Transparency and Communication within the FOMC: A computational linguistics approach*

Stephen Hansen† Michael McMahon‡ Andrea Prat§

March 14, 2014

Abstract

Deliberation is the primary activity of most expert committees, but there is little empirical work studying it. This paper exploits a natural experiment in the Federal Open Market Committee (FOMC) together with computational linguistic models (particularly Latent Dirichlet Allocation) to measure the effect of transparency on debate. As the career concerns literature predicts, difference-in-difference results show transparency induces the inexperienced to engage less in debate; prepare more; and be more willing to change positions between meetings. The net effect makes them more influential in shaping discussion. We argue central bank designers should thus consider the role of transparency in disciplining policymakers.

Work in Progress

Keywords: Monetary policy, deliberation, FOMC

JEL Codes: E52, E58, D78

*We would like to thank Dave Stockton, Rick Mishkin, Thomas Wood and Janet Yellen for insightful discussions. We are particularly grateful to Omiros Papaspiliopoulos for numerous helpful discussions on MCMC estimation. We have also benefited from comments and suggestions by seminar attendees at the San Francisco Federal Reserve, Warwick University, University of Manchester, INSEAD, Bank of England, the New York Federal Reserve and Columbia University. We thank Eric Hardy for excellent research assistance. Any errors remain ours alone.

†Universitat Pompeu Fabra and GSE. Email: stephen.hansen@upf.edu

‡University of Warwick, CEPR, CAGE (Warwick), CEP (LSE), CfM (LSE), and CAMA (ANU). Email: m.mcmahon@warwick.ac.uk

§Columbia University, and CEPR. Email: andrea.prat@columbia.edu

1 Introduction

In this paper we study how transparency, a feature of central bank institutional design, affects deliberation of monetary policy makers on the Federal Open Market Committee (FOMC). In other words, we ask what are the effects on "internal deliberation" of greater "external communication"? We focus mainly on deliberation because the majority of committee time is generally devoted to discussion and deliberation prior to making their decisions.[1] An ex-Governor of the Fed, Larry Meyer, saw the discussion as crucial for understanding the decisions.[2] Therefore understanding how policy committees, the dominant framework for setting monetary policy throughout the world (Fry, Julius, Mahadeva, Roger, and Sterne 2000), deliberate and share information, and understanding how different institutional design issues affect that deliberation, is important to understand monetary policy decisions. These issues have likely become even more important with the growing establishment of financial policy committees and the potential need to share information across central bank committees with different objectives.

However, information to analyse deliberation is often not available. As table 1 shows, there is heterogeneity across three major central banks in terms of the records of their policy meetings which are put on the public record.[3] Current ECB president, Mario Draghi, has said that "It would be wise to have a richer communication about the rationale behind the decisions that the governing council takes." (Financial Times 2013). But should the ECB move to only release minutes shortly after the meeting, or is there an additional benefit of also committing to release full transcripts of the discussion with a five-year lag as is the case with the FOMC meetings?

While greater transparency is generally accepted for accountability reasons, many people think that too much transparency about the deliberation will stifle the committee's discussions. In fact, before they had released transcripts, Alan Greenspan expressed his views to the Senate Banking Committee (our emphasis):

[1]Of course, policy makers' decisions remain an output of interest, and a growing complementary literature takes observed policy choices in both experimental (e.g. Blinder and Morgan 2005, Lombardelli, Proudman, and Talbot 2005) and actual committees (e.g. Hansen, McMahon, and Velasco 2012) and uses them to address central bank design questions.

[2]Meyer (2004) writes

> So was the FOMC meeting merely a ritual dance? No. I came to see policy decisions as often evolving over at least a couple of meetings. The seeds were sown at one meeting and harvested at the next. So I always listened to the discussion intently, because it could change my mind, even if it could not change my vote at that meeting. Similarly, while in my remarks to my colleagues it sounded as if I were addressing today's concerns and today's policy decisions, in reality I was often positioning myself, and my peers, for the next meeting.

