# Chinese monetary policy and text analytics: connecting words and deeds*

Jeannine Bailliu[2] , Xinfen Han[1], and Barbara Sadaba[†1]

[1] *Bank of Canada*
[2] *StatsCAN*[‡]

November 15, 2024

**Abstract**

We propose a novel approach to estimating the People's Bank of China (PBOC) monetary policy rule. Given China's complex monetary policy framework, the PBOC's actions are not well captured by standard monetary policy rules that can be used for its Western counterparts. This calls for new approach able to capture the complexities underlying the PBOC's observed behavior. This paper examines whether information provided in PBOC's official statements can help find their policy rule in practice. We use Latent Semantic Analysis to extract content information embedded in PBOC Monetary Policy Reports and examine whether these are significant in explaining Chinese monetary policy actions.

**JEL Classification:** E52, E58, C63
**Keywords:** monetary policy, central bank communication, latent semantic analysis

*The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

†Corresponding author: International Department Bank of Canada, 234 Wellington St W, Ottawa K1A 0G9, Canada. Email: BSadaba@bank-banque-canada.ca. Website: www.barbarasadaba.com

‡Statistics Canada, 150 Tunney's Pasture Driveway, Ottawa, Ontario, K1A 0T6, Canada.

# 1 Introduction

How does the People's Bank of China (PBOC) conduct its monetary policy in practice? Although a fairly good understanding exists regarding how monetary policy is run in advanced economies, Chinese monetary policy and its quantitative impact on the economy remains largely unknown to the outside world. This is due, in fact, to the complex nature of the monetary policy framework in China. The PBOC relies on multiple monetary policy instruments whose importance has been found to change over time, while simultaneously its monetary policy has been guided by multiple objectives. Additionally, monetary policy actions are collective decisions made by the State Council and not by an independent monetary policy committee as it is the case in western counterparts. The PBOC counts with a monetary policy committee, but it plays only an advisory role in monetary policy decisions. In practice, the role of the committee is to prepare monetary policy plans and then submit them to the State Council for approval.[1] Thus, relative to understanding an advanced economy, agents may find it challenging to infer the systematic pattern underlying the PBOC's monetary policy actions from its observed behaviour.

In essence, China's monetary policy is based on the "Law on the People's Bank of China" passed by the National People's Congress in 1995, which states that the objective of monetary policy is to maintain the stability of the currency and promote economic growth. In a speech in 2016, ex-PBOC Governor Xiaochuan Zhou added to this list the objectives of promoting employment and broadly maintaining the balance of payments, pointing out that the employment objective overlaps with that of economic growth and that the stability of the currency obeys the purpose of promoting economic growth (Zhou, 2016). The annual GDP targets are discussed in the State Council's Report on the Work of Government (RWG) together with other multiple macroeconomic targets of a more suggestive than mandatory nature and which are subordinate to the GDP growth target.[2] Among all these variables, inflation is the crucial indicator for monetary policy, which has been shown by Chen et. al (2016) together with GDP growth to move in tandem with monetary aggregates such as M2. This points to another institutional feature that importantly shapes monetary policy in China: the PBOC officially made monetary aggregates an explicit policy instrument to support economic growth goal. This is possible for the PBOC to implement due to the tight control on commercial banks operations that exists in China.

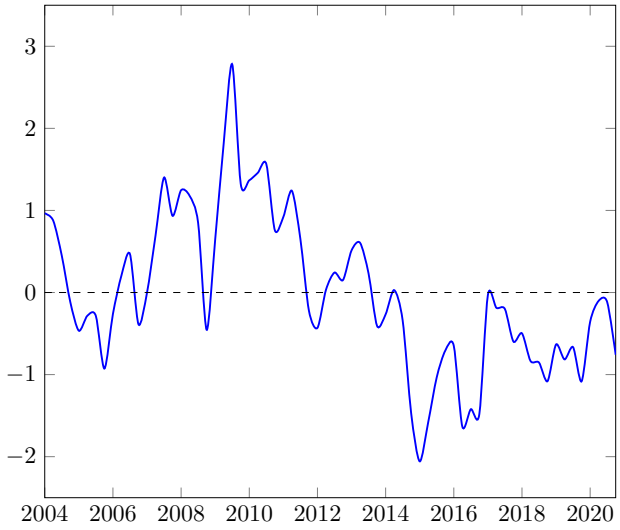
Taking all these institutional features together and following Taylor's (1993, p.197) principle that "...it is important to preserve the concept of a policy rule even in an environment where it is practically impossible to follow mechanically" the use of a quantity-based rule based on GDP and inflation could approximate PBOC systematic behavior in practice. We estimate such a rule and present the resulting residuals of such

[1] Its main monetary policy instruments in recent years include: reserve requirement ratios for banks, various benchmark interest rates, open market operations, and window guidance.

[2] These other targets are inflation, employment, trade balance, currency stability, household income growth and environmental protection.

a regression for China in Figure 1. Contrary to what is expected for the performance of a good policy rule, the plotted series are far from being just a noisy process, they present too much persistence. This shows the presence of some systematic component in the residuals that rules out their interpretation as just random monetary policy surprises. Therefore, the use of a standard rule to depict of the PBOC operational behavior seems to be missing out on some important systematic component.

Figure 1: The evolution of the residuals of a standard quantity-based monetary policy rule for China.




Notes: Blue line is the 3-months moving average of the residuals $\varepsilon_t$ in eq.: $\Delta TSF_t = \beta_0 + \beta_c \Delta TSF_{t-1} + \beta_y y_{t-1} + \beta_\pi(\pi_{t-1} - \pi^*) + \varepsilon_t$. $\Delta TSF_t$ are quarterly changes in Total Social Financing, $y_{t-1}$ is the output gap and $(\pi_{t-1} - \pi^*)$ are quarterly changes in CPI in deviations from target.

Nevertheless, the existing literature has focused on different forms of standard McCallum or Taylor rules or hybrid forms of both, to understand how monetary policy in China works in practice (see e.g., Burdekin and Siklos, 2008a, Fan et al., 2011, Sun et al., 2012, Chen et al., 2016, 2017, Klingelhöfer and Sun, 2018). To the best of our knowledge no existing framework has been able to accommodate all institutional particularities into a monetary rule for the PBOC. These particularities can potentially largely affect the systematic implementation of monetary policy in China.

Given China's prominent position in the world economy and the need to understand monetary policy in transition economies in general, improving our knowledge on this topic constitutes a very pressing matter. In this paper we propose a novel and tractable approach to fully characterize China's monetary policy rule able to handle all unique institutional arrangements in China. We make several contributions. First, we identify and estimate a missing component in the systematic part of a standard monetary policy rule for China. Our proposed framework is flexible enough to accommodate any complex unobservable influences unique to the Chinese monetary policy framework not accounted for in simple policy rules widely used for advanced economies. We start from an alternative quantity-based rule in the spirit of Taylor (1993, 2000) and

extended it to account for potential unobservable effects wrongly attributed to the residual term in standard approaches.[3] Quantity based rules for China are largely supported by the existing literature (Sun et al., 2012, Chen et al., 2016, 2017). The resulting unobserved component confirms the existence of a persistent element that do not constitute random surprises and is not included in growth and inflation measures.

Second, we present evidence that this missing component is linked to the role of monetary policy in supporting Chinese Government's economic policy agenda of supply-side structural reform (SSSR). To show this, we use text analytics techniques to derive empirical measures of the information content—so-called topics—present in official PBOC statements. That is, we learn about the PBOC's intentions and objectives through what they chose to communicate in its official Monetary Policy Reports (MPR). Due to the complexities attached to the extraction of information content from PBOC statements, our third contribution is methodological. We utilize a very powerful text analytics technique that can deal with complex textual structures, Latent Semantic Analysis, to measure the topics they are talking about in each particular document. We then isolate the set of topics that explain the largest share of the information contents in each statement and evaluate which significantly explain our estimated missing component. Our last contribution is related to the impact of this missing systematic component in the monetary policy transmission mechanism. To this end, we include this estimated component in a standard SVAR model together with the output gap, inflation gap and Total Social Financing growth.

Existing literature searching for appropriate simple monetary policy rules for China and/or developing better alternatives is not vast. Most existing works simply evaluate the performance of either a mix of price and quantity based rules or an expanded version of standard rules. No general consensus exists, however, as to which best approximates China's monetary policy rule. Li and Liu (2017) show that within a canonical New-Keynesian DSGE model for quarterly data from 1996Q2 to 2015Q4, an expanded Taylor rule obtains the best empirical fit to the data. The authors also find that money plays an important role as a monetary policy target in China. In a similar fashion, Liu and Zhang (2010) use a New-Keynesian model to show the inadequacy of both the Taylor and McCallum rules to capture PBOC monetary policy. The authors find that a hybrid rule combining both price and quantity based rules outperform other alternatives for the case of China. The hybrid-type rule has been favoured in the literature together with those that allow for structural breaks. In particular, Klingelhöfer and Sun (2018) find that allowing for structural breaks modelled through a regime-changing model is a good approximation to PBOC's behavior. Some other works, have explored which should be the target variable used in such a rule, i.e. which monetary aggregate is the most appropriate. For instance, Chen et al. (2016, 2017) and Burdekin and Siklos (2008b) find support for the use of M2 as the most appropriate monetary measure in a quantity-based type monetary rule for China. Other work have tried to infer PBOC behavior by focusing instead in obtaining a direct measure the

[3] See also, McCallum (1993, 1999, 2002)

PBOC's monetary policy stance (see e.g., He and Pauwels, 2008, Xiong, 2012).

Overall, one common feature of these studies is that they draw on the literature on estimating monetary policy rules in advanced economies and hence rely only on traditional indicators that usually serve as good predictors of monetary policy drivers in these countries—i.e., the output gap, GDP growth, deviations of inflation from target, and inflation expectations. However, this literature has yet to find a monetary policy rule for China that is successful in explaining PBOC's actions. As noted by Huang et al. (2019), "...empirical findings on China's monetary policy rules are inconclusive." Indeed, an element from the systematic component in the policy rule needed to accurately describe the PBOC's behaviour seems to be missing.

Our paper is also closely related to the literature on central bank communication. More specifically, to that that examines whether information extracted from official central bank statements can help agents to better understand central bank actions in practice. Previous studies have looked at PBOC communication but relied on manual approaches to extract the information content from official documents (Garcia-Herrero and Girardin, 2013, Sun, 2013, Shu and Ng, 2010). By relying on text mining techniques, our approach is superior in the sense that it is less likely to suffer from human bias attached to manual alternatives.

Our paper yields several interesting findings. First, our results support the view that words can help explain deeds in the case of the PBOC. We find evidence that some topics extracted from the PBOC's official documents help explain monetary policy actions in the context of a standard monetary policy rule. Thus, we find that official communication is an important tool to help explain the PBOC's actions to agents. Second, our estimate of the unobserved component points to a persistent and systematic element that is missing in standard monetary policy rules for China. Our topical analysis reveals that the hidden systematic component in the Chinese monetary policy rule is related to the government's supply-side structural reform agenda, including the containment of financial stability risks. Our results thus suggest that the PBOC uses monetary policy as a tool to achieve the government structural reform objectives. Finally, the results from our SVAR validate these findings by providing evidence of the importance of government's supply-side reform objectives for the conduct of monetary policy and for macroeconomic outcomes.

Our paper is structured as follows. In Section 2, we present the simple framework used to estimate the missing component in the monetary policy rule for China. Also, we show how incorporate PBOC communication to provide meaning to this component. Section 3 provides details on our dataset and an overview of our PBOC official document set. Section 4 describes the text analytics methodology used to extract information from the document set. In Section 5, we present and discuss key results. Section 6 discusses the analysis based on the SVAR framework, focusing on our key results. Section 7 concludes.

