (0.284)	1.511*** (0)	-0.0111 (0.321)	46	0.841
Δ 6-mo. Yield	0.903*** (3.83e-10)	0.0102 (0.324)	1.659*** (0)	-0.0154 (0.249)	46	0.780
Δ 1-yr. Yield	0.727*** (0.000304)	-0.0129 (0.397)	1.693*** (3.66e-10)	-0.0199 (0.174)	46	0.651
Δ 2-yr. Yield	0.284 (0.331)	-0.0264 (0.146)	1.533*** (4.12e-10)	-0.00838 (0.542)	46	0.564
Δ 3-yr. Yield	0.0411 (0.908)	-0.0296 (0.186)	1.222*** (1.40e-09)	0.0136 (0.274)	46	0.387
Δ 5-yr. Yield	-0.147 (0.707)	-0.0306 (0.160)	1.118*** (3.52e-10)	0.0206 (0.142)	46	0.347
Δ 10-yr. Yield	-0.345 (0.425)	-0.0177 (0.545)	0.768*** (4.97e-07)	0.0280* (0.0582)	46	0.206

P-values in parenthesis (based on heteroskedasticity-robust standard errors)

\*\*\* p&lt;0.01, \*\* p&lt;0.05, \* p&lt;0.1

6.2.Intraday Figures

6.2.1.  $\Delta FSO\_GSO^{19}$

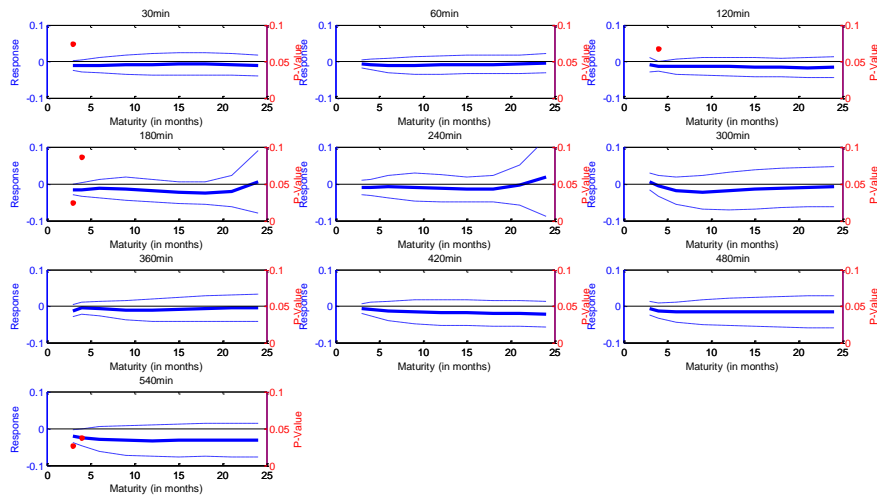


Figure 15 - Yield responses to the language for pre-Tombini period

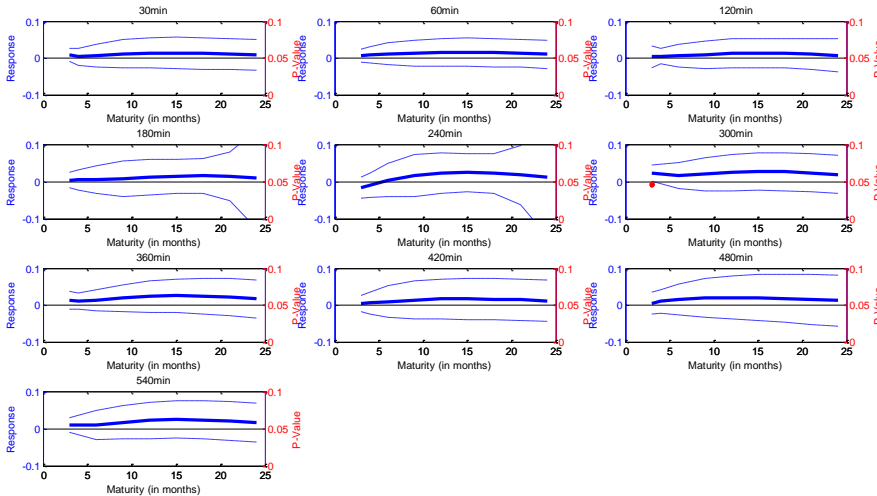
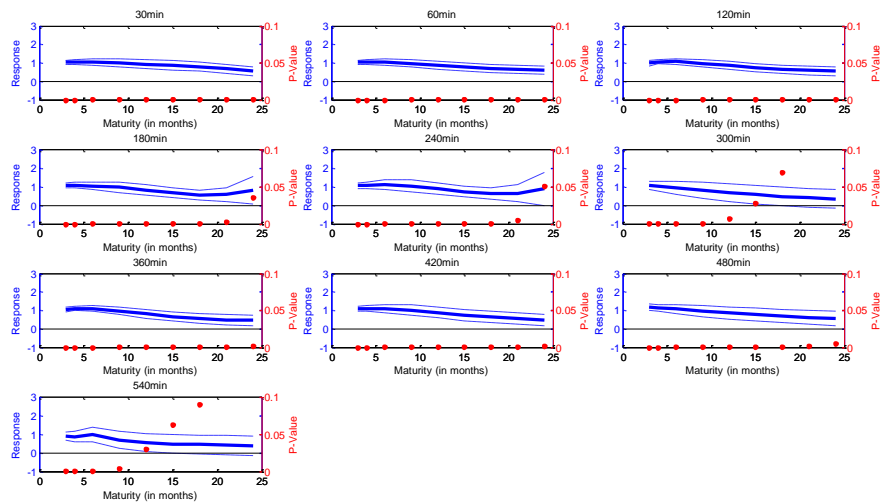
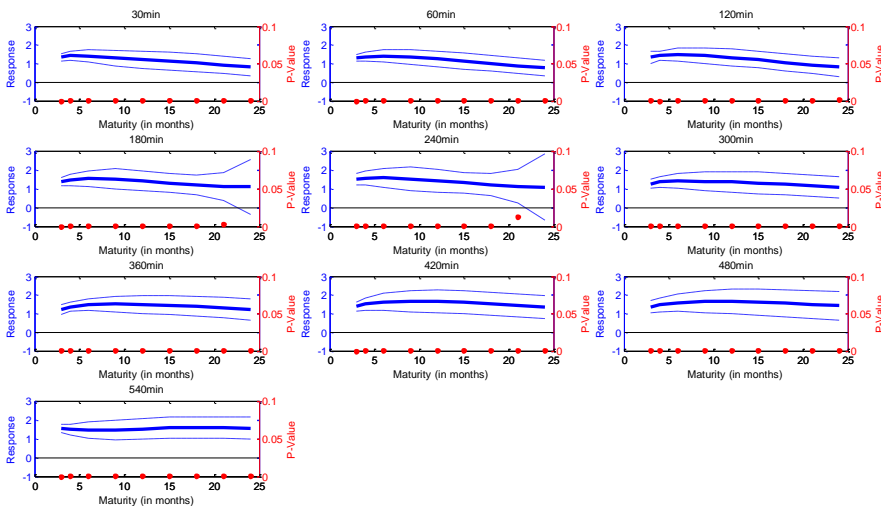


Figure 16 - Yield responses to the language for Tombini period

<sup>19</sup> Given that, our  $\Delta FSO\_GSO$  starts in 2003, all figures show the effects until the 2 year maturity, due to the fact that DI1 futures contracts were in general illiquid after the 2 year maturity back in 2003 and 2004.



