Disagreement in the FOMC Meeting Transcripts and Monetary Policy Effectiveness*

Xia Zizhe

Dated: November 20, 2018

Abstract

This paper uses the Federal Open Market Committee (FOMC) meeting transcripts to measure the levels of internal disagreement of the meetings and studies the impact of this disagreement level on the monetary policy effectiveness in terms of the responses of the macro variables, such as real GDP, inflation, unemployment, to monetary shocks. I use the GloVe algorithm and the seed word method to measure the disagreement level. A structural vector autoregression model with interaction terms is estimated using the Bayesian approach and the monetary shocks are identified by the recursive structure. I find that under a higher level of the FOMC internal disagreement, the monetary shocks are more influential to the macro variables.

*The codes for this paper are available at: `https://github.com/jordanxzz/Codes_for_Fed_Disagreement_Paper`

1 Introduction

The Federal Reserve Bank (Fed) of the US conducts monetary policy to maintain price stability and high employment level. The Federal Open Market Committee (FOMC) meets regularly eight times a year as the monetary policymaking body of the Fed, but disagreement can occur during the FOMC meetings. When this happens, the policy effectiveness should be concerned, since it is possible that the disagreement reflects the potential problems of the policy, or that the dissent central bankers can exert their influence in other ways, such as emphasizing a different aspect of the economy during their public speeches. It can also be the opposite situation where more disagreement brings more detailed analysis and a more appropriate and effective policy decision. In this study, I empirically examine whether this internal disagreement of FOMC could affect the effectiveness of monetary policy by constructing an index of the disagreement and sentiment of the central bankers using the FOMC verbatim transcripts and examining the effectiveness of the monetary policy tools during periods with different levels of disagreement.

The empirical study is based on an index of the internal disagreement level of the FOMC meetings. This index is constructed by an automatic textual analysis method applied to the verbatim transcripts of the FOMC meetings. This method is mainly based on the seed word approach proposed by Rheault et al. (2016) and global word vector representation algorithm (GloVe) by Pennington et al. (2014). It can learn the domain-specific word features and detect the words with high levels of positive and negative sentiment under the context of the FOMC meetings. The disagreement level can then be defined by how frequently the negative words would appear in the meetings compared to the positive words. To examine how this disagreement level could affect the monetary policy effectiveness, I estimate a structural vector autoregression (SVAR) model of the US macro variables with the uncertainty index as an exogenous variable with interactive effect. This interacted VAR methodology is developed by Towbin and Weber (2013); Aastveit et al. (2013). The Bayesian VAR estimation technique I use in the estimation is proposed by Koop et al. (2010). The monetary shocks are identified by the commonly used recursive strategy which is extensively used by many other monetary policy analysis studies [See, for instance, Carrillo and Elizondo (2015); Christiano et al. (1999)]. I define the policy effectiveness to be the level of responses of the macro variables, such as real GDP, unemployment and inflation, to a monetary policy shock, say, a rate hike by the Fed. The policy is considered more effective if the responses of macro variables are of a larger scale. For a robustness check, I also include the economic uncertainty level as an exogenous variable in the model, measured by the economic uncertainty (EPU) index constructed by Baker et al. (2016).

Several reasons lead to the choice of the FOMC meeting transcripts as the data input for the disagreement measure. First, they provide the most detailed information of the policymaking process than other information sources, such as the immediately published meeting minutes containing the summary and the formal voting results and the relatively short policy statements. Since 2012, FOMC started to disclose its members' individual assessment of the appropriate future rate path, but this data is only available after 2012 and may only reflect one aspect of their disagreement. Moreover, Lucca and Trebbi (2009) argue that more information can be extracted from the FOMC transcripts in their analysis of the transcripts to measure policy communication and predict the future rate changes. Second, since the FOMC has to make a policy decision after the meeting, sometimes the disagreement cannot be observed from the final policy outcomes. Meade (2005) collects the internal dissent during the FOMC discussions of the Fed Chairman Alan Greenspan's tenure and points out that the official dissent rate of 7.5 percent is much lower than the 30 percent internal disagreement rate. Besides, the transcripts will only be available after five years and they are not a policy tool themselves. Hence the transcripts are chosen to measure the internal disagreement of the FOMC meetings.

My main finding is that during periods of high levels of disagreement, the monetary shocks have a larger impact on the US macro variables. Under a high level of FOMC internal disagreement, the monetary shock of a one percent increase in the short-term interest rate could drag down the real GDP by more than twice as it could under a low disagreement level. The effect on inflation is less robust, but a larger response is still observed under a higher level of disagreement. The unemployment rate is shocked by only a-third when the disagreement level is low. After including the economic uncertainty level as an exogenous variable, the impact of the disagreement level becomes more significant. My result suggests that the designing of the Fed's regulations should not stand in the way of expressing dissent by the FOMC members, as it could make the interest rate tools less effective.

This study is most relevant to two branches of literature. The first branch is about analyzing the sentiment from the text documents by the policymakers, such as central bank meetings and the congressional debates documents. Lucca and Trebbi (2009) use the semantic orientation method to measure the hawkishness and dovishness of the central bank communication and examined the market reactions, while Nardelli et al. (2017) claim that supporting vector machine (SVM) could provide more stable measurements and predictions for the Fed's interest rate policy. Hansen et al. (2017) examine the impact of the effectiveness of the Federal Reserve's transparency policy by analyzing the FOMC meeting transcripts before and after the Fed agreed to publish the FOMC meeting transcripts. Apart from the analysis of central bank documents, sentiment analysis of the transcripts has also been adopted to measure the emotion of parliamentary debates by Rheault et al. (2016) to study the relation between the economic situation and the mood of politicians. Thomas et al. (2006) applied some textual analysis tools to the US congressional debate transcripts to determine the support and opposition of the debates. But to my knowledge, currently, there are only few studies directly analyzing the disagreement level by the FOMC meeting transcripts.