# 2 How words can help explain monetary policy actions: a simple framework

## 2.1 Quantity-based policy rule

The intended monetary policy variable of any central bank is a combination of a systematic component capturing the policy-makers' reaction function to their information set and an exogenous component reflecting unexpected changes in monetary policy, say, monetary policy surprises. This would translate into a rule that looks like this:

$$S_t = f(\Phi_t) + \varepsilon_t,$$

where $S_t$ is the monetary policy variable or intermediate target used to achieve the central bank goals, $f(\Phi_t)$ is the systematic monetary policy function with information set $\Phi_t$ and $\varepsilon_t$ should reflect unexpected movements in the policy variable purely driven by random occurrences, i.e. capture remaining noise. In this section, we present the empirical specification we consider to characterize the systematic component $f(\Phi_t)$ driving PBOC actions in such a way that the residual term, $\varepsilon_t$, is limited to resemble effective policy surprises.

Specifically, we estimate a tractable and flexible monetary policy rule able to capture the specificities of the Chinese economy. In the spirit of Taylor (1993, 2000), our specification consists of a quantity-based policy rule that we extend to accommodate an omitted component that captures all unobservable influences other than output and inflation gaps. To this end, we estimate a time-varying parameter model with stochastic volatility as follows:

$$S_t = \gamma_0 S_{t-1} + \gamma_{\pi t}(\pi_{t-1} - \pi^*_{t-1}) + \gamma_{yt} y_{t-1} + \mu_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, e^{h_t}), \tag{1}$$

$$h_t = \tau + \phi(h_{t-1} - \tau) + \beta S_{t-1} + \varepsilon^h_t \qquad \varepsilon^h_t \sim \mathcal{N}(0, \sigma^2), \tag{2}$$

where the intended monetary policy variable is equal to Total Social Financing such that $S_t = \Delta TSF_t$ for each period $t$. One unique feature of monetary policy in China is the use of credit growth as the key intermediate target to support the growth of gross domestic product (GDP) beyond annually targeted real GDP growth. Total Social Financing, therefore, constitutes a broad measure of aggregate credit in the Chinese economy able to capture the impact of the multiple monetary policy instruments used by the PBOC.[4] Furthermore, the use of aggregate credit measures such as TSF as an appropriate reference indicators for China has been explicitly supported by current PBOC Governor Mr. Yi Gang. He stated in a lecture at Chang'an Forum in Tsinghua University in Beijing on December 13th, 2018 that "In China, M2 is chosen as the intermediate goal

[4]This includes: reserve requirement ratios, benchmark interest rates, open market operations, targeted lending facilities, window guidance.

of monetary policy, as it can be measure and controlled. However, as an economy becomes more and more developed and market-oriented, the correlation between M2 and the real economy will decline. Therefore, in 2012 we adopted aggregate financing to the real economy as a reference indicator, which measures the financing offered by the financial sector the the real economy, covering loans, bonds, stock, trust, etc." Existing literature has, up to date, relied on monetary aggregates such as M2 as the preferred intermediate target. However, this was not without extensive criticism. The use of M2 as an appropriate intermediate target has been put into question due to the instability in the velocity of money during the past several decades in China (see e.g., Xia and Liao, 2001, Xie and Luo, 2002, Li and Liu, 2017). To acknowledge the widespread use of M2 in existing works, we conduct a robustness check using this monetary aggregate measure as our dependent variable in eq.(1). The exercise and results are discussed in Section 7.

On the right hand side of eq.(1), $\pi_{t-1}$ is the quarterly changes in Consumer Price Index (CPI) in differences from $\pi^*_{t-1}$, the annual inflation target interpolated for quarter $t-1$; $y_{t-1}$ is the output gap computed as the difference between annual real GDP target set by the central government (kept constant over the quarters) and real GDP in each quarter $t$. The error term $\varepsilon_t$ comes to represent the non-systematic part of central bank behavior, i.e. unexpected changes in monetary policy. The disturbances, $\varepsilon_t$ and $\varepsilon^h_t$, are mutually and serially uncorrelated. The log-volatility $h_t$ follows a stationary AR(1) process with $|\phi| < 1$ with unconditional mean and variance given by $h_1 \sim \mathcal{N}(\tau, \sigma^2/(1-\phi^2))$. The model in eq.(1)-(2) allows for all coefficients, i.e. $\gamma_{\pi t}$, $\gamma_{yt}$, $\mu_t$, to be time-varying. This feature is empirically important for the estimation of monetary policy rules in general and for our case in particular. First, it has been shown that the importance placed by monetary authorities to both the inflation and/or output gaps might be subject to non-negligible changes over time (see e.g., Chen et al., 2016, 2017, Klingelhöfer and Sun, 2018). Second, the time-varying intercept $\mu_t$ represents here the unobserved systematic component in $f(\Phi_t)$ that captures all other influences driving changes in $S_t$ different from the standard inflation and output gaps. We want to ensure that this time-varying measure is not reflecting any potential parameter instability over time.

The vector of time-varying coefficients $\boldsymbol{\theta}_t = (\gamma_{\pi t}, \gamma_{yt}, \mu_t)'$ in turn evolves according to a random walk process:

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\varepsilon}^{\boldsymbol{\theta}}_t \qquad\qquad \boldsymbol{\varepsilon}^{\boldsymbol{\theta}} \sim \mathcal{N}(0, \boldsymbol{\Omega}), \tag{3}$$

where $\boldsymbol{\Omega}$ is a 2x2 covariance matrix. Following Chan (2015), we allow for a generic correlation structure among the innovations to the random-walk coefficients. The process in eq.(3) is initialized with unconditional moments given by $\gamma_1 \sim \mathcal{N}(\gamma_0, \Omega_0)$.

### 2.2 Meaning of the unobserved component $\mu_t$

Next, we set out to investigate what is the meaning of this missing component $\mu_t$. To this end, we use state-of-the-arts text analytics techniques to extract the information content provided in official PBOC communication published in quarterly Monetary Policy Reports. The information is extracted in the form of topics, which reflect the main informational content of different sections of the reports. Armed with this topics, we then investigate which of these are significantly related to the unobserved component $\mu_t$ in order to provide this series with meaningful meaning. More specifically, we regress our estimated $\mu_t$ on the set of extracted topics from the Summary section of all PBOC's MPRs in our sample. This way, we explore whether official PBOC communication can be useful in helping agents better understand this additional systematic component in the PBOC monetary policy rule and, thus, how monetary policy in China works in practice. This is to say, we explore whether "words" can help help uncover the hidden meaning of "deeds" for the case of China. We estimate the following regression:

$$\mu_t = \beta_0 + \sum_{i=1}^{J} \beta_i topic_{t-1}^i + \boldsymbol{\delta} \boldsymbol{X}_{t-1}' + \epsilon_t, \tag{4}$$

where $topic_{t-1}^i$ (where $i = 1, ..., J$) represents the $ith$ topic (i.e., factor) of a set of $J$ topics extracted from the Summary section of official PBOC Monetary Policy Reports related to monetary policy actions and $\epsilon_t$ is the measurement error term. In order to rule out the impact of changes in employment conditions and currency fluctuations as potential explanations for $\mu_t$, we include $\boldsymbol{X}_t = (\Delta U_t, \Delta ER_t)'$ in eq.(4) where $U_t$ represents the unemployment rate and $ER_t$ the exchange rate. There is reason to believe that these two variables could be captured by $\mu_t$. On top of the two legal objectives of PBOC monetary policy framework of maintaining currency stability and promoting economic growth, a further breakdown present four more specific parts: maintaining price stability, boosting economic growth, promoting employment, and broadly maintaining balance of payment (Chen et al., 2016). The importance of employment and exchange rate for monetary policy in China has been shown in the literature. Bailliu et al. (2019) has shown that a newly estimated labour market indicator has important implications for credit growth decisions in China. Also, Lardy (2005) and Mehrotra and Sánchez-Fung (2010) highlight the relevance of exchange rate considerations for monetary policy decisions. To the extent that these two additional influences can have an impact in the monetary policy rule, we have to control for them in eq.(4).

## 3 Data

Our dependent variable is the Total Social Financing (TSF) variable, which is constructed using monthly levels and then computing the quarterly growth rate using a 3-month moving average. The series included in TSF are obtained from the PBOC official website and the resulting series is shown in Figure A-1. The

inflation rate is defined as the quarterly growth rate in the headline consumer price index (CPI). The output gap is defined as the difference between actual real gross domestic product (GDP) and real potential GDP (in percentage terms). GDP and inflation are obtained from the National Bureau of Statistics of China. The target inflation rate was obtained from the National People's Congress. All data was downloaded via Haver Analytics. Finally, in our topical regression we use data on unemployment rates (UR) kindly provided by the authors of Bailliu et al. (2019) obtained from own estimations and the exchange rate (ER) series are obtained from the PBOC official website through Haver.

Our document set consists of official monetary policy reports (MPRs) published on a quarterly basis by the PBOC over the period from 2003Q1 to 2019Q4. Each MPR is structured in a similar way consisting of a summary section as well as of five other sections covering the following topics: money, credit, and financial market developments; a description of monetary policy operations; an overview of recent macroeconomic developments; and an outlook of the Chinese economy and monetary policy. We provide a more detailed description of the documents in Table A-1 in the Appendix. We use the original version of the documents written in Chinese, as the English translations are of poor quality. Thus, we conduct our text analytics on Chinese-language documents.

The Chinese MPRs tend to be more backward-looking than similar documents produced by central banks in advanced economies. This is due to PBOC's communication being more restricted in nature given that the PBOC decision making process is subjected to the decisions made by the State Council. Because both the outcome and the timing of important monetary policy decisions are uncertain, the PBOC is more limited in the forward guidance that it can provide. Despite these drawbacks, official communication can nonetheless be useful via its role in explaining past monetary policy decisions; and in doing so, it can help to shed some light on the PBOC's monetary policy framework.

We note that the time range of the indicators data—TSF, CPI, GDP, UR, ER—goes from 2003Q1 until 2020Q4, which is a larger time span than that covered by the MPR documents due to availability. Therefore, all estimations in the paper are conducted using data until 2020Q4 and shown accordingly in the corresponding charts.However, the analysis conducted with information content from MPR documents is conducted using data until 2019Q4.

# 4 Estimation method

## 4.1 Estimation of $\mu_t$ and topical regression

The specification described in eqs.(1)-(3) constitutes a stochastic volatility time-varying parameters model. In this setting, we allow the coefficients in the conditional mean to be time-varying, which includes the intercept $\mu_t$. This model is then estimated using an efficient MCMC sampler—based on fast band and

sparse operations instead of the more traditional Kalman filter as in Koopman and Hol Uspensky (2002)—as proposed by Chan (2015). For more details we refer to Chan (2015) who provides a full description of the model along with the exact details on the Bayesian algorithm used in the estimation together and the steps involved.

The time-varying coefficients with bootstrapped standard errors in eq.(4) accounts for potential structural instability in the coefficients on inflation and output gaps. Allowing for time-varying changes in the importance given to fluctuations in inflation and output in the monetary policy rule has been shown to be relevant in the case of China (Chen et al., 2016, 2017, Klingelhöfer and Sun, 2018). Finally, modelling the stochastic volatility accounts for time-varying volatility in macroeconomic variables.

## 4.2 Measuring Topics using LSA

To discover the set of Topics for each section of the PBOC statements we use the Latent Semantic Analysis (LSA) algorithm. This is a popular technique used for knowledge representation developed by Landauer and Dumais (1997) and it is based on a mathematical matrix decomposition technique called singular value decomposition (SVD). LSA is a technique in natural language processing that involves analyzing relationships between a set of documents and the words contained therein. The product of this analysis is a set of concepts related to the documents and words contained in them. We selected the LSA methodology because it was found to outperform other commonly used text analytic alternatives in our particular setting.[5]

Our methodology consists of several steps that are described in detail in the subsections below. Figure 1 in the Appendix presents a graphical summary of the procedure.

## 4.3 Pre-processing documents

As described in Section 3, each MPR is organized in a similar way in that it includes a summary and five other sections. This structure provides a natural way to create six corpora from the MPR documents (see Table A-1). We believe that creating six corpora using the sections provides us with a richer set of information than would treating the MPR document as just one corpus.

