Incorporating the Textual Information of FOMC Meetings in Measuring Monetary Shocks

Shuyan Huang

1 Methodology

1.1 Latent Dirichlet allocation(LDA)

To analyze the information content of the Minutes of FOMC in a multidimensional way, I use the the Latent Dirichlet Allocation (LDA) algorithm first developed by Blei, Ng, and Jordan (2003)[1] to assign topic loadings to the Minutes.

I first extract the paragraphs from each FOMC Minutes document. Then the LDA algorithm views each paragraph as a mixture of various topics, whose distribution follows a Dirichlet distribution. For each word in the paragraph, the distribution over the words is a different Dirichlet distribution for each topic. I will first explain the the variable notations and the generative process, and then the inference step.

I index the Minutes documents by the date of the meeting, t, and the i-th paragraph in a specific document is indexed by ti. The variable names are defined as follows:

θ_{ti} is the topic distribution for the i-th paragraph in the minutes at date t;

φ_k is the word distribution for topic k;

α is the parameter of the Dirichlet prior on the per-document topic distributions;

β is the parameter of the Dirichlet prior on the per-topic word distribution;

w_{tij} is the j-th word in paragraph ti.

z_{tij} is the topic for the j-th word in paragraph ti;

K is the total number of topics;

I_t is the number of paragraphs in the minutes at date t;

N_{ti} is the number of words in paragraph ti.

V is the number of words in total.

The generative process is as follows:

1. For each paragraph ti, sample the topic distribution $\theta_{ti} \sim Dir(\alpha)$, where $Dir(\alpha)$ is a Dirichlet distribution with a symmetric parameter α typically smaller than 1. This step produces a Multinomial

distribution over the topics for each paragraph, i.e. a probability vector $\theta_{ti} = (p_1, p_2, ..., p_K)$ and $\sum_{k=1}^{K} p_k = 1$.

2. For each topic k, sample the word distribution $\varphi_k \sim Dir(\beta)$, and β is also smaller than 1. This step produces a Multinomial distribution over the words for each topic, i.e. a probability vector $\varphi_k = (p_1, p_2, ..., p_V)$ and $\sum_{v=1}^{V} p_v = 1$.

3. For each of the word positions ti, j, first sample a topic from the topic distribution of this paragraph: $z_{tij} \sim Multinomial(\theta_{ti})$. Then given this topic, sample a word from the associated word distribution: $w_{tij} \sim Multinomial(\varphi_{z_{tij}})$.

Given this generative process and the actual words in the paragraphs, we can calculate the likelihood. Although the posterior is computationally intractable, Bayesian inference can still be approximated by variational Bayes. I used the LDA implemented by Gensim[2], which is a realization of Hoffman, Bach and Blei(2010)[3], a online version of the original work by Blei, Ng, and Jordan (2003)[1]. The number of topics is set to $K = 8$, following Jegadeesh and Wu (2017)[4].

After training the LDA model by all the paragraphs, I use the trained model to estimate a topic loading vector for paragraph θ_{ti}. I weighted each paragraph's topic loading vector by its proportion of words in the document, and sum across all paragraphs to get a topic loading vector for a document. Specifically, for document t (the Minutes of the FOMC meeting happened at date t), the topic loading vector is:

$$\theta_t = \sum_{i=1}^{I_t} \theta_{ti} \frac{N_{ti}}{\sum_{i=1}^{I_t} N_{ti}} \tag{1}$$

Note that θ_t is a vector, and the topic k's loading of document t is the k-th element of θ_t, i.e. $\theta_{t,k}$.

1.2 Sentiment Analysis

The topic loadings from LDA provide us with the proportion of the Minutes that the Fed devotes to each topic. In addition to this, the Fed's attitudes towards these topics would be more informative. Therefore, I measure the tone of each paragraph using a bag-of-word approach similar to Jegadeesh and Wu (2017)[4], Tetlock (2007)[5], Loughran and McDonald (2011)[6] and Jegadeesh and Wu (2013)[7]. Specifically, I use the financial sentiment word lists developed by Loughran and McDonald (2011)[6], which contain a list of negative words and a list of positive words. For each paragraph, the number of positive words is Pos_{ti}, and the number of negative words is Neg_{ti}. Then the sentiment score of this paragraph is simply $Pos_{ti} - Neg_{ti}$.

