BANK OF CANADA
BANQUE DU CANADA

Staff Working Paper/Document de travail du personnel — 2021-3

Last updated: January 21, 2021

Chinese Monetary Policy and Text Analytics: Connecting Words and Deeds

by Jeannine Bailliu, Xinfen Han,[1] Barbara Sadaba[2] and Mark Kruger[3]

[1]Financial Markets Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

[2]International Economic Analysis Department
Bank of Canada, Ottawa, Ontario, Canada K1A 0G9

[3]Senior Fellow, Yicai Research Institute / Opinion Editor, Yicai Global
Shanghai, China 200041

xhan@bank-banque-canada.ca, BSadaba@bank-banque-canada.ca, krugermarksteven@yicaiglobal.com

ISSN 1701-9397

Acknowledgements

We thank Reinhard Ellwanger, Stefano Gnocci, Lin Shao, Gabriela Galassi and conference and seminar participants at the Bank of Canada for very useful comments. We thank Yu-Hsien Liu for outstanding research assistant work in the preparation of the data used in the present paper. The views expressed in this paper are those of the authors. No responsibility for them should be attributed to the Bank of Canada.

Abstract

Given China's complex monetary policy framework, the People's Bank of China's (PBOC) monetary policy rule is difficult to infer from its observed behaviour. In this paper, we adopt a novel approach, using text analytics to estimate and interpret the unknown component in the PBOC's reaction function. We extract the unknown component in a McCallum-type monetary policy rule for China through a state-space model framework using a set of summary topics extracted from official PBOC documents. Then, using a set of sectional topics extracted from the same set of PBOC documents, we provide this component with its rightful interpretation. Our results show that this unknown component is related to the Chinese government's agenda of supply-side structural reforms, suggesting that monetary policy is used as a tool to achieve structural reform objectives. Structural vector autoregression (SVAR) results confirm these findings by providing evidence of the importance of the government's supply-side reform objectives for the conduct of monetary policy.

Bank topics: Monetary policy communications; Monetary policy framework; Econometric and statistical methods; International topics
JEL codes: E52, E58, C63

1 Introduction

How does the People's Bank of China (PBOC) conduct its monetary policy? Although a fairly good understanding exists regarding how monetary policy is run in advanced economies, Chinese monetary policy remains a bit of a black box. This is due, in fact, to the complex nature of the monetary policy framework in China. In particular, the PBOC relies on multiple monetary policy instruments whose importance has been found to change over time, while its monetary policy has been guided by multiple objectives.[1] Additionally, monetary policy actions are collective decisions made by the State Council, not by an independent monetary policy committee.[2] 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.

Since the seminal work of Taylor (1993), a vast literature has emerged that seeks to characterize how a central bank conducts monetary policy by estimating monetary policy rules. For the case of China, existing works rely on best available approximations coming from Western counterparts, such as the Taylor or McCallum-type rules (He and Pauwels (2008), Xiong (2012) and Girardin et al. (2017), among others). These studies rely on a set of variables that are insufficient predictors of monetary policy instruments in China—i.e., the output gap, GDP growth, deviations of inflation from target, and inflation expectations. Therefore, a monetary policy rule for China is yet to be found that is successful in explaining the PBOC's actions. As noted by Huang et al. (2019, p. 56), ". . . empirical findings on China's monetary policy rules are inconclusive." This points to a missing element in the PBOC's reaction function needed to accurately describe its behaviour. Given China's prominent position in the world economy, finding this missing component constitutes a pressing matter. To this end, in this paper we adopt a novel approach to estimate the monetary policy reaction function for China by identifying this missing component and providing its rightful interpretation. This approach relies on state-space modelling and text analytics to extract and interpret the missing component in the reaction function from information embedded in official PBOC communication.

Our paper's contribution to the literature is threefold. First, our study distinguishes itself from others in its use of text mining techniques to extract information from PBOC official

[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] The PBOC does have a monetary policy committee, but it plays only an advisory role in monetary policy decisions. The role of the committee is to prepare monetary policy plans and then submit them to the State Council for approval.

documents to explore whether "words" can help explain "deeds." 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), and 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. Second, we provide an estimate of the missing component in the PBOC reaction function by estimating a linear state-space model for the PBOC monetary policy rule that allows for an unobserved component. The estimate of this unobserved component confirms the existence of a persistent and systematic element that is missing from standard monetary policy rules for China. Third, we find an alternative approximation of the PBOC reaction function that yields a unique insight into what drives monetary policy actions in China. Our analysis reveals that the hidden systematic component in the Chinese monetary policy is related to the government's agenda for supply-side structural reform. To the best of our knowledge, our paper is the first to uncover that the PBOC uses monetary policy as a tool to achieve the government's structural reform objectives. It is important to note that this new insight was discovered by exploiting unconventional data and could not have been obtained by relying solely on more traditional macroeconomic variables. Our paper thus highlights the potential of using unconventional data and associated techniques to gain new insights into important research questions about a complex economy like China's.