**Figure 17 - Yield responses to the interest rate surprise for pre-Tombini period**



**Figure 18 - Yield responses to the interest rate surprise for Tombini period**

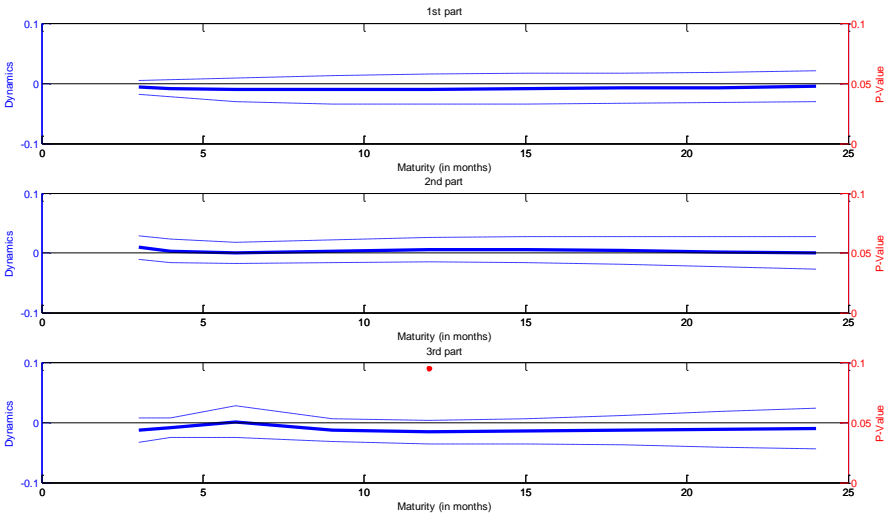


Figure 19 - Intraday dynamics of the language for pre-Tombini period

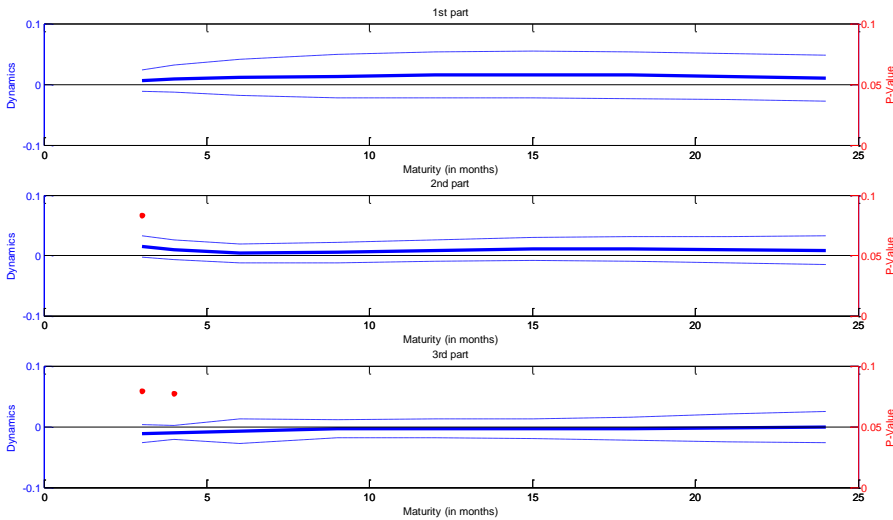
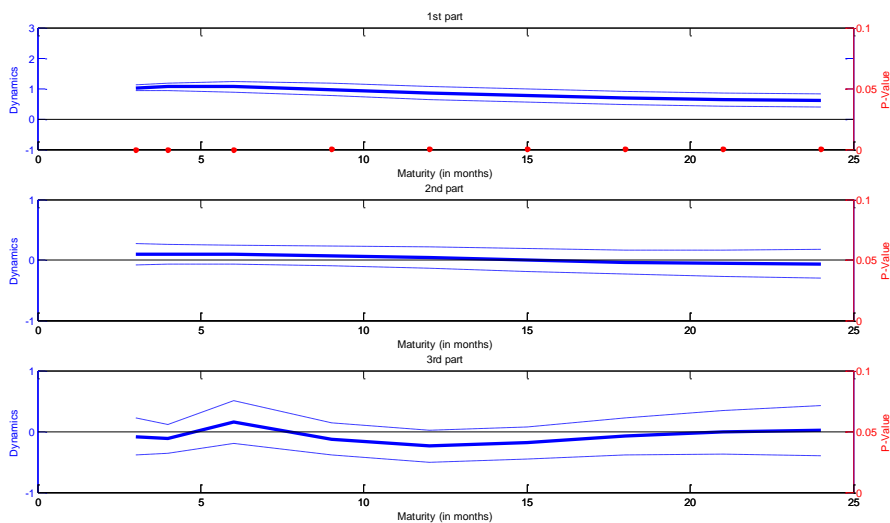
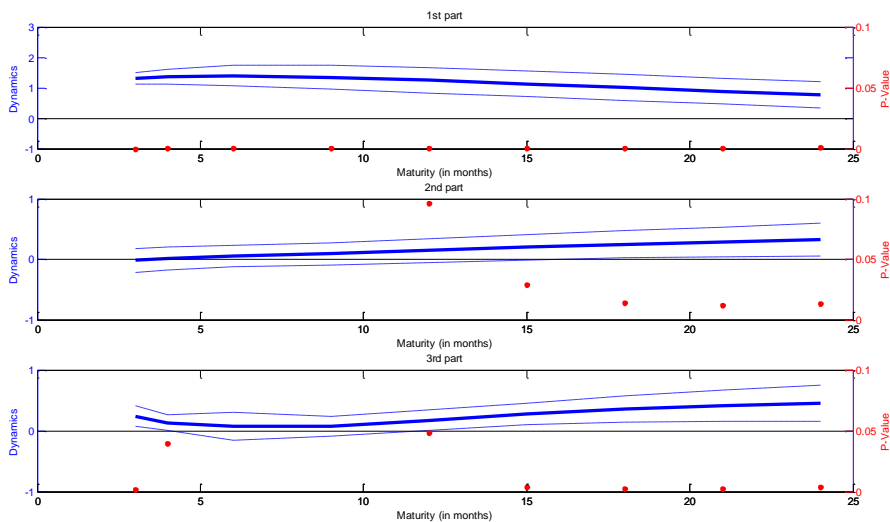