We then proceed with the pre-processing for each corpus, which includes removing stop words, punctuation, numbers, and special characters, as well as segmenting Chinese text into words. The text segmentation process is more involved than it would be for English text because Chinese text has no spaces between characters and a character, on its own, may not form a meaningful unit. Indeed, a large proportion of Chinese words are made up of two or more characters. In order to sort Chinese characters into words, we rely on a natural language processing software, Harbin LTP (Che et al., 2010). It is worth

[5] We found LSA to outperform Latent Dirichlet Allocation (LDA), given our particular setting and the small set of documents under analysis. LSA has been found to outperform LDA for smaller-scale databases (Cvitanic et al. (2016)). Additionally, our documents are in Chinese language and they are focused on a fairly narrow set of topics (i.e., topics related to monetary policy) can further explain LSA outperformance of LDA in our setting.

noting that we trained our own language model and data dictionary to extend the software's functionality to segment the Chinese text into meaningful words in this context (i.e., documents focused on monetary policy).

### 4.4 Transforming the text into a numerical matrix

Once the text has been pre-processed, it then needs to be transformed into a numerical matrix. Each document is first represented as a "bag-of-words" vector $[t_1, t_2, \ldots, t_j, \ldots t_m]$ that contains all $m$ unique words that are present in the corpus, where $t$ indicates how often the $jth$ word appears in the document. We use up to a 3-gram sequence to construct the bag-of-words vector.[6] The bag-of-words vector is then used to construct the term-frequency matrix $tf(n, m)$, where $n$ is the number of documents and $m$ is the number of unique words in the corpus. The term-frequency matrix essentially presents the distribution of unique words across all documents. To diminish the weight of words that occur frequently and increase the weight of those that appear rarely, the term-frequency matrix is multiplied by the inverse document frequency ($idf$) to obtain $tfidf$ matrix. The $idf$ measures the importance of a word in all documents in the corpus and is calculated as follows:

$$idf = \log \frac{Number\ of\ documents\ n\ in\ the\ corpus}{Number\ of\ documents\ in\ the\ corpus\ in\ which\ term\ j\ occurs}. \tag{5}$$

The re-weighting of $tf$ by $idf$ is to diminish the importance of words that occur very frequently in the documents but that carry little meaning. It increases the importance of words that appear rarely but are very meaningful.

### 4.5 Applying the LSA algorithm to extract topics

We construct the $tfidf$ matrix $Z$ for each corpus. We then use the LSA algorithm to transform the matrix into SVD components. SVD is a generalized form of principal component analysis. In SVD, the matrix $Z$ is decomposed into the product of three other matrices: $Z = USV^T$ (as shown in Figure A-3). The matrix $U$ describes words ($w$ rows) as vectors of the derived orthogonal factor values ($t$ columns) and the $V^T$ matrix describes the documents ($d$ columns) as vectors of the same factors ($t$ rows). These factors can be thought of as underlying topics that run through the documents. The meaning of each word or document can then be characterized by a vector of weights indicating the importance of each of these underlying topics. The $S$ matrix represents the importance of each topic for explaining the variance of meaning across the documents. With the elements of $S$ ordered by decreasing magnitude, the first topic is thus the most important one. If

[6] In other words, our bag-of-words vector for each document includes frequency counts of one word, two contiguous words, and three contiguous words.

the documents are ordered chronologically, then a row of $V^T$ represents a time series of a given topic and a column of $V^T$ denotes the weight of each topic in a given document.

LSA uses a k-dimensional approximation of the $Z$ matrix, $\hat{Z}_k$, by using the first $k$ columns of $U$ and $V$ and the k$x$k upper-left matrix of S. This approximation removes extraneous information that is in the document set and focuses only on those factors explaining the important variation in meaning across documents. The matrix $\hat{Z}_k$ is the least-squares best fit of $Z$. By performing the SVD and truncating it, we are able to capture the important underlying semantic structure of the words and documents while excluding the noise.

We derive $k$, the number of topics, in each corpus by using a topic coherence measure. More specifically, we use the topic coherence measure ($C_V$) as described in Röder et al. (2015) and implemented in the Python Gensim library (Řehůřek and Sojka, 2010). $C_V$ quantifies the relations between the top $n$ topic words and is computed as the sum of pairwise scores on the top $n$ words in a topic. Intuitively, $C_V$ captures how often the top $n$ topic words appear together in the corpus. The coherence score for each topic is then aggregated to an overall score for the topic model. The higher the overall score is, the better the topic model captures the semantic coherence. The coherence measure is intended to improve the interpretability of the topics so that they can be better understood by humans. We conduct a grid search on the number of topics for each corpus and keep the topic model that has the highest coherence score.

# 5 PBOC rule and topical analysis: key results

## 5.1 Monetary policy rule results

We estimate the model laid out in eq.(1)-(3) to obtain our unobserved component $\mu_t$ shown in Figure A-4. The estimation results are presented in detail in Table A-2. The series of estimated time-varying coefficients together with the stochastic volatility series are shown in Figure A-5. Overall, our results are in line with the existing literature exploring monetary policy rules for China (Chen et al., 2016, 2017, Girardin et al., 2017), with the coefficients on the inflation and output gaps being statistically significant and of the expected negative sign. This supports the claim that the PBOC has been following an anti-inflation policy since the early 2000s. In the context of our monetary policy reaction function, these results imply that the PBOC responds by tightening monetary policy when inflation moves above the official target or when output grows above its potential. Monetary policy can be tightened using a variety of instruments, but regardless of the instrument(s) used, the tightening will be reflected in a decline in the TSF growth gap.

## 5.2 MPR Summary section Topics in eq.(4)

Next, to analyse the information content of MPR documents and relate it to our unobserved component $\mu_t$ we proceed in two steps. First, we use the LSA methodology to extract the information content—extract

the topics—in the MPR Summary section. The Summary section provides a comprehensive overview of each MPR report and therefore, this corpus should contain sufficient information to identify the broader set of topics that capture the essence of the PBOC communication in official documents. We use the four extracted topics as right hand side variables in eq.(4) to identify which of these are statistically significant explanations for the unobserved component $\mu_t$. In a second step discussed in the next section, using the same methodology, we extract the information content embedded in Section three—financial market conditions—and Section four—macroeconomic conditions—of the MPR documents. The extracted topics of each MPR Section are used to conduct a more granular analysis and provide a better interpretation of the information content and thus, improve our understanding of the Summary topics found to be significant in eq.(4).

The results of estimating eq.(4) are presented in full in Table A-3. By looking at the coefficients of the MPR Summary topics we find that only Topic 1 and 4 are statistically significant in explaining $\mu_t$. In order to assess the underlying meaning of these two topics, we explore the corresponding word clouds shown in Figure A-6. These word clouds show the key words attached to these two topics and the size of each key word represents the weighted frequency each word (TFIDF) is found to have in the documents, i.e. the larger the font of the key word, the higher the frequency of occurrences of the key word in the analysed document. Although originally obtained in Chinese language, the word clouds are presented here translated to English language.

**MPR Summary Topic 1** The key words in the word cloud for Topic 1 in panel (a) show that this is a topic that focuses on supply side structural reforms (SSSR). The larger words found in the cloud are linked to this particular reform: *"structural supply side"*, *"social financing"*, *"structural adjustment"*. SSSR is a key component of China's economic policy agenda and is linked to its continued transition from a manufacturing-heavy economic model to one that is led by services and consumption. It consists of an overreaching suite of policies focused on reducing distortions in the supply side of the economy and upgrading the industrial sector. The reforms aim to promote advanced industries and innovation, reduce capacity in heavy industrial sectors (e.g., coal and steel), resolve zombie firms, and reduce property inventories (Boulter, 2018). The coefficient on Topic 1 is positive suggesting that monetary policy is loosened in response to a change in this topic.

**MPR Summary Topic 4** On the other hand, the key words in the word cloud for Topic 4 in panel (b) indicate that this is a topic that focuses on window guidance. The larger words found in the cloud are linked to regulation and guidance provided to commercial banks: *"according the State Council"*, *"commercial bank"*, *"moderately loose"*, *"window"*, *"policy targeting"*. Window guidance represents the ability of the PBOC to intervene in commercial banks lending activity and it constitute an important instrument of monetary policy by Chinese monetary authorities. Even after the Chinese economic reform program starting in 1978, the

financial system retained characteristics of a centrally planned economy. The PBOC employed a rigorous credit plan to control the extent and orientation of lending of the major state-owned banks in order to exert control over the amount of lending at each bank. As of 1995, a less binding scheme was introduced and the PBOC started to announce only an annual reference target for the total amount lent by the Chinese banks as a whole. The new target functioned as a reference value on which the commercial banks relied in their own budgetary planning. Subsequently, the PBOC started moving away to other instruments to head toward the goal of achieving price stability and economic growth including interest rate policy, open money market operations, and reserve requirements. However, it retained its ability to intervene in bank lending activity using the so-called "window guidance". The role of window guidance was stated in a China MPR in 2006 as follows: *"By communicating with the commercial banks and other financial institutions on a regular basis, the central bank is in a better position to have the market anticipate its monetary policy and thus to make its policy more effective. China's experience in recent years indicates that improving transparency through window guidance is not only conducive to reducing costs of monetary policy operations, but also to helping the central bank realize its policy objectives and enhance the effectiveness of monetary policy."* Lastly, the coefficient on Topic 4 has a negative sign implying that monetary policy is tightened in response to a change in this topic.

## 5.3 MPR Section 3 and 4 Topics

In order to obtain more granular information to improve our understanding of the meaning of these summary section topics, we conduct further analysis using the information content or topics of Sections 3 and 4 of the MPR documents. More specifically, we use the set of topics extracted from Sections 3 and 4 to conduct Granger causality tests on the Topics 1 and 4 of the summary section.[7] In order to account for changes over time in the way monetary policy was conducted in China during our period of interest, we extract topics of Section 3 and 4 for three different time periods: the pre-GFC period 2003-2007, the GFC period 2008-2014 and the post-GFC period 2015-2019. Finally, for each time period, we analyse the word clouds of the topics that are found to Granger cause the two summary section topics so that we improve our interpretation of the unobserved component $\mu_t$.

### 5.3.1 MPR Section Topics on Summary Topic 1: SSSR

The Granger causality results are presented in Table A-4. We first discuss the findings for MPR summary Topic 1. For the period before the GFC, 2003-2007, we find that while no topic related to Section 3 seems to be significantly related to Topic 1, only Topic 10 of Section 4 Granger causes MPR summary Topic 1 during this period. By looking at S4-Topic 10 word cloud in Figure A-7, we find that it focuses on the Severe

[7] We use one lag for each variable in the Granger causality tests since there is no valid reason to believe that same document information content should be related by larger time spans.

Acute Respiratory Syndrome (SARS). SARS was first discovered in Asia in February 2003. The outbreak lasted approximately six months as the disease spread to more than 31 countries around the world before it was stopped in July 2003. Although the virus ultimately affected multiple countries, China was among the three worst hit countries suffering non-negligible negative economic consequences (Xu, 2003). Given the massive impact of this virus outbreak, it is not surprising that it shows up as significantly related to our summary Topic 1 related to SSSR. However, we do not put too much weight on this finding since aside from this particular topic, none else was found significant. This is unsurprising since the SSSR was first officially announced in late 2015 by President Xi and in practice, these reforms were introduced in response to the Global Financial Crisis evolving thereafter into a major component of the Chinese Government's economic policy framework.

Results differ greatly for the GFC period, for which we find four MPR section topics to Granger cause MPR summary Topic 1. Two of these correspond to Section 3, Topic 2 and 3 are significantly related to summary Topic 1. The corresponding word clouds for these two topics are presented in Figure A-8. Analysing the key words for S3-Topic 2 we find it focuses on construction and institutional arrangements, while S3-Topic 3 focuses on growth and development prospects. In particular, key words for S3-Topic 2 are are *"institutional"*, *"analyse"*, *"construction"* while those for S3-Topic 3 are related to growth and development prospects: *"grow fast"*, *"stable development"*, *"insurance"*, *"steady and rising"*. The remaining two topics correspond to Section 4, Topic 3 and Topic 7 are significantly related to summary Topic 1. The word clouds are shown in Figure A-9. Based on the word clouds, both topics greatly overlap in terms of meaning both focusing on domestic and external economic conditions. Specifically, most relevant key words for S4-Topic 3 are *"economy real state"*, *"economic status"*, *"market trend"*, *"international economic situation"*, *"balance of payments"* while those for S4-Topic 7 *"economic situation"*, *"further improve"*, *"situation good"*, *"steady development"*, *"balance of payment"*. These topics clearly refer to the actions taken by Chinese authorities due to the outbreak of the GFC. In response to the sharp deterioration in external economic conditions and weakening demand in China during the GFC, several stimulus packages where introduced aimed at promoting infrastructure investment, i.e. construction. This was accommodated through an expansion of bank and non-bank lending. These measures constitute supply side discretionary reforms motivated by the desire to sustain and promote growth during times of economic turmoil.