[3]Minutes of the ECB's governing council meetings are not published though the monetary policy decision is explained at a press conference led by the ECB President after the meeting. The minutes are supposed to be released eventually after a 30 year lag.

Table 1: Information made available by different central banks

	Federal Reserve	Bank of England	European Central Bank
Release Minutes?	✓	✓	X
Release Transcripts?	✓	X	X

> "**A considerable amount of free discussion and probing questioning by the participants of each other and of key FOMC staff members takes place.** In the wide-ranging debate, new ideas are often tested, many of which are rejected ... **The prevailing views of many participants change as evidence and insights emerge.** This process has proven to be a very effective procedure for gaining a consensus ... It could not function effectively if participants had to be concerned that their half-thought-through, but nonetheless potentially valuable, notions would soon be made public. **I fear in such a situation the public record would be a sterile set of bland pronouncements scarcely capturing the necessary debates which are required of monetary policymaking.** A tendency would arise for one-on-one pre-meeting discussions, with public meetings merely announcing already agreed-upon positions or for each participant to enter the meeting with a final position not subject to the views of others." Greenspan (1993) as reported in Meade and Stasavage (2008).

Prat (2005) shows that transparency may induce conformism among experts in a vein similar to that predicted by Greenspan, economic theory also has other predictions such as disciplining experts to prepare more for their meetings, and even incentivising contrarian behaviour. While we will examine the predictions of economic theory more fully in the next section, it is an empirical question as to which of the different effects dominate on the FOMC.

We are able to explore these issues by exploiting the natural experiment that led to the release of the FOMC transcripts. FOMC members were unaware that transcripts of their meeting discussions were, since the 1970s, being filed in the archives before staff recorded over the tapes that were used to help write the minutes. They only learned of the existence of the records when Greenspan, under pressure from the US Senate Committee on Banking, Housing, and Urban Affairs (Senate Banking Committee hereafter), discovered and revealed their existence to the politicians and the rest of the FOMC. The Fed had to quickly agree to publish them as to not publish them would look like they had been trying to hide something they claimed they didn't have. This means that we have a complete record of meetings' policy discussion when members did know that their verbatim discussions were being kept on file let alone that such information would be made public (prior to November 1993), and the same verbatim record of their decision

making after that date when the members all knew with certainty that their discussions would eventually be made public.

Meade and Stasavage (2008) have previously used this natural experiment to analyze the effect of transparency on members' incentives to dissent in voice. We shall extend their analysis in two important dimensions. First, they compare only before and after the change in transparency (a pure "diff" approach); the problem is that other things that changed around the start of 1994 may have had an affect on committee voting. Instead, we will, for most of the analysis, employ a difference-in-difference ("diff-in-diff") approach to try to isolate the effects that transparency has through those with greater reputational concerns. While we confirm their result that dissent is lower after the change in transparency, there is no evidence to suggest it is driven by reputational concerns. Instead, lower dissent could result from a better (more convincing) discussion such that all members come to the same conclusion after hearing the debate.

Second, they did not analysis communication *per se* but rather used a narrative approach measure of whether an FOMC member dissents from the Chair's recommended policy decision. The use of dissent in voice is important because FOMC members often expressed reservations with the policy proposal in the discussion, but then vote in line with the Chair for the public record. However we are able to go a great deal further by using recently developed machine learning algorithms for natural language processing to analyse different dimensions of committee deliberation such as what topics members talk about, how much they talk about them, and whether they are able to influence what each other discuss.

