

# How do the Central Banks talk?: A Big Data approach to Turkey

## A Big Data Analysis for Turkish Central Bank Minutes & statements

Alvaro Ortiz / Tomasa Rodrigo / Joaquin Turina

October 17<sup>th</sup>, 2017

**Big Data analysis offers a huge potential for economic research. Natural Language Processing techniques allow us to read the policy documents of the Central Banks and to analyse the topics and sentiments inside. Previous literature focused on Developed countries. We extend the NLP analysis to the Central Banks of the Emerging Markets. This study sets out to close this gap and to go one step further in the computational approach to analyse the monetary policy evolution over time of the Central Bank of Turkey (CBRT). We show how the Monetary Policy complexity increased after the global financial crisis and how the different topics evolve over time. We also show that recent CBRT statements and minutes include an intensifying reference to inflation as well as a decisive commitment with a tight monetary policy stance to cope with it.**

## Sentiment Analysis and Central Bank communication<sup>1</sup>

Big Data allows us to measure and analyse text using natural language processing (NLP) techniques, also known as text mining or computational linguistics. The text information included in the media, blogs, economic and financial reports, etc. can be translated to indices and numbers, thus complementing the structured databases traditionally used in our economic research. In this document, we focus on monetary policy analysis to measure and monitor the Central Bank's communication.

Previous literature has used these techniques to analyse the Federal Reserve communication transparency strategy (Hansen et al, 2014) as well as the effects of this communication strategy on real economic variables (Hansen et al, 2015). From a methodological point of view, our main contribution to this research area is to extend the analysis of Central Bank communication for an emerging country like Turkey and to improve the computational approach to incorporate dynamics and topic evolution by applying not only Latent Dirichlet Allocation (LDA) algorithm, but also Dynamic Topic Models (DTM)<sup>2</sup>. The technical details are described in the appendix.

In this document, we analyse the minutes and statements published by the Central Bank of Turkey since March 2006 to the recent September 2017 releases. Through these monetary policy documents, we can extract the tone or sentiment on references included on the global conditions, national business cycle conditions, labour market, inflation and monetary policy in general. Beyond the information content, we provide a sentiment measure of the stance of monetary policy.

---

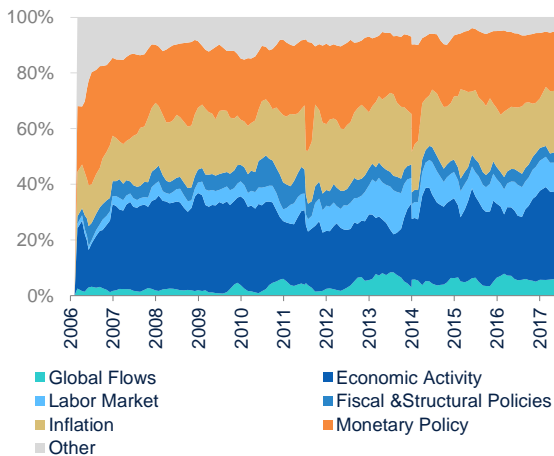
1: A more detailed analysis will be found in a forthcoming paper by Iglesias J. Ortiz A and Rodrigo T. How do the EM Central Banks talk?: a Big Data approach for Turkey.

2: Dynamic topic models are a particular case of Structural Topic Models (see Roberts et al, 2013).



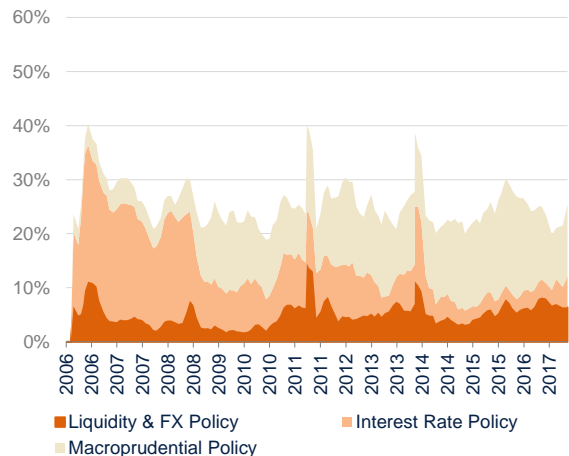
Once the groups of words have been identified, we can estimate the prevalence of these topics. Thus, we can analyse the evolution of the discussions on the minutes over time. The key results can be observed in the Figure 2 and 3:

**Figure 2** CBRT Topics evolution (in %)



Source: BBVA Research

**Figure 3** CBRT Monetary Policy topics evolution (in %)



Source: BBVA Research and CBRT

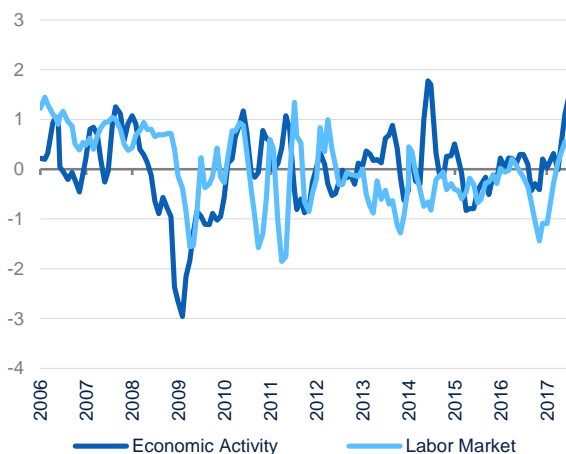
- On average (2007-17) the topics with greater influence (or weight) are *economic activity (26%) monetary policy (26%) and inflation (21%)*. Note that there is some margin for error as some of the topics are difficult to identify and we have decided to include them within the group other (10% average, but declining over time).
- The Economic activity references weight has been fluctuating. With an initially higher weight it shrank somehow from 2010 to 2014 (displaced by the Employment topic) to increase its importance after the 2016 slump in the economy. Employment has also been fluctuating, but its relevance has been increasing since 2012.
- Fiscal and structural policies normally have a low, but constant weight. This is not difficult to understand as the Central Banks make references to these policies in the sense that they can affect monetary policy.
- The importance of the monetary policy component has been navigating around 26% of the document. However, the nature of monetary policy has been changing over time. In Figure 3, we can appreciate that while liquidity and FX policies remain more or less constant during the first period, they increased marginally during the second half. However, there is a big transition between more traditional interest rate policies and macroprudential ones, with the former especially important during the first years (the disinflation period) and the Macro prudential policy gaining momentum since 2009 first and 2014. Here, maybe, it is difficult to disentangle the traditional interest rate policy

from the macroprudential since the implementation of the interest corridor, which has sometimes been associated as a macroprudential tool<sup>3</sup>.

- The inflation topic has been gaining momentum in the documents, especially after periods in which supply shocks rapidly passed through to consumer prices.
- Last but not least, although still with a low weight (4% average), the Global Capital Flows discussions have been gaining presence, thus reflecting the increasing importance of Global conditions and Capital flows since the financial crisis. This is particularly true in Post-Lehman years and after Bernanke’s tapering (6% average of the topic). This increase has been parallel to the macroprudential monetary policy increase.

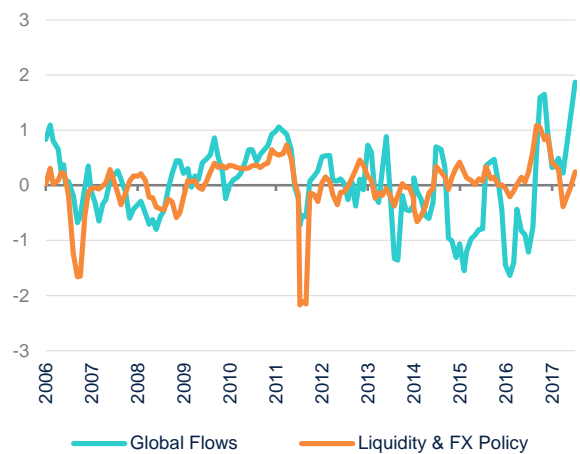
Beyond describing the influence on different topics, using the Lexicon approach we track the sentiment evolution in any of the vectors or group of words. Figures 4 and 5 are illustrative of the sentiment of some of the topics:

**Figure 4** CBRT Activity sentiment evolution



Source: BBVA Research

**Figure 5** CBRT Monetary Policy and flows sentiment



Source: BBVA Research and CBRT

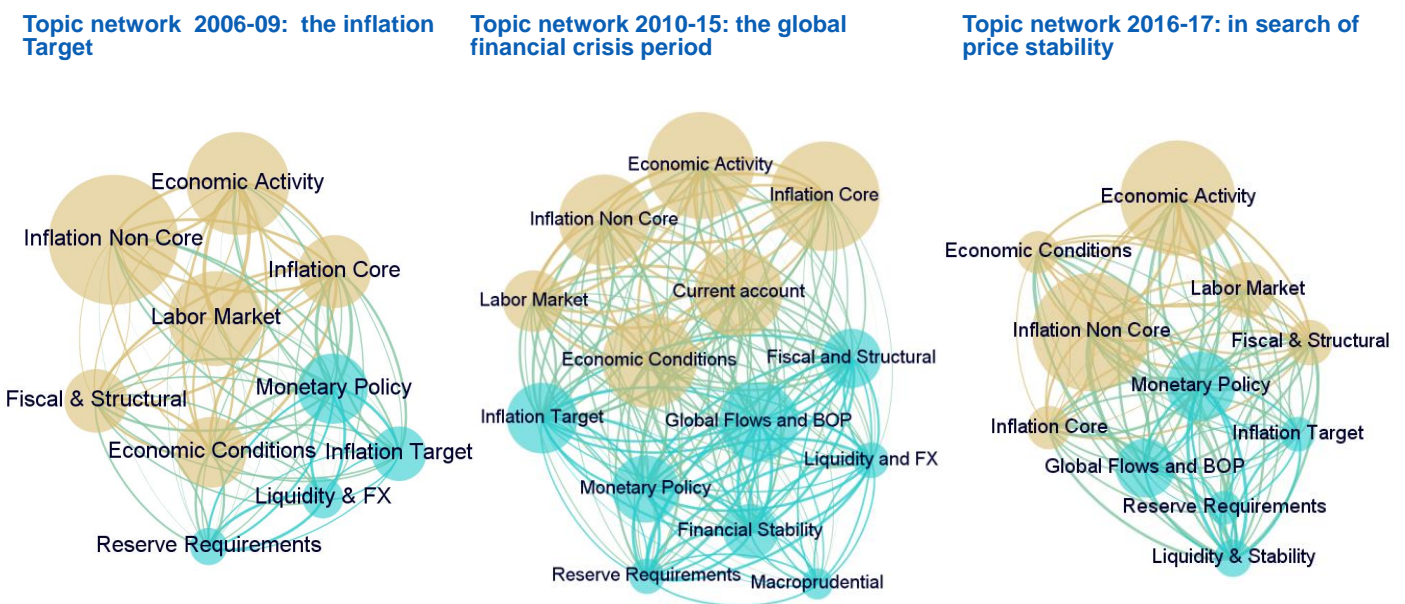
- The first important aspect is the divergence trend of economic activity and inflation. As depicted in Figure 4, while the economic conditions sentiment has been improving after the slump in the 3Q of 2016, the labour market sentiment performance remains lagging behind. This is particularly true at the end of the sample when the economic activity is jumping (in line with our monthly indicator of GDP growth, which has been growing by nearly 8% in 3Q) but the labour market, although improving, is still lagging.