Combining this with the topic loading vector of this paragraph, we can get the sentiment score for each topic in this paragraph:

$$Score_{ti,k} = \theta_{ti,k}(Pos_{ti} - Neg_{ti}) \tag{2}$$

Then weighting each paragraph by the reciprocal of its length (because longer paragraphs are more difficult to process, and the strength of the sentiment expressed by longer paragraphs is weaker), I sum the sentiment score over all the paragraphs in each FOMC Minutes and get the sentiment scores for the Minutes:

$$Score_{t,k} = \sum_{i=1}^{I_t} Score_{ti,k} \frac{1}{N_{ti}} \tag{3}$$

1.3 Estimating Monetary Policy Shocks

After I get the topic loadings and sentiment scores of the Minutes, they can be included in the original regression in Romer and Romer(2004)[8], the residual of which is a new estimate of monetary policy shocks. The motivation for this regression is to overcome the problem of anticipatory movements in traditional monetary policy measures. Although the problem of endogeneity can be largely resolved by using the Fed's targets for the funds rate or non-borrowed reserves as the measure of monetary policies, these targets still contain the Fed's efforts to smooth out fluctuations based on its predictions on the future movement of the economy. So a regression using the Fed's forecast information as explanatory variables should produce residuals that are unaffected by anticipatory movements. The original regression model in Romer and Romer(2004)[8] is:

$$\begin{aligned} \Delta ff_t = \alpha + \beta ffb_t + \sum_{i=-1}^{2} \gamma_i \widetilde{\Delta y}_{ti} + \sum_{i=-1}^{2} \lambda_i \left(\widetilde{\Delta y_{ti}} - \widetilde{\Delta y_{t-1,i}} \right) \\ + \sum_{i=-1}^{2} \varphi_i \tilde{\pi}_{ti} + \sum_{i=-1}^{2} \theta_i \left(\tilde{\pi}_{ti} - \tilde{\pi}_{t-1,i} \right) + \rho \tilde{u}_{t0} + \varepsilon_t \end{aligned} \tag{4}$$

Δff_t is the change in the intended funds rate around FOMC meeting t. ffb_t is the level of the intended funds rate before any changes associated with meeting t, which is included to capture any tendency behavior. $\widetilde{\Delta y}$, $\tilde{\pi}$, and $\tilde{u}$ refer to the forecasts of real output growth, inflation, and the unemployment rate. The subscripts t refer to the FOMC meeting, and i refer to the horizon of the forecasts: -1 is the previous quarter; 0 is the current quarter; and 1 and 2 are one and two quarters ahead, respectively. The previous quarter "forecasts" are often actual data, instead of forecasts. These lagged information is included because recent history is substantially discussed at the meetings and thus is considered in the Fed's decision

making. For real output growth and inflation, the changes in the forecast since the previous meeting are also included because they are also important determinants of the Fed's behavior.

In the original model of Romer and Romer(2004)[8], the proxies for the forecast information of the Fed is the quantitative variables – the forecasts of real output growth, inflation, and the unemployment rate. In addition to these, I include in the regression the topic loadings and tone scores extracted from the texts of the Minutes, as proxies for textual information. Specifically, the residuals ε_t from the regression (4) is used as the dependent variable in the regression model below:

$$\varepsilon_t = \sum_{k=1}^{K} \eta_k \theta_{t,k} + \sum_{k=1}^{K} \rho_k Score_{t,k} + \nu_t \tag{5}$$

The residuals ν_t are my new measure of monetary policy shocks.

1.4 vector autoregression (VAR)

To show how the new measure of monetary policy shocks can be applied to macroeconomic studies and to see the implications of the new measure, I examine the impact of monetary policy on output and price in a vector autoregression (VAR) framework.

For output, the VAR model is:

$$\Delta y_t = a_0 + \sum_{i=1}^{24} b_i \Delta y_{t-i} + \sum_{j=1}^{36} c_j S_{t-j} + e_t \tag{6}$$

Δy is the difference of the log of industrial production, S is the new measure of monetary policy shocks. Following Romer and Romer(2004)[8], I include 24 lags of output growth and 36 lags of the monetary policy measure. And the contemporaneous shock is not included following the convention that monetary policy doesn't affect contemporaneous output. Since I use seasonly adjusted industrial production data, the monthly dummies are not included here.