While we also use some more primitive text analysis tools such as counting numbers of questions and lengths of each statement, our key methodological innovation is the use of Latent Dirichlet Allocation (LDA) to study economic questions of interest. LDA is widely used in linguistics, computer science, and other fields; the article that introduced it, Blei, Ng, and Jordan (2003), has nearly 7,000 citations in 10 years. While computational linguistic models are beginning to appear in the political science literature, their use is still mainly descriptive; for example, Quinn, Monroe, Colaresi, Crespin, and Radev (2010) use a topic model similar to LDA to study congressional speeches to see what congress is talking about. We hope that our approach of using computational linguistics to create measures from large databases of text data to use in more conventional econometric analysis can help to bring monetary economics into the increasingly important world of "Big Data" and open the way for future inter-disciplinary work.

We are, of course, not the first to use the transcripts. Romer and Romer (2004) use the transcripts to derive a narrative-based measure of monetary policy shocks. Chappell, McGregor, and Vermilyea (2005) and Meade (2005) use the information in the transcripts to try to measure preferences. Our work also complements the analysis in Bailey and

Schonhardt-Bailey (2008) and Bailey and Schonhardt-Bailey (2013) who also use text analysis to examine the FOMC transcripts. Their differs from ours in that they focus on arguments and persuasive strategies adopted by policymakers (using a computer package called "Alceste"). Moreover, while they examine three periods of interest to their questions, each of three years in length (1979-1981, 1991-1993, and 1997-1999); we will examine the behaviour using a diff-in-diff approach around our natural experiment and estimate our topic models using all data in the Greenspan era.[4]

Overall, we find results that are in line with the predictions of the career concerns literature. Our difference-in-difference results show transparency induces the inexperienced to engage less in debate; prepare more; and be more willing to change positions between meetings. The net effect makes them more influential in shaping discussion. We therefore believe that central bank designers should consider the role of transparency in disciplining policymakers.

2 Transparency and Career Concerns

Since agreeing to release transcripts in 1993, the Fed has done so with a five-year lag. The main channel through which one expects transparency to operate at this time horizon is career concerns rather than, for example, communication with financial markets to shift expectations about future policy. By career concerns, we mean that policymakers derive utility from signalling their expertise in monetary policy to outsiders, who could include the broader Fed community, financial market participants, politicians, etc. A well-developed literature contains several theoretical predictions on the effects of career concerns, so instead of constructing a formal model, we summarize how we expect career concerns to operate on the FOMC and how transparency should modify them.

Discipline The canonical reference in the literature is Holmström (1999), who shows that career concerns motivate agents to undertake costly, non-contractible actions ("effort") to improve their productivity. We consider the key dimension of effort exertion on the FOMC to be the acquisition of information about economic conditions. Members choose how much time to spend analyzing the economy in the weeks between each meeting. Clearly performing independent analysis rather than relying on staff briefings incurs a higher opportunity cost of time, but also leads a member to having more information on the economy.

As for transparency, Holmström (1999) predicts that effort exertion increases as the noise in observed output decreases. If one interprets transparency as increasing the precision of observers' information regarding member productivity, one would expect transparency

[4]For other projects using these methods, we can analyse discussion from the mid-1970s until 2008.

to increase incentives to acquire information prior to meetings. Gersbach and Hahn (2012) provide a formalization of this idea applied to monetary policy committees.

Conformity/Non-conformity Scharfstein and Stein (1990) show that agents with career concerns unsure of their expertise tend to herd on the same action, thereby avoiding being the only one to take an incorrect decision. Interpreted broadly, such conformity would appear on the FOMC as any behavior consistent with members seeking to fit in with the group rather than standing out. On the other hand, models in which agents know their expertise such as Prendergast and Stole (1996) and Levy (2004) predict the opposite. There is a reputational value for an agent who knows he has an inaccurate signal to take unexpected actions in order to appear smart.

In general, the effect of transparency is to amplify whatever the effect of career concerns is. When agents do not know their expertise, transparency increases incentives to conform, as shown by Prat (2005) for a single agent and Visser and Swank (2007) for committees. On the other hand, Levy (2007) has shown that transparency leads committee members who know their expertise to take contrarian actions more often. We will therefore leave as an open question whether transparency leads to more conformity or less non-conformity on the FOMC, and let data resolve the issue.