For the post-GFC period, 2015-2019, we find two significant section topics: Topic 5 of Section 3 and Topic 7 of Section 4. Word clouds presented in Figure A-10 show that S3-Topic 5 focuses on market liquidity and foreign exchange and S4-Topic 7 points instead to structural industry reforms and innovation. In particular, most important keywords for S3-Topic 5 are *"interest rate shock"*, *"forex trading"*, *"repo transactions"*, *"institutional"*, while for S4-Topic 7 we find *"investment slow"*, *"external debt"*, *"industrial robot"*, *"care services"*, *"industry"*. These topics refer to two important aspects of the SSSR, the financial reforms and

the industrial policy.

The essence of SSSR is to help stir Chinese economy into a more market-driven economy. This involves both financial sector as well as economic reforms. As of its China's 2013 reform agenda, the country pledged more market-driven allocation of resources and pricing of capital, a more flexible exchange rate regime, and openness to foreign investors. Although administrative controls remain in one form or another, the interest rate regime was changed fundamentally since 2013 and interest rate liberalization was well under way as of 2015. However, interest rates are not yet fully liberalized hampering the relationship between the demand and supply of liquidity on the one hand and benchmark interest rates on the other. In contrast, interest rates for repo transactions are commonly regarded as the most accurate indicator of the true cost of liquidity. Indeed, repos represent the largest, most efficient and most liquid segment of financial markets. A repurchase agreement, or repo, is a contract for the sale of a security with a commitment by the seller to buy the same security back from the buyer at a specified price on a designated future date.[8] Repo is an important investment option for RMB (Renminbi/Chinese Yuan) money market funds and has been a key driver of increased flexible liquidity in Chinese money market funds. The growth of money market funds assets in China hit a record high of 2.5 trillion Chinese Yuan as of April 2015 driven by the ability to invest in repo (Frank et al., 2002).

Turning to the industrial side, once the effects of the GFC eased and the resulting policy stimulus faded, the industrial sector was left with significant excess productive capacity, declining prices for industrial products and falling profits. By 2015, there was a pressing need to address an unprofitable industrial sector, turning away from heavy industries into the service sector and fostering innovation. As of 2002, the government work report pointed out that high and new technologies and advanced applicable technologies should be adopted to transform and upgrade traditional industries, and must actively develop the tertiary industry, especially the modern service industry (Rongji, 2011). These efforts pay off as China became the world's largest market for industrial robots for eight consecutive years by 2020. In 2020, the manufacturing robot density, a metric used to measure a country's level of automation, nearly doubled the global average. The involvement of local and central institutions in driving all these supply-side reforms was stated clearly in the March 2016 government work report: *"Focus on steel, coal and other difficulty-ridden industries to cut overcapacity. Enterprises as the main body, should be encouraged by the market, organized by the local, and supported by the central government..."* The report also point to the involvement of the authorities in strategically slowing investments in undesired sectors such as heavy-industries like coal and steel.

[8]Repo is effectively a short-term, interest-bearing loan against a pool of collateral, but while repos resemble collateralized loans, their treatment under bankruptcy law is more beneficial to the buyer. That is because a buyer can sell the collateral quickly in the event of a counterparty default.

### 5.3.2 MPR Section Topics on Summary Topic 4: window guidance

Next, we look into the findings for the MPR Summary Topic 4 related to window guidance. For the run up period before the GFC—2003-2007—we find three significant section topics related to the financial and banking sectors and the growth outlook. We find Topic 3 and 4 from Section 3 and Topic 3 of Section 4 to Granger cause Summary Topic 4. By looking at the word clouds in Figure A-11 and A-12, the following keywords stand out: *"RMB interest rate"*, *"derivative product"*, *"forward transaction"*, *"pledge bonds"*, *"outlook"*, *"substantial growth"*, *"continue expand"*, *"income expenditures"*, *"balance of payments"*.

This information content makes sense for the role of window guidance during the period under consideration. At this time, the provision of liquidity was encouraged by Chinese authorities with the aim of meeting the growth target largely by supporting an export led economic model. At the same time, a set of reforms were well underway to continue the transition to a more market-oriented economy. One example of this was the liberalization of interest rates while at the same time to retain some control over credit expansion and allocation, window guidance grew in relevance. For most of this period, window guidance was directed to continue to expand credit growth to support income and expenditures supporting substantial growth. Things took a turn by the end of the period—2007-2008—due to the risk of overheating in the economy. This caused a shift in the aim of window guidance to that of limiting credit growth through the introduction of credit limits. Simultaneously, during 2005-2007 financial and banking transactions were facilitated by the introduction of several types of interest rate derivatives. For instance, RMB interest rate derivatives help financial institutions to smooth out fluctuations in these key economic indicators. Also, several types of FOREX swaps agreements were introduced against the main trading currencies contributing to channel credit to support expanding trade (Yan, 2012). The utilization of financial derivatives by China's commercial banks grew considerably during this years and therefore, it played a key role in supporting the role of the banking sector in the monetary policy transmission mechanism to the economy, heavily guided by window guidance.

For the GFC period we find four MPR section topics to Granger cause summary Topic 4. Specifically, these correspond to Topic 1 and 2 of Section 3 and 4 with corresponding word clouds presented in Figure A-13 and A-14. Overall, these section topics focus on banking and real state. Main keywords are the following: *"bank room"*, *"bank large"*, *"company"*, *"bond"*, *"business trade"*, *"interbank certificate of deposit"*, *"institutional"*, *"construction"*, *"economy real state"*, *"dollar"*, *"international economic situation"*. The clear reference to external issues such as the USD dollar and international situation are unsurprising given the outbreak of the GFC and its influence in domestic policies. As previously mentioned, in response to the sharp deterioration in external economic conditions and weakening demand in China, several stimulus packages where introduced aimed at promoting infrastructure investment, i.e. real state, construction. This was accommodated through an expansion of bank as well as non-bank lending largely guided by window

guidance.

Finally, for the post-GFC period—2015-2019—we find just one topic to Granger cause Summary Topic 4: Topic 2 of Section 3. Word cloud shown in Figure A-15 highlight the following keywords: *"interest rate options"*, *"insurance"*, *"social service"*, *"coupon trading"*, *"principal"*, *"cash"*, *"billion"*. The reference to interest rate options is very interesting here. Short-term interbank rates have assumed greater importance in Chinese monetary policymaking over time. New facilities and operations have been introduced in recent years to improve the provision of central bank liquidity and limit the volatility of interbank rates—e.g. one-day PBOC Standing Lending Facility, the one-year Medium Lending Facility, and seven-day reverse repos (Lee, 2017, Zhao et al., 2017). We also highlight the keyword "insurance" since the insurance financial industry rapidly developed during this period as a key part of China's financial reforms dating back to 1978. Coupon trading took place in China around 2017 due to cash constrained companies trying to raise billions to stay liquid. As more market-based instruments were introduced and th banking sector was liberalized, it is said that the effectiveness of window guidance in China has decreased considerably. This search for market capitalization of large companies in China can be taken as a proof of that given that in the past, this would have been channelled through guided allocation of credit, i.e. window guidance. Finally, in recent years, window guidance has been used to steer credit in the Chinese economy to sectors that have gain relevance due to social interest such as promoting green economies (Dikau and Volz, 2023).

Taken together, the granular information provided by the MPR Section's Topics offer support for the meaning found for Summary Topics 1 and 4, which provide an interpretation of our estimate of $\mu_t$, the unobserved component missing in the Chinese monetary policy rule presented in eq.(1). Our results suggest that the this hidden systematic component is related to the government's agenda of supply-side structural reform and the implementation of discrete interventions through window guidance. This points to the PBOC using other motives rather than standard monetary policy tools as part of their systematic monetary rule.

# 6 SVAR analysis

Next, we investigate the effects of the missing dimension of monetary rule in China that we capture with the unobserved component $\mu_t$. To this end, we estimate a simple four-equation Structural Vector Autoregression (SVAR) comprising the output gap, the unobserved component $\mu_t$, the inflation gap, and the TSF growth gap. We include the variable $\mu_t$ because we believe that it is a good proxy for the relevant topics that were identified. This way, we explore the role of this missing component in the transmission mechanism of monetary policy and we attempt to validate our results from the topic analysis and estimation of the monetary policy reaction function.

The specification of the SVAR model is as follows:

$$Y_t = \Phi(L)Y_{t-1} + \upsilon_t, \tag{6}$$

$$B_0 z_t = k + B_1 z_{t-1} + ... + B_p z_{t-p} + \upsilon_t, \tag{7}$$

where $z_t$= ($y_t$, $\mu_t$, $\pi_t^g$, $S_t$), and $y_t$ is the output gap, $\mu_t$ is the missing dimension unobserved component, $\pi_t^g$ = ($\pi_{t-1}$ - $\pi_{t-1}^*$) is the inflation gap, and $S_t = \Delta TSF_t$ is TSF growth. The vector of disturbances, $\upsilon_t$, which are assumed to be serially and mutually uncorrelated. The matrix $B_0$ governs the contemporaneous relations among the variables in the system.

To identify the structural innovations, we specify a set of restrictions on the contemporaneous interactions among the variables. We achieve this identification by ordering the variables from most exogenous to least exogenous and then by imposing restrictions on $B_0$ to be lower triangular (i.e., a Cholesky decomposition). Thus, $B_0$ will be written as follows:

$$B_0 = \begin{bmatrix} \beta_{11} & 0 & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix}. \tag{8}$$

We estimated the SVAR using two lags of each variable. Overall, the results of our SVAR analysis seem reasonable. This is shown in Figure A-16, which displays the impulse response functions (IRFs) of all the variables to shocks from the other variables in the system. As expected, a positive inflation shock or GDP gap shock would lead to a tightening in monetary policy via the TSF variable—a decline in the TSF growth rate relative to trend. Also, as anticipated, a positive monetary policy shock—a shock to the detrended TSF growth rate equation—would result in an increase in the inflation gap and an increase in the output gap.

The IRFs shown in the second column provide evidence that shocks to the unobserved component, $\mu_t$, equation also impact the other variables in the system consistent with our priors about the monetary policy transmission process. Notably, a positive shock to the unobserved component would lead to a loosening in monetary policy, a widening of the output gap, and an increase in the inflation gap. These results suggest that shocks that affect the government's supply-side reform agenda are important enough to impact monetary policy, the output gap, and the inflation gap. Moreover, the variance decomposition exercise suggests that the government's objectives of supply-side structural reform play an important role in explaining the conduct of Chinese monetary policy. Notably, we find that $\mu_t$ explains over 30% of the TSF gap variable (see Figure A-17 for more details). This is in line with the share explained by the output gap but is more than that explained by the inflation gap. Thus, our results from the SVAR analysis validate our findings from the topic

analysis and estimation of the monetary policy reaction function by providing evidence of the importance of supply-side reform objectives for the conduct of Chinese monetary policy and for macroeconomic outcomes.

# 7 Robustness analysis with M2

The results reported in the previous sections are based on the use of Total Social Financing as our preferred intended monetary policy variable. This is based on several reasons we repeat here for the sake of completeness. First, TSF captures the impact of the multiple monetary policy instruments used by the PBOC. Second, aggregate credit measures of the Chinese economy, rather than monetary aggregates, have been explicitly supported by current PBOC Governor Mr. Yi Gang. Finally, during the past several decades the unstable velocity of money have rendered the growth rates of monetary aggregates poor intermediate targets (Li and Liu, 2017).