Our paper is also related to the literature on central bank communication that examines whether information extracted from official central bank statements can help agents to better understand central bank actions. As highlighted by Blinder et al. (2008), communication is an important part of a central bank's tool kit as it can enhance the effectiveness of monetary policy. Because agents are forward-looking, central bank communication can have an impact on the economy via its influence on expectations. Monetary policy will thus become more effective if the public better understands the central bank's actions and intentions. Or in other terms, communication of their words—via official statements—can help agents to better understand their deeds (i.e., central bank actions). There is reason to believe that even more importance should be attached to words in the case of China. Given China's complex monetary policy framework, PBOC communication can provide the missing link for agents to better identify monetary policy actions. Such communication could raise the signal-to-noise ratio by more clearly explaining the PBOC's actions and intentions.

Our empirical strategy combines two different approaches. We obtain an estimate of the unobserved missing component in the PBOC monetary policy rule, and we also provide

the rightful interpretation for this component. First, we use a univariate linear state-space approach to extract the unobserved component in a standard monetary policy rule for China. Second, we augment the state equation to include the set of topics that we extracted from the official PBOC documents using a latent semantic analysis (LSA) technique. Our corpus covers official PBOC documents on monetary policy decisions over the period from 2003Q2 to 2018Q4. In this way, we are able to assess which of the topics are significant to explain this missing component in the monetary rule. In our second approach, we use our identified set of significant topics to examine whether they play a relevant role in the transmission of Chinese monetary policy. To this end, we estimate a four variable structural vector autoregression (SVAR) model. This approach enables us to validate the results found in the first part of our analysis.

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 the 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 that we use to incorporate communication into a monetary policy reaction function for China. Section 3 provides an overview of our official document set. Section 4 describes the methodology that we use to extract the topics from the document set. In Section 5, we present the key results related to the estimation of the alternative PBOC reaction function and the topical analysis. 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

Given the complex nature of the PBOC's monetary policy framework, it is challenging to capture the conduct of monetary policy with a single standard reaction function. Previous studies have generally found that in the case of China, a McCallum-type rule seems to be more appropriate than a Taylor-type rule (Huang et al. (2019)). McCallum (1988) proposed a base money rule to inform the conduct of monetary policy in the U.S. by prescribing settings for the monetary base to keep nominal GDP growing at a non-inflationary rate. Therefore, in the present paper, we use a general form of the McCallum rule and extend it to accommodate an additional unobserved component that captures all potential omitted variables in the equation. Furthermore, we incorporate communication into our monetary policy reaction function. By including the information extracted from PBOC official statements, we explore whether communication can be useful in helping agents better understand this missing component and, thus, monetary policy actions in practice. In other words, we explore whether "words" can help explain "deeds" in the case of China.

We define our monetary policy rule as follows:

$$\Delta TSF_t^{gap} = \alpha_0 + \alpha_1 \Delta TSF_{t-1}^{gap} + \alpha_2(\pi_{t-1} - \pi_{t-1}^*) + \alpha_3 y_{t-1} + \mu_t + \varepsilon_t, \tag{1}$$

where ΔTSF_t^{gap} is the gap in total social financing (TSF) in period t, defined as the difference between TSF growth rate and its trend; π_{t-1}^* is the inflation target; π_{t-1} is the inflation rate; y_{t-1} is the output gap; and μ_t is the unobserved component. The error term ε_t is added to account for measurement error.[3]

We use TSF, a broad measure of credit in the Chinese economy, as it is a key intermediate target for monetary policy in China. As such, it can be thought of as a summary variable that captures the impact of all of the monetary policy instruments used

[3]The TSF variable was constructed using monthly levels, and then a quarterly growth rate was created using a 3-month moving average. The trend of TSF growth was calculated using a Hodrick-Prescott (HP) filter with λ=1600. The TSF series were obtained from the PBOC via Haver. The inflation rate is defined as the year-over-year 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). Potential GDP is constructed using an HP filter with λ=1600. The GDP and inflation data were obtained from the National Bureau of Statistics of China via Haver. The target inflation rate was obtained from the National People's Congress via Haver.

by the PBOC. Over our sample period, the PBOC has used the following instruments to conduct monetary policy: reserve requirement ratios, benchmark interest rates, open market operations, targeted lending facilities, and window guidance. We assume that the impact of these instruments will be reflected in TSF.

We model the unobserved component, μ_t, as an AR(1) process with a set of exogenous explanatory variable such that

$$\mu_t = \beta_0 \mu_{t-1} + \sum_{i=1}^{J} \beta_i topic_{t-1}^i + \eta_t, \tag{2}$$

where $topic_{t-1}^i$ (where $i = 1, ..., J$) represents the ith topic (i.e., factor) of a set of J topics extracted from official PBOC documents related to monetary policy actions and η_t is the measurement error term.

Chinese monetary policy has been guided by multiple objectives over our sample period. To the extent that there are other important determinants of the TSF growth gap besides the inflation target $(\pi_{t-1} - \pi_{t-1}^*)$ and the economic cycle (y_{t-1}), they would be captured in eq.(1) by the systematic, unobserved component (μ_t). In this way, we are able to investigate whether a systematic component can be estimated from eq.(1) and if so, whether insights can be gleaned about its drivers by extracting key topics from the PBOC's official communication.

