Does the use of big data variables improve the prediction of Monetary Policy decisions? The case of Mexico

Luis Alberto Delgado de la Garza 192414
Gonzalo Adolfo Garza Rodríguez 352119
Daniel Alejandro Jacques Osuna 523025
Alejandro Múgica Lara 345409

Director: Carlos Alberto Carrasco Sánchez

Department of Economics, Universidad de Monterrey

Abstract

We analyzed the predictive power of a market attention variable, generated using big data, for Banco de Mexico's (Mexican central bank, hereby "Banxico") monetary policy decisions. The novelty of this paper relies on the lack of previous research that incorporates a non-conventional variable that uses big data analysis in monetary policy research. We used a binary probit approach and contrasted different models to identify whether the proposed variable improved the prediction. Our general results show there is significant evidence that the variable improves the prediction, as it helps reduce information criteria and it stays significant across the different models. We consider that further research is necessary to determine the scope of big data in monetary policy analysis prediction.

Keywords: Monetary policy decisions, big data, Banxico, interest rates
JEL Classifications: C55, E47, E52, E58

1. Introduction

Traditionally, one of the most important topics for economists has been to understand how monetary policy is transmitted to the economy in terms of financial and economic indicators. This has been examined extensively in papers for the case of United States and Europe, but using conventional variables, such as economic activity, inflation and the exchange rate. However, there have not been attempts to use a big data non-conventional variable to approximate Banxico's monetary policy decisions.

We analyze the usefulness of a market attention variable, constructed using big data, for predicting Banxico's monetary policy decisions. Our approach is a pioneer in the use of big data in monetary policy for the case of Mexico.

Economic agents are significantly affected by monetary policy decisions on the central bank's instrumental interest rate. Thus, accurately predicting movements over the interest rate would allow these agents to incorporate this information into their decision-making process in advance. Undoubtedly, this task poses significant complications as there is uncertainty on both the determinants of monetary policy and the soft-sided considerations of policy makers. However, this has not stopped researchers from proposing and estimating predictive models, ranging from the use of purely economic and financial variables as predictors to modern papers like De Hann (2011) and Pereira (2018) incorporating communication and sentiment variables. For the Mexican economy, foreign economic decisions are particularly significant as the economic activity is heavily influenced by the country's trade and closeness with the United States, as has been shown by numerous research such as Moreno et. al (2004), that analyze how the Mexican economy has become more dependent on US economic conditions over the years.

Having established the value of a predictive model for monetary policy decisions, this paper attempts to broaden the type of variables used for this purpose. Since the invention of the Internet, the quantity and depth of information available for the general public has reached new frontiers. Specifically, we are now able to capture the behavior of the masses through tools such as Google Trend, which generates indexes that contain massive searches of specific terms. Through this, we can now use variables that model individuals or markets attention for economic research purposes through the identification of economic terms that reflect economic agents' sentiments. Using big data, this paper aims to propose a new way of capturing the markets attention that can better predict changes in Banxico's interest rate of reference. Thus, besides this paper's obtained results on the analyzed variables, our main contribution is the introduction of a big data approach to measure market attention for predicting monetary policy decisions.

The paper is organized as follows. Section 2 presents a brief literature review on monetary policy prediction and big data. Section 3 defines the methodology, the model and the construction of the big data variable. In Section 4 we present the empirical results and discuss robustness. Finally, Section 5 concludes.

2. Literature Review on monetary policy prediction and the use of big data

Monetary policy prediction on economic literature

Central bank decisions on monetary policy have long been a subject of interest for researchers and policy makers. One of the pioneer papers regarding this topic was Taylor's (1993) "Discretion versus policy rules in practice," which proposes that algebraic formulations should not be followed mechanically by central banks, but rather rule-like behaviors that give certainty to economic agents. Furthermore, theoretical papers such as the widely cited paper from Clarida et al (1999) "The New Keynesian Science of Monetary Policy", developed a baseline framework that has guided the modern study of the decision-making processes of central banks. The paper proposes a model that incorporates three major equations of macroeconomics: the IS curve (IS), the Phillips curve (PC) and a monetary rule for interest rate (MR). This model shows how interest rates defined by the central bank affects the real economy and the money market.

Derived from the importance of interest rates in the economy, understanding and predicting interest rates has been part of the agenda of economic agents. Enders and Granger (1998) developed pioneering research about the dynamics of interest rates and how its term structure interacts in the short and long-term equilibriums.

On the predictive side of the research, many different econometric methods have been used to forecast monetary policy decisions. McMillan (2009) compared the use of both linear and non-linear forecasting models for interest rates in the US and Australia and reported significant evidence of non-linear behavior of interest rates in recent years. Baghestani and Danila (2014) compared analysts' forecast for next month interest rates, published by the Czech Republic National Bank, to a random walk benchmark and found that the analysts' forecasts have a higher directional accuracy. Nyberg (2017) forecasted interest rates of federal funds using the business cycle in a QR-VAR model and found positive evidence of an improved forecast compared to a regular non-adjusted VAR.

In addition, recent papers have analyzed the usefulness of incorporating a variable that reflects the communication efforts of central banks into conventional predictive models for interest rates. Jansen and De Hann (2011) compared the performance in predicting changes in ECB interest rates for both a conventional ordered probit based on the Taylor rule and one that incorporated a communication variable, finding that the statements of high-level policy makers in the European Central Bank (ECB) have not helped the market predict changes in interest rates. Pereira (2018) studied the impact of the Central Bank of Brazil's press releases in the interest rate curve and found evidence that the volatility of the interest rate curve was higher on days when new publications were uploaded to the website of the central bank. Hayo and Neuenkirch (2009) also employed an ordered probit approach to compare a conventional model based on the Taylor's rule to one that incorporated a communication variable, finding that the latter variable improved the predictive power of the model.

For Mexico, some work has already been done around this topic. Tellez and Venegas (2013) used ordered probits to identify the main determinants in the monetary policy decisions of Banxico, finding that general macroeconomic variables (such as production, inflation, market interest rates, inflation expectations) were useful for the prediction of changes in the target interest rate. Prior to this, Cuevas (2003) used a binary probit to analyze the behavior of contractive monetary policy in Mexico and found that inflation and exchange rate pressures were useful to predict the appearance of a contractive policy. Elizondo (2017) showed that the use of affine models could help improve the term structure of interest rates in Mexico compared to AR (1), VAR(1) and random walk models.

As described in the papers above, most predictive models for interest rate in the literature incorporate economic and communication variables in a time series framework. Leveraging the expansion of information availability, there are new frontiers to be explored, both in terms of methodology and non-conventional variables for predictive models.

Nowadays, financial variables (such as the differential between long-term bonds and inflation-indexed bonds) and polls are used as proxies for inflation expectations (Aguilar et al, 2016). However, with the exponential rise of technological resources in the last decade, there is a new methodology to explore that can predict monetary policy in a potentially more accurate manner: big data.

Big data can be defined as the large and diverse data generated from economic transactions and social media interactions (Armah, 2013). It has 4 main characteristics: volume, variety, velocity and value. Volume is described as the element that defines big data. It states that for information to be considered as big data, it needs to be significantly larger than traditional data sets. Variety refers to the fact that big data mainly englobes information that can not be stored in a typical structured form. This information presents itself in the form of emails, social media posts, audiovisual data and road traffic information, among others, amounting to approximately 90% of it. Velocity references to the update capacity of the data sources that contain big data, mainly the Internet and government and private enterprises' data systems. However, value is the most important element of big data. This is due to its essence, where it is not an outcome but rather acts as a means for the creation of knowledge. Such is the case for this paper, where big data will be used as a means of predicting monetary policy, specifically interest rate changes.

Big data has had clear, important and significant results in fields such as meteorology, biology, physics, chemistry, and astronomy (Schintler and Kulkarni, 2014). However, the use of big data beyond these areas has been limited, mainly due to the controlled access to technological developments. Yet, the unprecedented amount of massive, detailed data that has been spreading outside of the traditional disciplines has allowed this to change.

The novelty topic that is big data implies that there is limited research information available, especially for a specific topic such as monetary policy. Even then, the papers that use big data as part of their methodology are quite recent and are not referenced enough. One such case is Silverstovs and Wochner (2018), that attempt to find how reliable is Google Trends' behavior to reflect real economic conditions in Switzerland. Through this, they conclude that Google Trends can be treated as an effective tool to guide economic policy decisions.

One of the pioneers in applying a Monetary Policy Attention Index using Google Trends is Wohlfarth (2018). With the objective of analyzing the transmission and spillover mechanisms of monetary policy in the United States and Europe, he applies the obtained index to fixed income data. As a result, he found that the impact of monetary policy provides evidence for an international channel of monetary transmission on capital and money markets.

Lucca and Trebbi (2009) are one of the precursors of the use of a search index to measure changes in monetary policy. Using statements released by the Federal Open Market Committee, they seek to measure the Federal Reserve's future interest rate decisions based on Google specific word searches. As a result of this, they find that yields in long-term bonds respond better to changes in communication than yields in short-term bonds.

In this regard, Choi and Varian (2009) are one of the first authors that apply Google Trends specifically in their monetary policy investigation. Using auto-regressive models,they attempt to observe if Google Trends helps improve the prediction of future unemployment levels, measured through an initial unemployment claims proxy variable. They conclude that there is a significant positive correlation between Google Trends searches for jobs and welfare related terms and unemployment. As this is a major macroeconomic indicator, this could be taken as a starting point to predict future interest rates, more so considering Google Trends responds favorably to inflection points (i.e. Banxico's board meetings).

In the specific case of big data being implemented in Mexico, Durán, Hernandez and Ortiz (2018) try to demonstrate the advantages of using Google Trends to predict Mexican peso- U.S. dollar exchange rate volatility in the short-term. However, they concluded that Google Trends only partially explain the behavior of volatility as it does not capture all financial decisions taken in the currency markets.