However, the use of TSF could still be contested in favour of a commonly used credit measure in existing literature: M2. In order to address this issue, in this section we present a robustness check replacing our intended monetary policy variable, $TSF_t$ by M2. We estimate the exact same model described by eqs.(1)-(3). First, we investigate whether our estimated time-varying parameters $\mu_t$, $\gamma_{yt}$, $\gamma_{\pi t}$ and the stochastic volatility $h_t$ differ considerably when using M2. These results are shown in Figure A-18 in the Appendix. Next, we check the significance of our set of topics on the $\mu_t$ from the M2 regression. That is to say, we explore whether the SSSR interpretation still holds for M2 as the policy variable. We find that the baseline estimations are supported since results do not vary substantially with the use of M2 as the dependent policy variable. The persistence in the unobserved component is still observed in the right-top panel in Figure A-18 although the trend is more pronounced. In Table A-5 we show the results of eq.(4) for the robustness exercise. Parameters do not vary substantially from baseline estimates for TSF. Therefore, the use of M2 confirms the SSSR interpretation found in our baseline estimations.

# 8 Concluding remarks

In this paper we propose a novel approach that relies on text analytics to extract and interpret topics from official PBOC communication and examine whether these can help us to better understand Chinese monetary policy in the context of a standard quantity-based monetary policy reaction function. Our empirical strategy combines two different approaches. First, we obtain an estimate of the missing component in the PBOC monetary policy rule and, second, we use information content in official PBOC MPRs to find its rightful interpretation. To extract the missing unobserved component, we use a linear state-space approach applied to an otherwise standard monetary policy rule. To analyse the information content of PBOC MPR documents, we use text analytics through a Latent Semantic Analysis model and obtain a set of topics that capture

essential information embedded in the documents. Finally, we use an SVAR framework to examine whether the extracted topics play a role in the transmission of Chinese monetary policy.

Our paper yields several interesting findings. First, our results support the view that central bank communication can help explain deeds in the case of the PBOC. We find evidence that topics extracted from the PBOC's official documents help explain monetary policy actions in China providing the missing content in a standard monetary policy rule. Second, by digging deeper into the topics that we find are relevant and attempting to interpret them, our topical analysis reveals that the unobserved component in the Chinese monetary policy is related to the supply-side structural reforms agenda and window guidance. Thus, we find evidence that China uses monetary policy as a tool to achieve the government's objectives. Finally, the SVAR results offer support to our findings by providing evidence of the importance of the government's supply-side reform and window guidance objectives for the conduct of monetary policy and for macroeconomic outcomes.

By using state-of-the-arts text analytics techniques on unconventional data, our paper finds a much needed approximation of the PBOC monetary policy rule yielding unique insights into what drives monetary policy actions in China. It is important to note that this uncovered role for government objectives as an important determinant of monetary policy actions could not have been obtained by relying solely on traditional macroeconomic indicators readily available for the general public. Given the complexity of the Chinese economy, we believe that unconventional data and associated techniques have the potential to generate new insights for the Chinese economy and constitute a promising avenue for future research.

## Acknowledgements

We thank Narayan Bulusu, Reinhard Ellwanger, Gabriela Galassi, Stefano Gnocci, Evi Pappa, Lin Shao, Luis Henrique Uzeda Garcia and conference and seminar participants at Bank of Canada and the 6$^{\text{th}}$ Annual Bank of Canada-Tsinghua PBCSF-University of Toronto Conference on the Chinese Economy for very useful comments. We also thank Yu-Hsien Liu for outstanding research assistant work in the preparation of the data used in this paper.

# References


Bailliu, J., Han, X., Kruger, M., Liu, Y.-H., and Thanabalasingam, S. (2019). Can media and text analytics provide insights into labour market conditions in china? *International Journal of Forecasting*, 35(3):1118–1130.

Boulter, J. (2018). China's supply side structural reform. Reserve Bank of Australia Bulletin, December 2018.

Burdekin, R. and Siklos, P. (2008a). What has driven chinese monetary policy since 1990? investigating the people's bank's policy rule. *Journal of International Money and Finance*, 27(5):847–859.

Burdekin, R. and Siklos, P. (2008b). What has driven Chinese monetary policy since 1990? investigating the people's bank's policy rule. *Journal of International Money and Finance*, 27(5):847–859.

Che, W., Li, Z., and Liu, T. (2010). LTP: A Chinese language technology platform. *Coling 2010: Demonstration Volume*, pages 13–16.

Chen, K., Higgins, P., Waggoner, D., and Zha, T. (2016). China pro-growth monetary policy and its asymmetric transmission. *NBER Working Paper No. 22650*.

Chen, K., Higgins, P., Waggoner, D., and Zha, T. (2017). The asymmetric transmission of china's monetary policy. In *2017 Meeting Papers*, number 516. Society for Economic Dynamics.

Cvitanic, T., Lee, B., Song, H. I., Fu, K., and Rosen, D. (2016). LDA v. LSA: A comparison of two computational text analysis tools for the functional categorization of patents. Georgia Institute of Technology.

Dikau, S. and Volz, U. (2023). Out of the window? green monetary policy in china: window guidance and the promotion of sustainable lending and investment. *Climate Policy*, 23(1):122–137.

Fan, L., Yu, Y., and Zhang, C. (2011). An empirical evaluation of china's monetary policies. *Journal of Macroeconomics*, 33(2):358–371.

Frank, J. F., Steven, V. M., and Moorad, C. (2002). Finance investment-the global money markets.

Garcia-Herrero, A. and Girardin, E. (2013). China's monetary policy communication: Money markets not only listen, they also understand. HKIMR Working Paper no. 02/2013.

Girardin, E., Lunven, S., and Ma, G. (2017). China's evolving monetary policy rule: From inflation-accommodating to anti-inflation policy. BIS Working Papers No. 641.

He, D. and Pauwels, L. (2008). What prompts the People's Bank of China to change its monetary policy stance? Evidence from a discrete choice model. *China and World Economy*, 16:1–21.

Huang, Y., Ge, T., and Wang, C. (2019). Monetary policy framework and transmission mechanism. In M. Amstad, S. G. and Xiong, W., editors, *Handbook on China's Finance System*.

Klingelhöfer, J. and Sun, R. (2018). China's regime-switching monetary policy. *Economic Modelling*, 68:32–40.

Koopman, S. J. and Hol Uspensky, E. (2002). The stochastic volatility in mean model: empirical evidence from international stock markets. *Journal of applied Econometrics*, 17(6):667–689.

Landauer, T. and Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104:211–240.

Lardy, N. R. (2005). Exchange rate and monetary policy in china. *Cato J.*, 25:41.

Lee, J. (2017). China's monetary policy is looking like alphabet soup. https://www.bloomberg.com/.

Li, B. and Liu, Q. (2017). On the choice of monetary policy rules for china: A Bayesian DSGE approach. *China Economic Review*, 44:166–185.

Liu, L. and Zhang, W. (2010). A New Keynesian model for analysing monetary policy in mainland china. *Journal of Asian Economics*, 21:540–551.

McCallum, B. (1993). Specification and analysis of a monetary policy rule for Japan. *NBER Working Paper No. 4449*.

McCallum, B. (1999). Issues in the design of monetary policy rules. *Handbook of macroeconomics*, 1:1483–1530.

McCallum, B. (2002). Use of policy rules in monetary policy analysis.

Mehrotra, A. and Sánchez-Fung, J. R. (2010). China's monetary policy and the exchange rate. *Comparative Economic Studies*, 52:497–514.

Řehůřek, R. and Sojka, P. (2010). Software framework for topic modelling with large corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, pages 45–50, Valletta, Malta. ELRA. `http://is.muni.cz/publication/884893/en`.

Röder, M., Both, A., and Hinnenburg, A. (2015). Exploring the space of topic coherence measures. *In Proc. of the 8th ACM International Conference on Web Search and Data Mining*, pages 399–408.

Rongji, Z. (2011). The State Council of the People's Republic of China Order no. 321. *Chinese Law & Government*, 44(4):42–48.

Shu, C. and Ng, B. (2010). Monetary stance and policy objectives in China: A narrative approach. *Hong Kong Monetary Authority, China Economic Issues no. 1/10*.

Sun, R. (2013). Does monetary policy matter in China? A narrative approach. *China Economic Review*, 26:56–74.

Sun, S., Gan, C., and Hu, B. (2012). Evaluating mccallum rule as a policy guideline for china. *Journal of the Asia Pacific Economy*, 17(3):527–545.

Taylor, J. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy*, 39:195–214.

Taylor, J. (2000). Using monetary policy rules in emerging market economies. *Unpublished manuscript, Standford University*.

Xia, B. and Liao, Q. (2001). Money supply is no longer suitable to serve as intermediary target of china's current monetary policy. *Journal of Economic Research*, 8:33–43.

Xie, P. and Luo, X. (2002). Taylor rule and its empirical test in china's monetary policy. *Economic Research Journal*, 3(3):12.

Xiong, W. (2012). Measuring the monetary policy stance of the People's Bank of China: An ordered probit analysis. *China Economic Review*, 23:512–533.

Xu, X. (2003). The impact of sars on china. *Seton Hall J. Dipl. & Int'l Rel.*, 4:45.

Yan, J. (2012). Development and utilisation of financial derivatives in china. In *China and the Irving Fisher Committee in Zhengzhou on 27–29 September 2010.*, page 20. Citeseer.

Zhao, Y., Xie, H., and Wu, Z. (2017). PBOC trying to keep it simple with spotlight on two rates. https://www.bloomberg.com/news/articles/2017-05-17/pboc-trying-to-keep-it-simple-withnew-spotlight-on-two-rates.

Zhou, X. (2016). Managing multi-objective monetary policy: From the perspective of transitioning Chinese economy. Michel Camdessus Central Banking Lecture at the Iinternational Monetary Fund.

# Appendix: Figures and tables

Table A-1: Structure of PBOC Monetary Policy Reports

| Section name | Period covered | Number of docs |
|---|---|---|
| Summary | 2002Q3–2019Q4 | 65 |
| Section 1: Money and credit | 2002Q4–2019Q4 | 64 |
| Section 2: Monetary policy | 2001Q1–2019Q4 | 76 |
| Section 3: Financial market conditions | 2001Q1–2019Q4 | 71 |
| Section 4: Macroeconomic conditions | 2001Q1–2019Q4 | 71 |
| Section 5: Monetary policy trends | 2001Q1–2019Q4 | 71 |

Table A-2: Posterior distributions of the parameters of the system given by eqs.(1)-(3)

| parameter | mean | standard deviation | 10%-tile | 90%-tile |
|---|---|---|---|---|
| $\gamma_0$ | 0.14 | 0.04 | 0.08 | 0.20 |
| $\tau$ | 0.23 | 1.28 | -2.1 | 1.96 |
| $\beta$ | -3E-04 | 0.01 | -0.01 | 0.01 |
| $\phi$ | 0.94 | 0.05 | 0.83 | 0.99 |
| $\sigma^2$ | 0.01 | 0.04 | 0.03 | 0.06 |
| $\boldsymbol{\omega}_y^2$ | 0.10 | 0.09 | 0.04 | 0.23 |
| $\boldsymbol{\omega}_\tau^2$ | 5E-02 | 0.11 | -0.09 | 0.11 |
| $\boldsymbol{\omega}_{\pi_g}^2$ | 0.23 | 0.22 | 0.08 | 0.52 |

*Notes:* Reported are the results of the estimation of eqs.(1)-(3). Data used in the estimation is reported in Section 3. Columns 2-5 present the means, standard deviations and the 90% highest posterior density intervals of the posterior distributions of hyper parameters in the model.

Table A-3: OLS estimates for Topics regression in eq.(4)

| parameter | coefficient | 10%-tile | 90%-tile |
|---|---|---|---|
| $\beta_0$ | 0.22 | 0.12 | 0.32 |
| $\beta_{T1t-1}$ | -2.03 | -2.95 | -1.16 |
| $\beta_{T2t-1}$ | 0.08 | -0.53 | 0.72 |
| $\beta_{T3t-1}$ | 0.55 | -0.03 | 1.10 |
| $\beta_{T4t-1}$ | 0.49 | 0.11 | 0.84 |
| $\delta_{ERt-1}$ | 0.03 | -0.10 | 0.14 |
| $\delta_{URt-1}$ | -0.20 | -0.62 | 1E-03 |
| $R^2$ | | 0.31 | |

*Notes:* Reported are the results of the estimation of eq.(4). Data used in the estimation is reported in Section 3. Columns 2-4 present the bootstrapped least-squares coefficients and corresponding 90% upper and lower confidence interval mean values. All variables in eq.(4), expect that of Topic 1, are included in first differences since, for these series, the presence of a unit root could not be rejected.