Figure 20 - Intraday dynamics of the language for Tombini period



**Figure 21 - Intraday dynamics of the interest rate surprise for pre-Tombini period**



**Figure 22 - Intraday dynamics of the interest rate surprise for Tombini period**

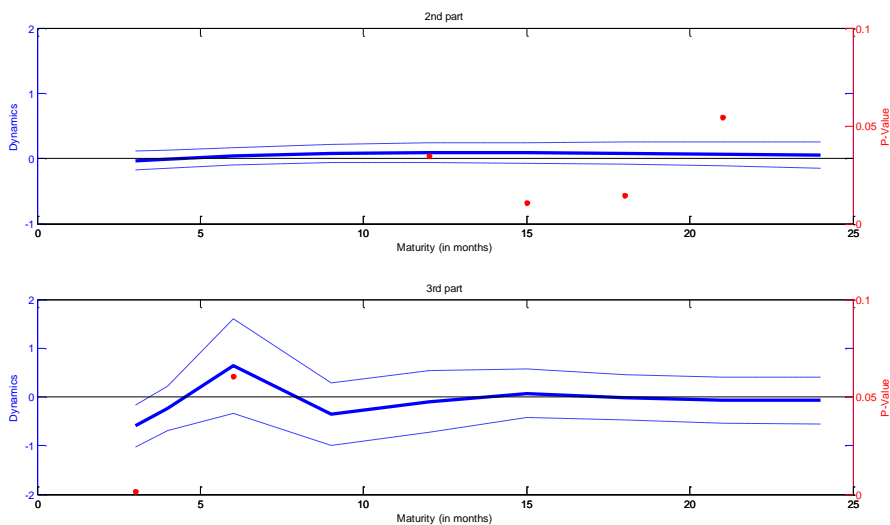


Figure 23 - Intraday dynamics of residuals for pre-Tombini period

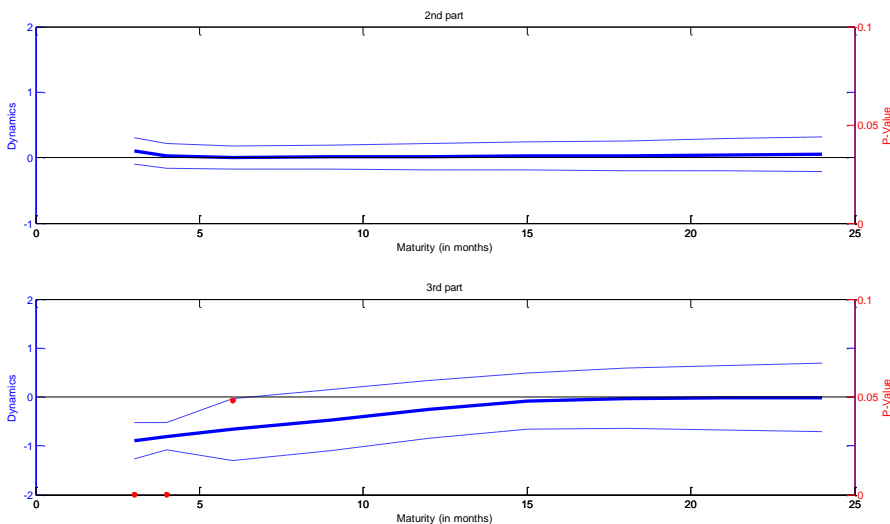


Figure 24 - Intraday dynamics of the residuals for Tombini period

6.2.2. GSO Intraday Results (Alternative Interest Rate Surprise)

- Daily Interest Surprise Measure

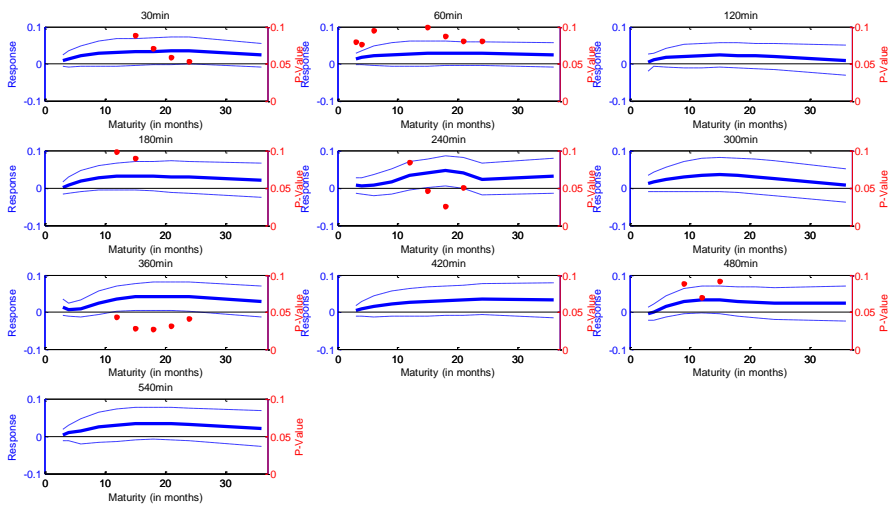