To obtain the estimate for the unobserved component (μ_t) and for the parameters in the model, we first put the system described by eq.(1)–(2) in state-space form. In particular, we estimate a linear Gaussian state-space model. The parameters in the system are estimated using maximum likelihood.

3 Data: official PBOC documents

Our document set comprises quarterly monetary policy reports (MPR) published by the PBOC over the period from 2003Q1 to 2018Q4. Each MPR includes a summary as well as five other sections covering an analysis of 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 (see Table 1 for more details). 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 MPR tends to be more backward-looking than similar documents produced by advanced economies. This is because the PBOC's communication is more constrained given that the central bank does not have full decision-making power. 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.

4 Methodology for topic extraction

In order to extract topics from official documents related to Chinese monetary policy, we use LSA. LSA is a technique in natural language processing that involves analyzing relationships between a set of documents and the words they contain by producing a set of concepts related to the documents and words. It was developed into a theory of knowledge representation by Landauer and Dumais (1997) and is based on a mathematical matrix decomposition technique called singular value decomposition (SVD).

We selected the LSA methodology because we found that it performed better than other techniques used to extract topics from documents, such as the Latent Dirichlet Allocation (LDA). LSA may be a better option in our case given that we have a small set of documents and LSA has been found to outperform LDA for smaller-scale databases (Cvitanic et al. (2016)). The facts that our documents are in Chinese and that they are focused on a fairly narrow set of topics (i.e., topics related to monetary policy) may also explain why LSA performs betters than LDA in our setting.

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.

4.1 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 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 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).

4.2 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.[4] 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:

$$idf = \log \frac{Number\ of\ documents\ n\ in\ the\ corpus}{Number\ of\ documents\ in\ the\ corpus\ in\ which\ term\ j\ occurs}. \tag{3}$$

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

4.3 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 2). 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 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 kXk 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. Table 2 summarizes the number of topics for each MPR section.

5 PBOC reaction function and topical analysis: key results

Given the large number of topics identified, we decided to focus first on the eight topics obtained from the MPR summaries. Since the summary section provides a comprehensive overview of each report, this corpus should be sufficient to identify the broader topics in the official documents. In a second step, we use the information contained in the remaining sections to improve our interpretation of the topics found to be significant in our estimation of eq.(1).

We thus estimate the model laid out in eq.(1)–(2) including the eight identified MPR summary topics. The estimation results are summarized in Table 7. Overall, our results are reasonable and in line with what has been found for China in the literature (Girardin et al. (2017)). Notably, we find the coefficients on the inflation and output gaps to be statistically significant and of the expected negative sign. This suggests that the PBOC has been following an anti-inflation policy since the early 2000s. In the context of our monetary policy reaction function, this would 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.

Turning to the coefficients on the MPR summary topics, we find Topics 3 and 5 to be statistically significant. The key words corresponding to these topics are presented as word clouds in Figures 3 and 4. In each figure, the word cloud in Chinese is shown in panel (a) and the word cloud in English is shown in panel (b). The key words in the word cloud for Topic 3 suggest that this topic is linked to structural policies and supply-side reforms. Supply-side structural reform 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. 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 3 is positive, which suggests that monetary policy is loosened in response to a change in this topic. On the other hand, the key words in the word cloud for Topic 5 indicate that this topic is linked to regulation and guidance provided to commercial banks (i.e., window guidance). The coefficient on Topic 5 is negative, which implies that monetary policy is tightened in

response to a change in this topic.

In order to obtain more granular information to help improve our interpretation of these topics, we conduct further analysis using the information contained in other MPR sections. More specifically, we use the topics found in the other MPR sections to conduct Granger causality tests on Topics 3 and 5.[5] Finally, we analyze the word clouds for any of the topics that are found to Granger cause Topics 3 and 5 to further develop our understanding of our two main topics.

The Granger causality results for Topic 3 suggest that Topic 3 from Section 3 and Topics 6 and 9 from Section 4 are Granger-causing Topic 3 from the MPR summary (see Tables 3 and 4 for more details). We then examine the word clouds from these relevant topics in Sections 3 and 4 (see Figures 5, 6, and 7 for more details). In examining the key words from these word clouds, several seem related to Topic 3 (i.e., supply-side policies and structural reforms). In particular, several key words are related to industries that the government has targeted as wanting to reduce in capacity (e.g., coal and textiles industries). On the other hand, several key words are also related to the internet, online retail, and consumption—in line with the government's desire to grow the consumption's share of GDP and to promote innovation.

We conduct a similar exercise for MPR summary Topic 5 and find that Topic 4 from Section 3 and Topic 5 from Section 4 Granger-cause Topic 5 from the MPR summary (see Tables 5 and 6 for more details). We then examine the word cloud from these relevant topics in Sections 3 and 4 (see Figures 8 and 9 for more details). In examining the key words from these word clouds, we find several key words that are related to overcapacity in specific sectors (such as cement) and to reform of state-owned enterprises (SOEs) (i.e., recombination). Moreover, we believe that the prominent key words "rural credit union" in Figure 9 relate to containing financial stability risks as these institutions have traditionally been viewed as financial stability risks. All of these key words are associated with the government's supply-side reform agenda—i.e., reducing credit growth to zombie firms, reforming SOEs, downsizing overcapacity sectors, and alleviating financial stability risks.