Persistence Prendergast and Stole (1996) provide one of the only treatments in the literature on the behavior of an expert with career concerns for more than one period. They find a tendency for such experts to persist with their previous actions, even when they receive new information indicating this action is wrong. Changing one's mind is an admission one initially received an inaccurate signal, which lowers reputation.

While there is no theory of how transparency modifies the dynamic behavior of experts, we hypothesize that, as with the previous two effects, it should tend to increase the relevance of the basic mechanism. We therefore expect the effect of transparency to be an increase in the extent to which experienced FOMC members persist in their actions.

Overall, the effect of increased transparency can be positive (through increased discipline) or negative (through increased conformity/non-conformity and persistence). More relevant for this paper, which uses a Diff-in-Diff approach, is whether the behavior of FOMC members with greater career concerns—whom we identify as being inexperienced in a way made precise below—becomes relatively more efficient than that of members with less career concerns after transparency. Again, the direction is ambiguous. Discipline and persistence both favor the inexperienced becoming relatively more efficient, with conformity/non-conformity working the other way. After presenting our empirical results, we will attempt to measure the overall effect of transparency on the inexperienced relative to the experienced through a particular kind of influence score.

3 The FOMC and their meetings

3.1 FOMC Membership

The FOMC, which meets 8 times per year to formulate monetary policy (by law it must meet at least 4 times) and to determine other Federal Reserve policies, is composed of 19 members; there are seven Governors of the Federal Reserve Board (in Washington DC) of whom 1 is the Chairperson (of both the Board of Governors and the FOMC) and there are twelve Presidents of Regional Federal Reserve Banks with the President of the New York Fed as Vice-Chairman of the FOMC.[5]

The US president nominates members of the Board of Governors which are then subject to approval by the US Senate; once appointed, they are given terms of 14 years (with an expiry at the end of January every even-numbered year). Regional Fed presidents are appointed by their own bank's board of nine directors (which is appointed by the Banks in the region (6 of the members) and the Board of Governors (3 of the members))and are approved by the Board of Governors; these members serve 5 year terms.

The main policy variable of the FOMC is a target for the Federal Funds rate (Fed Funds rate), as well as, potentially, a bias (or tilt) in future policy.[6] At any given time, only twelve of the FOMC have policy voting rights though all attend the meetings and take part in the discussion. All seven Governors have a vote (though if there is a Governor vacancy then there is no alternate voting in place); the president of the New York Fed is a permanent voting member (and if absent, the first vice president of the New York Fed votes in his/her place); and four of the remaining eleven Fed Presidents vote for one year on a rotating basis.[7]

3.2 The Structure of FOMC Meetings

Most FOMC meetings last a single day except for the meetings that precede the Monetary Policy Report for the President which last two days.[8] Before FOMC meetings, the

[5]Federal Reserve staff also attend the meeting and provide briefing in it.

[6]Over time, this has changed quite a bit. Now, the FOMC states whether the risks are greater to price stability or sustainable growth, or balanced. Between 1983 and December 1999, the FOMC included in its monetary policy directive to the Open Market Trading Desk of the New York Fed a signal of the likely direction of future policy. In 2000, these signals were just made more explicit. Moreover, there was never a clear understanding of why the bias was even included; Meade (2005) points to transcript discussions in which FOMC members debate the point of the bias though Thornton and Wheelock (2000) conclude that it is used most frequently to help build consensus.

[7]Chicago and Cleveland Fed presidents vote one-year on and one-year off, while the remaining 9 presidents vote for 1 of every 3 years.

[8]The FOMC also "meets" via conference calls in between the official face-to-face meetings. These meetings can be about monetary policy (and can even change the Fed Funds target after these calls - see October 1998 and September 2001 as examples). We do not currently analyse the discussion in the conference calls.

members receive briefing in advance (such as the 'Green Book" (staff forecasts), "Blue Book" (staff analysis of monetary policy alternatives) and the "Beige Book" (Regional Fed analysis of economic conditions in each district).