These papers show the usefulness of big data variables, such as Google Search indexes, in modern research. However, due to the recent availability of this data sets, few economic studies have incorporated them. Our research resembles the investigation of Duran, Hernandez and Ortiz (2018), as we analyze the usefulness of the results of Google Search indexes on improving the prediction of an economic variable. Specifically, we analyzed whether the accuracy of a predictive model for the Mexican interest rate that uses conventional variables is improved by adding a Google trend index composed of economic key terms extracted from the minutes of Banxico's meetings.

3. Methodology

For the purpose of this paper, we have decided to employ the use of probit models, as authors such as Téllez & Venegas (2013), Jansen & De Haan (2011) and Jung (2016) have done when predicting monetary policy decisions made by a central bank. Our objective was determining whether the inclusion of a non-conventional variable that measured the population's interest over time on the current macroeconomic context yielded better results in a monetary-policy-predicting probit model. As this paper is a pioneer in monetary policy prediction with big data, there is no literature that suggests how to make this comparison. For this reason, we opted to estimate and assess eight different probit models for predicting Banxico's monetary policy decisions.

3.1. Functional form of probit models

Before detailing each model, it is convenient to first define the functional form and uses of a probit model. This model is useful when the dependent variable is of binary nature, and the researcher's interest lies on determining how unit changes in the explanatory variables affect the probability of one of the two events of the dependent variable happening. However, before defining a probit model, we must establish the concept of a linear probability model. According to Brooks (2014), the model assumes a linear relation between the probability of a binary event occurring and an n number of explanatory variables of the form:

$$P(Y = 1 \mid x_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n + u_i$$

which can be estimated with least squares.

Nevertheless, as this model assumes there is a linear relation between the probability of occurrence and the independent variables, the obtained result may fall outside of the logical 0 to 1 probability spectrum. To overcome this limitation, the probit model transforms the regression using a cumulative normal distribution:

$$P(Y = 1 \mid z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z_i} e^{\frac{-(z_i)^2}{2}} dz$$

Where $z_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots \beta_n x_n + u_i$. This transformation produces probabilities of occurrence that lie between the (0,1) interval. In the probit model, the marginal impact of a unit change in each independent variable x_i is given by $\beta_i P(Y = 1 \mid z_i)$.

3.2. Definition of dependent conventional explanatory variables

Having defined the functional form of probit models, now we define the dependent and independent variables used for each of the eight models we employed, as well as explain how these variables were measured and their expected behavior within the models. It is important note that all observations of each variable are measured monthly, with the time period spanning February of 2008 to December of 2018, for a total of 130 observations. Also, when the variables are implemented in the model, we use one lag for all explanatory variables, as we argue that the prediction of the current month's monetary policy decision depends on the most recent information of all variables.

Starting with the dependent variable, this is one of binary nature that takes the value of 1 when Banxico raises the interest rate of reference, and 0 when it does not. In the case of the explanatory variables for the baseline models, these are a set of macroeconomic variables that are conventional among monetary-policy-predicting literature, and the added value of this paper: a non-conventional variable that captures the populations' interest about the current state of the economy.

The first of the macroeconomic variables of the Mexican economy is inflation. This variable is measured with monthly interannual inflation rates of Mexico. Laopodis (2006) establishes that, if there is a favorable interest rate climate, inflation will grow as asset prices do as well, thus cueing the central bank to contain inflationary pressures by raising the reference interest rate. Because of this, we expected to find that inflation had a positive correlation with Banxico's interest rate of reference. This would imply that the central bank has incentives to raise the interest rate to avoid inflation from spiraling out of control.

The second macroeconomic variable is inflation expectations, measured with the median of Banxico's monthly survey to the private sector. Barro and Gordon (1983) contributed with a theoretical paper describing how the market's expectations for future inflation affect the decisions of the central bank, as credibility plays an important role in the transmission mechanisms. Considering this, we expected that inflation expectations were positively correlated with a rise in interest rates, suggesting that when expectations rise, the central bank's commitment to control current inflation is reflected by rising interest rates.

We included the output gap as the third macroeconomic variable. To obtain data for measuring it we used the Hodrick-Prescott filter to analyze the Mexican economic activity index's trend and cycle. This economic activity index (IGAE) is measured monthly by the National Institute of Statistics and Geography (INEGI). As described by Maravall (2001) using the Hodrick-Prescott filter allows for decomposing the trend and cycle of the GDP, and so it can be possible to analyze the output gap. Avdjiev and Zeng (2014) find that the effect of monetary policy shocks is strongest when economic activity is slow. This implies that monetary policy reacts more aggressively to economic shocks when the economy is underperforming. Thus, as an aggressive expansive monetary policy suggests that the central bank wants to incentivize economic growth, it can be expected that Banxico lowers its interest rate of reference, consequently having a positive correlation with economic activity.

The fourth macroeconomic variable included in our models is the nominal USD/MXN exchange rate. Bjornland and Halvorsen (2014) observe a strong interaction between monetary policy and exchange rate variation. They find that shocks that cause the exchange rate to depreciate increase the interest rate in an immediate effect. Due to this, we would expect to find that exchange rate depreciations are positively related with Banxico's interest rate of reference.

The influence of the United States' monetary policy on Mexican monetary policy is the next macroeconomic variable considered. This influence is measured by the differential of the interest rate of reference between the American and the Mexican economies. Crespo et al. (2016) attempt to find the spillover effects of the United States´ monetary policy over various countries, one of them being Mexico, from 1979 to 2013. They find that Mexico's monetary policy tends to show stronger responses throughout the analyzed period to unexpected rises in the United States' reference interest rate. They argue that this is due to the strong trade links between both countries. Thus, given the positive correlation between interest rates in both countries, we would expect that, as the interest rate differential rises (implying that United States rose their rates), Banxico would also raise their instrumental interest rate.

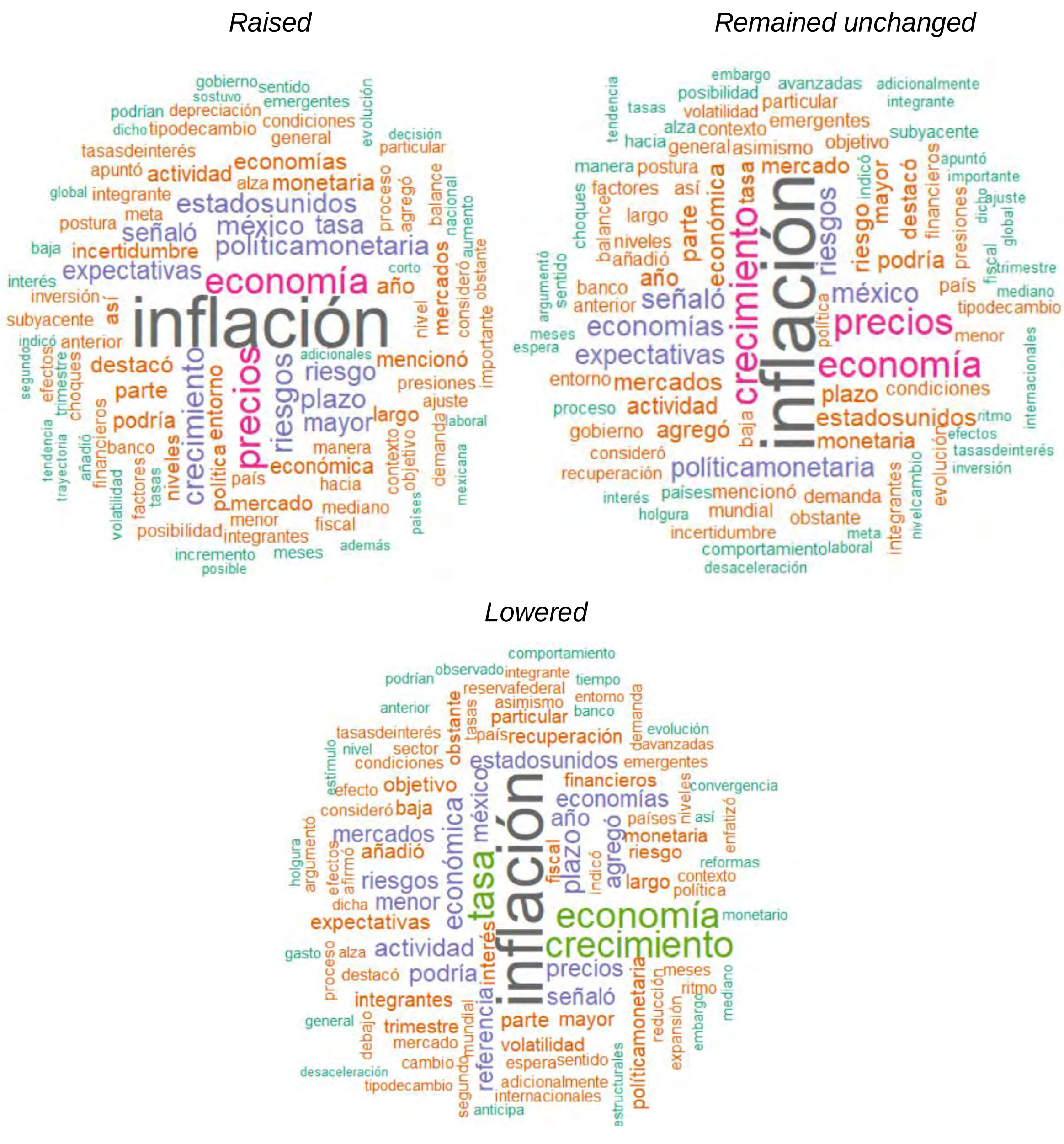
The sixth and final macroeconomic variable considered for the models is the expectation sovereign debt market participants have over future macroeconomic conditions (referred to as Market Expectations in the models). This is measured with the slope of the Mexican yield curve, as it has been used for its empirical predictive power when analyzing economic variables. Estrella and Hardouvelis (1991) found robust evidence that the slope of the yield curve performs even better than a group of leading economic indicators when forecasting GDP, while Wright (2008) found that uncertainty about future inflation substantially explains the positive slope of the yield curve. Considering the findings from previous research, we would expect that an increase in the slope of the yield curve is correlated with increases in the reference interest rates. For this variable we used the differential of 1 month and 1-year CETES (the Mexican Treasury Bonds).