Figure 25 - Yield responses to the language for pre-Tombini period

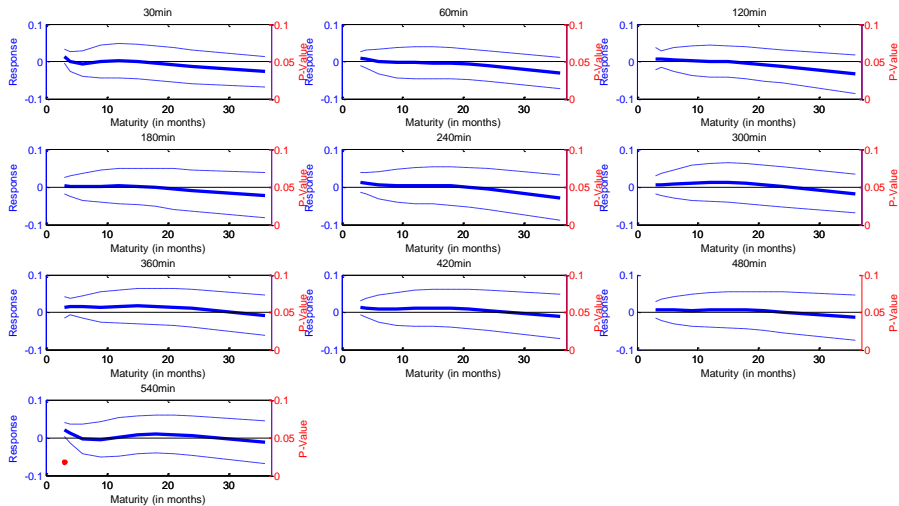
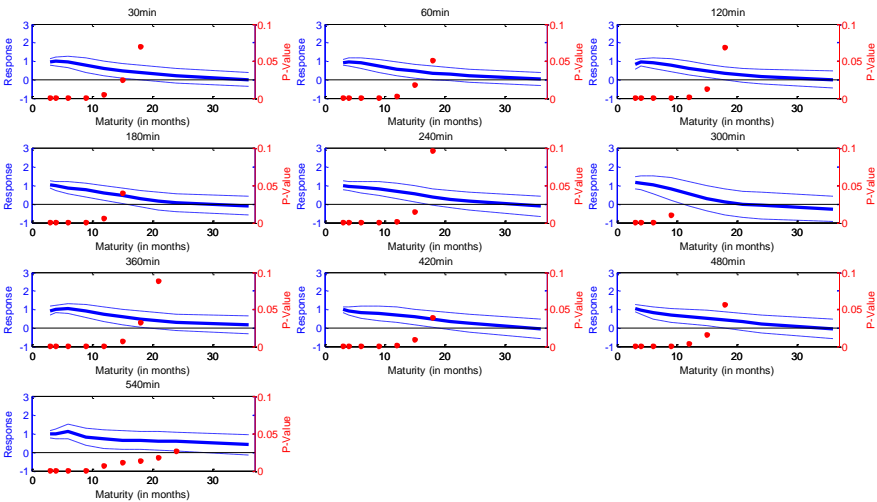
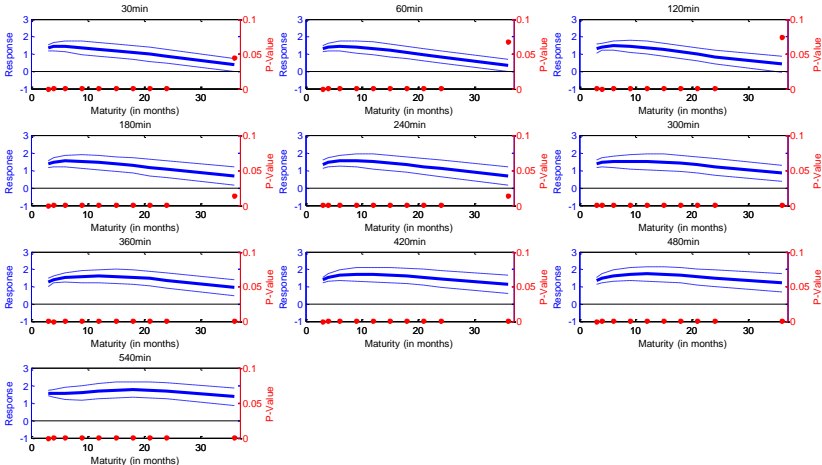


Figure 26 - Yield responses to the language for Tombini period



**Figure 27 - Yield responses to the interest rate surprise for pre-Tombini period**



**Figure 28 - Yield responses to the interest rate surprise for Tombini period**

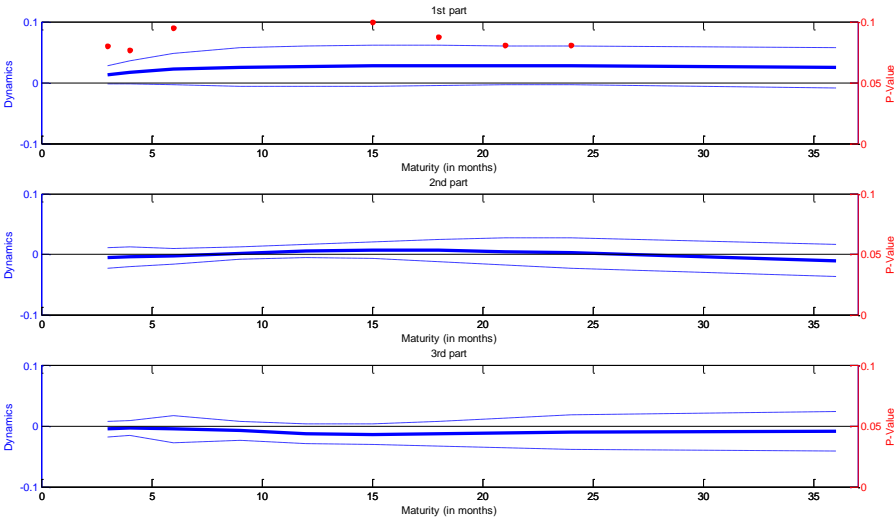


Figure 29 - Intraday dynamics of the language for pre-Tombini period

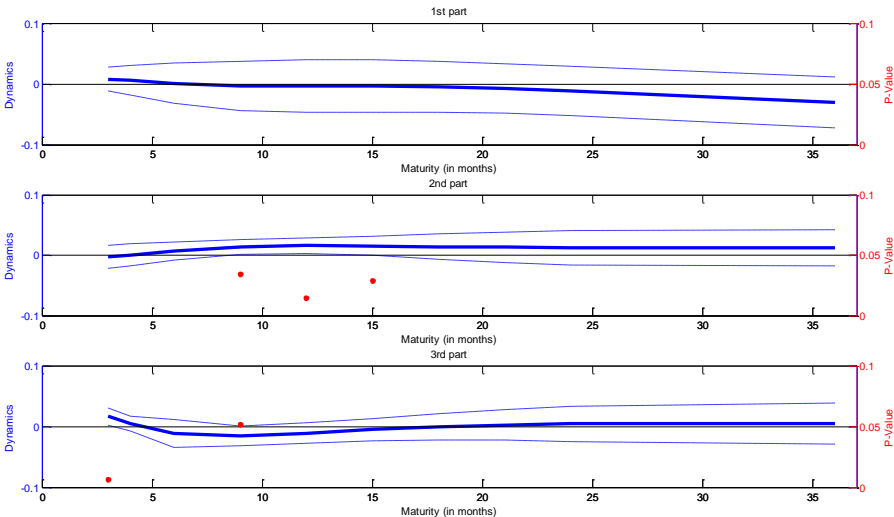
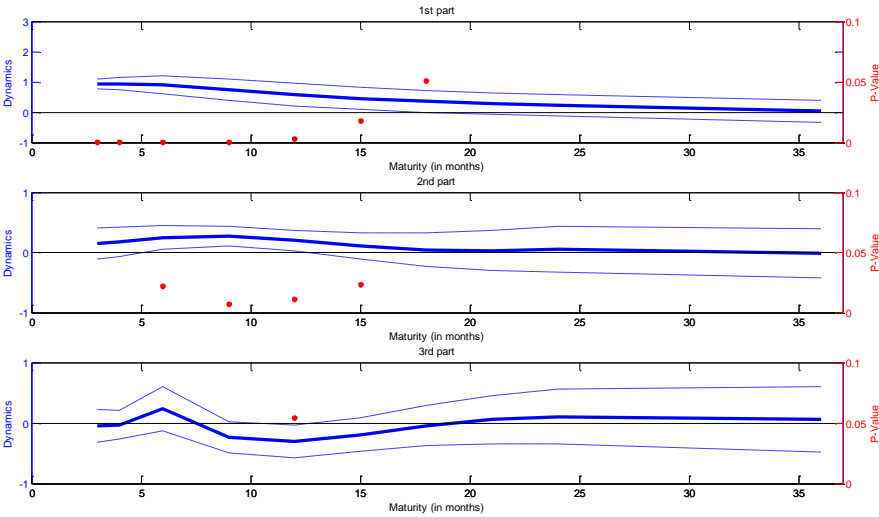
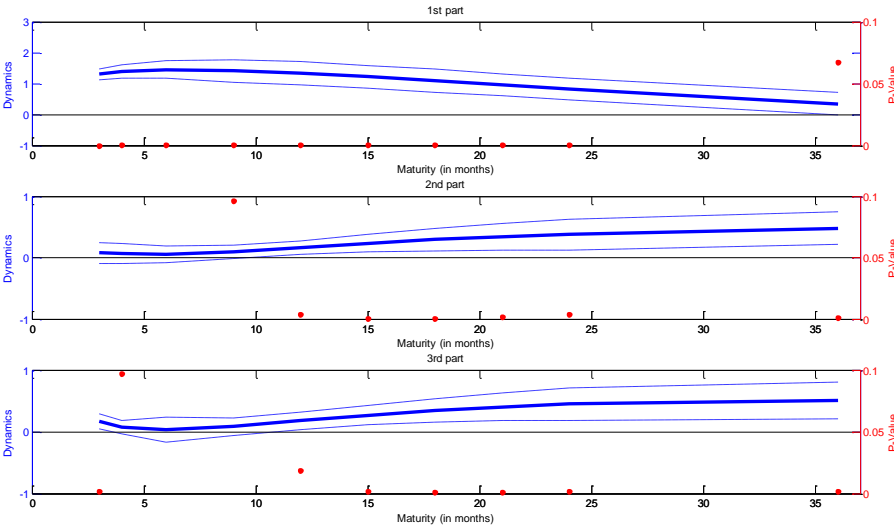