Table A-4: Granger Tests MPR Sections' Topics on Summary Topics 1 and 4

| | Sec. 3 on T1 | Sec. 3 on T4 | Sec. 4 on T1 | Sec. 4 on T4 |
|---|---|---|---|---|
| Pre-GFC period 2003-2007 | | | | |
| Topic 1 | 0.564 | 0.219 | 0.501 | 0.672 |
| Topic 2 | 0.365 | 0.687 | 0.504 | 0.989 |
| Topic 3 | 0.891 | 0.068 | 0.214 | 0.067 |
| Topic 4 | 0.879 | 0.075 | 0.727 | 0.170 |
| Topic 5 | 0.408 | 0.839 | 0.945 | 0.197 |
| Topic 6 | 0.912 | 0.608 | 0.299 | 0.894 |
| Topic 7 | 0.613 | 0.534 | 0.600 | 0.144 |
| Topic 8 | 0.646 | 0.256 | 0.304 | 0.501 |
| Topic 9 | 0.159 | 0.332 | 0.292 | 0.775 |
| Topic 10 | 0.052 | 0.114 | 0.083 | 0.755 |
| GFC period 2008-2014 | | | | |
| | Sec. 3 on T1 | Sec. 3 on T4 | Sec. 4 on T1 | Sec. 4 on T4 |
| Topic 1 | 0.309 | 0.006 | 0.242 | 0.065 |
| Topic 2 | 0.097 | 0.106 | 0.710 | 0.024 |
| Topic 3 | 0.002 | 0.135 | 0.020 | 0.635 |
| Topic 4 | 0.497 | 0.395 | 0.846 | 0.198 |
| Topic 5 | 0.993 | 0.563 | 0.223 | 0.180 |
| Topic 6 | 0.598 | 0.549 | 0.554 | 0.455 |
| Topic 7 | 0.805 | 0.119 | 0.087 | 0.761 |
| Topic 8 | 0.699 | 0.612 | 0.925 | 0.676 |
| Topic 9 | 0.963 | 0.834 | 0.469 | 0.868 |
| Topic 10 | 0.726 | 0.252 | 0.293 | 0.668 |
| Topic 11 | 0.639 | 0.429 | 0.358 | 0.611 |
| Topic 12 | 0.279 | 0.837 | 0.459 | 0.332 |
| Topic 13 | 0.462 | 0.743 | NA | NA |
| Post-GFC period 2015-2019 | | | | |
| | Sec. 3 on T1 | Sec. 3 on T4 | Sec. 4 on T1 | Sec. 4 on T4 |
| Topic 1 | 0.590 | 0.375 | 0.992 | 0.475 |
| Topic 2 | 0.391 | 0.061 | 0.628 | 0.372 |
| Topic 3 | 0.396 | 0.334 | 0.220 | 0.166 |
| Topic 4 | 0.764 | 0.937 | 0.747 | 0.744 |
| Topic 5 | 0.085 | 0.566 | 0.997 | 0.720 |
| Topic 6 | 0.672 | 0.987 | 0.212 | 0.600 |
| Topic 7 | 0.413 | 0.954 | 0.086 | 0.428 |
| Topic 8 | 0.663 | 0.307 | 0.508 | 0.552 |
| Topic 9 | 0.251 | 0.884 | 0.847 | 0.420 |
| Topic 10 | 0.126 | 0.531 | NA | NA |
| Topic 11 | 0.443 | 0.254 | NA | NA |

*Notes:* Reported are the p-values of bi-variate Granger causality tests with 1-lag. Direction of causality goes from the Section topic to the Summary topic.

Table A-5: Robustness check with M2: OLS estimates for Topics regression in eq.(4)

| parameter | coefficient | 10%-tile | 90%-tile |
|---|---|---|---|
| $\beta_0$ | 0.12 | 0.05 | 0.19 |
| $\beta_{T1t-1}$ | -1.25 | -1.80 | -0.67 |
| $\beta_{T2t-1}$ | -0.15 | -0.54 | 0.26 |
| $\beta_{T3t-1}$ | 0.18 | -0.13 | 0.48 |
| $\beta_{T4t-1}$ | 0.15 | -0.05 | 0.38 |
| $\delta_{ERt-1}$ | -0.06 | -0.15 | 0.02 |
| $\delta_{URt-1}$ | -0.14 | -0.33 | -0.05 |
| $R^2$ | | 0.31 | |

*Notes:* Reported are the results of the estimation of eq.(4) using M2 as dependent variable. Data used in the estimation is reported in Section 3. Columns 2-4 present the bootstrapped least-squares coefficients and corresponding 90% upper and lower confidence interval mean values. All variables in eq.(4), expect that of Topic 1, are included in first differences since, for these series, the presence of a unit root could not be rejected.