For price, the VAR model is:

$$\Delta p_t = a_0 + \sum_{k=1}^{11} a_k D_{kt} + \sum_{i=1}^{24} b_i \Delta p_{t-i} + \sum_{j=1}^{48} c_j S_{t-j} + e_t \tag{7}$$

Δp is the difference of the log of PPI for finished goods. Following Romer and Romer(2004)[8], I include 24 lags of price growth and 48 lags of the monetary policy measure, since the impact of monetary policy on price is more persistent. The contemporaneous shock is still not included. Since I use PPI data that are not seasonly adjusted, I also include the monthly dummies D_k.

2 Data and Empirical Results

2.1 Data

From the early 1980s, the FOMC holds eight regularly scheduled meetings per year, during which members discuss the economic outlook and formulate monetary policy. Any policy change decided at the meeting is implemented through open market operations. Detailed records of the discussions during each meeting are summarized in the form of meeting minutes and released to the public after a delay.

From the Federal Reserve's website, I download all FOMC Minutes from February 1993 to June 2012, and extract the paragraphs from the Minutes. Only paragraphs than contain more than 350 characters are counted as meaningful paragraphs. Our sample contains 156 documents and 5748 paragraphs. Then I preprocess the words in the paragraphs by removing the digits, punctuation and stop words developed by Loughran and McDonald (2011)[6]. Collecting all the preprocessed words, I get a word collection of 7815 unique words.

To estimate the regression model (5), I used residual of the regression (4) updated by Basil Halperin[9] using the method of Romer and Romer(2004)[8], from February 1993 to June 2012, since the original work only estimated the monetary policy shocks before 1996.

The industrial output data I use in the VAR is the monthly Industrial Production Index released by the Federal Reserve, from January 1993 to June 2012, since I need the differences of monthly data in my VAR. For the price, I use the monthly Producer Price Index for All Commodities released by U.S. Bureau of Labor Statistics, from January 1993 to June 2012.

2.2 LDA Result

Using the corpus of 5748 paragraphs, I train a LDA with 8 topics, the most important words for all 8 topics are shown in Table 1. The words "prices", "inflation", "energy" in the first topic indicate that this topic is about inflation. In the second topic, the words "exports", "foreign", "domestic" show relevance to international trade. The words "spending", "consumer", "housing" indicate that the third topic is about consumption. The words "dollar", "currencies", "foreign" show that the fourth topic is about the currency market. The fifth topic contain many general terms related to the economic growth. In the sixth topic, the words "bank", "york", "securities" show relevance to the financial market. According to the words "policy", "federal", "meeting", the senventh topic is about the policies implemented by the Fed. Finally, the eighth topic is about production, judge from the words "production", "manufacturing" and "motor".

Topic 1 (Inflation):	Topic 2 (Trade):
prices inflation labor consumer energy price core growth costs increases months remained market markets expectations participants continued markets expectations participants compensation employment economic pace higher increase productivity expected business level rose food rise	growth economic exports market foreign federal domestic activity markets monetary securities financial debt trade imports policy pace continued employment open conditions economy expansion developments economies remained price countries meeting goods
Topic 3 (Consumption):	Topic 4 (Currency):
spending business growth consumer sales activity housing investment equipment continued pace construction expected demand remained capital high prices sector expansion suggested expenditures level appeared homes economic mortgage low real months	dollar currencies foreign inter-meeting market markets interest major securities yields meeting economic exchange declined ratio balance currency slightly terms trade-weighted inventory-sales rose treasury stocks data policy participants fell yen months
Topic 5 (Growth):	Topic 6 (Financial market):
inflation participants markets financial economic outlook noted risks expectations credit growth conditions pressures prices price labor risk slack current economy policy generally remained fiscal meeting remain possibility anticipated measures expected	foreign bank federal market open reserve york operations securities transactions system currency account banks accounts international agreements arrangements paragraph government currencies imports authorization repurchase manager pursuant goods swap amount undertaken
Topic 7 (Policy):	Topic 8 (Production):
policy economic inflation growth federal reserve meeting monetary conditions financial funds price markets risks market intermeeting economy time outlook directive consistent discussion expansion developments stability current agreed prices sustainable expected	production months manufacturing remained business continued consumer motor output prices market equipment industrial increased spending rose real pace vehicles goods securities gains inventories economic sales level decline appeared sector low

Table 1. Topics and the 30 most important words in each topic

Figure 1 shows the topic loadings of each FOMC meeting Minutes in chronological order, which represent the proportions of Minutes devoted to each topic. As can be seen from Figure 1, over my sample period, the proportion of Minutes devoted to currency was low and stable. The proportion devoted to financial market was highly volatile, and was generally high around the recession in 2008. The proportions devoted to production, trade and growth increased after the recession in 2008, while the proportion devoted to policy decreased over the years. The proportion devoted to consumption was high and grew steadily before 2004, but experienced a sharp decrease since 2005. The proportion devoted to inflation saw a major jump in 2005 but fell a little afterwards.