The other branch of literature consists of studies assessing the monetary policy effectiveness. Pancrazi and Vukotic (2016) estimate a medium-scale New-Keynesian model for two time periods to study how the effectiveness of the conventional monetary policy changes over the past few decades. Smets and Wouters (2003) estimate a dynamic stochastic general equilibrium (DSGE) model with sticky prices and wages using the Euro area data, and Smets and Wouters (2007) extend this analysis to a more complicated Bayesian DSGE model for the US business cycles with the Bayesian vector autoregression (BVAR) approach and study the monetary policy shocks. Bloom (2009) also apply the VAR model to study how the uncertainty shocks would affect the economy and point out that high levels of uncertainty can make the policy shock less effective. An interacted VAR model with respect to panel data is developed and used by Towbin and Weber (2013); Sá et al. (2014) to study the interaction effect of economic uncertainty, treating it as an exogenous variable. Aastveit et al. (2013) apply a similar interacted VAR approach without using the panel data formulation. The VAR approach is standard in the related macroeconomic empirical research.

The rest of the paper is organized as follows. Section 2 describes the measurement method of the disagreement index in details. Section 3 describes the macro data and the estimation and identification strategy for the BVAR model. Section 4 presents the results and some discussion. Section 5 concludes the paper.

2 Mining the Disagreement

To study whether the internal disagreement of FOMC would impact the monetary policy effectiveness, I first construct a measure of the internal disagreement from the FOMC verbatim transcripts from 1993 to 2012. As mentioned in Section 1, this method is based on the seed word method to create domain-specific lexicons by Rheault et al. (2016) and the GloVe algorithm for word representation by Pennington et al. (2014). The basic idea of this method is to first select two lists of commonly used words as the positive and negative seed words. Then the unsupervised learning algorithm called GloVe is applied to obtain the vector representations of the words based on the contextual word-word co-occurrences. Such vector representations can capture the similarity between different words under a certain context. Words with highest similarities to the positive (or, negative) seed words are considered to be mostly related to positive (or, negative) meanings in the FOMC meeting context, which can reflect the agreement (or, disagreement) level.

2.1 The Text Corpus

My text corpus consists of the FOMC verbatim transcripts from 1993 to 2012 for all regular meetings each year with eight meetings per year, covering a total of 160 meetings. An automatic script is used to download all FOMC verbatim transcripts from the Fed's website[1]. Only the transcripts of the regular meetings are included in the analysis since some conference calls may not be directly about monetary policy and those transcripts are not as complete as the regular meeting transcripts. The starting point of 1993 is chosen mainly for consistency since the Fed has agreed to publish these verbatim transcripts of the FOMC meetings except only a small part of the text for confidential reasons (for instance, involving foreign officials) since 1993. Though the records before that are available, they might not

[1] https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm

be consistent with the later documents, as Hansen et al. (2017) point out that the transparency policy itself can affect the discussion of the FOMC meetings. Besides, as described by the Fed's website, no present or past members of the Committee reviewed the edited pre-1994 transcripts, while later the transcripts are produced by the FOMC Secretariat shortly after each meeting from an audio recording of the proceedings. To make sure the result is not affected by the time period chosen, I also extend the starting point to 1980 and redo the analysis. The result is not heavily affected by the starting year. However, the main focus is the transcripts since 1993.

The textual documents first go through a series of standard pre-processing steps including lemmatization , part-of-speech (PoS) tagging[2], and filtering. The lemmatization step reduces the words to its lemma. This ensures that different forms of the same word are considered the same. For instance, "increases" and "increase". On the other hand, PoS tagging distinguishes the possible different meanings of the same words by attaching the words their corresponding part-of-speeches. An example would be the word "right" in the phrases "the right candidate for the job" and "the right to disagree". Some PoS tags are similar, for instance, singular and plural nouns, verbs of different tenses, and comparative and superlative adjectives and adverbs, which are grouped together as nouns, verbs, adjectives and adverbs. I then filter out all the proper nouns, cardinal numbers, existential *there*, foreign words, symbols, and *to*, as well as a list of the common stopwords in English[3] since these words may appear very frequently but carry little information. Including them may contaminate the result. These steps help avoid substantial problems. The documents are then reduced to lists of the word lemma and PoS tag pairs to be used in the word vectorization, which I refer to as "lemmas" in later sections. After the pre-processing, the corpus contains a vocabulary of 38,310 different lemmas.

2.2 Methodology

After obtaining the lemmas, I then use the seed word method with the GloVe algorithm to measure the disagreement. Many approaches have been developed in the literature to mine the emotions and opinions, but it is usually unreliable to apply most of them across domains directly. For instance, a polarity lexicon created for social media opinion mining may not perform well in the formal and professional discussion context of the FOMC meetings. Some sentiment lexicons are created for general purposes [for instance, Wiebe et al. (2005)], yet those results are not tailored to the special environment of FOMC meetings.

Rheault et al. (2016) propose the seed word method to overcome these problems. I adopt their positive and negative seed word lists[4], as presented in Table 1 and Table 2 which consist of only 100 words for both sides as the seed word lists

[2] A full list of tags is available at `https://www.ling.upenn.edu/courses/Fall_2003/ling001/penn_treebank_pos.html`

[3] `http://snowball.tartarus.org/algorithms/english/stop.txt`

[4] Adopted from the technical appendix of Rheault et al. (2016) at `https://journals.plos.org/plosone/article/file?type=supplementary&id=info:doi/10.1371/journal.pone.0168843.s006`

in this study. The lists include only very basic and common words with an unambiguous emotional orientation but they are not linked to any procedure or topics of the FOMC meetings. For instance, the word "unemployment" is excluded since mentioning it during an FOMC meeting is common and does not necessarily have a positive or negative sentiment linked to it. Those words are used as the starting point to generate the FOMC context specific lexicons.