<sup>3</sup> See Binici, M., Erol, H, Kara, Ozlu, P. & Unalmis, D (2013). Interest Rate Corridor: A New Macroprudential Tool? CBRT Research Notes in Economics

- With this analysis, we can assess the relationship between Global Capital flows and some specific monetary policies such as the ones we have identified as FX and Liquidity policies. Figure 5 shows how the CBRT has fine tuned this policy to temper capital flows of sharp FX movements. However, given the higher volatility of capital flows since the Bernanke’s tapering, the ability to stem flows has been challenged.

Using network analysis, we can also analyse the interrelation between the topics and how this relational structure can change over time. Figure 6 shows the networks between the main topics identified in the CBRT texts and how they are interconnected each other depending on the challenges that the monetary authority faces at every period of time:

**Figure 6** Network analysis: topic relationships and evolution over time



Source: BBVA Research

- The *full fledge inflation targeting period before the financial crisis (2006-09)* shows the centrality and relevance of both core and non core inflation nodes. These are clearly related to the economic conditions and labor markets as key determinants of inflation pressures. During this period, monetary policies are clearly aligned with the prominence of the standard monetary policy.
- *This Global Financial Crisis period (2010-15)* shows the increase of complexity in the monetary policy of the Turkish Central Banks and the Emerging Markets in general. As the recession hit the economy, the economic activity conditions increased their presence and inflation worries reduced its weight. External factors appeared in the network, especially the Global capital flows and current account balances worried. Monetary policy increased in complexity, and the traditional and standard monetary policy was divided between standard monetary policy,

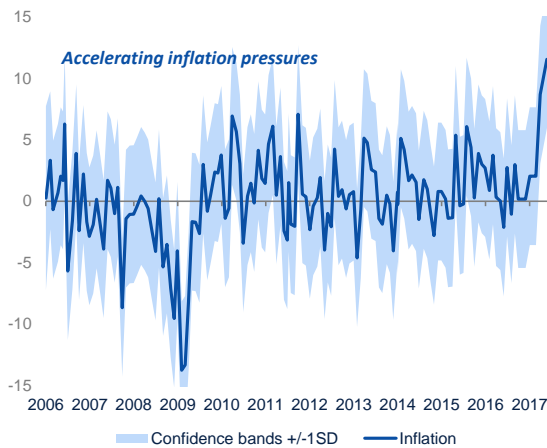
liquidity and exchange rate management and macroprudential policies. On top of this, financial stability topics gained momentum in the discussions.

- There is a *return to price stability discussions in the recent period (2016-2017)*. While economic activity and labor market remain in the network, inflation increased its presence in the policy texts. This time, the focus is on non core inflation woes, which is consistent with cost push supply shocks and structural problems to cope with food price volatility. Policies are somehow more normalized and aligned than during the crisis period, but global capital flows remained present as the FED tapering is also generating capital flows volatility.

### What is the current stance of the Central Bank of Turkey?

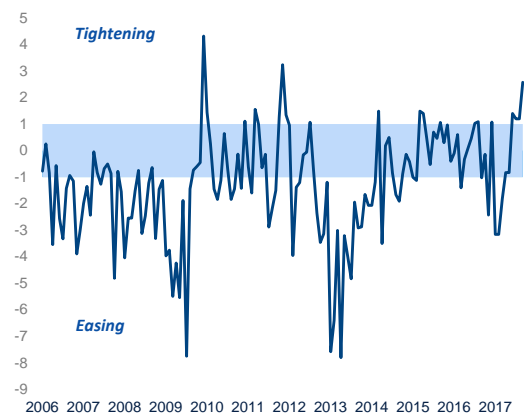
A final and straightforward use of NLP techniques has been employed to assess the monetary policy stance. In this sense, using the lexicon approach, we analyse the tone or sentiment of the latent topics, which should be an adequate tool to understand the stance of monetary policy included in the texts. The next two examples (Figure 7 and 8) can help to understand the current stance of monetary policy:

**Figure 7** CBRT inflation sentiment



Source: BBVA Research

**Figure 8** CBRT monetary policy sentiment



Source: BBVA Research

- Figure 7 shows the sentiment about inflation in the CBRT texts. The stance is normalized so after the post-Lehman deflationary period the inflation sentiment remains above average (this is consistent with the average high inflation during the period and difficulties to achieve the inflation target since 2011). Beyond this, the most salient feature of the graph is the recent spike in inflation sentiments, with no precedents during the period 2006-17. Thus, the CBRT texts show the worries regarding current inflation momentum and that the CBRT is very aware of the situation.

- As a consequence, the stance of monetary policy is really tight according to the sentiment analysis and it constitutes a reversion of the policy implemented during most of 2016. It has no precedents (only just after the international financial crisis to avoid the second round effects of the sharp depreciations of Emerging Market currencies) and is very unusual in the Post-Lehman period. In short, monetary policy remains about being clearly above neutral levels in a tightening mode.

## Conclusion

We have applied Big Data and Data Science techniques to the analysis of the monetary policy texts of the Central Bank of Turkey (CBRT). We found that the NLP techniques are an adequate tool to the analysis of monetary policy in Turkey since they help us to understand the main challenges and the evolution of Turkish Monetary Policy. In this sense, these new methods should be considered as a relevant tool to complement the traditional ones.

There are a couple of significant insights from the policy texts. First, topics under discussion have been changing over time as the global economic conditions have changed substantially. The traditional discussions on inflation were replaced by economic conditions and capital flows volatility during the crisis, but they are now returning to the fore. Second, the monetary policy strategy has increased complexity as the financial crisis amplified capital flows volatility; and the standard monetary policy has to be complemented with macroprudential and liquidity policies to deal with inflation and financial stability simultaneously.

The analysis shows some insights on actual juncture. The prominence on inflation has increased in both core and non core inflation. In fact, the inflation topic is increasingly important again. Besides, the commitment to a tight monetary policy stance to deal with this is now at maximum levels.

## Appendix: Methodology

We analyse Turkish Central Bank wording on monetary policy using communication reports (statements and minutes). Documents in the analysis are defined as paragraphs of these communication reports. To clean the text, documents with less than 200 characters are excluded (i.e. titles, contents sections or other documents which have little economic content) from the corpus (the structured set of texts to be analysed). Then, words are stemmed (the reduction of words to their semantic root) to generate tokens. Common stop words and words with a length of 3 or less are removed and the remaining words are stemmed. Very common or uncommon words are removed, thereby filtering out words with low semantic content.

Using matching learning algorithms, we use computational linguistics and different specifications of topic modelling. Latent Dirichlet Allocation (LDA) (Blei, Ng, and Jordan 2003) is a Bayesian model with a prior distribution on the document-specific mixing probabilities where the count of terms within documents are independent and identically distributed given a Dirichlet prior distribution. To introduce time-series dependencies into the data generating process, we use the dynamic topic model (DTM), a particularization of the Structural Topic Models (STM) where each time period has a separate topic model and time periods are linked via smoothly evolving parameters. STM (Roberts et. al. 2016) explicitly introduces covariates into a topic model allowing us to estimate the impact of document-level covariates on topic content and prevalence as part of the topic model itself. STM models text as a set of documents (each document being a list of words) being generated by a mixture of a (user defined) number of latent topics (each topic being a distribution over the whole collection of words in the text) drawn from a Multinomial distribution.

To measure the tone or sentiment, we rely on Lexicon methods using the Loughran-McDonald dictionary (Loughran McDonald 2011). This dictionary was created specifically to analyse financial texts, and solved a misclassification issue of certain specific financial or economic words in standard sentiment analysis dictionaries. We also apply the FED dictionary for financial stability (Correa et al, 2017). Using the negative and positive words of this dictionary, the average “tone” of a given document is computed by:

$$\text{Average tone} = 100 * \frac{\sum \text{Positive words} - \sum \text{Negative words}}{\sum \text{Total words}}$$

The score ranges from -100 (extremely negative) to +100 (extremely positive) but common values range between -10 and +10, with 0 indicating neutral. A neutral sentiment can be the result of neutral language or a balancing of some extreme positive sentiments compensated by negative ones. Once negative and positive words are identified with each dictionary, we construct a tone variable based on the balance between the number of positive and negative words that appear in a given document divided by the total number of words included in the document.