Taken together, these results allow us to provide an interpretation of our estimate of μ_t, the unobserved component missing in the Chinese monetary policy rule presented in eq.(1). The estimate of this unobserved component, shown in Figure 10, points to a persistent and systematic element. Our results suggest that the this hidden systematic component is related to the government's agenda of supply-side structural reform. Thus we can conclude

[5]We use two lags for each variable in the Granger causality tests.

that the PBOC uses monetary policy as a tool to achieve the government's structural reform objectives, including the containment of financial stability risks.

6 SVAR analysis

In the second part of our empirical work, we estimate a SVAR model to validate our results from the topic analysis and estimation of the monetary policy reaction function. We use a simple four-equation SVAR comprising the output gap, the unobserved component μ_t, the inflation gap, and the TSF growth gap. We include the variable μ_t because we believe that it is a good proxy for the relevant topics that were identified. It is also a better option than to include all of the relevant topics (and their lags) in the SVAR given the limit on degrees of freedom.

The specification of the SVAR model is written as follows:

$$B_0 z_t = k + B_1 z_{t-1} + ... + B_p z_{t-p} + \upsilon_t, \tag{4}$$

where z_t= (y_t, μ_t,($\pi_t - \pi_t^*$), ΔTSF_t^{gap}), and y_t is the output gap, μ_t is the unobserved component, (π_{t-1} - π_{t-1}^*) is the inflation gap, and ΔTSF_t^{gap} is the TSF growth gap. The vector of disturbances, υ_t, represents the structural innovations; these disturbances are assumed to be serially uncorrelated and uncorrelated with each other. 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:

$$B_0 = \begin{bmatrix} \beta_{11} & 0 & 0 & 0 \\ \beta_{21} & \beta_{22} & 0 & 0 \\ \beta_{31} & \beta_{32} & \beta_{33} & 0 \\ \beta_{41} & \beta_{42} & \beta_{43} & \beta_{44} \end{bmatrix}. \tag{5}$$

We estimated the SVAR using two lags of each variable. Overall, the results of our SVAR analysis seem reasonable. This is shown in Figure 11, 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 (i.e., a decline in the TSF growth rate relative to trend). Also, as anticipated, a positive monetary policy shock (i.e., 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 (μ_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 μ_t explains over 30% of the TSF gap variable (see Figure 12 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 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 reaction function. Our empirical strategy combines two different approaches to, first, obtain an estimate of the missing component in the PBOC monetary policy rule and, second, find its rightful interpretation. First, we use a univariate linear state-space approach to estimate the unobserved component in a standard monetary policy rule for China. In the second part of our empirical work, 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

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, by digging deeper into the topics that we find are relevant and attempting to interpret them, our topical analysis reveals that the hidden systematic component in the Chinese monetary policy is related to supply-side structural reforms. Thus, we find evidence that China uses monetary policy as a tool to achieve the government's objectives for structural reform. And finally, the results from our SVAR analysis validate these findings by providing evidence of the importance of the government's supply-side reform objectives for the conduct of monetary policy and for macroeconomic outcomes.

By using unconventional data that we analyze using text analytics, our paper finds an alternative approximation of the PBOC reaction function, which yields a unique insight into what drives monetary policy actions in China. It is important to note that this new insight (i.e., that government structural reform objectives are an important determinant of monetary policy actions in China) was found by exploiting unconventional data and could not have been obtained by relying solely on more traditional macroeconomic variables. Given the complexity of the Chinese economy, we believe that unconventional data and associated techniques have the potential to generate new insights for other important research questions about the Chinese economy and constitute a promising avenue for future research.

References

Blinder, A., Ehrmann, M., Fratzscher, M., Haan, J. D., and Jansen, D. (2008). Central bank communication and monetary policy: A survey of theory and evidence. *Journal of Economic Literature*, 46(4):910–945.

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

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

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.

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*.

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.

McCallum, B. (1988). Robustness properties of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy*, 29:173–203.

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.

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.

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

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.