3.3. Non-conventional explanatory variable, text mining and principal component analysis

For the non-conventional variable, we focused on the Mexican populations' interest about the current state of the economy (referred to as Population's Interest in the models). Since this work pioneers the use of big data in the context of monetary policy, we needed to come up with an approach to measure such variable. To do so we used Google Trends indexes, which is a Google tool that creates an index of the behavior of the monthly total searches of a specific term relative to the monthly total overall searches. We also considered using Twitter Crowds and Wikipedia visits, where the problem with the former was the cost of information and with the latter the lack of data before 2015. Because of the innovative nature of this paper, we were unable to find literature on the expected relationship between the population's interest in macroeconomic context and monetary policy decisions using big data. Therefore, we hypothesized that when the population's interest of the economy is high, it is because the current state of the economy gives rise to uncertainty and is therefore a signal of general preoccupation. It is then that the central bank intervenes to raise interest rates and ease this preoccupation

At this point, the next step was to select the key terms we were going to lift the Google Trends Index for. To select these terms, a look at the literature shows that authors such as Chague et al. (2015) and Luca & Trebbi (2009) employ a methodology of revising the central banks' public statements for the most-used words and determine their overall impact on monetary policy decisions. For this reason, we made a quantitative analysis of the content of the Banco de México's meetings minutes focusing on the repetition of economic terms. This was done through a process known as text mining. Berry (2004) defines text mining as the process of efficiently organizing, classifying and extracting relevant information from textual sources using software and/or algorithms. We gathered all the minutes available on Banxico's public website and divided them into three separate documents: minutes of meetings where interest rates were raised, minutes of meetings where interest rates were lowered, and minutes of meetings where interest rates remained unchanged. Then, we analyzed the frequency of the most-used terms across all three types of documents. By doing this, we obtained the following results presented in Figure 1, where the top left word cloud represents the cases when the interest rate was raised, the top right when the interest rate remained unchanged and the bottom when the interest rate was lowered.

Figure 1. Cloud of words[1]



Our objective here was to construct an index that measured the Mexican population's' interest over the current macroeconomic context using Google Trends indexes

[1]Cases when the interest rate was raised, remained unchanged and was lowered

and define a base for which to select those terms. In order to do so, we estimate a principal component analysis, as Chague et al. (2015), to create a single variable of central bank communication in their paper. Abdi & Williams (2010) point out that the goal of principal component analysis is to extract the important information from several variables that are inter-correlated and express it in a “set of new orthogonal variables called principal components”. This methodology also orders the new variables by measuring how well they explain the variance of the original information. Our proposal is to use the first component of the Google Search indexes previously mentioned as a covariate in our analysis of Banco de México’s reference rate.

The next step was to choose which terms to include in the principal component analysis in the case of Mexico. As our paper is the first of its kind to combine monetary policy prediction with big data, we opted to employ nine different methodologies to construct the population’s interest variable. These methodologies consider the most frequently used terms obtained from the text mining algorithm, as well as the top related queries (given by Google Trends itself) of these frequently used terms. The nine methodologies are letter-coded by two characters (see Figure 2). Both characters can only be letters from A to C. The first character represents the number of Banxico terms, where A is including the top 25% of the most repeated terms, B is including the top 5 most frequent terms, and C is the same as B, but with the top 3 terms. On the other hand, the second character represents the number of related queries included, where A represents no related queries included, B represents the top 5 related queries included, and C represents the top 3 related queries. This information is summarized in the matrix presented in Figure 2.

Figure 2. Methodology Matrix

		Related Queries		
		None	**5**	**3**
Bank of Mexico's Terms	**Top 25%**	AA (19 terms)	AB (82 terms)	AC (59 terms)
	Top 5	BA (5 terms)	BB (24 terms)	BC (18 terms)
	Top 3	CA (3 terms)	CB (18 terms)	CC (12 terms)

This paper presents the results of methodologies CA, BA and CB. The rest are presented in the appendix

It is important to note that before including the terms in the principal component analysis, it was necessary to remove repeated terms and remove terms that had negative correlation with more than 20 other terms in the sample. The former because given the nature of related queries, some queries that were related to more than one term found themselves duplicated in our Google Trends index database. This was done in order to avoid including terms in the analysis that exhibited an overall different behavior than most of the terms.

To justify that the first component gives an accurate representation of the variables used to create it, we measured the percentage of the variance that it represents. We did so for each methodology proposed and consider that the first component must explain more than 60% of the variance for it to significantly represent the overall behavior of the terms used to create it. A scree plot for every methodology is shown in the appendix.

4. Results

The analysis of results can be split in two sections: an analysis of estimations for the baseline models (which did not employ the non-conventional variable), and an analysis of the estimations of baseline models with the non-conventional variable added to each one. The average marginal effects of the variables in the models will be shown in the appendix, as they have the same significance and sign as the coefficients in the estimations. Starting with the baseline models, Table 1 shows the results obtained.

Table 1: Baseline Model Estimations

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Constant	-2.913	-0.015	-1.701	1.394
	(3.05)	(3.59)	(3.27)	(3.86)
Inflation (-1)	17.175	-15.559	29.286	-4.268
	(15.73)	(23.40)	(18.01)	(25.17)
Inflation Expectations (-1)	-0.781	-1.78	-1.408	-2.436**
	(0.87)	(1.11)	(1.00)	(1.24)
Output Gap (-1)	0.204***	0.183***	0.217***	0.193***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.247***	0.226***	0.239***	0.215***
	(0.06)	(0.06)	(0.06)	(0.06)
United States' Influence (-1)	--	0.547*	--	0.553*
		(0.28)		(0.29)
Market Expectations (-1)	--	--	1.804*	1.754*
			(1.06)	(1.05)
Observations	130	130	130	130
Log Likelihood	-36.398	-34.541	-34.838	-33.073
Akaike Inf. Crit.	82.796	81.083	81.676	80.146
Residual Deviance	72.796 (df = 125)	69.083 (df = 124)	69.676 (df = 124)	66.146 (df = 123)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562
Note:				*p<0.1; **p<0.05; ***p<0.01

We observed that results of the output gap and exchange rate are robust, considerably significant (a p-value lower than 0.01) and in line with what is expected from economic theory. Avdjiev and Zeng (2014) show that a positive correlation can be expected from empirical data of economic activity and restrictive monetary policy. Bjornland and Halvorsen (2014) show strong positive correlations between exchange rate depreciations and restrictive monetary policy. Both findings are reflected with the positive sign of the coefficients, strong significance and robustness of the output gap and exchange rate as explanatory variables of Mexico's monetary policy.

The influence of the United States' monetary policy and market expectations of future macroeconomic conditions also showed robust and significant results, albeit being significant with 90% confidence. Authors such as Estrella and Hardouvelis (1991) and Wright (2008) found that increases in the slope of the yield curve are correlated with restrictive monetary policy, which can be seen in our models by the constantly positive sign. In addition, Crespo et al. (2016) explain a positive correlation between interest rates of both the United States and Mexico, which is also reflected by the positive sign in our models.

In contrast, inflation and inflation expectations showed results that were not robust, as their coefficients seemed to depend on which model was considered. Inflation remained non-significant across all models. The sign of the coefficient for inflation presented changes within the models. Inflation expectations showed one case of significance (to the 95% confidence level, in model 4), as well as a consistently negative sign. This negative sign was not expected, as Barro and Gordon (1983) show positive correlations between inflation expectations and restrictive monetary policy. This negative sign may be explained by the period we chose to analyze, as Banxico increased the interest rate consistently from 2016 and onwards. During this period, inflation expectations were higher due to the macroeconomic context of the time. As the interest rate increased, inflation expectations began lowering (as they were anchored to Banxico’s inflation target). This could be why a negative coefficient for inflation expectations is obtained across all base models.

Moving on to models with the non-conventional variable, we found that the hypothesized relationship between the populations’ interest variable and monetary policy decisions is not rejected in all cases, with robustness and consistency across specifications. Estimations of models using methodologies BA, CA and CB are shown below, while the remaining six methodologies are presented in the appendix.