Finally, to build the sentiment indices, we use the topic mixture that combines dictionary methods with the output of LDA to weight words counts by topic, following the approach proposed by Hansen and McMahon (2015). This allows different sentiment measures to be generated from a set of text, and that sentiment to be focused on the topics of interest.



## References

- Blei, D.M. and Ng, A.Y., Jordan, M.I. and Lafferty, J. (2003) Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3.
- Binici, M., Erol, H, Kara, Ozlu, P. & Unalmis , D (2013). Interest Rate Corridor: A New Macroprudential Tool? CBRT Research Notes in Economics.
- Blondel V. D., Guillaume J.-L., Lambiotte R., Lefebvre E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech. Theory E* 10, P10008.10.1088/1742-5468/2008/10/P10008.
- Correa, Ricardo, Keshav Garud, Juan M. Londono, and Nathan Mislant (2017). Sentiment in Central Banks' Financial Stability Reports. *International Finance Discussion Papers* 1203.
- Hansen, S, McMahon, M and Prat, A (2014), 'Transparency and Deliberation within the FOMC: a Computational Linguistics Approach', CEP Discussion Papers DP1276, Centre for Economic Performance, LSE.
- Hansen, S. and M. McMahon (2015). Shocking Language: Understanding the macroeconomic effects of central bank communication. *Discussion Papers* 1537, Centre for Macroeconomics (CFM).
- Hansen S. and McMahon M., (2016), Shocking language: Understanding the macroeconomic effects of central bank communication, *Journal of International Economics*, 99, (S1), S114-S133.
- Iglesias J, Ortiz A and Rodrigo T (Forthcoming). How the Turkish Central Bank talks?: A Big Data Approach, work in progress.
- Loughran, T., McDonald, B., 2011. When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *Journal of Finance* 66, 35-65.
- Roberts M.E., B. M. Stewart, and E. M. Airoldi. A model of text for experimentation in the social sciences. *Journal of the American Statistical Association*, 111(515): 988-1003, 2016.

## **DISCLAIMER**

This document has been prepared by the BBVA Research Department and is provided for information purposes only and expresses data, opinions or estimations regarding the date of issue of the report, prepared by BBVA or obtained from or based on sources we consider to be reliable, and have not been independently verified by BBVA. Therefore, BBVA offers no warranty, either express or implicit, regarding its accuracy, integrity or correctness.

Estimations this document may contain have been undertaken according to generally accepted methodologies and should be considered as forecasts or projections. Results obtained in the past, either positive or negative, are no guarantee of future performance.

This document and its contents are subject to changes without prior notice depending on variables such as the economic context or market fluctuations. BBVA is not responsible for updating these contents or for giving notice of such changes.

BBVA accepts no liability for any loss, direct or indirect, that may result from the use of this document or its contents.

This document and its contents do not constitute an offer, invitation or solicitation to purchase, divest or enter into any interest in financial assets or instruments. Neither shall this document nor its contents form the basis of any contract, commitment or decision of any kind.

In regard to investment in financial assets related to economic variables this document may cover, readers should be aware that under no circumstances should they base their investment decisions in the information contained in this document. Those persons or entities offering investment products to these potential investors are legally required to provide the information needed for them to take an appropriate investment decision.

The content of this document is protected by intellectual property laws. Its reproduction, transformation, distribution, public communication, making available, extraction, reuse, forwarding or use of any nature by any means or process is forbidden except in cases where it is legally permitted or expressly authorized by BBVA.