Appendix: Figures and tables

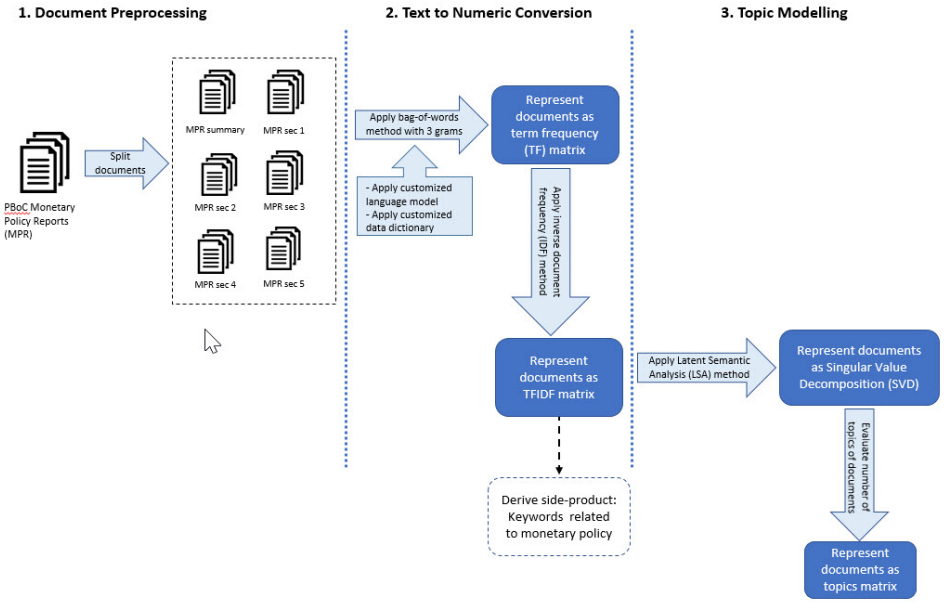

Figure 1: Methodology: Text Analytics

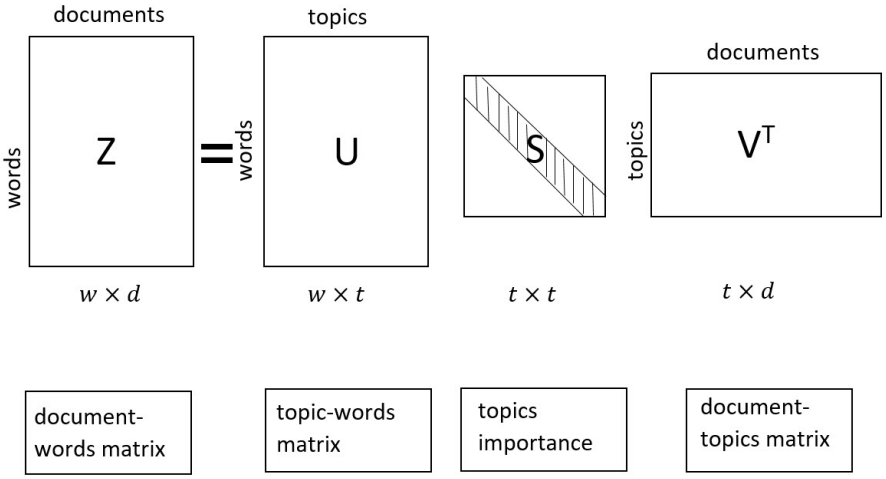

Figure 2: Singular Value Decomposition

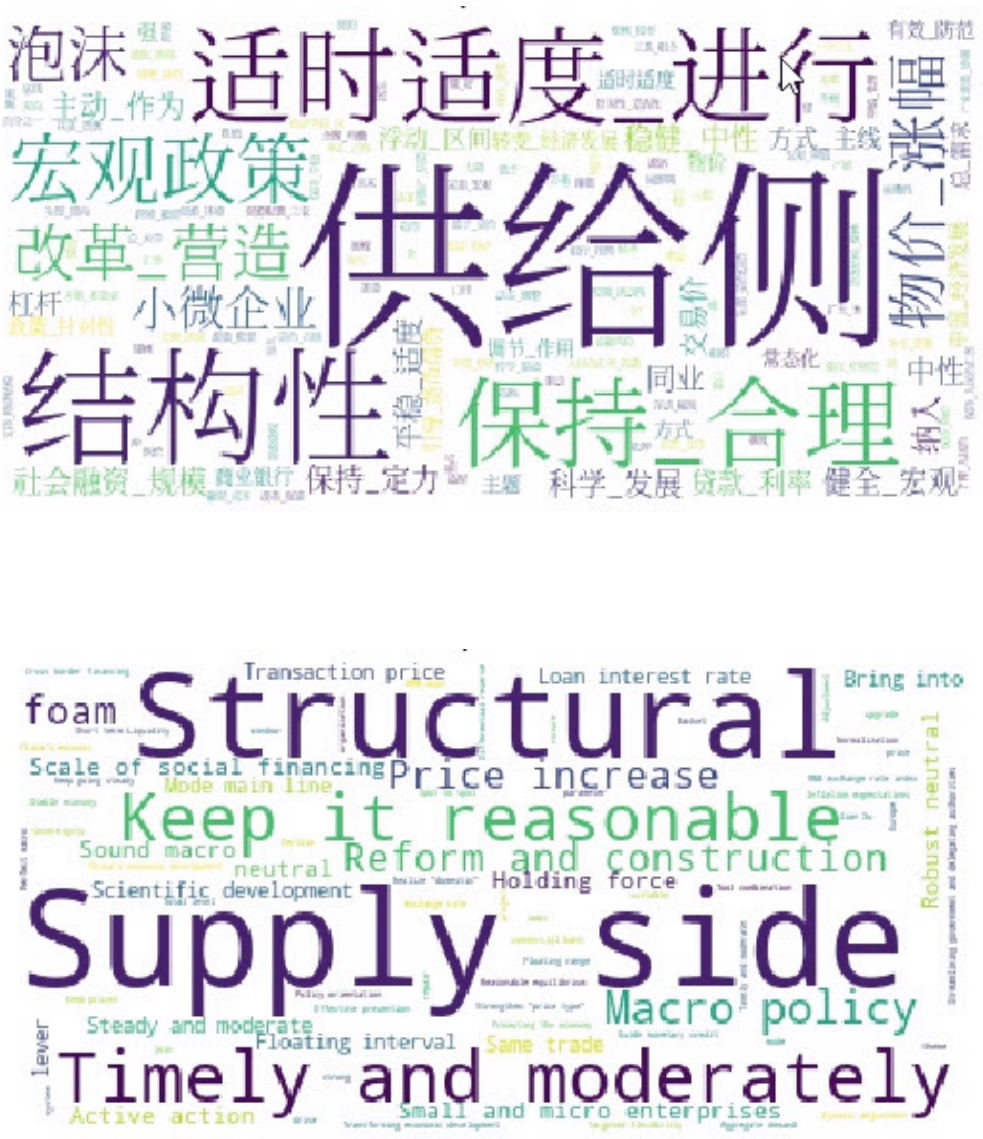

Figure 3: Word Clouds for Monetary Policy