The GloVe algorithm is applied to get the vector representation of all lemmas in the vocabulary. The details of the algorithm are described by Pennington et al. (2014). I use a symmetric contextual window of 15 words and compute the word vectors of 100 dimensions for each lemma. The number of dimensions of the vector representation is a hyperparameters of the model defining the size of the word vectorization results and is chosen arbitrarily here, since it is merely an input needed for the algorithm to calculate the vector representation and it does not have a meaningful interpretation. Other choices of this parameter are also used to examine whether the hyperparameters can influence the measure and the result to be presented in Section 2.3 and Figure 1 shows that this dimension would not substantially influence the disagreement measure. The sentiment score of each word in the vocabulary is given by its similarity to the positive and negative seed words in terms of the word vector representations by the formula:

$$Score_i = \sum_{p=1}^{P} \frac{\boldsymbol{v_i} \cdot \boldsymbol{v_p}}{\|\boldsymbol{v_i}\| \|\boldsymbol{v_p}\|} - \sum_{n=1}^{N} \frac{\boldsymbol{v_i} \cdot \boldsymbol{v_n}}{\|\boldsymbol{v_i}\| \|\boldsymbol{v_n}\|} \tag{1}$$

where $\boldsymbol{v_i}$ is the vector corresponding to lemma i, and $p \in \{1, 2, 3, ...P\}, n \in \{1, 2, 3, ..., N\}$ indexes the positive and negative seed words respectively. With the scores for all the individual lemmas in the vocabulary, I construct the domain-specific lexicons using the lemmas with the highest and lowest-scored 1000 words.

After obtaining the lexicons L, the disagreement levels for each FOMC meeting are measured by summing up the score of all the lemmas of a transcript that appear in the lexicons. To further account for the negative use of the words, I flip the signs of the scores of the words in a negative sentence between any one of the words *not, no, never, neither* and *nor*, and the sentence ending. The disagreement index for time t is thus given by:

$$x_t = \sum_{i=1}^{N_t} k_{it} \cdot Score_i \cdot \mathbf{1}\{w_{it} \in L\} \tag{2}$$

where x_t is the sentiment index for meeting transcript at time t, w_{it} is any word in transcript t, k_{it} indicates whether the word is in a negative clause and equals -1 if the usage is negated and 1 otherwise.

2.3 Disagreement Measure

Since the macro data to be used in the empirical study is quarterly, the raw disagreement scores are then converted into a quarterly index by taking the average scores of the meetings within the same quarter. There are eight regular meetings per year with two in each quarter approximately. Sometimes the meetings are dated at the beginning of the next quarter (for example, 3 April), but I still divide the eight meetings equally into the four quarters and calculate the average scores for simplicity. The result is presented in Figure 1. I calculate the scores using both dimension 100 and 200 to ensure that the calculation is not dependent on the choice of the hyperparameter. Figure 1 shows that indeed the dimension has little influence on the scores since the shapes of the two curves are almost the same. The 100-dimension scores are selected for the regression. To provide a intuitive cross-validation of the disagreement measure, I also calculate a naïve count index as presented in Figure 2. This is obtained by simply count all the interrogative marks and the *wh*-words, which can be considered some rough representations of the disagreement and doubts. This naïve count is in general consistent with the disagreement index with similar peaks and troughs, except that during the first several years the naïve count is relatively higher.

The result also shows that the FOMC generally disagrees more during the periods of crisis, such as the dot.com bubble crash and the financial crisis in 2007 - 2009. But the disagreement level is more volatile than the economy. Even during periods of economic expansion, the disagreement level might not necessarily decrease. The reason might be that the FOMC is concerned about a rate hike during expansion and the timing for the hike is also debatable. Besides, the oscillation pattern also reveals that there might be a seasonal effect on the disagreement level of the FOMC meetings.

3 Empirical Model

The US quarterly macro data of output, inflation, unemployment rate and short-term interest rate from 1993Q1 to 2012Q4, together with the disagreement index, are used to estimate the empirical model. The output, inflation, and the short-term interest rate are measured by real GDP, CPI, and effective federal funds rate, respectively.

3.1 Model Specification

I use a VAR model with interaction terms to study how different levels of disagreement would affect the effectiveness of monetary policy. The interaction term enables the evaluation of the impact of the disagreement level. This interacted VAR method is developed by Aastveit et al. (2013). I assume the disagreement index x_t to be exogenous and let it interact with only the short-term interest rate since in this study I want to examine how the other macro variables would respond to a shock in the short-term interest rate. The lagged interaction between the macro variables is accounted for by a maximum of $p = 8$ quarters time lag. The model is presented as follows:

$$\boldsymbol{y_t} = \boldsymbol{a_0} + \boldsymbol{b_0} x_t + \sum_{l=1}^{p} (A_l \boldsymbol{y_{t-l}} + B_l x_t \boldsymbol{y_{t-l}}) + C \boldsymbol{z_t} + \boldsymbol{e_t} \tag{3}$$

where x_t is the disagreement index, $\boldsymbol{y_t}$ is the vector of the dependent variables consisting of output, inflation, unemployment rate and short-term interest rate, $\boldsymbol{z_t}$ is the exogenous variables other than x_t, $\boldsymbol{e_t}$ is the residual, and $\boldsymbol{a_0}, \boldsymbol{b_0}, A_l, B_l, C$ are the coefficient matrices to be estimated. Since x_t can only interact with part of $\boldsymbol{y_t}$, the matrices B_l are assumed to have some zero columns.

The exogenous variables $\boldsymbol{z_t}$ are added for the robustness check of the model. One important exogenous variable to be included is the economic uncertainty. As suggested by many theoretical and empirical works, higher uncertainty level may trigger the cautiousness effects which can delay the investment decision and cause the ineffectiveness of monetary policy [Aastveit et al. (2013); Bloom et al. (2007); Bloom (2009); Bloom et al. (2018); Bekaert et al. (2013)]. Since the economic uncertainty and the disagreement level may be correlated, I also include the uncertainty as an exogenous variable as a robustness check. The economic uncertainty is measured by the US economic policy uncertainty (EPU) index[5] composed by Baker et al. (2016). The EPU index consists of the frequencies of ten large newspaper coverage of news related to economic policy uncertainty, the number of federal tax code provisions scheduled to expire over the next ten years, and the disagreement among the inflation and government purchases forecasters from the Federal Reserve Bank of Philadelphia's Survey of Professional Forecasters. The original monthly index is averaged to be a quarterly data in the regression.