Figure 30 - Intraday dynamics of the language for Tombini period



**Figure 31 - Intraday dynamics of the interest rate surprise for pre-Tombini period**



**Figure 32 - Intraday dynamics of the interest rate surprise for Tombini period**

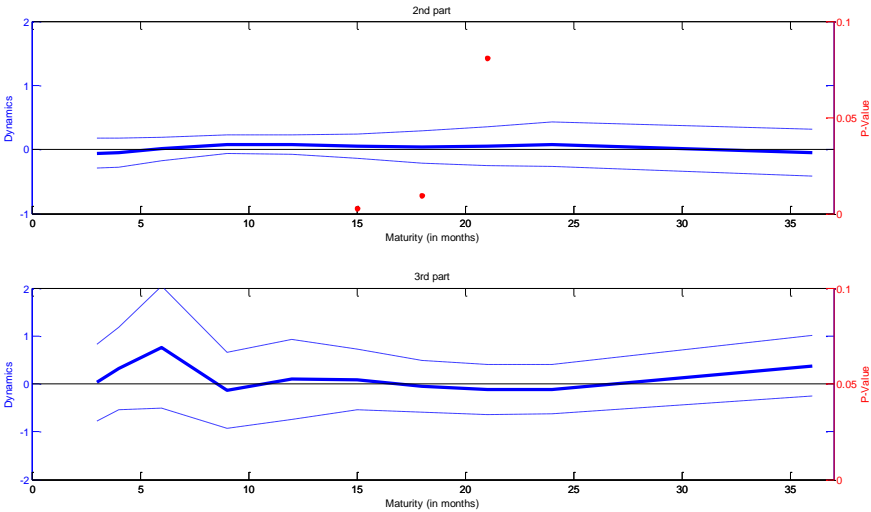


Figure 33 - Intraday dynamics of the residuals for pre-Tombini period

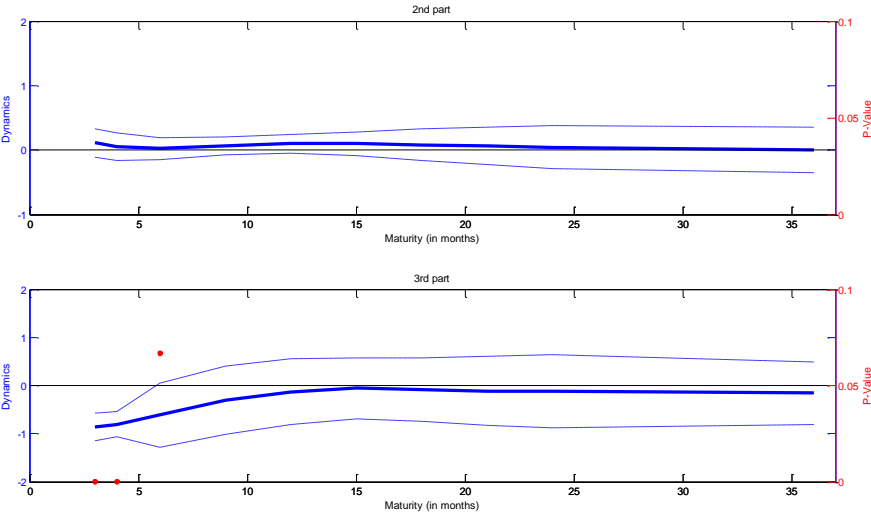


Figure 34 - Intraday dynamics of the residuals for Tombini period

- Intraday Interest Surprise Measure (without replacing the missing values)