Report, Summary Topic 3, in Chinese and English

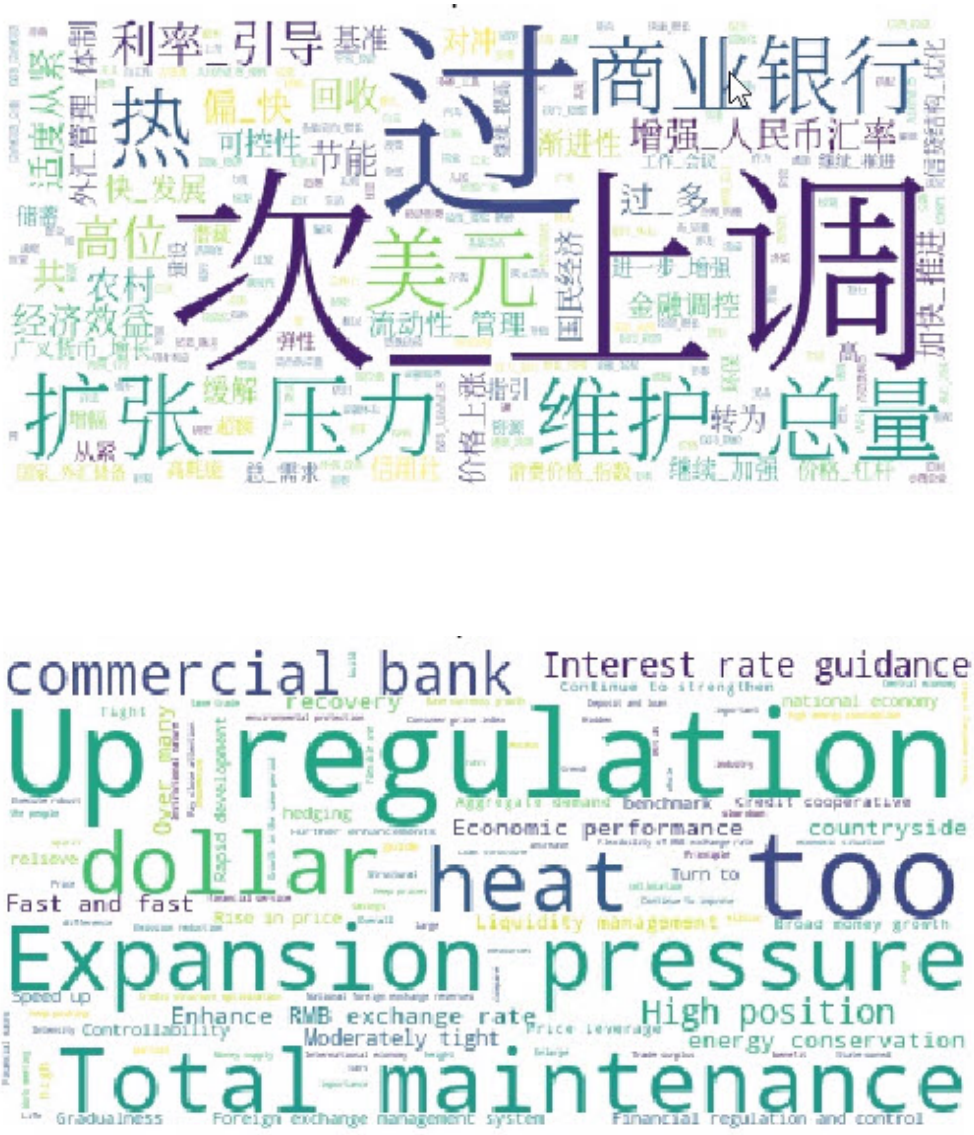

Figure 4: Word Clouds for Monetary Policy Report, Summary Topic 5, in Chinese and English

Figure 5: Word Clouds for Monetary Policy Report, Section 3, Topic 3, in Chinese and English