3.2 Estimation and Identification

The Bayesian estimation techniques are applied to estimate the model. The prior distributions of the regression coefficients and the covariance matrix are imposed to be of the Normal-Wishart distribution, as suggested by Koop et al. (2010). The prior is set to be uninformative and the posterior sample is jointly drawn from the posterior distribution using a Gibbs sampler. More technical details are described in Appendix A.

To identify a monetary shock, a recursive structure of the VAR model is assumed. By a Cholesky decomposition of the estimated variance matrix, I assume that the short-term interest rate can respond to all the macro variables within the same period, while the macro variables respond to monetary shocks with a lag. This recursive structure is well-established in the SVAR literature and is applied in many studies [for instance, Stock and Watson (2001)].

The impact of different disagreement levels on the monetary policy effectiveness is evaluated by the estimated impulse response function of the macro variables to monetary shocks at different levels of disagreement. The empirical 90^{th} and 10^{th} percentiles, denoted by x_H and x_L, are set to be the value of the exogenous

[5] A detailed description is available at `http://www.policyuncertainty.com/methodology.html`

interaction variable x_t for the estimated VAR model:

$$\begin{aligned} \boldsymbol{y_t^H} &= \hat{\boldsymbol{a}}_{\boldsymbol{0}} + \hat{\boldsymbol{b}}_{\boldsymbol{0}} x_H + \sum_{l=1}^{p} (\hat{A}_l \boldsymbol{y_{t-l}} + \hat{B}_l x_H \boldsymbol{y_{t-l}}) + \hat{C} \boldsymbol{z_t} + \hat{\boldsymbol{e}}_{\boldsymbol{t}} \\ \boldsymbol{y_t^L} &= \hat{\boldsymbol{a}}_{\boldsymbol{0}} + \hat{\boldsymbol{b}}_{\boldsymbol{0}} x_L + \sum_{l=1}^{p} (\hat{A}_l \boldsymbol{y_{t-l}} + \hat{B}_l x_L \boldsymbol{y_{t-l}}) + \hat{C} \boldsymbol{z_t} + \hat{\boldsymbol{e}}_{\boldsymbol{t}} \end{aligned} \tag{4}$$

After estimating the impulse response functions using the reduced models above, I calculate the differences of the impulse responses for each of the 4000 draws from the posterior distribution. To study whether the responses under different levels of disagreement are different from each other, I estimate the empirical distribution of the difference and obtained the 90% and 68% probability bands of the difference. Hence if the whole probability band is above or below zero, the two impulse effects can be considered statistically different.

4 Results and Discussion

The main purpose of this study is to evaluate the impact of the internal disagreement on the monetary policy effectiveness. Figure 3 displays the 12-quarter responses of the macro variables including output, inflation, and unemployment to a one percent increase in the short-term interest rate during periods of high disagreement level. This interest rate shock can be interpreted as a one percent unanticipated policy rate hike by the Fed. Figure 5 shows the same responses under low level of disagreement. The probability bands of the empirical distribution of the differences in the responses are reported in Figure 7. From the figures, it can be observed that an increase in the short-term interest rate would lower the real GDP level and the inflation, and increase the unemployment regardless of the disagreement level. From Figure 3 and 5, responses of the real GDP is almost tripled to a two percent decrease after three years of the shock when the disagreement level increases from the lower decile to the upper decile. A similar phenomenon can be observed for the unemployment rate and inflation rate. The monetary shock does not have a clear effect on the inflation rate under a low disagreement level, while the impact turns out to be more negative when the disagreement level increases. The responses of the unemployment rate are more sensitive to the change of the disagreement level with a three percent more increase under the high disagreement level. It should also be pointed out that despite disagreement level has a sizable effect on the macro variables, the impact is small during the first year after the shock. Figure 7 shows that under a higher level of disagreement, the responses of the macro variables are larger, yet the 68% probability bands are not always above or below zero.

As mentioned in Section 3.1, economic uncertainty measured by the EPU index is included as an exogenous variable for a robustness check. The result is reported in Figure 4, 6 and 8. The model is robust to the inclusion of the EPU index as an exogenous variable. As shown in Figure 8, the difference of the impact can be considered statistically different at 68% level. Compared with the previous model specification, the differences betweem the two impulse responses become slightly larger after adding this additional variable. This indicates that the effect of the disagreement level is robust to the economic uncertainty. But the magnitudes of both responses are still small during the first several quarters after the shock.

In general, higher levels of internal disagreement of the FOMC meetings can increase the effects of monetary shocks. This increase may result from more different ideas expressed during the meeting and that the policy decisions are made after considering the different opinions of different members. The empirical result shows that the disagreement during the meeting needs to be encouraged. Since the meeting transcripts are not directly available to the general public, expressing different ideas may not cause more policy uncertainty in the economy compared to casting a dissenting vote, and it can benefit the policy effectiveness empirically. An important policy implication of the result is that the FOMC's organization structure and regulations should encourage the disagreement during the meetings instead of suppressing it. For instance, some studies are concerned about the side effect of the Fed's transparency policy, arguing that such kind of policy can have some negative effect by incentivizing conformity. Theoretical models by Schultefrankenfeld (2017); Prat (2005) suggest that transparency increases conformity level if the agents do not know their expertise, trying to avoid being on the wrong side. Meade and Stasavage (2008) also discover that the members' career concern may discourage them from expressing dissent under some transparency policy. The disagreement hence might not be expressed during the meeting due to those transparency regulations, which undermines the effectiveness of the interest rate tools. However, the result is not against the transparency policy. Rather, it suggests that the Fed's regulations should not discourage internal disagreement.