During the meeting there are a number of stages (including 2 discussion stages):[9]

1. A NY Fed official presents financial and foreign exchange market developments
2. staff present the staff economic and financial forecast
3. **Economic Situation Discussion (FOMC 1):**
   - Board of Governors' staff present the economic situation (including forecast).
   - There are a series of questions on the staff presentations.
   - FOMC members (including those not currently voting) present their views of the economic outlook.
   - Chairman tended to speak reasonably little during this round and didn't take a dedicated slot in the "go around".
4. In two-day meetings when the FOMC had to formulate long-term targets for money growth, a discussion of these monetary targets took place in between the economic and policy discussion rounds.
5. **Policy Discussion (FOMC 2):**
   - The Board's director of monetary affairs then presents a variety of monetary policy alternatives (without a recommendation).
   - Another potential round of questions.
   - The Chairman (1st) and the other FOMC discuss their policy preferences.
6. The FOMC votes on the policy decision - FOMC votes are generally unanimous (or close to) but there is more dissent in the discussion.

Our analysis focuses mainly on the part of the meeting relating directly to the economic situation discussion which we call FOMC 1, and the part related to the discussion of the monetary policy decision which we call FOMC 2. However, we estimate our models using the whole sample and we even carry out, for robustness, some analysis using all meeting parts.

[9]See `http://www.newyorkfed.org/aboutthefed/fedpoint/fed48.html` and Chappell, McGregor, and Vermilyea (2005) for more details.

3.3 FOMC Discussions outside the meeting?

The Government in Sunshine Act, 1976, aims to ensure that Federal bodies make their decisions in view of the public and require the Federal bodies to follow a number of strict rules about disclosure of information, announcement of meetings, etc... While the FOMC is not obliged to operate under the rules of the Sunshine Act, they maintain a position that is as close to consistent with it though with closed meetings[10] This position suggests that the Committee takes very seriously the discussion of its business in formal meetings and accords with what we have been told by staff and former members of the FOMC, as well as behaviour discussed in the transcripts discussing how to notify the public that they chose to start a meeting a day early. As such, we can take as given that the whole FOMC does not meet outside the meeting to discuss the decision.

4 Empirical strategy

The main element of our empirical strategy is the use of the natural experiment used by Meade and Stasavage (2008). The natural experiment is the result two main things; first, diligent staff archiving, and second, external political pressure on the FOMC to be more transparent.

Specifically on the first, the Fed staff had recorded meetings to assist with the preparation of the minutes. As far as FOMC members were concerned, the staff would record over the tapes for subsequent meeting recordings once the minutes were released. While the staff did record over the older tapes, unknown to FOMC members, the staff first typed up and archived a verbatim text of the discussion.

The existence of these archives was not highlighted to FOMC members, including the Chair, until political pressure from Henry B. Gonzalez forced the Fed to discuss how it might be more transparent. In October 1993, Chairman Greenspan made the existence of the transcripts known to the Senate Banking Committee hereafter. While Greenspan argued that he didn't want to release any verbatim information as it would stifle the discussion, the pressure on the Fed was growing and so the Fed quickly moved to release not just the future transcripts with a five-year lag, but also to release the previous years. This means that we have transcripts from prior to November 1993 in which the discussion took place under the assumption that individual statements would not be on the public record, and transcripts after November 1993 in which each policy maker knew that every spoken word would be public within five years.