Figure 6: Word Clouds for Monetary Policy Report, Section 4, Topic 6, in Chinese and English

Figure 7: Word Clouds for Monetary Policy Report, Section 4, Topic 9, in Chinese and English

Figure 8: Word Clouds for Monetary Policy Report, Section 3, Topic 4, in Chinese and English

Figure 9: Word Clouds for Monetary Policy Report, Section 4, Topic 5, in Chinese and English

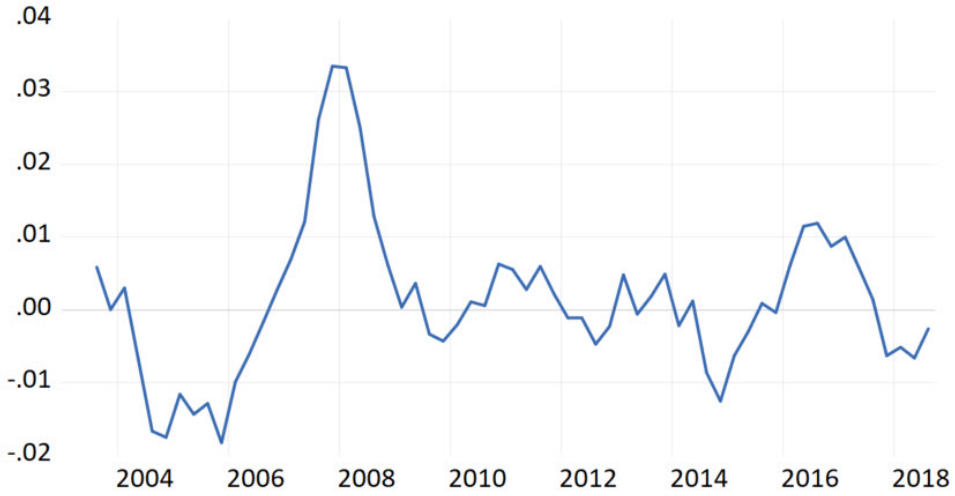

Figure 10: Unobserved Component (μ_t)