5 Conclusion

In this study I use the Federal Open Market Committee (FOMC) meeting transcripts to measure the levels of internal disagreement of the meetings and study the impact of this disagreement level on the monetary policy effectiveness in terms of the scale of the responses of the macro variables, such as real GDP, inflation, unemployment, to monetary shocks. I use the GloVe algorithm and the seed word method to measure the disagreement level and compose an index. A structural vector autoregression model with interaction terms is estimated using the Bayesian approach and the monetary shocks are identified by the recursive structure assumption. I find that with a higher level of the FOMC internal disagreement, the monetary shocks are more influential to the macro variables. After adding the EPU index, the reaction of the macro variables can be considered statistically larger when the disagreement level is higher. My result supports some studies concerning the negative effect of the Fed's transparency policy. If the policy suppresses the internal disagreement, the interest rate tools would be less effective. The policy implication here is that the Fed's organization design should not discourage the members to express their dissent. Further research could focus on the theoretical models of the internal disagreement and the policymaking process of FOMC. Moreover, more advanced econometric tools can be developed and applied to treat the disagreement level as an endogenous variable since it may also be affected by the economic situation.

References

Aastveit, K. A., Natvik, G. J. J., and Sola, S. (2013). Economic uncertainty and the effectiveness of monetary policy.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Bekaert, G., Hoerova, M., and Duca, M. L. (2013). Risk, uncertainty and monetary policy. *Journal of Monetary Economics*, 60(7):771–788.

Bloom, N. (2009). The impact of uncertainty shocks. *econometrica*, 77(3):623–685.

Bloom, N., Bond, S., and Van Reenen, J. (2007). Uncertainty and investment dynamics. *The review of economic studies*, 74(2):391–415.

Bloom, N., Floetotto, M., Jaimovich, N., Saporta-Eksten, I., and Terry, S. J. (2018). Really uncertain business cycles. *Econometrica*, 86(3):1031–1065.

Carrillo, J. A. and Elizondo, R. (2015). Identification of a monetary policy shock in a small open economy. Technical report, Technical report.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (1999). Monetary policy shocks: What have we learned and to what end? *Handbook of macroeconomics*, 1:65–148.

Hansen, S., McMahon, M., and Prat, A. (2017). Transparency and deliberation within the fomc: a computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Koop, G., Korobilis, D., et al. (2010). Bayesian multivariate time series methods for empirical macroeconomics. *Foundations and Trends® in Econometrics*, 3(4):267–358.

Lucca, D. O. and Trebbi, F. (2009). Measuring central bank communication: an automated approach with application to fomc statements. Technical report, National Bureau of Economic Research.

Meade, E. E. (2005). The fomc: preferences, voting, and consensus. *Federal Reserve Bank of St. Louis Review*, 87(March/April 2005).

Meade, E. E. and Stasavage, D. (2008). Publicity of debate and the incentive to dissent: evidence from the us federal reserve. *The Economic Journal*, 118(528):695–717.

Nardelli, S., Martens, D., Tobback, E., et al. (2017). Between hawks and doves: measuring central bank communication. In *IFC-Bank Indonesia Satellite Seminar on Big Data at the ISI Regional Statistics Conference, 21 March, 2017, Bali, Indonesia*, pages 1–12.

Pancrazi, R. and Vukotic, M. (2016). On the effectiveness of conventional monetary policy.

Pennington, J., Socher, R., and Manning, C. D. (2014). Glove: Global vectors for word representation. In *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543.

Prat, A. (2005). The wrong kind of transparency. *American economic review*, 95(3):862–877.

Rheault, L., Beelen, K., Cochrane, C., and Hirst, G. (2016). Measuring emotion in parliamentary debates with automated textual analysis. *PloS one*, 11(12):e0168843.

Sá, F., Towbin, P., and Wieladek, T. (2014). Capital inflows, financial structure and housing booms. *Journal of the European Economic Association*, 12(2):522–546.

Schultefrankenfeld, G. (2017). Appropriate monetary policy and forecast disagreement at the fomc.

Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175.

Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American economic review*, 97(3):586–606.

Stock, J. H. and Watson, M. W. (2001). Vector autoregressions. *Journal of Economic perspectives*, 15(4):101–115.

Thomas, M., Pang, B., and Lee, L. (2006). Get out the vote: Determining support or opposition from Congressional floor-debate transcripts. In *Proceedings of EMNLP*, pages 327–335.

Towbin, P. and Weber, S. (2013). Limits of floating exchange rates: The role of foreign currency debt and import structure. *Journal of Development Economics*, 101:179–194.

Wiebe, J., Wilson, T., and Cardie, C. (2005). Annotating expressions of opinions and emotions in language. *Language resources and evaluation*, 39(2-3):165–210.

Appendix A BVAR Estimation

I use the independent Normal-Wishart prior as proposed by Koop et al. (2010). It is a general framework for VAR modelling compared to other priors such as the natural conjugate or the Minnesota priors. All the right-hand-side terms of Equation 3 can be stacked into a single data matrix Z_t and a coefficient vector β, which is the vectorization of the coefficient matrices. Hence the model can be rewritten as:

$$\boldsymbol{y_t} = Z_t\beta + \boldsymbol{e_t} \tag{5}$$

where $\boldsymbol{e_t}$ is assumed to be i.i.d. $N(0, \Sigma)$. Further stacking all vectors and matrices for different time period vertically, it becomes:

$$\boldsymbol{y} = Z\beta + \boldsymbol{e}, \boldsymbol{e} \sim N(0, I \otimes \Sigma) \tag{6}$$

The general form of the independent Normal-Wishart prior is given by:

$$\begin{aligned} \mathbb{P}(\beta, \Sigma^{-1}) &= \mathbb{P}(\beta)\mathbb{P}(\Sigma^{-1}) \\ \beta &\sim N(\bar{\beta}, \bar{V_\beta}) \\ \Sigma^{-1} &\sim W(\bar{S}^{-1}, \bar{\nu}) \end{aligned} \tag{7}$$

and an uninformative prior is to set $\bar{\beta} = \mathbf{0}, \bar{V_\beta} = I, \bar{\nu} = M + 1, \bar{S} = I$, where M is the number of the dependent variables.