Since the decision to change transparency was not driven by FOMC concerns about

[10]See `http://www.federalreserve.gov/monetarypolicy/files/FOMC_SunshineActPolicy.pdf` and `http://www.federalreserve.gov/aboutthefed/boardmeetings/sunshine.htm` for the Fed's official position.

the nature or style of the FOMC deliberation, and the change came as a surprise (as we will show later), the most straightforward empirical strategy to identify the effects of transparency on deliberation is to estimate a baseline "Diff" regression of the following form: DIFF

$$y_t = \alpha + \beta D(Trans) + \lambda X_t + \varepsilon_t \tag{1}$$

where y_t is the output variable of interest, $D(Trans)$ is a transparency dummy (1 after November 1993), and X_t is a vector of macro controls for the meeting at time t.

However, one of the potential risks of a simple "Diff" analysis is that the timing of other changes may have coincided with the change in transparency. As such, the β estimated in (1) may capture the effects of the other changes and it is impossible to disentangle the different effects. In order to more clearly attribute the changes with transparency, we propose a "Diff-in-Diff" analysis in which we argue that the effects of transparency should be greatest on those people who have the greatest career concerns.

To measure the extent of career concerns we take an idea from proposition 1 of Holmström (1999) which argues that the strength of an experts career concerns increases in the uncertainty of the principal's belief about that members type. Given that we believe the natural channel is that FOMC members want to signal expertise about monetary policy, we measure the strength of career concerns as declining with prior experience within the Federal reserve. In particular we define a variable, $FedExp_{i,t}$, that captures the number of years that a person has spent working in the Fed system (prior to FOMC), as well as their years on the FOMC. We then estimate our "diff-in-diff" regression:

$$y_{it} = \alpha_i + \delta_t + \beta D(Trans)_t + \eta FedExp_{i,t} + \phi \left(D(Trans)_t \times FedExp_{i,t} \right) + \epsilon_{it} \tag{2}$$

ADD COEFFICIENT OF INTEREST AND WHY + INVERT THE DIRECTION...

Given the inclusion of time and member fixed effects, the identification comes mostly off those members who served before and after the change in transparency. For the analysis presented below, we will focus on a sample which uses eight years around our natural experiment, and we will separately analyse the changes in behaviour in the two main monetary policy parts of the meetings (FOMC1 and FOMC2 in our language).

5 Data

This section describes the raw data we use in our empirical exercise.

5.1 FOMC Transcripts

Nearly verbatim transcripts of FOMC meetings are available for download from `http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm`. Apart from minor redactions relating, for example, to maintaining confidentiality of certain participants in open market operations, they provide a complete account of every FOMC meeting from the mid-1970's onwards. For this paper, we use the set of transcripts from the tenure of Alan Greenspan—August 1987 through January 2006, inclusive, a total of 149 meetings. During this period, the FOMC also engaged in numerous conference calls for which there are also verbatim accounts, but as many of these were not about directly about monetary policy we do not use them in our analysis.

The transcripts available from the Fed website are not machine-readable, and also contain non-spoken text such as footnotes, page headers, and participant lists. After converting the raw files into machine-readable format (using OCR where appropriate), we used regular expressions to identify and remove these extraneous parts, but no doubt some minor errors remain. There are also several apparent transcription errors relating to speaker names, which always have an obvious correction. For example, in the July 1993 meeting a "Mr. Kohn" interjects dozens of times, and a "Mr. Koh" interjects once; we attribute the latter statement to Mr. Kohn. Finally, from July 1997 backwards, staff presentation materials were not integrated into the main transcript. We used a social enterprise data service to extract staff statements from appendices, which we then matched into the main transcripts. The final dataset contains 46,502 unique interjections along with the associated speaker.

After assembling the final database, we coded the statements in each meeting that corresponded to the economic situation discussion (FOMC1) and the policy discussion (FOMC2), as described in section 3. These make up around 31% and 26% of the total number of statements.

5.2 FOMC member characteristics

In the analysis we use every member of the FOMC. We also, in our analysis, have all of the interjections of the staff. However, while we treat each individual member of the FOMC as a separate person in the estimation for the entire duration of their time on the FOMC, we treat all members of staff as a single homogenous group. Moreover, in our main regressions we only consider staff reactions to changes in transparency.