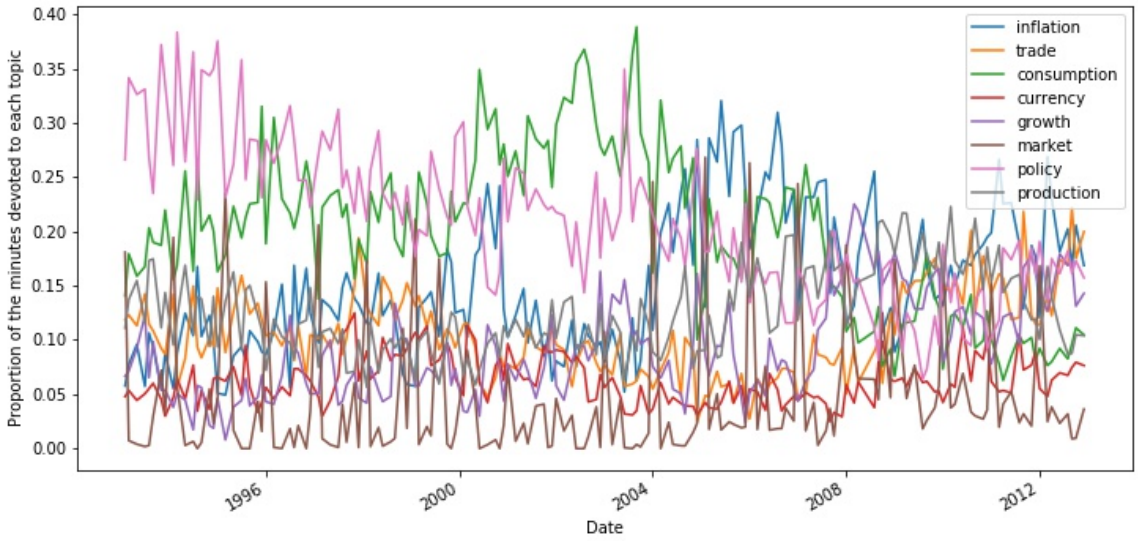

Figure 1. Proportions of Minutes Devoted to Each Topic Over Time

2.3 Sentiment Analysis Result

Figure 2 shows the topic sentiment scores of each FOMC meeting Minutes in chronological order, which represent the tone related to each topic over time. The sentiment scores are not standardized, so the tone was usually negative. Tones of all topics experienced major declines during the 2001 and 2008 recessions, and the magnitude of decline was significantly larger in 2008. This shows the sentiment scores are generally procyclical, becoming more positive during boom periods and turning sharply negative during recessions.