Since the posterior distribution does not have a convenient form, I use a Gibbs sampler to estimate the posterior distribution. A total of 20,000 draws are taken and the first 4,000 are discarded for burn-in. I retain every four draws from the remaining 16,000 to avoid correlations.

Appendix B Disagreement Measure

B.1 Top-Rated Words and Seed Words

Lemma	PoS	Lemma	PoS	Lemma	PoS	Lemma	PoS
well	Adv	happy	Adj	wonderful	Adj	lovely	Adj
good	Adj	perfect	Adj	friendly	Adj	splendid	Adj
important	Adj	gain	Verb	pleasant	Adj	sympathetic	Adj
best	Adj	excellent	Adj	creative	Adj	generous	Adj
better	Adj	superior	Adj	worthy	Adj	vigorous	Adj
TRUE	Adj	fairly	Adv	friendship	Noun	perfection	Noun
love	Verb	reasonable	Adj	sympathy	Noun	appreciate	Verb
able	Adj	secure	Verb	nice	Adj	loving	Adj
help	Verb	efficiency	Noun	honour	Noun	magnificent	Adj
strong	Adj	valuable	Adj	comfort	Noun	integrity	Noun
solution	Noun	properly	Adv	honest	Adj	talent	Noun
importance	Noun	improvement	Noun	genuine	Adj	kindly	Adv
respect	Noun	safe	Adj	healthy	Adj	fortunately	Adv
truth	Noun	desirable	Adj	intelligent	Adj	grateful	Adj
strength	Noun	satisfactory	Adj	welcome	Adj	glorious	Adj
effective	Adj	wise	Adj	helpful	Adj	fortunate	Adj
success	Noun	protect	Verb	encourage	Verb	clever	Adj
freedom	Noun	truly	Adv	praise	Noun	sincere	Adj
significant	Adj	satisfaction	Noun	dignity	Noun	confident	Adj
interesting	Adj	efficient	Adj	prosperity	Noun	delightful	Adj
useful	Adj	joy	Noun	comfortable	Adj	strengthen	Verb
successful	Adj	improve	Verb	reliable	Adj	respected	Adj
beautiful	Adj	enjoy	Verb	succeed	Verb	admirable	Adj
appropriate	Adj	happiness	Noun	delight	Noun	smart	Adj
fair	Adj	glad	Adj	merit	Noun	satisfying	Adj

Table 1: Positive Seed Words

Lemma	PoS	Lemma	PoS	Lemma	PoS	Lemma	PoS
problem	Noun	fail	Verb	hate	Verb	wicked	Adj
death	Noun	sick	Adj	complaint	Noun	disadvantage	Noun
difficult	Adj	unfortunately	Adv	painful	Adj	disappointment	Noun
loss	Noun	confusion	Noun	worry	Verb	unfair	Adj
bad	Adj	burden	Noun	unfortunate	Adj	nonsense	Noun
fear	Noun	anxiety	Noun	neglect	Verb	ridiculous	Adj
failure	Noun	terrible	Adj	prejudice	Noun	undesirable	Adj
enemy	Noun	suffer	Verb	disaster	Noun	imperfect	Adj
wrong	Adj	fault	Noun	distress	Noun	harmful	Adj
difficulty	Noun	anxious	Adj	hatred	Noun	horrible	Adj
pain	Noun	destroy	Verb	tragic	Adj	disastrous	Adj
ill	Adj	worst	Adj	shame	Noun	unsatisfactory	Adj
risk	Noun	excessive	Adj	breach	Noun	hopeless	Adj
danger	Noun	threat	Noun	contempt	Noun	complain	Verb
error	Noun	mistake	Noun	unhappy	Adj	fearful	Adj
evil	Adj	inferior	Adj	frightened	Adj	unjust	Adj
criticism	Noun	weakness	Noun	regret	Noun	irrelevant	Adj
FALSE	Adj	anger	Noun	corruption	Noun	corrupt	Adj
weak	Adj	hurt	Verb	restriction	Noun	unreasonable	Adj
dangerous	Adj	angry	Adj	poorly	Adv	restrict	Verb
excess	Noun	tragedy	Noun	fraud	Noun	careless	Adj
damage	Noun	abuse	Noun	miserable	Adj	grim	Adj
lose	Verb	inadequate	Adj	stupid	Adj	wretched	Adj
worse	Adj	sad	Adj	injustice	Noun	discomfort	Noun
afraid	Adj	harm	Verb	ugly	Adj	brutal	Adj

Table 2: Negative Seed Words

Lemma	PoS	Lemma	PoS	Lermma	PoS	Lemma	PoS
forward	Noun	consensus	Noun	setting	Verb	possible	Adj
addition	Noun	discussed	Verb	clearly	Adv	comfortable	Adj
exception	Noun	keep	Verb	often	Adv	helpful	Adj
evidence	Noun	expect	Verb	likely	Adv	effective	Adj
power	Noun	providing	Verb	actually	Adv	willing	Adj

Table 3: Some Top-Rated Positive Lemmas

Lemma	PoS	Lemma	PoS	Lermma	PoS	Lemma	PoS
markdown	Noun	rebate	Noun	confined	Verb	driven	Adv
down	Noun	outside	Noun	regarded	Verb	poorly	Adj
scarcity	Noun	requiring	Verb	accounting	Verb	military	Adj
deadline	Noun	questioned	Verb	rebuilding	Verb	unknown	Adj
worsening	Noun	avoiding	Verb	imposed	Verb	whenever	WRB