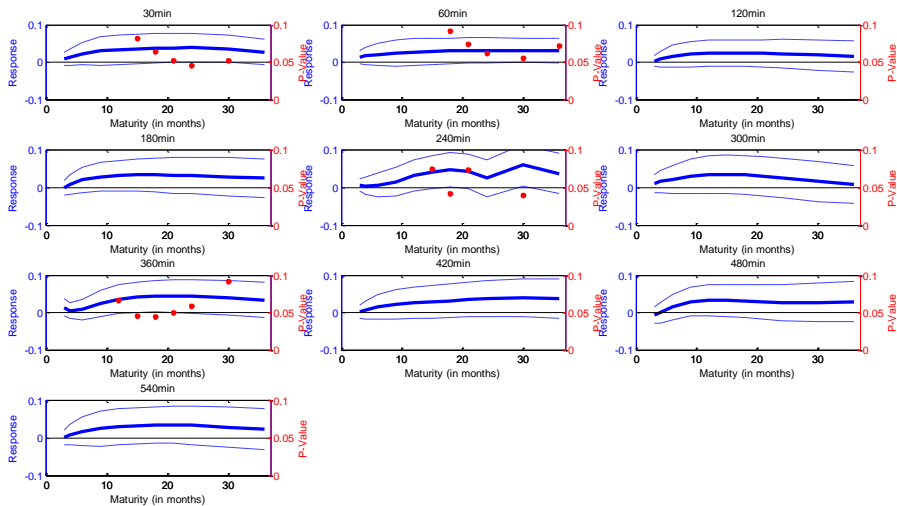


Figure 35 - Yield responses to the language for pre-Tombini period

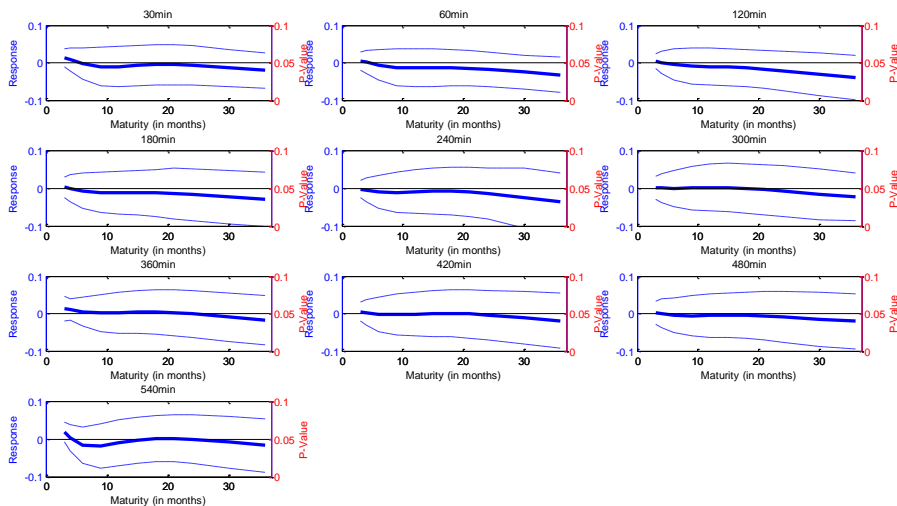
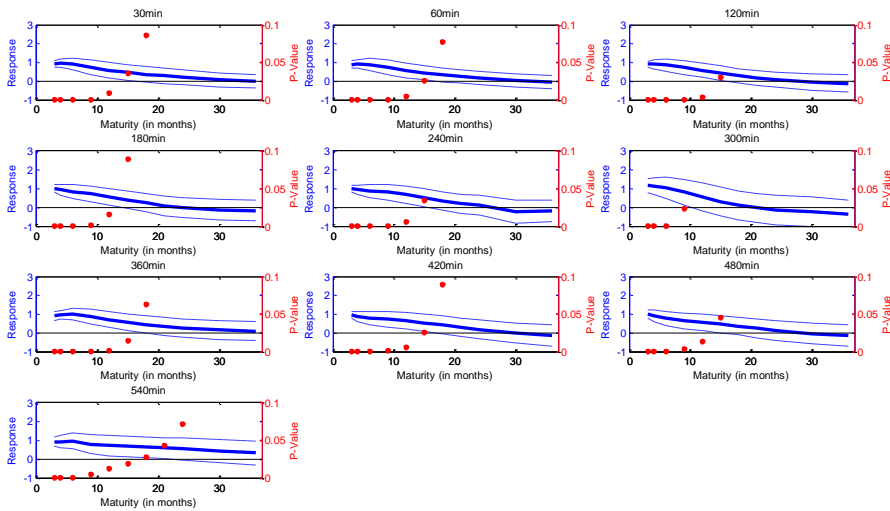
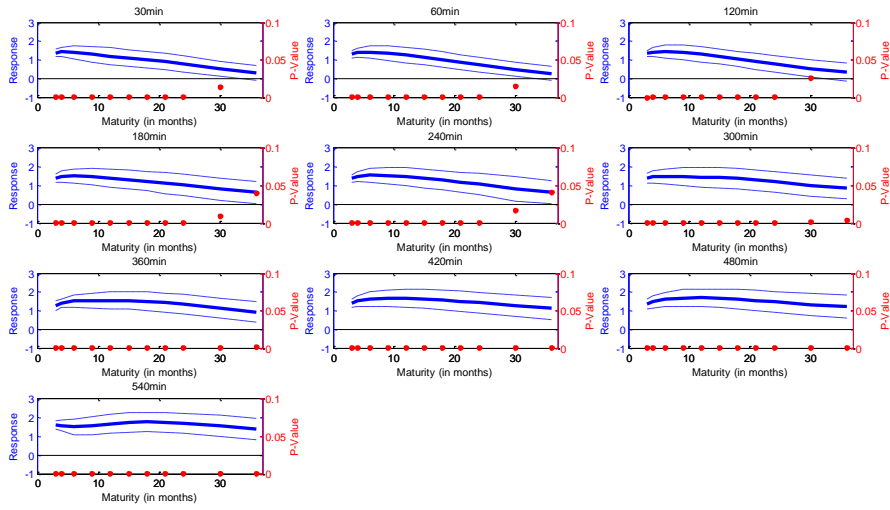


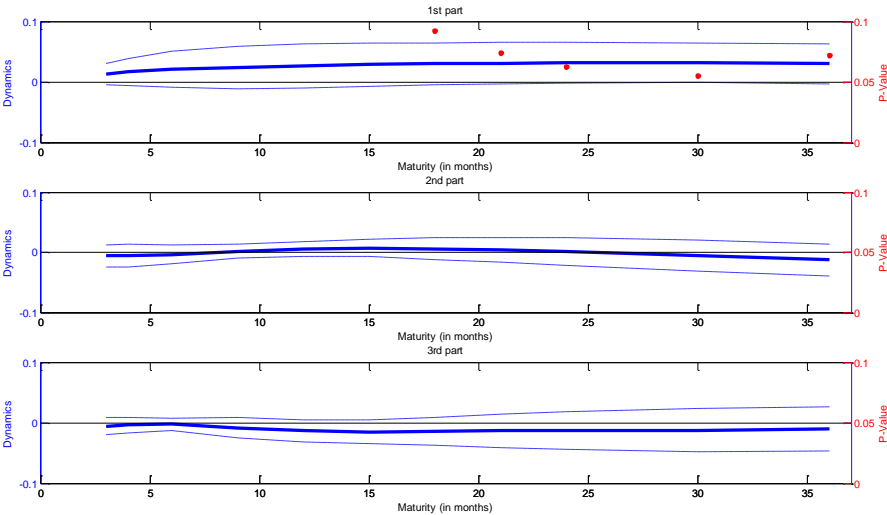
Figure 36 - Yield responses to the language for Tombini period



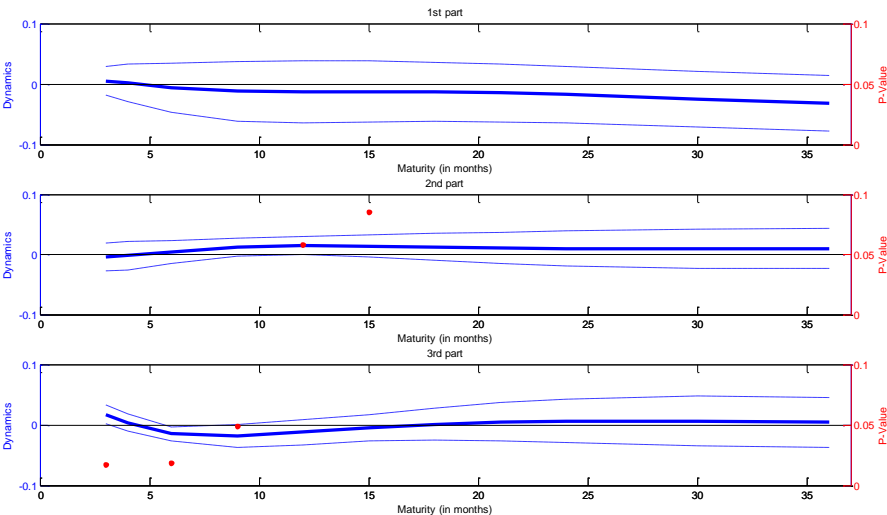
**Figure 37 - Yield responses to the interest rate surprise for pre-Tombini period**