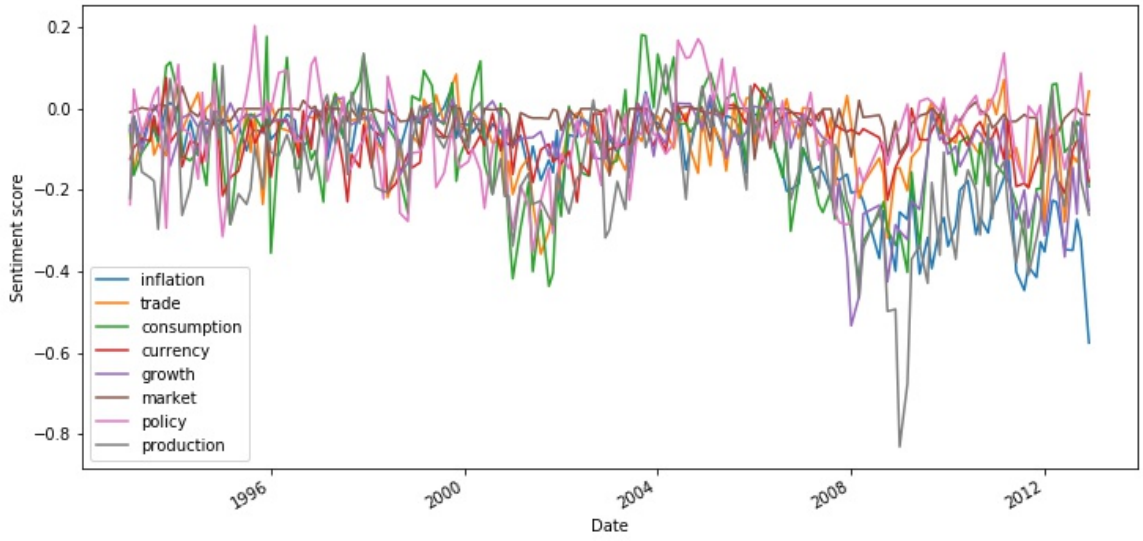

Figure 2. Sentiment Scores of Each Topic Over Time

2.4 New Measure of Monetary Policy

From the residuals of the regression model (5), I can construct the new monthly measure of monetary policy shocks. Specifically, for each month, if there's a meeting then the shock of that month is the residual of the corresponding meeting, and if not, the shock of that month is set to 0. The new measure is shown in Figure 3, together with the measure of Romer and Romer(2004)[8]. The two series of monetary policy shocks are close to each other, which implies that the quantitative and narrative records used by Romer and Romer(2004)[8] are quite accurate proxies for the Fed's information. But compared to my more comprehensive proxy set, their measure generally underestimates the magnitude of the monetary policy shocks, but overestimates several major negative shocks before 2004.

2.5 VAR Results

Using the new monthly measure of monetary policy shocks, I estimate the VAR models (4) and (5), and plot the impulse responses to one percentage point of monetary policy shock for 48 months, as shown in Figure 4 and 5 (the impulse responses to the monetary policy shock measured by Romer and Romer(2004)[8] are also plotted). My new measure reveals a weaker response of the industrial output after 7 months. The impact after 48 months is -0.16 percent, as opposed to -0.23 percent when the the measure of Romer and Romer(2004)[8] is used. The price level rises for the first 9 months after the shock, under both measure.

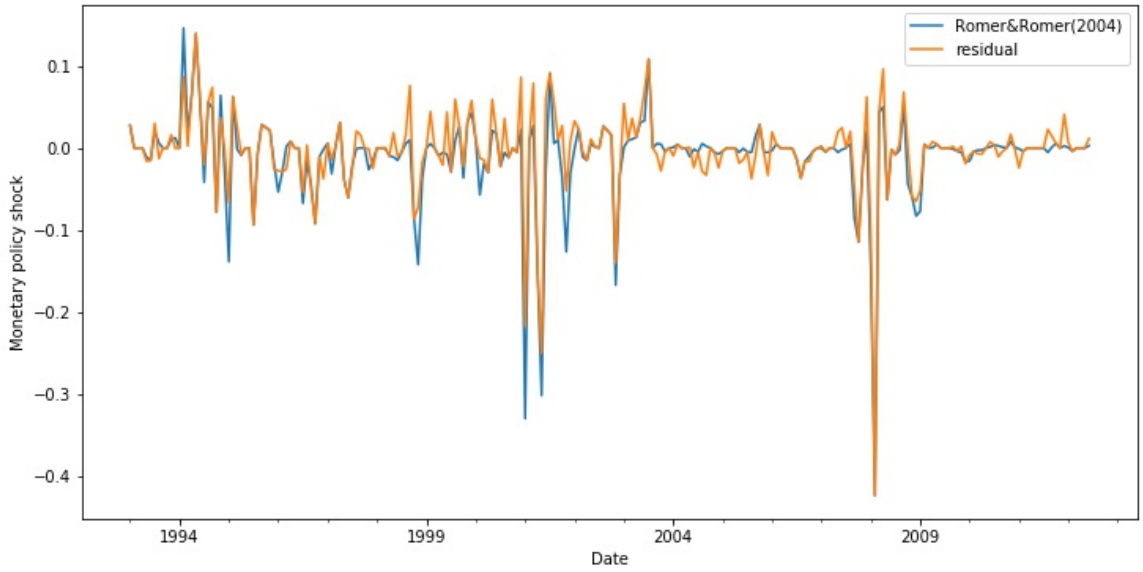

Figure 3. New Measure of Monetary Policy

This phenomenon is referred to as "price puzzle", which is common when traditional measures such as the Federal funds rate are used. The presence of "price puzzle" might suggest that both two measures are still contaminated by endogenous and anticipatory movements. But since the price level rises less when my new measure is used, this problem might be mitigated by including additional textual information from the FOMC Minutes.