Table 2: Methodology BA Model Estimations (No. Of Terms in Index: 5)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-4.851	-2.821	-3.445	-1.171
	(3.42)	(4.44)	(3.61)	(4.77)
Inflation (-1)	11.53	-3.176	23.011	6.84
	(16.94)	(26.03)	(19.12)	(27.88)
Inflation Expectations (-1)	-1.749*	-1.915*	-2.397**	-2.564**
	(1.05)	(1.10)	(1.19)	(1.24)
Output Gap (-1)	0.201***	0.188***	0.213***	0.198***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.391***	0.332***	0.382***	0.314**
	(0.10)	(0.13)	(0.10)	(0.14)
United States' Influence (-1)	--	0.277	--	0.304
		(0.37)		(0.40)
Market Expectations (-1)	--	--	1.66	1.681
			(1.04)	(1.05)
Population Interest (-1)	-0.032**	-0.022	-0.032*	-0.02
	(0.02)	(0.02)	(0.02)	(0.02)
Observations	130	130	130	130
Log Likelihood	-34.393	-34.123	-33.053	-32.758
Akaike Inf. Crit.	80.787	82.246	80.106	81.515
Residual Deviance	68.787 (df = 124)	68.246 (df = 123)	66.106 (df = 123)	65.515 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 73.3%

Table 3: Methodology CA Model Estimations (No. Of Terms in Index: 3)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-5.883	-6.518	-4.445	-4.998
	(3.69)	(4.57)	(3.86)	(4.85)
Inflation (-1)	4.512	8.626	15.823	19.353
	(18.36)	(25.53)	(20.29)	(27.46)
Inflation Expectations (-1)	-1.940^{*}	-1.871^{*}	-2.564^{**}	-2.510^{**}
	(1.10)	(1.12)	(1.23)	(1.26)
Output Gap (-1)	0.196^{***}	0.200^{***}	0.205^{***}	0.209^{***}
	(0.07)	(0.07)	(0.07)	(0.08)
Exchange Rate (-1)	0.450^{***}	0.468^{***}	0.439^{***}	0.455^{***}
	(0.11)	(0.14)	(0.12)	(0.15)
United States' Influence (-1)	--	-0.085	--	-0.073
		(0.37)		(0.39)
Market Expectations (-1)	--	--	1.552	1.55
			(1.07)	(1.07)
Population Interest (-1)	-0.054^{***}	-0.057^{**}	-0.053^{**}	-0.056^{**}
	(0.02)	(0.03)	(0.02)	(0.03)
Observations	130	130	130	130
Log Likelihood	-32.333	-32.309	-31.222	-31.206
Akaike Inf. Crit.	76.667	78.618	76.444	78.413
Residual Deviance	64.667 (df = 124)	64.618 (df = 123)	62.444 (df = 123)	62.413 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Percentage of Variance Explained by First Component: 86.2%

Table 4: Methodology CB Model Estimations (No. Of Terms in Index: 18)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-5.96	-5.269	-4.551	-3.704
	(3.66)	(4.64)	(3.86)	(4.94)
Inflation (-1)	8.983	4.339	20.625	14.987
	(17.49)	(25.85)	(19.65)	(27.86)
Inflation Expectations (-1)	-1.872*	-1.928*	-2.561**	-2.618**
	(1.06)	(1.09)	(1.22)	(1.24)
Output Gap (-1)	0.192***	0.188***	0.204***	0.199***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.474***	0.451***	0.470***	0.440***
	(0.12)	(0.16)	(0.13)	(0.17)
United States' Influence (-1)	--	0.089	--	0.106
		(0.36)		(0.38)
Market Expectations (-1)	--	--	1.643	1.647
			(1.06)	(1.06)
Population Interest (-1)	-0.020**	-0.019*	-0.021**	-0.019
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-33.321	-33.294	-32.046	-32.01
Akaike Inf. Crit.	78.642	80.588	78.091	80.02
Residual Deviance	66.642 (df = 124)	66.588 (df = 123)	64.091 (df = 123)	64.020 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 74.6%

Our results show that overall, the inclusion of the non-conventional variable improves all baseline models based on the minimization of the Akaike Information Criterion (AIC), regardless of the specification implemented. Also, our estimations are robust across models with high levels of significance on output gap and exchange rate. However, we noted that, regardless of the specification, the results shown by the United States' influence variable in the baseline models change drastically when adding the non-conventional variable, suggesting the existence of conflict. Relating to this, model 7 (which does not consider the United States' influence but considers all other variables) is shown to be the one that consistently minimizes the AIC of the eight models considered. This shows that excluding the United States' influence on Mexican monetary policy and including all other variables yields better predictions of Mexican monetary policy. For this reason, model 7 will hereafter have an important part in the analysis.

Another overall result is that the non-conventional variable is considerably significant to at least a 95% confidence level in most cases. In fact, when excluding the models that employ the United States' influence variable (for the reason previously stated), the non-conventional variable is always significant to at least a 90% confidence level. This shows that the Mexican population's interest in the macroeconomic context of their country is a relevant covariate for predicting monetary policy decisions of Banxico. The non-conventional variable is also consistent in both the magnitude of the coefficient and its negative sign, regardless of the number of related queries included. This means that related queries will not significantly alter the results obtained.

However, what indeed improves model performance is the number of Banxico's terms chosen to be included in the Google Trends index search. It is important to note that the variance explained by the first component of every methodology A index was in the 50% to 55% range. This shows that those methodologies failed to capture the overall behavior of the terms that composed them, which makes the models' coefficients unreliable at the moment of interpreting their results. Considering this, we found that models tend to present better results when less Banxico terms are included in them, as evidenced by the results of methodologies BA, CA and CB.

For example, models estimated using methodology BA tend to perform worse than their counterparts in the baseline models when the non-conventional variable is non-significant at the 90% confidence level, as evidenced by their greater AIC. This is not true for models in methodology CB, where even though the inclusion of the non-conventional variable in model 8 proved to be non-significant, its AIC is lower than the one in model 4.

Another example regarding the improved performance when decreasing the number of Banxico terms can be seen in methodology CA, where the non-conventional variable proves to be significant at the 95% confidence level across all models. Also, it is important to note that the lowest AIC out of all the models in this paper is model 7 from the CA methodology (in which the non-conventional variable was built using only the three most frequently used terms in Banxico's minutes). From here, we observed that the AIC gets progressively higher as more terms are included, as evidenced by the AIC of model 7 from categories CB and CC and all the categories from methodology B.

Conclusions

In this paper, we propose the construction and inclusion big data variable that captures the Mexican populations' interest in macroeconomic topics. Our results show that the Google searches variable of macroeconomic topics in Mexico was relevant in predicting restrictive monetary policy decisions made by the Mexico's central bank. Additionally, we observed clear benefits of including a non-conventional variable in a model that predicts Banxico's monetary policy decisions. These benefits are significance in the variable measured with big data, and minimization of the Akaike Information Criterion. The consistency of the results of all variables across the models, and the consistency of the results with economic theory indicate the usefulness the model has in explaining the determinants of monetary policy for the case of Mexico.

In addition, our results suggested that including less terms in the construction of the non-conventional variable will yield better results for predicting Mexican monetary policy. A possible explanation for this is that the population pays more attention to the most frequently used macroeconomic terms in Banxico's communications. Therefore, measuring a population interest variable with only the terms that make the most presence in the central bank's communication will result in more concise predictions of monetary policy.

One key thing to note is that the Google Trend Index for a given term tends to be slightly different every time it gets lifted. We noticed that in both the official page for Google Trends and in the *gtrendsR* package, when lifting the exact same term, the system consistently gave us the same index around 60% of the time and another index around 40% of the time. Nevertheless, the correlation coefficient between both indices was more than 0.9 and the overall inferences given by the models' coefficients did not change significantly.

Since this paper pioneers the use of big data in a monetary policy context for the case of Mexico, we encountered several limitations, which include difficulties accessing to information, lack of methodologies to analyze big data information and overall structure of the data. The first one we mitigated by using Google Trends, and the second one we tried to resolve it by utilizing principal component analysis. However, the third one posed a challenge, since it caused some results that did not go according to the literature (as with inflation and inflation expectations) and generated some doubts when asserting the validity of our findings. With this in mind, we propose that this methodology to analyze big data gets revisited in the future, with more time-sample data, in order to confirm or negate the results shown in this paper.

5. References

Armah, N. (2013). Big Data Analysis: The Next Frontier. Bank of Canada Review.

Avdjiev, S., & Zeng, Z. (2014). Credit Growth, Monetary Policy, and Economic Activity in a Three-Regime TVAR Model. Bank for International Settlements, 449.

Baghestani, H., & Danila, L. (2014). Interest rate and exchange rate forecasting in the Czech Republic: Do analysts know better than a random walk? Finance a Uver - Czech Journal of Economics and Finance.

Balcilar, M. (2019). mFilter: Miscellaneous Time Series Filters (Version 0.1-5) [R package] Retrieved from: https://CRAN.R-project.org/package=mFilter

Banco de México (2019), Estadísticas México, BM. Retrieved from: http://www.anterior.banxico.org.mx/estadisticas/index.html

Banco de México. 2008-2019. Informe de política monetaria, several years, México. Retrieved from: https://www.banxico.org.mx/publicaciones-y-prensa/minutas-de-las-decisiones-de-politica-monetaria/minutas-politica-monetaria-ta.html

Bjornland, H., & Halvorsen, J. (2014). How does Monetary Policy Respond to Exchange Rate Movements? New International Evidence. Oxford Bulletin of Economics and Statistics, 76(2). doi: 10.1111/obes.12014

Bouchet-Valat, M. (2019). SnowballC: Snowball Stemmers Based on the C 'libstemmer' UTF-8 Library (Version 0.6.0) [R package]. Retrieved from: https://CRAN.Rproject.org/package=SnowballC

Chague, F., De-Losso, R., Giovannetti, B., & Manoel, P. (2015). Central Bank Communication Affects the Term-Structure of Interest Rates. Revista Brasileira de Economia, 69(2), 147–162.

Choi, H., & Varian, H. (2009). Predicting Initial Claims for Unemployment Benefits. Google Inc, 1-5.

Clarida, R., Galí, J., & Gertler, M. (1999). The science of monetary policy: A new Keynesian perspective. Journal of Economic Literature. https://doi.org/10.1257/jel.37.4.1661

Crespo-Cuaresma, J., Doppelhofer, G., Feldkircher, M., & Huber, F. (2016). US monetary policy in a globalized world (No. 5826). CESifo Working Paper

Durán-Bustamante, M., Hernandez-del Valle, A., & Ortiz-Ramírez, A. (2018). The Google Trends Effect on the behavior of the exchange rate Mexican peso - US dollar. Contaduría y Administración, 0(0).