Table 4: Some Top-Rated Negative Lemmas

B.2 Disagreement Index

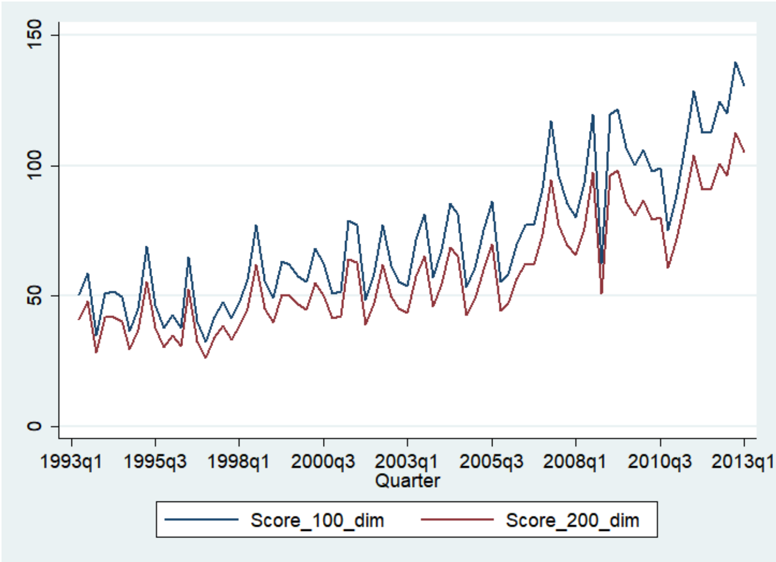

Figure 1: Disagreement Measure since 1993

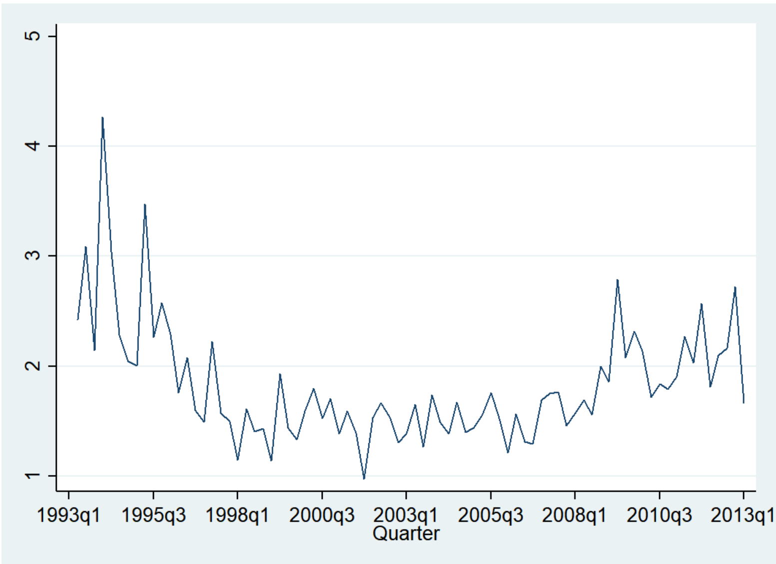

Figure 2: Naive Count since 1993

Appendix C Figures

C.1 Impulse Responses

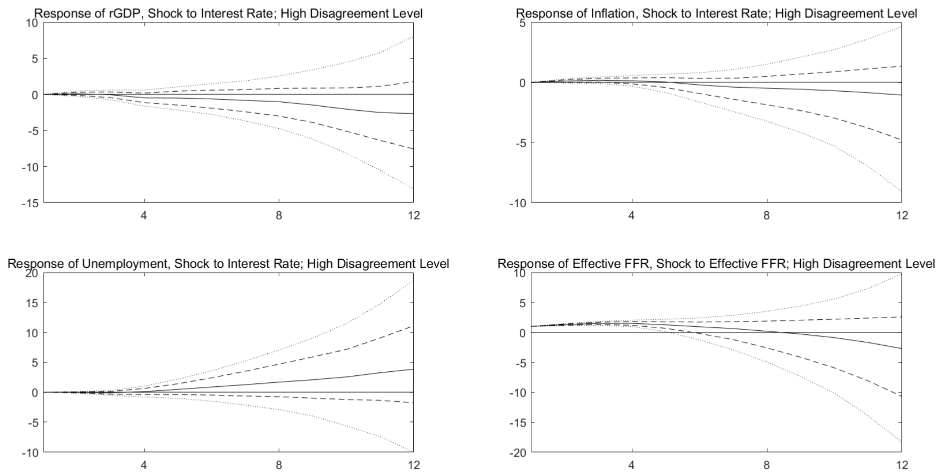

Figure 3: Responses to a 1% Increase in Interest Rate at High Level of Disagreement

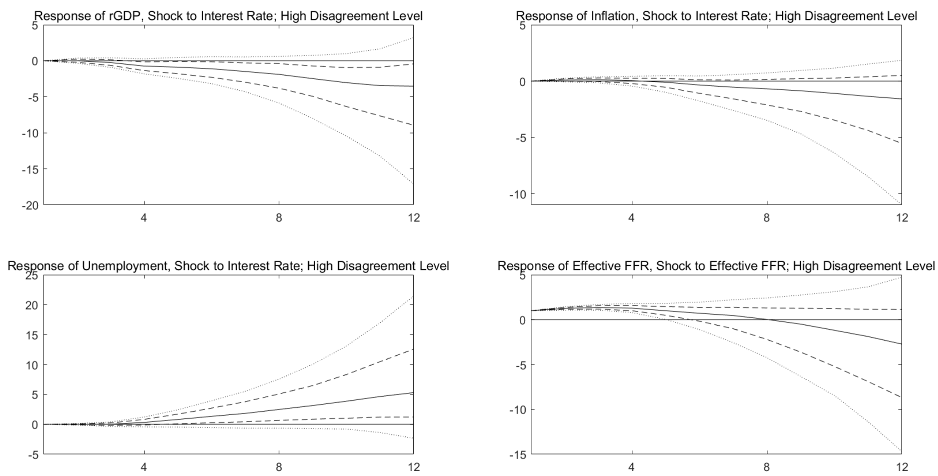

Figure 4: Responses to a 1% Increase in Interest Rate at High Level of Disagreement with EPU Index as an Exogenous Variable