In terms of each FOMC member, we gathered biographical information using "Who's who" and online sources. From this information, we are able to measure how much experience the member had in the Fed before their appointment, as well as how much experience they have built up serving on the committee. Adding these two together gives

our main "Fed Experience" variable.

6 Measuring Communication on the FOMC

A major challenge for the analysis is to convert the raw text in the transcript files into meaningful quantities for the dependent variables in the regressions described in section 4. The first step in the text processing is to *tokenize* each statement, or break it into its constituent linguistics elements: words, numbers and punctuation.[11] One can easily then count the number of occurrences of a given token in each statement. Using such an approach, we construct three measures of language per statement

1. Number of questions (count of token '?')
2. Number of sentences (count of tokens '?', '!', and '.')
3. Number of words (count of alpha-numeric tokens; 5,594,280 in total).

From these counts, one can then measure the total number of questions/sentences/words at various aggregate levels of interest. In addition, we also use the total number of statements as a fourth count-based measure of communication within meetings.

While the simplicity of count-based analysis is appealing, a basic problem for determining what FOMC members talk about is what tokens one should count. For example, one might count the number of times 'growth' appears in statements to create an index of focus on economic activity. But clearly other words are also used to discuss activity, and knowing which list to choose is not obvious and would involve numerous subjective judgements. Moreover, the word 'growth' is also used in other contexts, such as in describing wage growth as a factor in inflationary pressures. The topic modelling approach addresses these issues by adopting a flexible statistical structure that groups words together to form "topics," and by allowing the same word to appear in multiple topics.

The rest of the section describes our implementation of the LDA model. It first lays out the underlying statistical model, and then describes how we estimate it. Finally, it discusses how to transform the output of the estimation into measures of communication.

6.1 Statistical model

Consider a collection of D documents, where a document d is a list of words $\mathbf{w}_d = (w_{d,1}, \ldots, w_{d,N_d})$. In our dataset, a document is a particular statement. Here "word" does not necessarily mean an English word, and should be understood as simply an abstract element.

[11] For tokenization and some other language processing tasks outlined below, we used the Natural Language Toolkit developed for Python and described in Bird, Klein, and Loper (2009).

Let V be the number of unique words across all documents. These words form K topics, where a *topic* $\beta_k \in \Delta^V$ is a distribution over these V words. The vth element of topic k β_k^v represents the probability of a given word appearing in topic k. Suppose each β_k is drawn from a symmetric Dirichlet(η) prior. Each document is made up of potentially different topics. Let $\theta_d \in \Delta^K$ be the distribution of topics in document d, where θ_d^k represents the probability of topic k appearing in document d. Suppose each θ_d is drawn from a symmetric Dirichlet(α) prior.

The words in document d are assumed to be generated in a two-step process. First, for each word $w_{d,n}$ choose a topic $z_{d,n}$ from θ_d. Second, $w_{d,n}$ is drawn from $\beta_{z_{d,n}}$. The probability of document d conditional on the hyperparameters α and η is therefore

$$\Pr[\,\theta_d \mid \alpha\,] \prod_{k=1}^{K} \Pr[\,\beta_k \mid \eta\,] \prod_{n=1}^{N_d} \sum_{z_{d,n}} \Pr[\,z_{d,n} \mid \theta_d\,] \Pr\left[\, w_{d,n} \;\middle|\; \beta_{z_{d,n}} \,\right] \tag{3}$$

where the summation is over all possible topic assignments for word $w_{d,n}$. The total probability of observing all D documents is the product of the individual probabilities.