Enders, W., & Granger, C. W. J. (1998). Unit-Root Tests and Asymmetric Adjustment with an Example Using the Term Structure of Interest Rates. Journal of Business & Economic Statistics. https://doi.org/10.2307/1392506

Feinerer, I. & Hornik, K. (2018). tm: Text Mining Package (Version 0.7-6) [R package]. Retrieved from: https://CRAN.R-project.org/package=tm

Fellows, I. (2018). wordcloud: Word Clouds (Version 2.6) [R package]. Retrieved from: https://CRAN.R-project.org/package=wordcloud

Hayo, B., & Neuenkirch, M. (2013). Do Federal Reserve presidents communicate with a regional bias? Journal of Macroeconomics. https://doi.org/10.1016/j.jmacro.2012.10.002

Horvath, R., & Karas, P. (2013). Central Bank Communication and Interest Rates: The Case of the Czech National Bank. Finance a Uver/Czech Journal of Economics and Finance, 63(5), 454–464

INEGI. (2019). Banco de Información Económica. Retrieved from: http://www.inegi.org.mx/sistemas/bie/

Jansen, D.-J., & De Haan, J. (2009). Has ECB Communication Been Helpful in Predicting Interest Rate Decisions? An Evaluation of the Early Years of the Economic and Monetary Union. Applied Economics, 41(16–18), 1995–2003.

Jung, A. (2016). Have monetary data releases helped markets to predict the interest rate decisions of the European Central Bank?

Jung, A. (2018). Have money and credit data releases helped markets to predict the interest rate decisions of the European Central Bank? Scottish Journal of Political Economy. https://doi.org/10.1111/sjpe.12143

Kulkarni, R., & Schintler, L. (2014). Big Data for Policy Analysis: The Good, The Bad, and The Ugly. Review of Policy Research, 31 (4)

Lamla, M. J., & Sturm, J.-E. (2013). Interest Rate Expectations in the Media and Central Bank Communication. In P. L. Siklos & J.-E. Sturm (Eds.), Central Bank Communication, Decision Making, and Governance: Issues, Challenges, and Case Studies (pp. 101–111). CESifo Seminar Series. Cambridge and London: MIT Press.

Lanne, M., & Nyberg, H. (2016). Generalized Forecast Error Variance Decomposition for Linear and Nonlinear Multivariate Models. Oxford Bulletin of Economics and Statistics. https://doi.org/10.1111/obes.12125

Laopodis, N. (2006). Dynamic Interactions among the Stock Market, Federal Funds Rate, Inflation, and Economic Activity. The Financial Review, 41 (pp. 513-545)

Leeper, T. (2018). margins: Marginal Effects for Model Objects (Version 0.3.23) [R package]

Lucca, D. O., & Trebbi, F. (2009). Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements. In SSRN. https://doi.org/10.2139/ssrn.1470443

Maravall, A., & Del Rio, A. (2001). Time aggregation and the Hodrick-Prescott filter (No. 0108). Banco de España.

Marek, H. (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables. R package. (Version 5.2.1) [R package]. Retrieved from: https://CRAN.R-project.org/package=stargazer

Massicotte, P. & Eddelbuettel, D. (2019). gtrendsR: Perform and Display Google Trends Queries (Version 1.4.4) [R package] Retrieved from: https://CRAN.R-project.org/package=gtrendsR

McMillan, D. G. (2009). Non-linear interest rate dynamics and forecasting: Evidence for US and Australian interest rates. International Journal of Finance and Economics. https://doi.org/10.1002/ijfe.358

Moreno-Bird, J. C., Nápoles, P. R., & Valdivia, J. C. (2004). NAFTA and the Mexican Economy: A Look Back on a Ten-Year Relationship. NCJ Int'l L. & Com. Reg., 30, 997.

Neuwirth, E. (2014). RColorBrewer: ColorBrewer Palettes (Version 2.6) [R package]. Retrieved from: https://CRAN.R-project.org/package=RColorBrewer

Owens, R.E, & Webb, R.H. Using the Federal Funds Futures Market to Predict Monetary Policy Actions. Federal Reserve Bank of Richmond Economic Quarterly. 2001;87(2):69-77

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Retrieved from: https://www.R-project.org/

Sax C, & Eddelbuettel D. (2018). Seasonal Adjustment by X-13ARIMA-SEATS in R. Journal of Statistical Software, 87(11), 1-17. doi: 10.18637/jss.v087.i11

Silva, J. & Iqbal, A. (2015). “An Ordered Probit Approach to Predicting the Probability of Inflation/Deflation”. Business Economics, 50(1), 12-19. https://doi.org/10.1057/be.2015.1

Silverstovs, B., & Wochner, D. S. (2018). Google Trends and Reality: Do the Proportions Match? Appraising the Informational Value of Online Search Behavior: Evidence from Swiss Tourism Regions. Journal of Economic Behavior and Organization, 145, 1–23.

Taylor, J. B. (1993). Discretion versus policy rules in practice. Carnegie-Rochester Confer. Series on Public Policy. https://doi.org/10.1016/0167-2231(93)90009-L

Tellez Leon, I. E., & Venegas Martínez, F. (2013). Principales determinantes en las decisiones de política monetaria de México: Un análisis econométrico. (With English summary.) Estudios Económicos, 28(1), 79–107.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Wohlfarth, P. (2018). Measuring the impact of monetary policy attention on global asset volatility using search data. Economics Letters, 173, 15–18. https://doi.org/10.1016/j.econlet.2018.08.009.

6. Appendix

A1. Model estimations for different methodologies

Methodology AA Model Estimations (No. Of Terms in Index: 19)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-5.704	-6.637	-4.218	-5.032
	(3.68)	(4.85)	(3.84)	(5.13)
Inflation (-1)	8.029	14.065	20.053	25.388
	(17.57)	(27.35)	(19.63)	(29.86)
Inflation Expectations (-1)	-1.735	-1.637	-2.430**	-2.357*
	(1.06)	(1.10)	(1.21)	(1.24)
Output Gap (-1)	0.179***	0.183***	0.188***	0.191***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.412***	0.436***	0.404***	0.426***
	(0.10)	(0.13)	(0.11)	(0.14)
United States' Influence (-1)	--	-0.119	--	-0.103
		(0.41)		(0.43)
Market Expectations (-1)	--	--	1.613	1.61
			(1.07)	(1.08)
Population Interest (-1)	-0.022***	-0.025**	-0.022**	-0.025*
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-32.582	-32.544	-31.394	-31.368
Akaike Inf. Crit.	77.165	79.088	76.789	78.737
Residual Deviance	65.165 (df = 124)	65.088 (df = 123)	62.789 (df = 123)	62.737 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 56.7%

Methodology AB Model Estimations (No. Of Terms in Index: 82)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-6.564*	-6.749	-5.135	-5.253
	(3.81)	(4.81)	(3.98)	(5.08)
Inflation (-1)	8.697	9.893	20.998	21.774
	(17.53)	(26.25)	(19.64)	(28.52)
Inflation Expectations (-1)	-2.008*	-1.994*	-2.767**	-2.760**
	(1.10)	(1.11)	(1.27)	(1.28)
Output Gap (-1)	0.186***	0.187***	0.197***	0.198***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.487***	0.493***	0.488***	0.492***
	(0.13)	(0.16)	(0.14)	(0.17)
United States' Influence (-1)	--	-0.023	--	-0.014
		(0.37)		(0.38)
Market Expectations (-1)	--	--	1.681	1.681
			(1.08)	(1.08)
Population Interest (-1)	-0.013***	-0.013**	-0.014**	-0.014*
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-32.528	-32.527	-31.251	-31.251
Akaike Inf. Crit.	77.057	79.053	76.502	78.501
Residual Deviance	65.057 (df = 124)	65.053 (df = 123)	62.502 (df = 123)	62.501 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 50%

Methodology AC Model Estimations (No. Of Terms in Index: 59)

Variable	Model 5	Model 6	Model 7	Model 8
	Dependent variable for all Models: Rate Increases or Not			
Constant	-6.327*	-6.642	-4.873	-5.086
	(3.78)	(4.77)	(3.94)	(5.03)
Inflation (-1)	8.281	10.351	20.497	21.907
	(17.53)	(26.39)	(19.67)	(28.69)
Inflation Expectations (-1)	-2.040*	-2.018*	-2.776**	-2.764**
	(1.10)	(1.11)	(1.27)	(1.28)
Output Gap (-1)	0.185***	0.187***	0.196***	0.197***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.477***	0.487***	0.475***	0.482***
	(0.12)	(0.16)	(0.13)	(0.17)
United States' Influence (-1)	--	-0.04	--	-0.026
		(0.37)		(0.39)
Market Expectations (-1)	--	--	1.662	1.662
			(1.08)	(1.08)
Population Interest (-1)	-0.016**	-0.016**	-0.016**	-0.016*
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-32.488	-32.483	-31.245	-31.243
Akaike Inf. Crit.	76.976	78.966	76.49	78.486
Residual Deviance	64.976 (df = 124)	64.966 (df = 123)	62.490 (df = 123)	62.486 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 51.1%

Methodology BB Model Estimations (No. Of Terms in Index: 24)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-5.34	-3.629	-3.952	-2.094
	(3.48)	(4.59)	(3.68)	(4.86)
Inflation (-1)	11.141	-0.872	23.1	10.021
	(16.92)	(26.16)	(19.18)	(28.08)
Inflation Expectations (-1)	-1.680*	-1.827*	-2.381**	-2.524**
	(1.01)	(1.07)	(1.17)	(1.22)
Output Gap (-1)	0.191***	0.182***	0.204***	0.193***
	(0.07)	(0.07)	(0.07)	(0.07)
Exchange Rate (-1)	0.429***	0.373**	0.425***	0.362**
	(0.11)	(0.15)	(0.12)	(0.15)
United States' Influence (-1)	--	0.224	--	0.24
		(0.37)		(0.39)
Market Expectations (-1)	--	--	1.707	1.713
			(1.06)	(1.05)
Population Interest (-1)	-0.015**	-0.011	-0.015**	-0.011
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-34.123	-33.952	-32.733	-32.548
Akaike Inf. Crit.	80.247	81.903	79.465	81.097
Residual Deviance	68.247 (df = 124)	67.903 (df = 123)	65.465 (df = 123)	65.097 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 65.7%