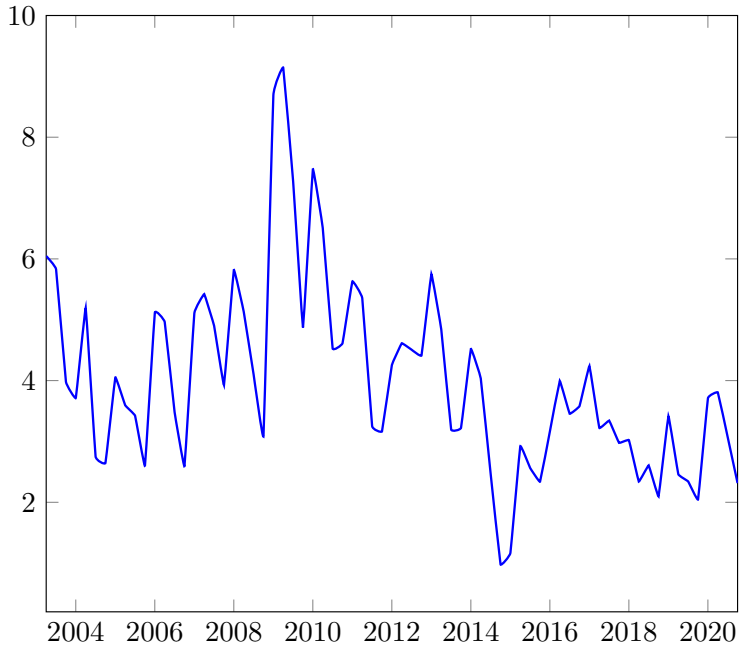



Figure A-1: China Total Social Financing quarterly changes from 2003Q2 to 2020Q4

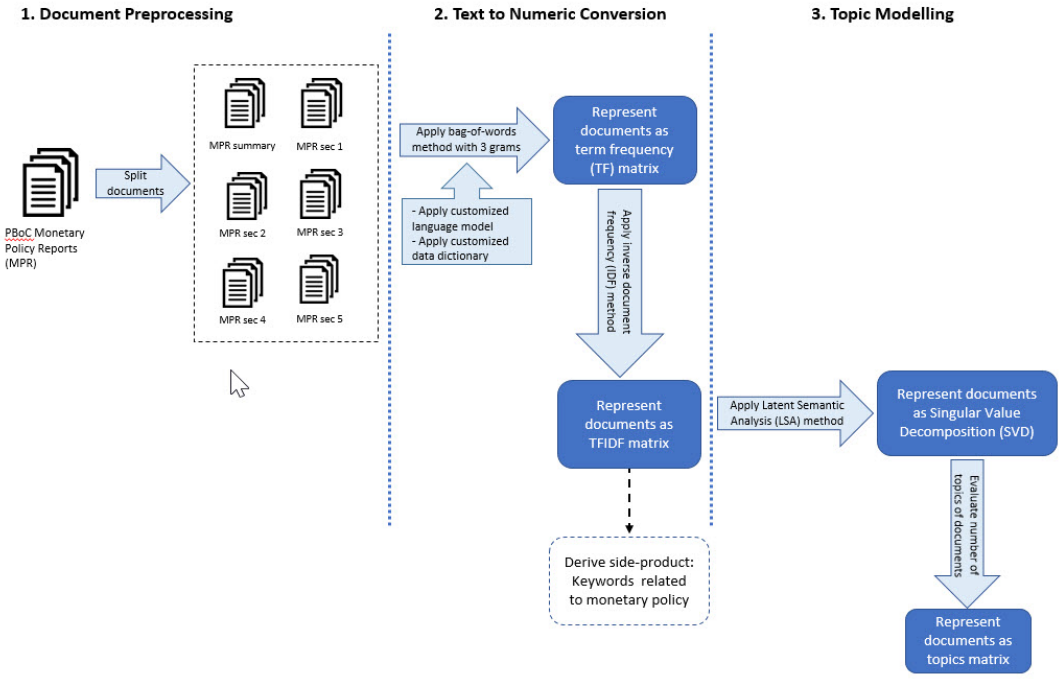



Figure A-2: Methodology: Text Analytics

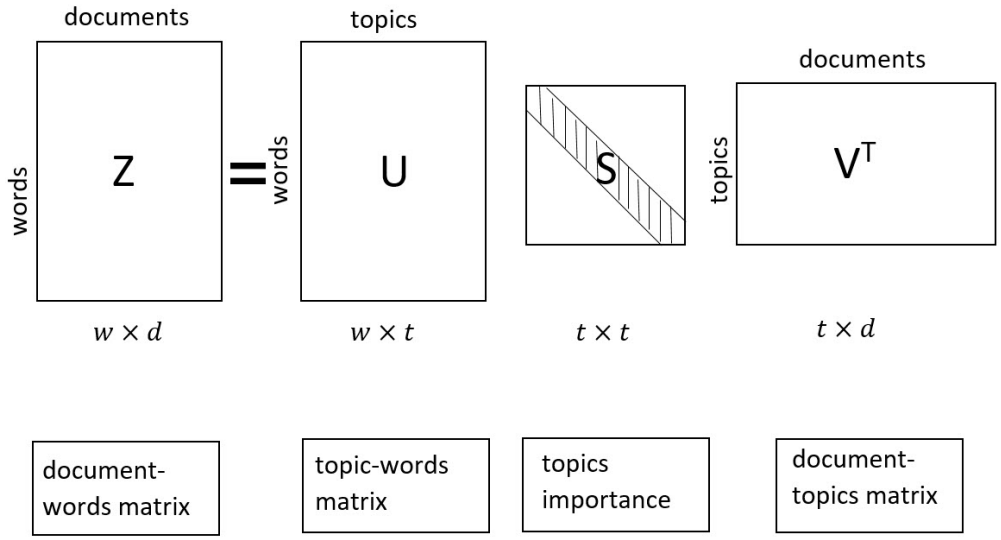



Figure A-3: Singular Value Decomposition

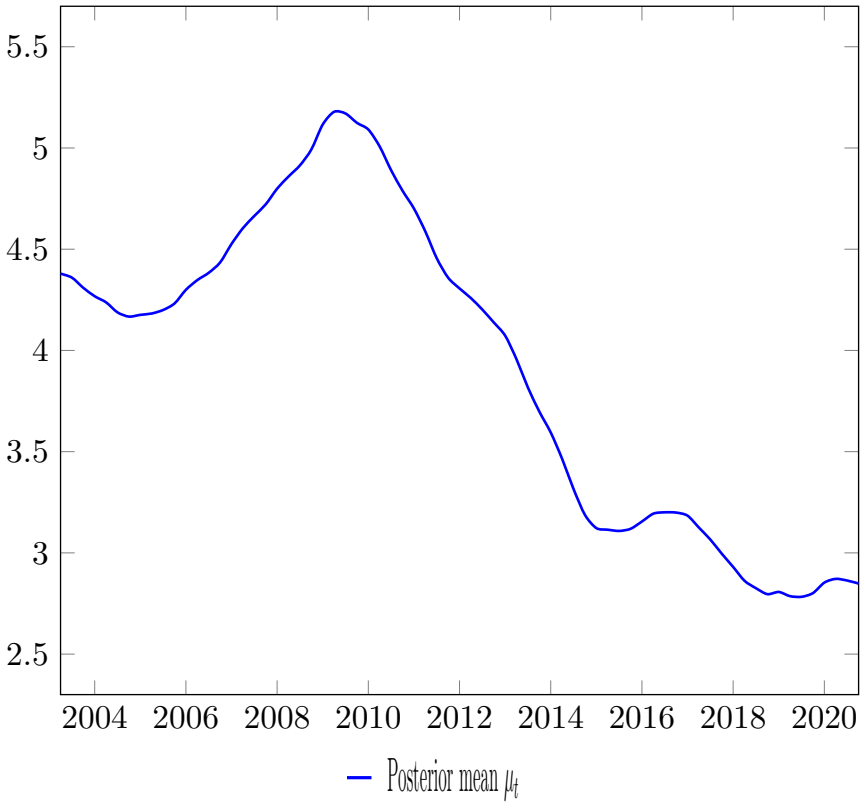
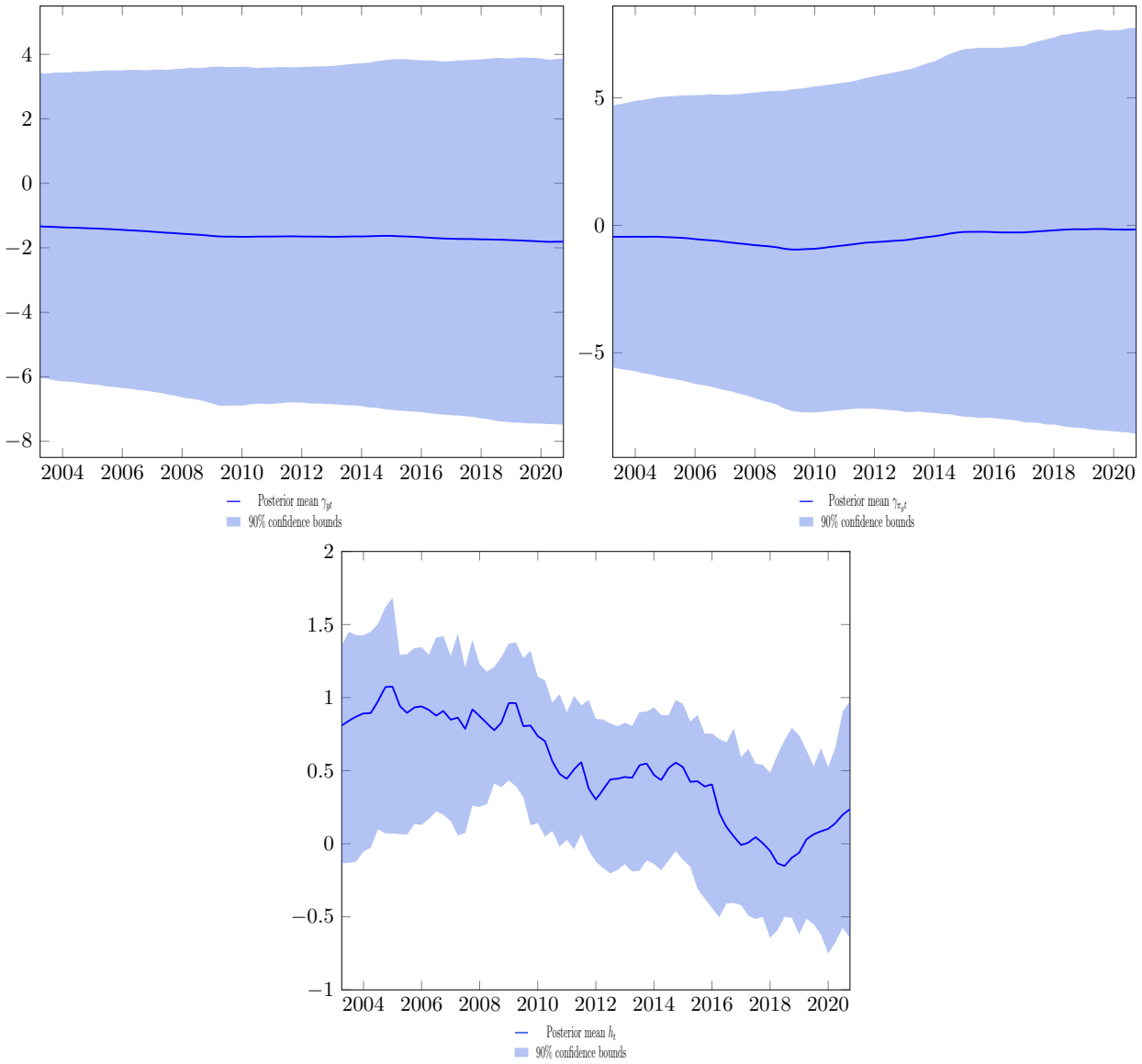



Figure A-4: Unobserved component $\mu_t$ in eq.(1)

Figure A-5: Time-varying coefficients $\gamma_{\pi t}$, $\gamma_{yt}$ and stochastic volatility component $h_t$ in eqs.(1)-(3)






Topic 1




Topic 4

Figure A-6: Topics covering MPR's Summary Section

Notes: These figures show estimated topics using LSA. Each Topic's word distributions are represented as word clouds for visual analysis.




Figure A-7: Pre GFC period 2003-2007 MPR Section 4 (Macroeconomic conditions) Topic 10

Notes: The word cloud belong to the sectional topics found to Granger cause Summary Topic 1: SSSR. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.




Figure A-8: GFC period 2008-2014 MPR Section 3 (Financial market conditions) Topic 2 (top) and Topic 3 (bottom)

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: SSSR. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.




Figure A-9: GFC period 2008-2014 MPR Section 4 (Macroeconomic conditions) Topic 3 (top) and Topic 7 (bottom)

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: SSSR. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.

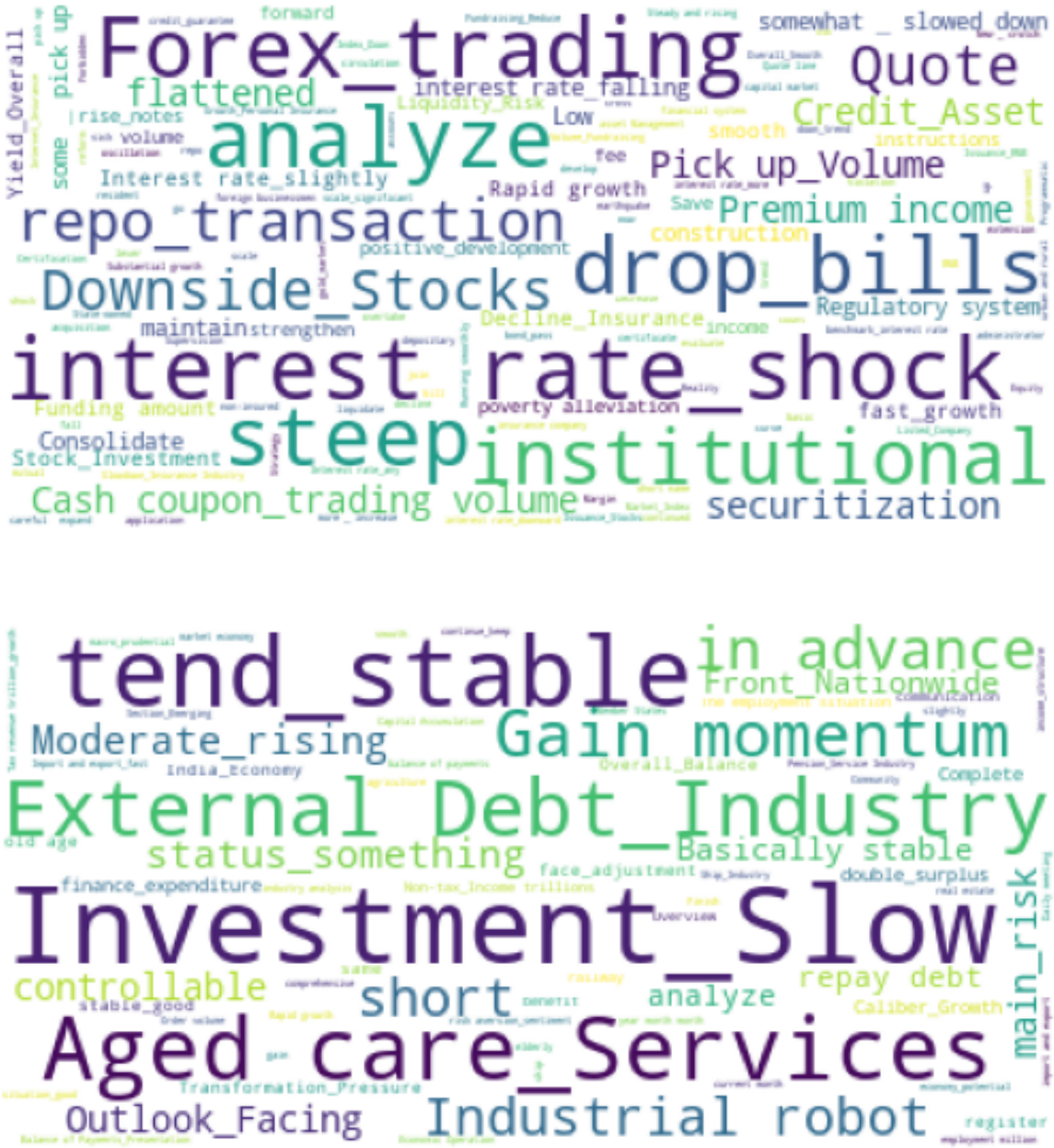



Figure A-10: Post GFC period 2015-2019 MPR Section 3 (Financial market conditions) Topic 5 and MPR Section 4 (Macroeconomic conditions) Topic 7

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: SSSR. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.