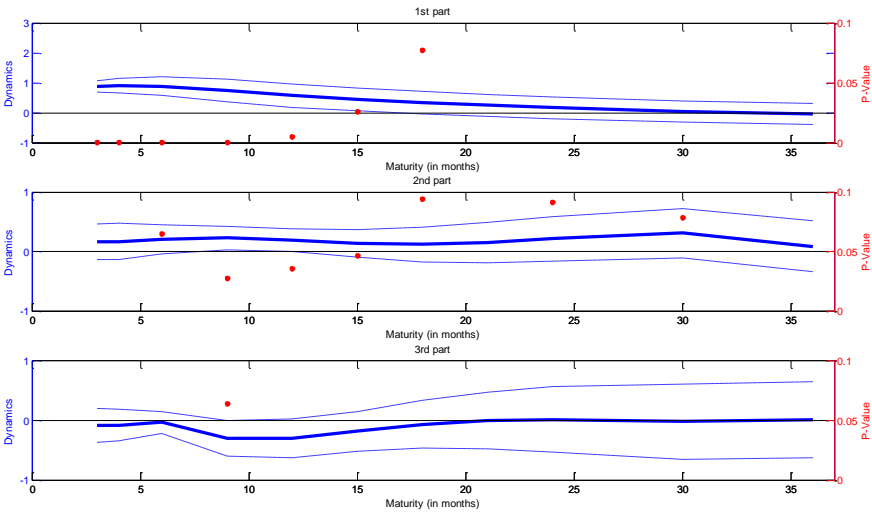
**Figure 38 - Yield responses to the interest rate surprise for Tombini period**



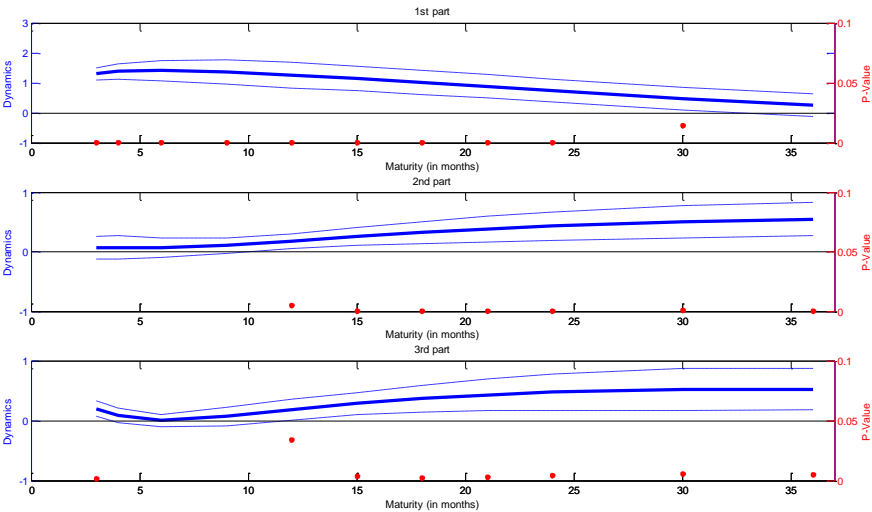
**Figure 39 - Intraday dynamics of the language for pre-Tombini period**