Methodology BC Model Estimations (No. Of Terms in Index: 18)

Variable	Model 5	Model 6	Model 7	Model 8
	Dependent variable for all Models: Rate Increases or Not			
Constant	-6.079*	-5.251	-4.671	-3.61
	(3.65)	(4.48)	(3.84)	(4.79)
Inflation (-1)	9.428	3.489	20.598	13.234
	(17.39)	(25.22)	(19.55)	(27.00)
Inflation Expectations (-1)	-2.095*	-2.147*	-2.720**	-2.781**
	(1.09)	(1.11)	(1.24)	(1.26)
Output Gap (-1)	0.202***	0.196***	0.214***	0.205***
	(0.07)	(0.07)	(0.07)	(0.08)
Exchange Rate (-1)	0.488***	0.459***	0.477***	0.440***
	(0.13)	(0.16)	(0.13)	(0.17)
United States' Influence (-1)	--	0.111	--	0.138
		(0.34)		(0.36)
Market Expectations (-1)	--	--	1.59	1.602
			(1.05)	(1.05)
Population Interest (-1)	-0.025**	-0.023*	-0.025**	-0.022
	(0.01)	(0.01)	(0.01)	(0.01)
Observations	130	130	130	130
Log Likelihood	-33.276	-33.228	-32.073	-32.005
Akaike Inf. Crit.	78.552	80.455	78.147	80.01
Residual Deviance	66.552 (df = 124)	66.455 (df = 123)	64.147 (df = 123)	64.010 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 62.6%

Methodology CC Model Estimations (No. Of Terms in Index: 12)

	Dependent variable for all Models: Rate Increases or Not			
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Constant	-5.906	-5.52	-4.486	-3.974
	(3.66)	(4.67)	(3.85)	(4.99)
Inflation (-1)	8.619	6.08	20.278	16.925
	(17.60)	(26.06)	(19.74)	(28.17)
Inflation Expectations (-1)	-1.906*	-1.935*	-2.594**	-2.625**
	(1.07)	(1.09)	(1.23)	(1.24)
Output Gap (-1)	0.193***	0.191***	0.205***	0.202***
	(0.07)	(0.07)	(0.07)	(0.08)
Exchange Rate (-1)	0.468***	0.455***	0.462***	0.445***
	(0.12)	(0.16)	(0.13)	(0.17)
United States' Influence (-1)	--	0.049	--	0.064
		(0.37)		(0.39)
Market Expectations (-1)	--	--	1.643	1.644
			(1.06)	(1.06)
Population Interest (-1)	-0.026**	-0.024*	-0.026**	-0.024
	(0.01)	(0.01)	(0.01)	(0.02)
Observations	130	130	130	130
Log Likelihood	-33.214	-33.206	-31.945	-31.933
Akaike Inf. Crit.	78.427	80.412	77.889	79.865
Residual Deviance	66.427 (df = 124)	66.412 (df = 123)	63.889 (df = 123)	63.865 (df = 122)
Null Deviance (df = 129)	104.562	104.562	104.562	104.562

Note: *p<0.1; **p<0.05; ***p<0.01

Percentage of Variance Explained by First Component: 77.2%

A2. Model marginal effects

Base Models Average Marginal Effects				
Variable	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Inflation (-1)	2.608	-2.245	4.225*	-0.586
	(2.361)	(3.366)	(2.527)	(3.457)
Inflation Expectations (-1)	-0.119	-0.257	-0.203	-0.334**
	(0.132)	(0.157)	(0.142)	(0.163)
Output Gap (-1)	0.031***	0.026***	0.031***	0.026***
	(0.01)	(0.009)	(0.009)	(0.009)
Exchange Rate (-1)	0.038***	0.033***	0.035***	0.03***
	(0.007)	(0.007)	(0.007)	(0.007)
United States' Influence (-1)	--	0.079**	--	0.076*
		(0.04)		(0.039)
Market Expectations (-1)	--	--	0.26*	0.241*
			(0.148)	(0.138)
Note:				*p<0.1; **p<0.05; ***p<0.01

Methodology AA Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.095	1.915	2.63	3.328
	(2.392)	(3.708)	(2.533)	(3.868)
Inflation Expectations (-1)	-0.237*	-0.223	-0.319**	-0.309**
	(0.142)	(0.147)	(0.15)	(0.155)
Output Gap (-1)	0.024***	0.025***	0.025***	0.025***
	(0.008)	(0.008)	(0.008)	(0.008)
Exchange Rate (-1)	0.056***	0.059***	0.053***	0.056***
	(0.011)	(0.015)	(0.011)	(0.017)
United States' Influence (-1)	--	-0.016	--	-0.014
		(0.055)		(0.056)
Market Expectations (-1)	--	--	0.212	0.211
			(0.135)	(0.136)
Population Interest (-1)	-0.003***	-0.003**	-0.003***	-0.003*
	(0.001)	(0.002)	(0.001)	(0.002)
Note:				*p<0.1; **p<0.05; ***p<0.01

Methodology AB Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.183	1.346	2.745	2.846
	(2.379)	(3.561)	(2.523)	(3.691)
Inflation Expectations (-1)	-0.273*	-0.271*	-0.362**	-0.361**
	(0.145)	(0.147)	(0.156)	(0.158)
Output Gap (-1)	0.025***	0.025***	0.026***	0.026***
	(0.008)	(0.008)	(0.008)	(0.008)
Exchange Rate (-1)	0.066***	0.067***	0.064***	0.064***
	(0.014)	(0.019)	(0.014)	(0.02)
United States' Influence (-1)	--	-0.003	--	-0.002
		(0.05)		(0.05)
Market Expectations (-1)	--	--	0.22	0.22
			(0.136)	(0.136)
Population Interest (-1)	-0.002***	-0.002**	-0.002***	-0.002**
	(0.001)	(0.001)	(0.001)	(0.001)
Note:				*p<0.1; **p<0.05; ***p<0.01

Methodology AC Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.125	1.405	2.678	2.863
	(2.376)	(3.573)	(2.527)	(3.712)
Inflation Expectations (-1)	-0.277*	-0.274*	-0.363**	-0.361**
	(0.145)	(0.147)	(0.156)	(0.158)
Output Gap (-1)	0.025***	0.025***	0.026***	0.026***
	(0.008)	(0.008)	(0.008)	(0.008)
Exchange Rate (-1)	0.065***	0.066***	0.062***	0.063***
	(0.013)	(0.018)	(0.014)	(0.019)
United States' Influence (-1)	--	-0.005	--	-0.003
		(0.05)		(0.051)
Market Expectations (-1)	--	--	0.217	0.217
			(0.136)	(0.136)
Population Interest (-1)	-0.002***	-0.002**	-0.002***	-0.002*
	(0.001)	(0.001)	(0.001)	(0.001)
Note:				*p<0.1; **p<0.05; ***p<0.01

Methodology BA Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.654	-0.452	3.161	0.93
	(2.419)	(3.705)	(2.577)	(3.787)
Inflation Expectations (-1)	-0.251*	-0.272*	-0.329**	-0.349**
	(0.147)	(0.153)	(0.157)	(0.16)
Output Gap (-1)	0.029***	0.027***	0.029***	0.027***
	(0.009)	(0.009)	(0.009)	(0.009)
Exchange Rate (-1)	0.056***	0.047***	0.052***	0.043**
	(0.012)	(0.017)	(0.012)	(0.017)
United States' Influence (-1)	--	0.039	--	0.041
		(0.053)		(0.053)
Market Expectations (-1)	--	--	0.228	0.229*
			(0.139)	(0.137)
Population Interest (-1)	-0.005**	-0.003	-0.004**	-0.003
	(0.002)	(0.003)	(0.002)	(0.003)
Note:				*p<0.1; **p<0.05; ***p<0.01

Methodology BB Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.582	-0.123	3.145	1.356
	(2.393)	(3.7)	(2.561)	(3.789)
Inflation Expectations (-1)	-0.238*	-0.258*	-0.324**	-0.341**
	(0.141)	(0.149)	(0.152)	(0.157)
Output Gap (-1)	0.027***	0.026***	0.028***	0.026***
	(0.008)	(0.009)	(0.008)	(0.009)
Exchange Rate (-1)	0.061***	0.053***	0.058***	0.049**
	(0.013)	(0.019)	(0.013)	(0.019)
United States' Influence (-1)	--	0.032	--	0.032
		(0.053)		(0.052)
Market Expectations (-1)	--	--	0.232*	0.232*
			(0.139)	(0.137)
Population Interest (-1)	-0.002**	-0.002	-0.002**	-0.002
	(0.001)	(0.001)	(0.001)	(0.001)
Note:				*p<0.1; **p<0.05; ***p<0.01