References

1. Blei, D. M., Ng, A. Y. & Jordan, M. I. Latent dirichlet allocation. *J. machine Learn. research* **3**, 993–1022 (2003).

2. Řehůřek, R. & Sojka, P. Software Framework for Topic Modelling with Large Corpora. In *Proceedings of the LREC 2010 Workshop on New Challenges for NLP Frameworks*, 45–50 (ELRA, Valletta, Malta, 2010). http://is.muni.cz/publication/884893/en.

3. Hoffman, M., Bach, F. R. & Blei, D. M. Online learning for latent dirichlet allocation. In *advances in neural information processing systems*, 856–864 (2010).

4. Jegadeesh, N. & Wu, D. A. Deciphering fedspeak: The information content of fomc meetings. (2017).

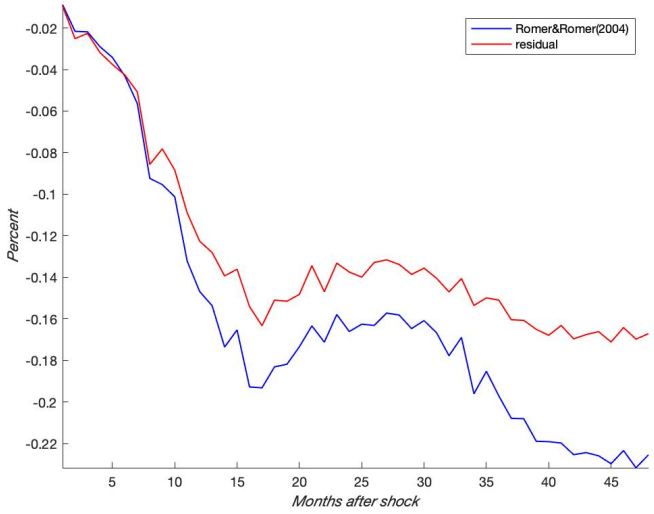

Figure 4. Impulse Response Function of Log Industrial Output

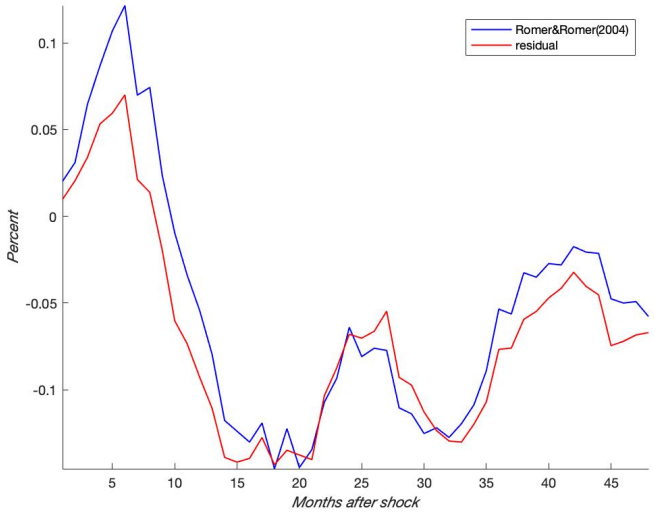

Figure 5. Impulse Response Function of Log PPI

5. Tetlock, P. C. Giving content to investor sentiment: The role of media in the stock market. *The J. finance* **62**, 1139–1168 (2007).

6. Loughran, T. & McDonald, B. When is a liability not a liability? textual analysis, dictionaries, and 10-ks. *The J. Finance* **66**, 35–65 (2011).

7. Jegadeesh, N. & Wu, D. Word power: A new approach for content analysis. *J. Financial Econ.* **110**, 712–729 (2013).

8. Romer, C. D. & Romer, D. H. A new measure of monetary shocks: Derivation and implications. *Am. Econ. Rev.* **94**, 1055–1084 (2004).

9. Updated romer and romer (2004) measure of monetary policy shocks. http://www.basilhalperin.com/blog/2013/12/updated-romer-and-romer-2004-measure-of-monetary-policy-shocks/.