**Figure 40 - Intraday dynamics of the language for Tombini period**



**Figure 41 - Intraday dynamics of the interest rate surprise for pre-Tombini period**



**Figure 42 - Intraday dynamics of the interest rate surprise for Tombini period**

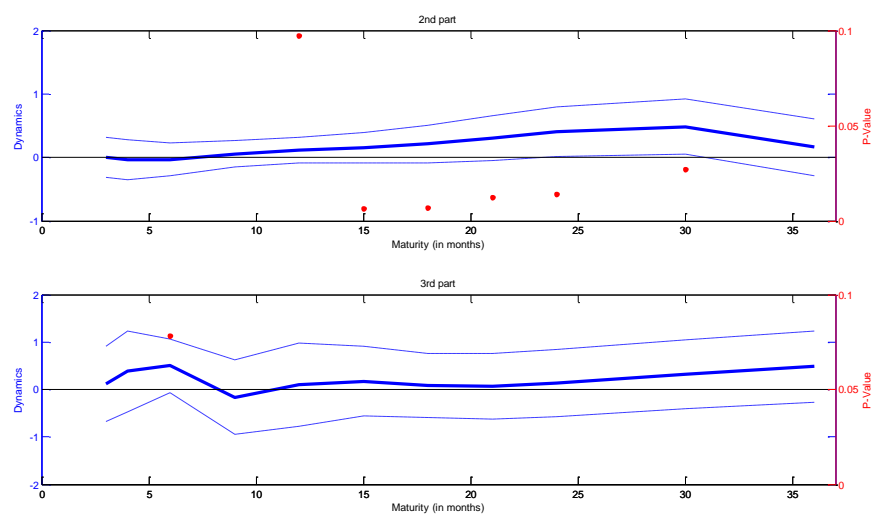


Figure 43 - Intraday dynamics of the residuals for pre-Tombini period

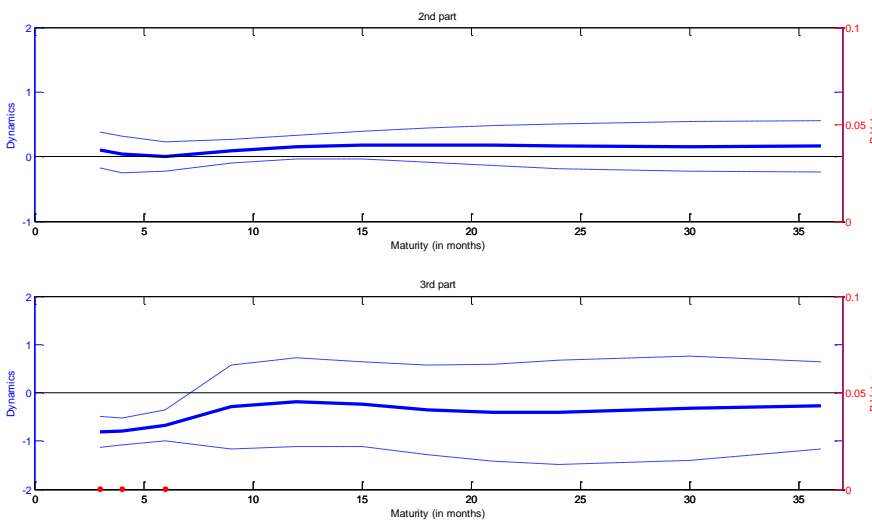


Figure 44 - Intraday dynamics of the residuals for Tombini period

6.2.3. Intraday Dynamics Specification (without the residuals)

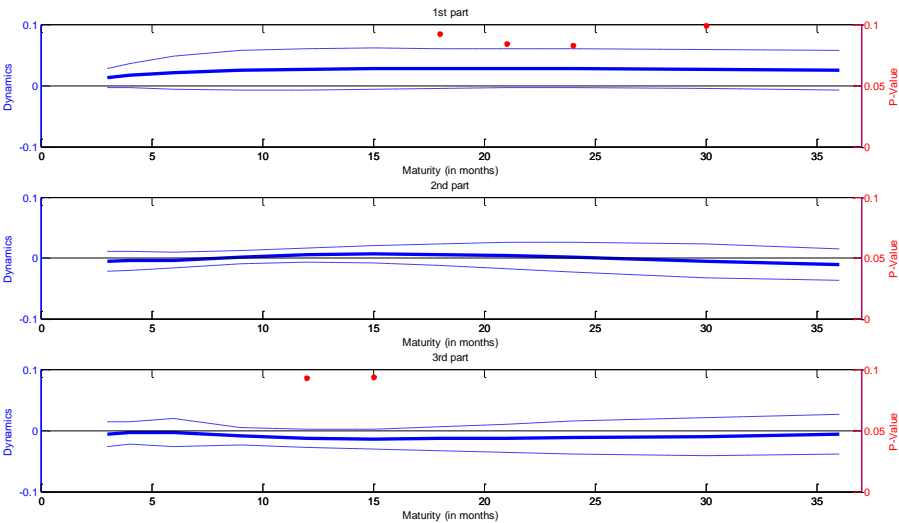


Figure 45 - Intraday dynamics of the language for pre-Tombini period

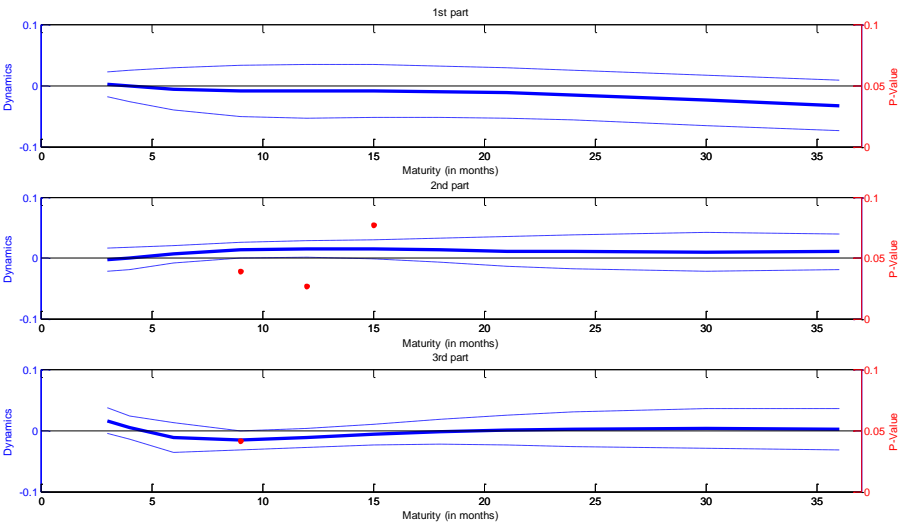
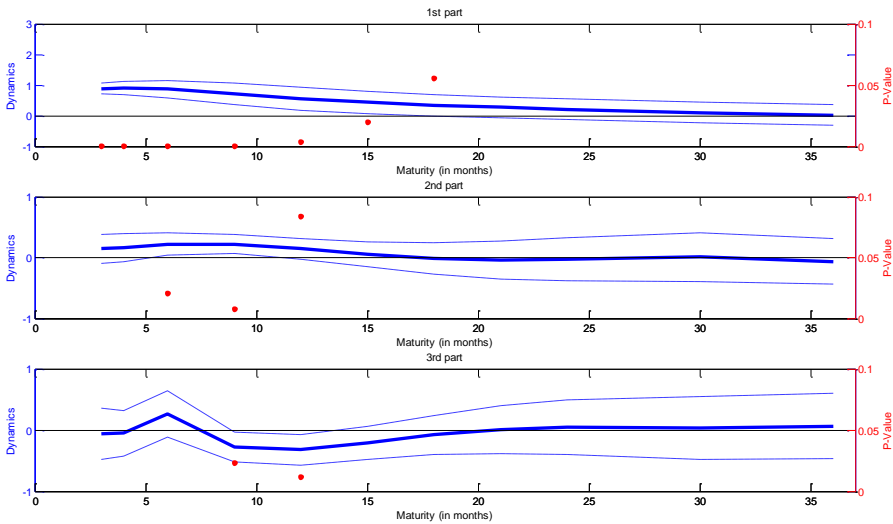
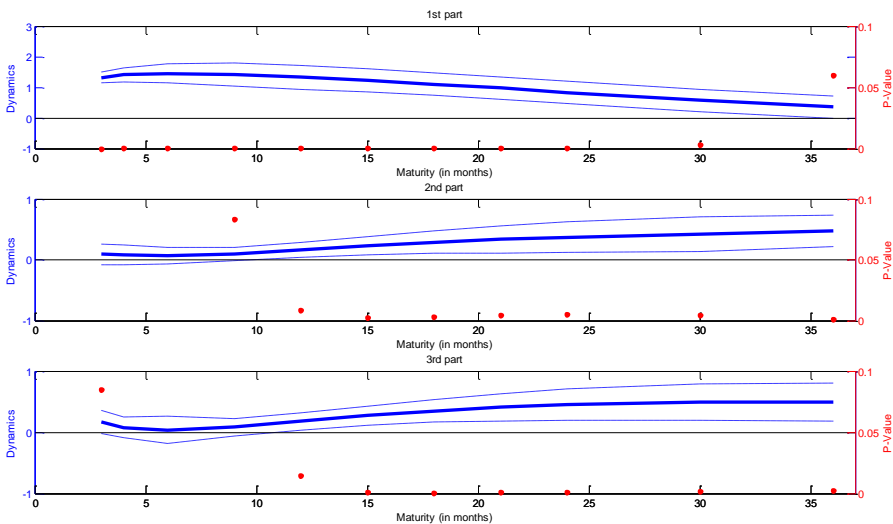


Figure 46 - Intraday dynamics of the language for Tombini period



**Figure 47 - Intraday dynamics of the interest rate surprise for pre-Tombini period**



**Figure 48 - Intraday dynamics of the interest rate surprise for Tombini period**