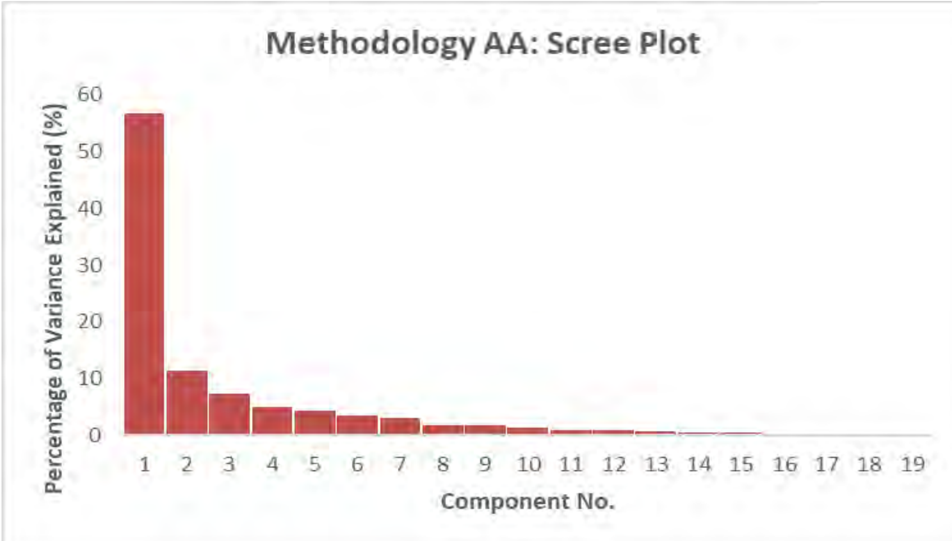
Methodology BC Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.313	0.485	2.757	1.766
	(2.416)	(3.504)	(2.576)	(3.588)
Inflation Expectations (-1)	-0.292**	-0.298**	-0.364**	-0.371**
	(0.147)	(0.15)	(0.157)	(0.159)
Output Gap (-1)	0.028***	0.027***	0.029***	0.027***
	(0.008)	(0.009)	(0.008)	(0.009)
Exchange Rate (-1)	0.068***	0.064***	0.064***	0.059***
	(0.015)	(0.02)	(0.015)	(0.02)
United States' Influence (-1)	--	0.015	--	0.018
		(0.048)		(0.048)
Market Expectations (-1)	--	--	0.213	0.214
			(0.136)	(0.135)
Population Interest (-1)	-0.003**	-0.003*	-0.003**	-0.003
	(0.001)	(0.002)	(0.001)	(0.002)
Note:	*p<0.1; **p<0.05; ***p<0.01			

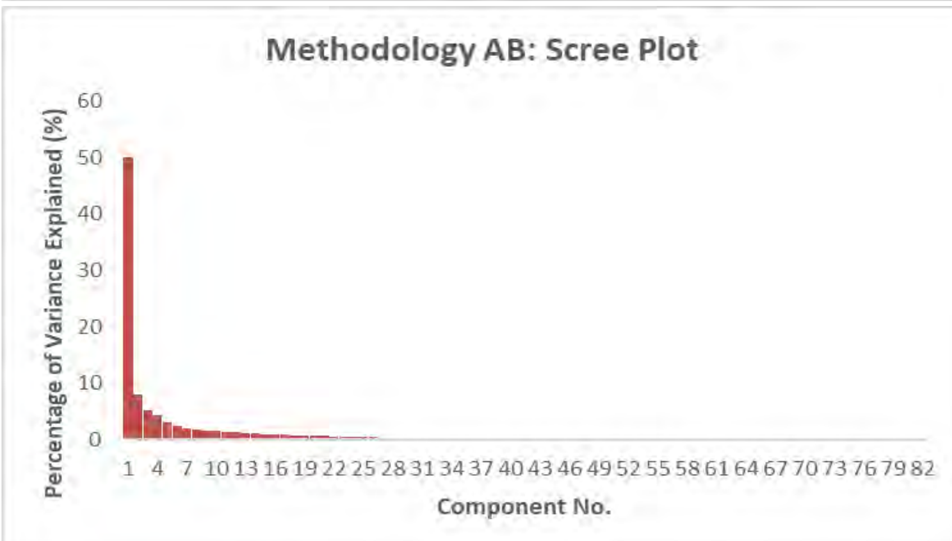
Methodology CA Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	0.613	1.171	2.072	2.535
	(2.492)	(3.456)	(2.63)	(3.565)
Inflation Expectations (-1)	-0.263*	-0.254*	-0.336**	-0.329**
	(0.144)	(0.149)	(0.152)	(0.157)
Output Gap (-1)	0.027***	0.027***	0.027***	0.027***
	(0.008)	(0.008)	(0.008)	(0.009)
Exchange Rate (-1)	0.061***	0.063***	0.057***	0.06***
	(0.012)	(0.015)	(0.012)	(0.016)
United States' Influence (-1)	--	-0.012	--	-0.01
		(0.05)		(0.05)
Market Expectations (-1)	--	--	0.203	0.203
			(0.134)	(0.135)
Population Interest (-1)	-0.007***	-0.008**	-0.007***	-0.007**
	(0.003)	(0.003)	(0.003)	(0.003)
Note:	*p<0.1; **p<0.05; ***p<0.01			

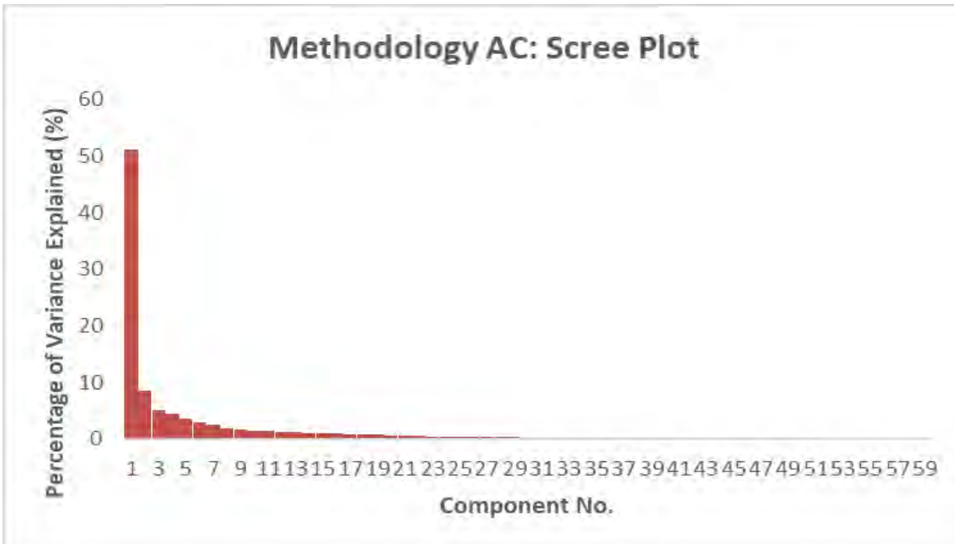
Methodology CB Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.253	0.605	2.762	2.003
	(2.433)	(3.6)	(2.587)	(3.705)
Inflation Expectations (-1)	-0.261*	-0.269*	-0.343**	-0.35**
	(0.144)	(0.148)	(0.155)	(0.157)
Output Gap (-1)	0.027***	0.026***	0.027***	0.027***
	(0.008)	(0.009)	(0.008)	(0.009)
Exchange Rate (-1)	0.066***	0.063***	0.063***	0.059***
	(0.014)	(0.02)	(0.015)	(0.021)
United States' Influence (-1)	--	0.012	--	0.014
		(0.051)		(0.051)
Market Expectations (-1)	--	--	0.22	0.22
			(0.137)	(0.136)
Population Interest (-1)	-0.003**	-0.003*	-0.003**	-0.002
	(0.001)	(0.001)	(0.001)	(0.002)
Note:				*p<0.1; **p<0.05; ***p<0.01

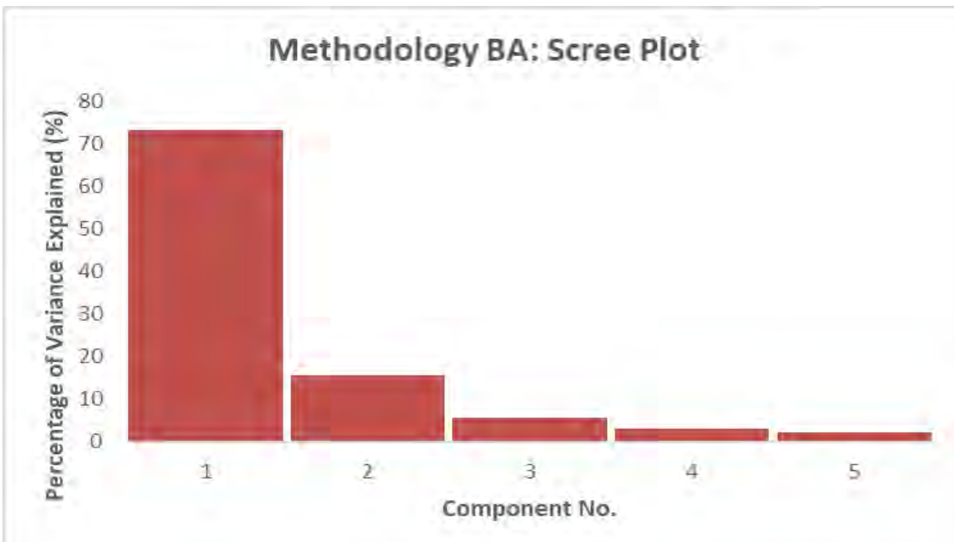
Methodology CC Average Marginal Effects				
Variable	**Model 5**	**Model 6**	**Model 7**	**Model 8**
Inflation (-1)	1.198	0.845	2.707	2.257
	(2.44)	(3.616)	(2.592)	(3.735)
Inflation Expectations (-1)	-0.265*	-0.269*	-0.346**	-0.35**
	(0.145)	(0.148)	(0.155)	(0.157)
Output Gap (-1)	0.027***	0.027***	0.027***	0.027***
	(0.008)	(0.009)	(0.008)	(0.009)
Exchange Rate (-1)	0.065***	0.063***	0.062***	0.059***
	(0.014)	(0.019)	(0.014)	(0.02)
United States' Influence (-1)	--	0.007	--	0.009
		(0.052)		(0.052)
Market Expectations (-1)	--	--	0.219	0.219
			(0.137)	(0.136)
Population Interest (-1)	-0.004***	-0.003*	-0.003**	-0.003
	(0.001)	(0.002)	(0.001)	(0.002)
Note:				*p<0.1; **p<0.05; ***p<0.01