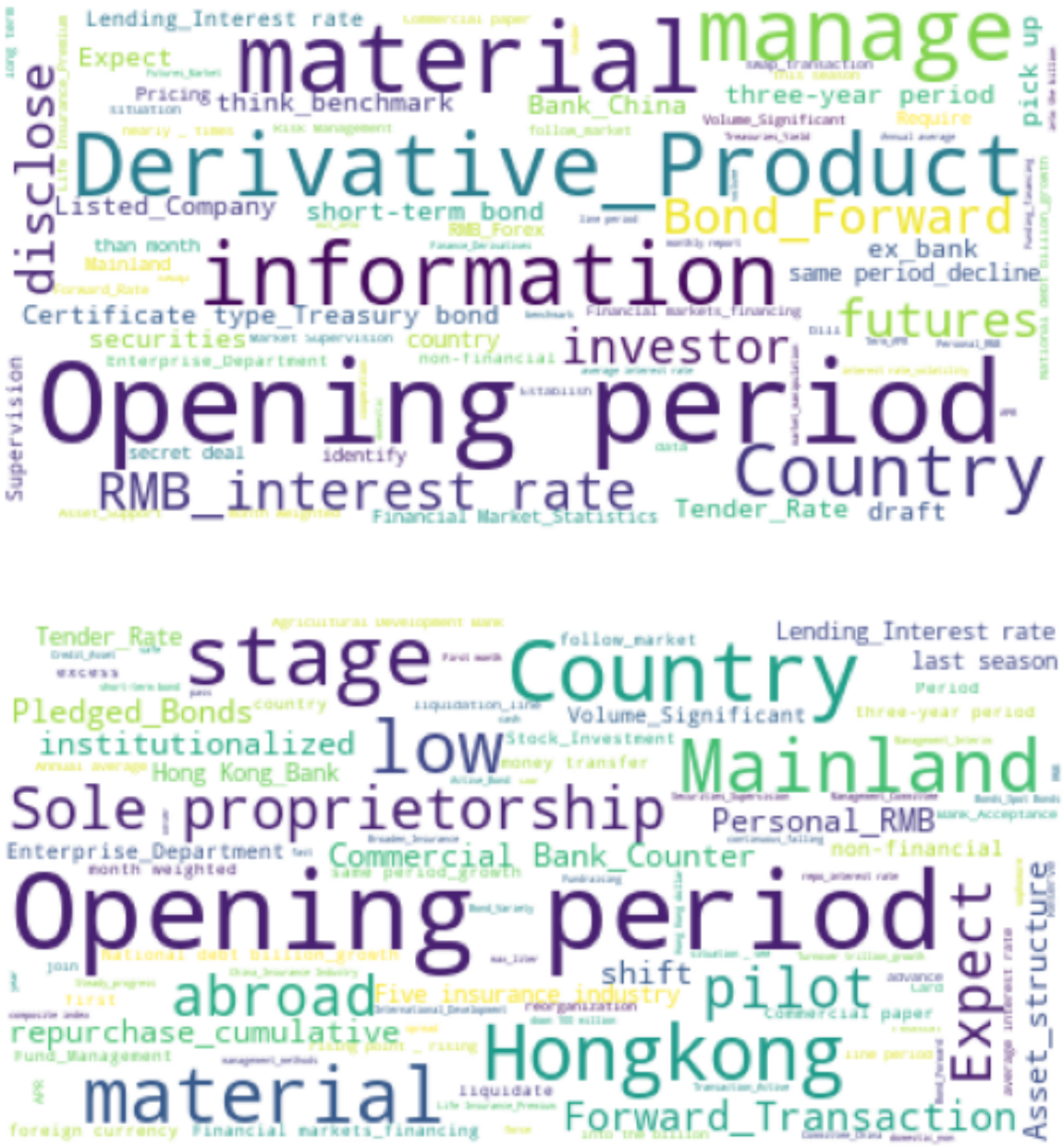



Figure A-11: Pre GFC period 2003-2007 MPR Section 3 (Financial market conditions) Topic 3 (top) and Topic 4 (bottom)

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: window guidance. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.

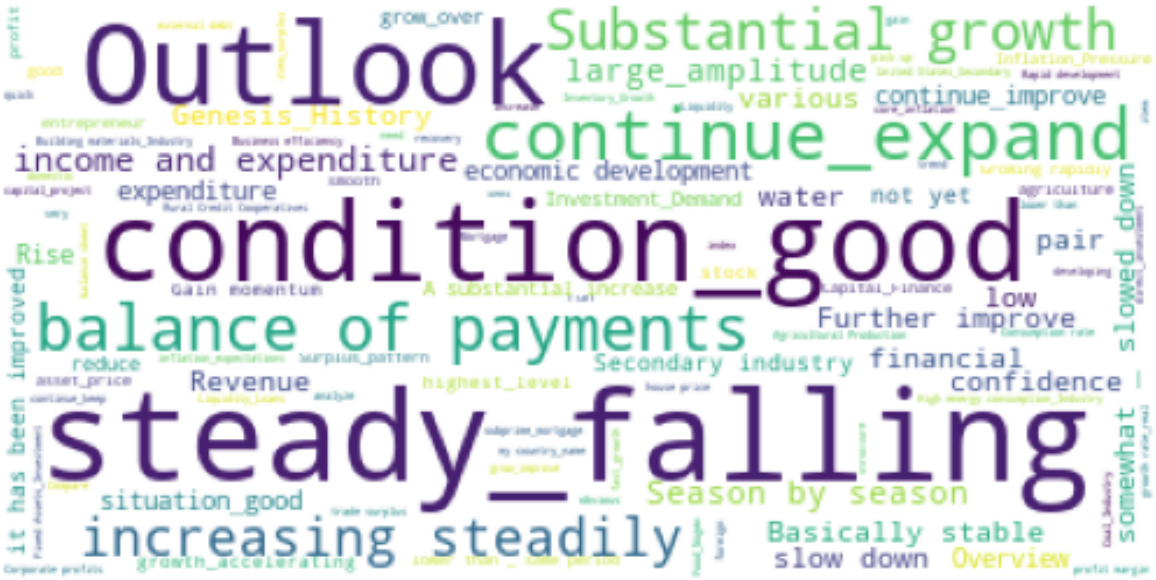



Figure A-12: Pre GFC period 2003-2007 MPR Section 4 (Macroeconomic conditions) Topic 3

Notes: The word cloud belongs to the sectional topics found to Granger cause Summary Topic 1: window guidance. It corresponds to estimated topic using LSA technique. The topic's word distributions is represented as a word cloud for visual analysis.




Figure A-13: GFC period 2008-2014 MPR Section 3 (Financial market conditions) Topic 1 (top) and Topic 2 (bottom)

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: window guidance. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.




Figure A-14: GFC period 2008-2014 MPR Section 4 (Macroeconomic conditions) Topic 1 (top) and Topic 2 (bottom)

Notes: The word clouds belong to the sectional topics found to Granger cause Summary Topic 1: window guidance. They correspond to estimated topics using LSA technique. Each topic's word distributions are represented as word clouds for visual analysis.

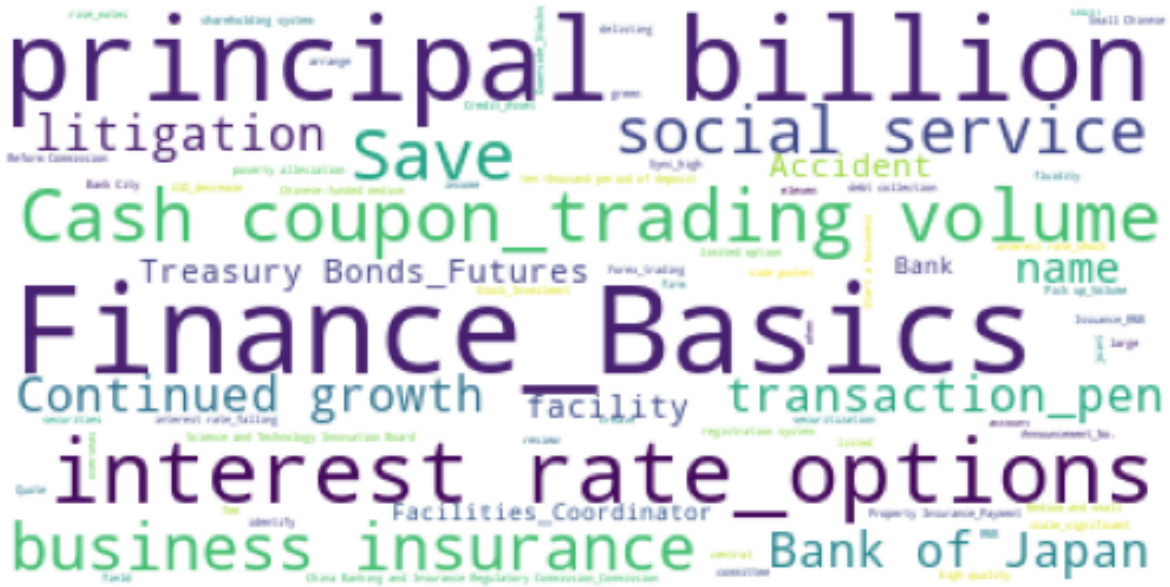



Figure A-15: Post GFC period 2015-2019 MPR Section 3 (Financial market conditions) Topic 2

Notes: The word cloud belongs to the sectional topics found to Granger cause Summary Topic 1: window guidance. It corresponds to estimated topic using LSA technique. The topic's word distributions is represented as a word cloud for visual analysis.

Figure A-16: Responses of Variables to Different Shocks

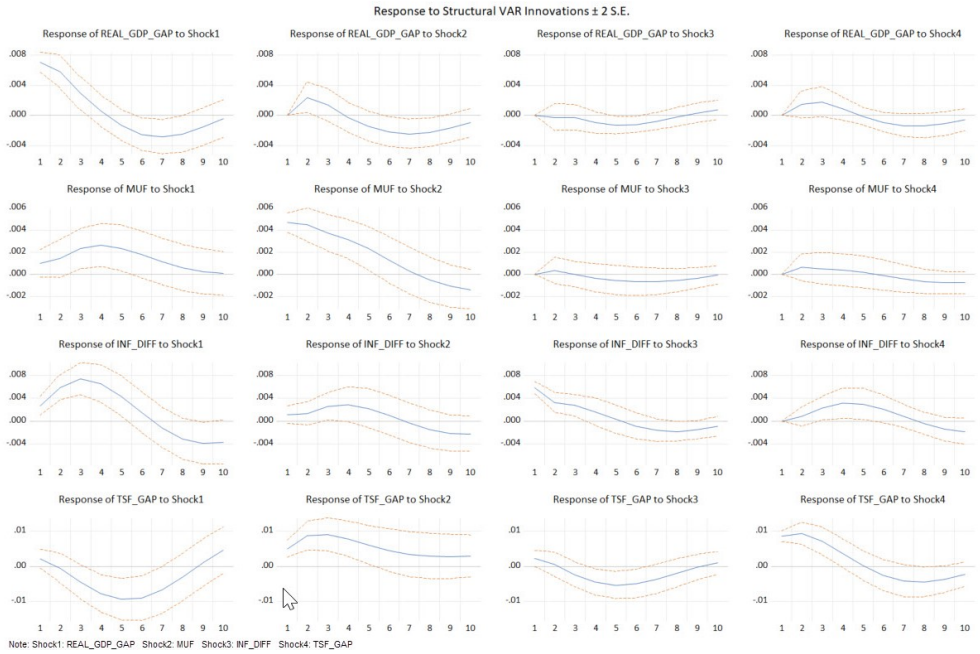



Figure A-17: Variance Decomposition

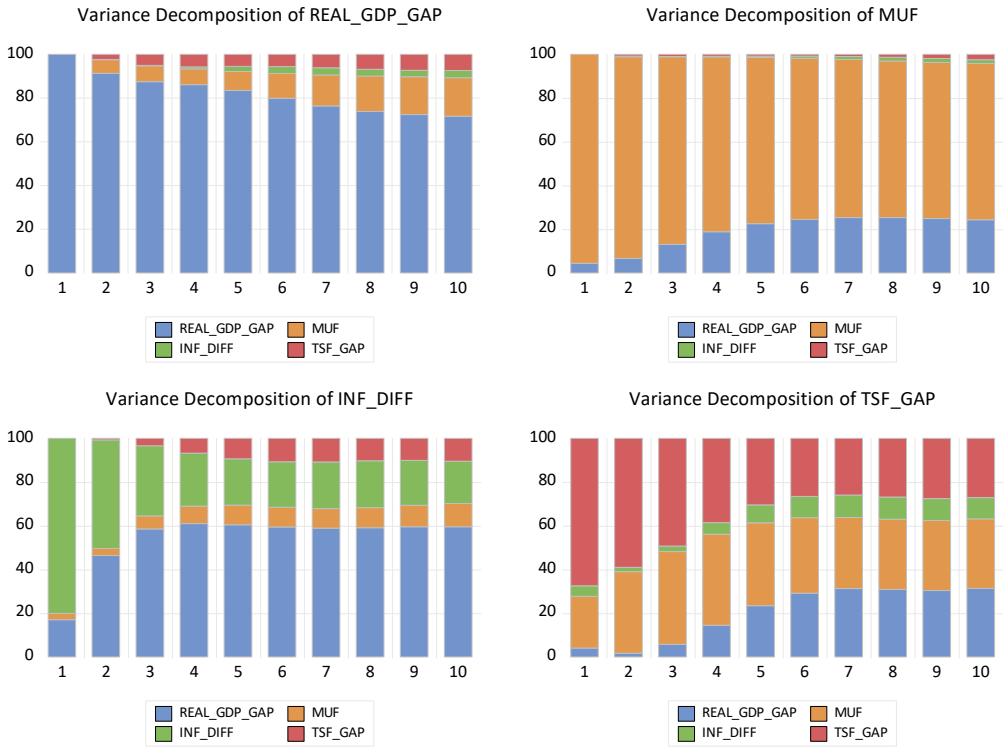


Figure A-18: Robustness check: M2 series and time-varying coefficients estimates $\gamma_{\pi t}$, $\gamma_{yt}$, $\mu_t$ in eqs.(1)-(3)