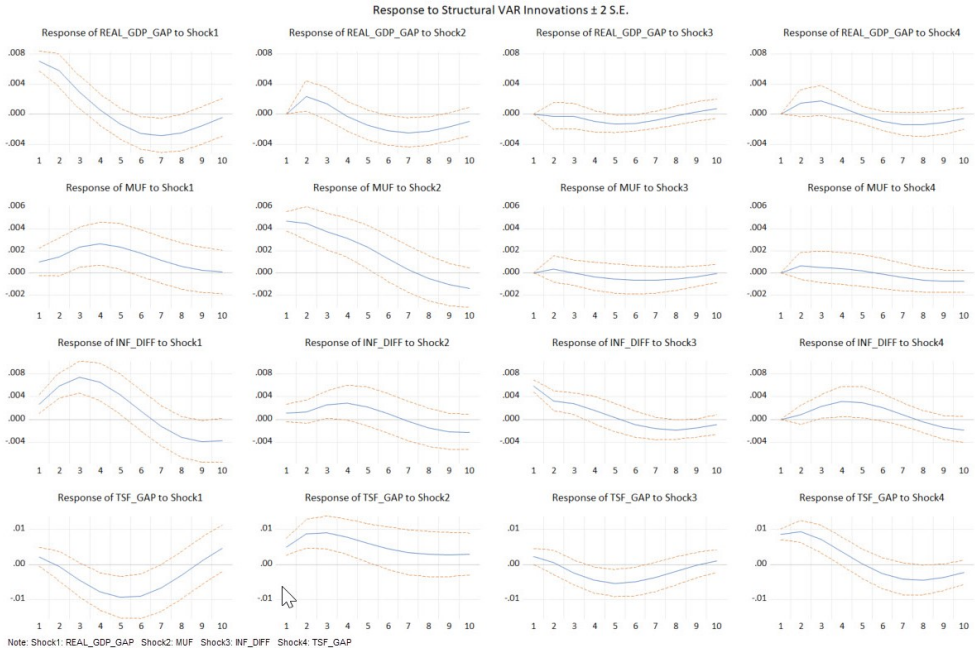

Figure 11: Responses of Variables to Different Shocks

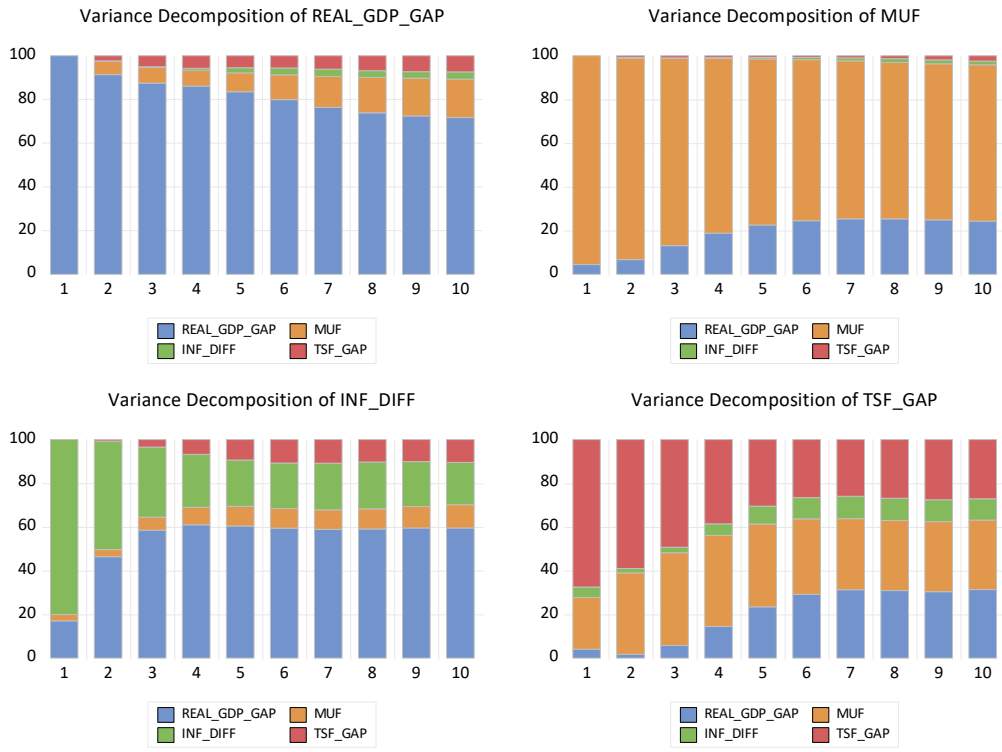

Figure 12: Variance Decomposition

Table 1: Official Documents: Structure of Monetary Policy Reports

Section	Period covered	Number of docs
Summary	2002Q3–2018Q3	65
Section 1: Money and credit	2002Q4–2018Q3	64
Section 2: Monetary policy	2001Q1–2018Q3	71
Section 3: Financial market conditions	2001Q1–2018Q3	71
Section 4: Macroeconomic conditions	2001Q1–2018Q3	71
Section 5: Monetary policy trends	2001Q1–2018Q3	71

Table 2: Number of Topics for Monetary Policy Report Sections

MPR section	Number of topics
Summary	8
Section 1: Money and credit	2
Section 2: Monetary policy	6
Section 3: Financial market conditions	6
Section 4: Macroeconomic conditions	11
Section 5: Monetary policy trends	12

Table 3: Granger Causality Test Results (MPR, Section 3 Topics on Topic 3)

Topics from MPR Section 3	p-values
Topic 1	0.9320
Topic 2	0.6051
Topic 3	0.0145
Topic 4	0.9021
Topic 5	0.5462
Topic 6	0.2321

Table 4: Granger Causality Test Results (MPR, Section 4 Topics on Topic 3)

Topics from MPR Section 4	p-values
Topic 1	0.1280
Topic 2	0.6342
Topic 3	0.2986
Topic 4	0.8451
Topic 5	0.4608
Topic 6	0.0782
Topic 7	0.4176
Topic 8	0.8612
Topic 9	0.0093
Topic 10	0.1713
Topic 11	0.2312

Table 5: Granger Causality Test Results (MPR, Section 3 Topics on Topic 5)

Topics from MPR Section 3	p-values
Topic 1	0.8339
Topic 2	0.3828
Topic 3	0.871
Topic 4	0.0136
Topic 5	0.2842
Topic 6	0.712

Table 6: Granger Causality Test Results (MPR, Section 4 Topics on Topic 5)

Topics from MPR Section 4	p-values
Topic 1	0.6191
Topic 2	0.106
Topic 3	0.9915
Topic 4	0.2375
Topic 5	0.0037
Topic 6	0.3941
Topic 7	0.462
Topic 8	0.8021
Topic 9	0.1922
Topic 10	0.8084
Topic 11	0.7159

Table 7: Maximum Likelihood Estimation Results eq.(1)

		Coefficient
α_0	constant	-0.008
		[0.000]
α_1	ΔTSF_{t-1}^{gap}	0.834
		[0.000]
α_2	$(\pi_{t-1} - \pi_{t-1}^{*})$	-0.764
		[0.000]
α_3	y_{t-1}	-0.291
		[0.089]
β_0	μ_{t-1}	0.364
		[0.037]
β_1	$topic_{t-1}^{1}$	0.078
		[0.645]
β_2	$topic_{t-1}^{2}$	-0.067
		[0.344]
β_3	$topic_{t-1}^{3}$	0.056
		[0.092]
β_4	$topic_{t-1}^{4}$	-0.031
		[0.353]
β_5	$topic_{t-1}^{5}$	-0.060
		[0.000]
β_6	$topic_{t-1}^{6}$	-0.011
		[0.370]
β_7	$topic_{t-1}^{7}$	0.003
		[0.758]
β_8	$topic_{t-1}^{8}$	0.008
		[0.371]

Notes: The dependent variable is the growth rate of the gap of total social financing (ΔTSF_t^{gap}). The sample comprises quarterly data over the period 2003Q2–2018Q4. The p-values are reported in square brackets.