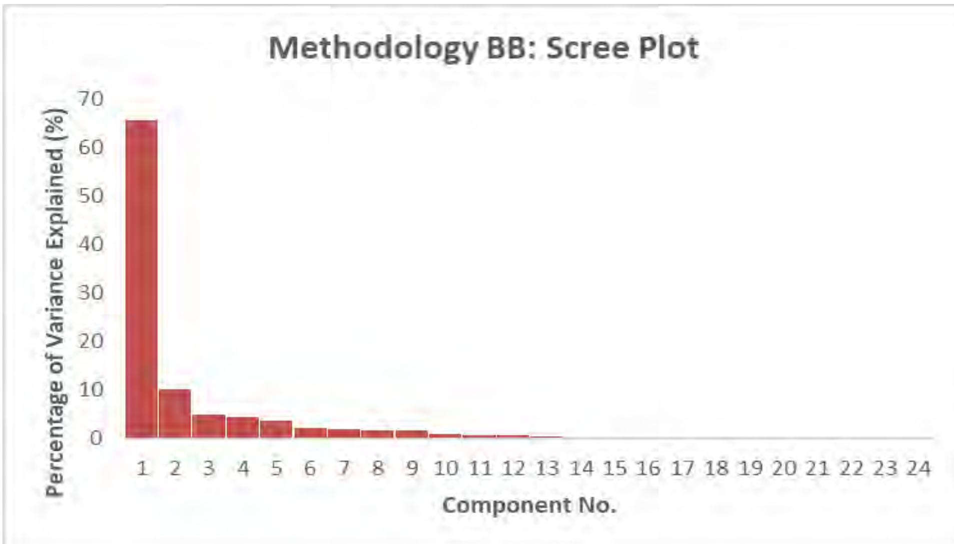
A2. Scree Plots

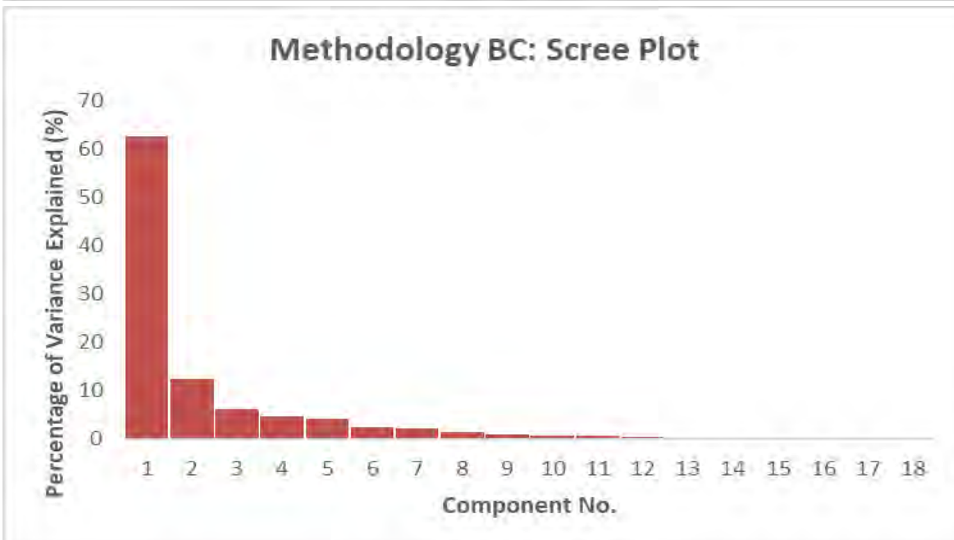


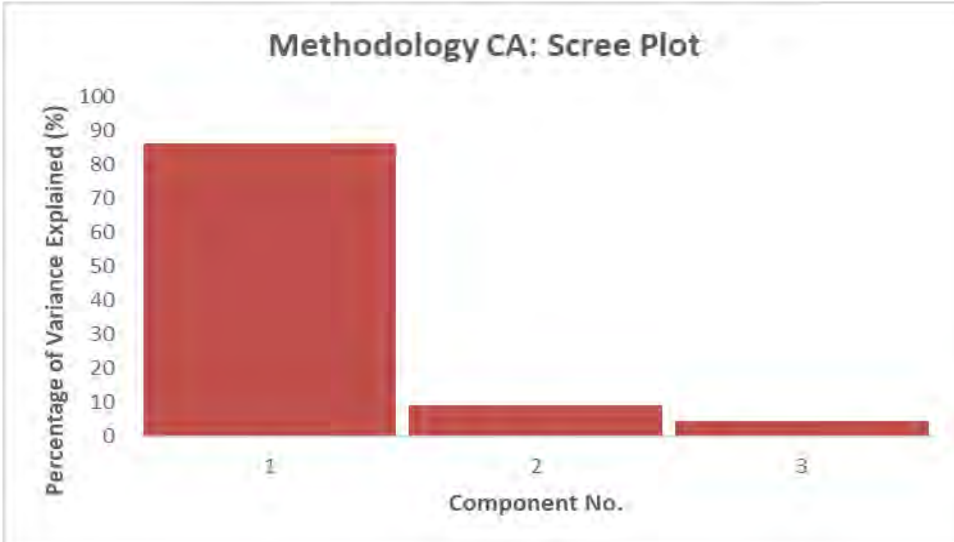


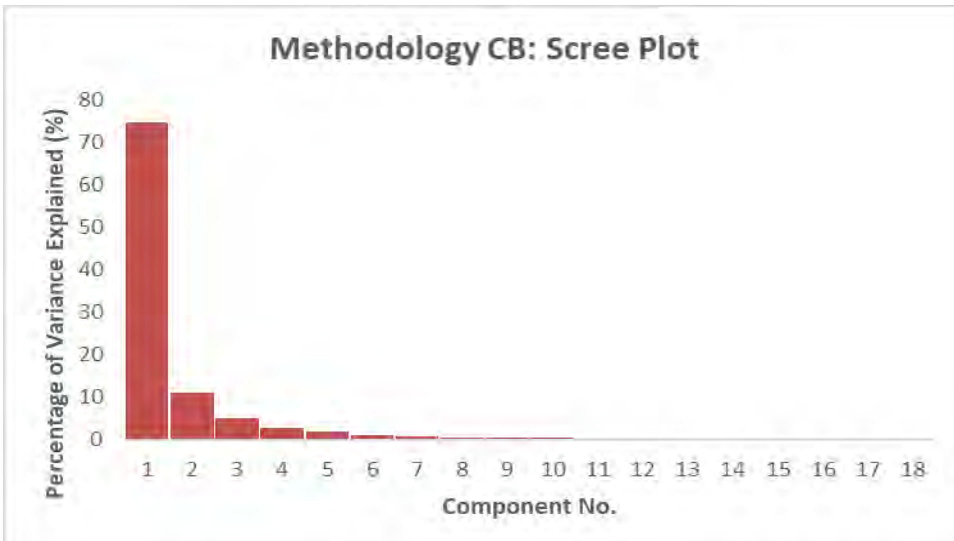

Methodology AC: Scree Plot
Percentage of Variance Explained (%)
60
50
40
30
20
10
0
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29 31 33 35 37 39 41 43 45 47 49 51 53 55 57 59
Component No.

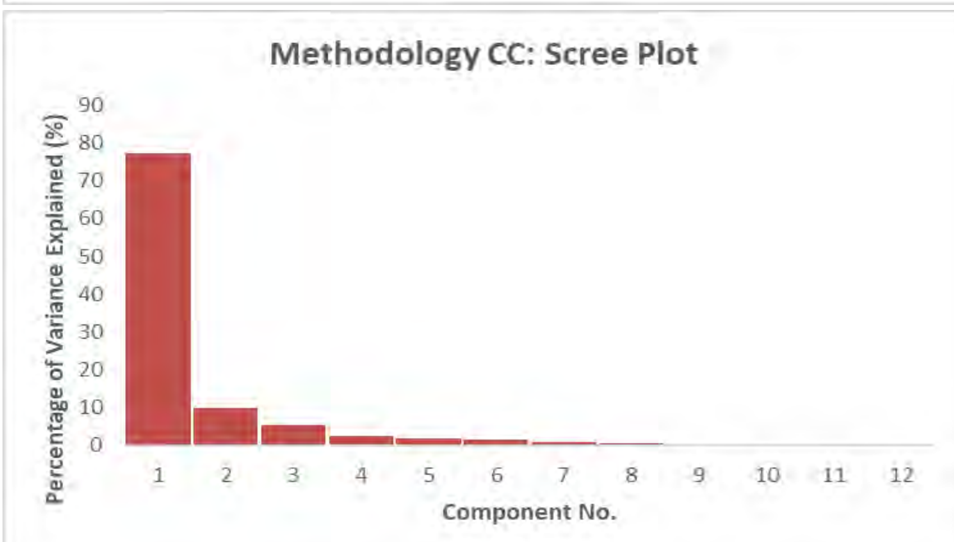

Methodology BA: Scree Plot
Percentage of Variance Explained (%)
80
70
60
50
40
30
20
10
0
1
2
3
4
5
Component No.


Methodology BB: Scree Plot
Percentage of Variance Explained (%)
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24
Component No.


Methodology BC: Scree Plot
Percentage of Variance Explained (%)
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Component No.


Methodology CA: Scree Plot
Percentage of Variance Explained (%)
100
90
80
70
60
50
40
30
20
10
0
1
2
3
Component No.


Methodology CB: Scree Plot
Percentage of Variance Explained (%)
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
Component No.


Methodology CC: Scree Plot
Percentage of Variance Explained (%)
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12
Component No.