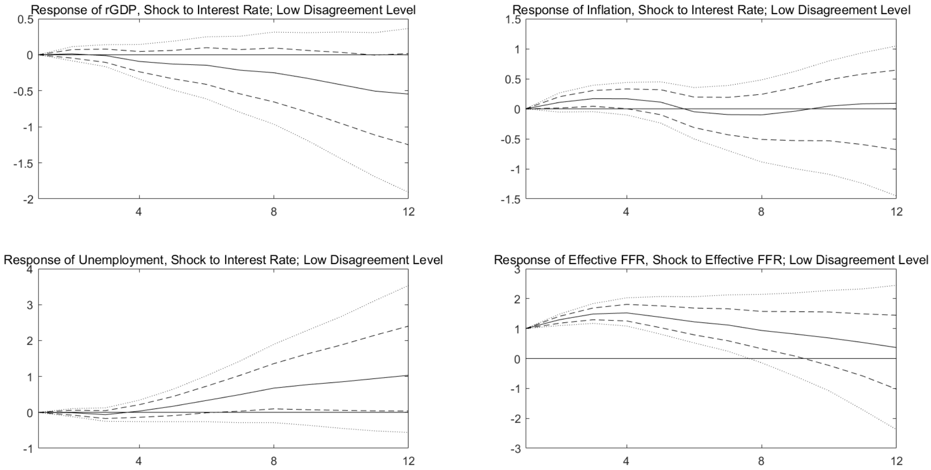

Figure 5: Responses to a 1% Increase in Interest Rate at Low Level of Disagreement

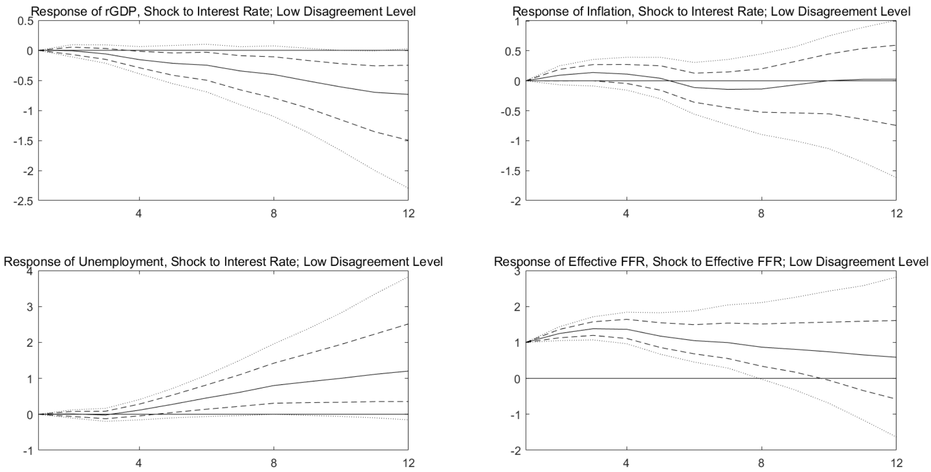

Figure 6: Responses to a 1% Increase in Interest Rate at Low Level of Disagreement with EPU Index as an Exogenous Variable

C.2 Probability Band

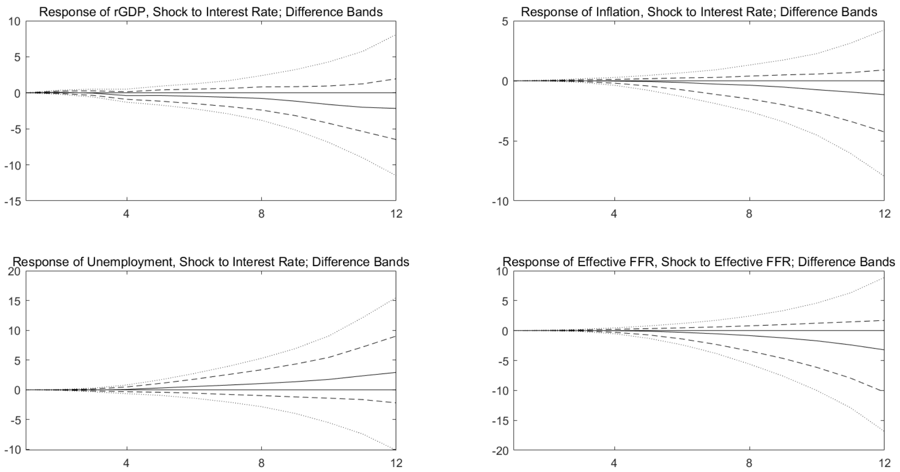

Figure 7: Probability Band of the Differences

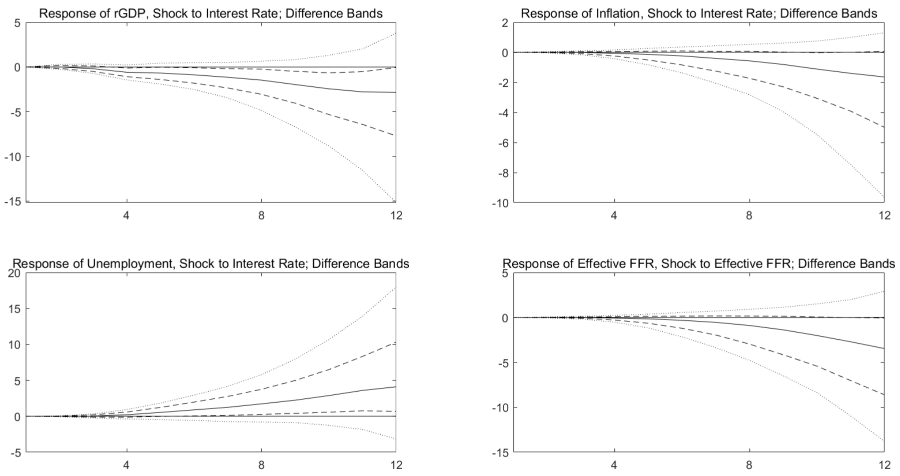

Figure 8: Probability Band of the Differences with EPU Index as an Exogenous Variable