Two assumptions of LDA are worth noting. First, LDA is a *bag-of-words* model in which the order of words does not matter, just their frequencies. While this assumption clearly throws away information, it is a useful simplification when the primary consideration is to measure *what* topics a document covers. Word order becomes more important when the goal is sentiment analysis, or *how* a document treats a topic. Second, documents are assumed to be iid. LDA can be extended to model various dependencies across documents.[12] Dynamic topic models allow β_k to evolve over time. These are particularly important when documents span many decades. For example, Blei and Lafferty (2006) study the evolution of scientific topics during the 20th century. In contrast our sample covers roughly 20 years, and we use a much smaller window to study the effect of transparency. Author-topic models (Rosen-Zvi, Chemudugunta, Griffiths, Smyth, and Steyvers 2010) model a document as being generated by it author(s), essentially substituting authors for documents in the generative statistical model. Since we expect the same speaker to use different topic distributions across and within meetings, we prefer to conduct the analysis at the document level.

6.2 Estimation

The parameters of interest of the model are the topics β_k and document-topic distributions θ_d. Direct maximization is not feasible due to the intractability of the computation of (3), so adopting alternative estimation techniques is necessary. We use the Gibbs

[12]For a discussion of extensions to LDA, see Blei and Lafferty (2009) or the lectures given by David Blei at the Machine Learning Summer School in 2009 (Blei 2009).

sampling approach introduced into the literature by Griffiths and Steyvers (2004) (see also Steyvers and Griffiths 2006). Their approach directly estimates the posterior distribution over topic assignments ($z_{d,n}$) given the observed words. Full details of the approach are in appendix A. The algorithm begins by randomly assigning topics to words, and then updating topic assignments by repeatedly sampling from the appropriate posterior distribution. For estimation we adapt the C++ code of Phan and Nguyen (2007) available from `gibbslda.sourceforge.net`.

As with all Markov Chain Monte Carlo methods, the realized value of any one chain depends on the random starting values. For each specification of the model, we therefore run 3,000 iterations beginning from 10 different starting values and choose for analysis the chain that achieves the best fit of the data based on its final value as measured by *perplexity*, a common measure in the natural language processing literature.[13] For robustness, we also analyze the chain that achieves the second best fit. In practice the differences in perplexity at the 3,000th iteration across chains are marginal, indicating that the estimates are not especially sensitive to starting values.

6.2.1 Vocabulary selection

Before sampling the model, one must choose the list of words to include. For each document, we:

1. Remove all tokens not containing solely alphabetic characters. This strips out punctuation and numbers.

2. Remove all tokens of length 1. This strips out algebraic symbols, copyright signs, currency symbols, etc.

3. Convert all tokens to lowercase.

4. Remove *stopwords*, or extremely common words that appear in many documents. Our list is rather conservative, and contains all English pronouns, auxiliary verbs, and articles.[14]

[13] The formula is

$$\exp\left[-\frac{\sum_{d=1}^{D}\sum_{v=1}^{V} N_{d,v}\log\left(\sum_{k=1}^{K}\widehat{\theta}_d^k\widehat{\beta}_k^v\right)}{\sum_{d=1}^{D} N_d}\right]$$

where $N_{d,v}$ is the number of times word v occurs in document d.

[14] The list is available from `http://snowball.tartarus.org/algorithms/english/stop.txt`. The removal of length-1 tokens already eliminates the pronoun 'I' and article 'a'. The only other length-1 English word is the exclamation 'O'. We converted US to 'United States' so that it was not stripped out with the pronoun 'us'. We also converted contractions into their constituent words (e.g. 'isn't' to 'is not') so that words such as 'not' were not stripped out.

5. Stem the remaining tokens to bring them into a common linguistic root. We use the default Porter Stemmer implemented in Python's Natural Language Toolkit. For example, 'preferences', 'preference', and 'prefers' all become 'prefer'. The output of the Porter Stemmer need not be an English word. For example, the stem of 'inflation' is 'inflat'.

We then tabulate the frequencies of all two- and three-word sequences in the data, known as *bigrams* and *trigrams*, respectively. For the most frequently occurring, we construct a single token and replace it for the sequence. For example, 'fed fund rate' becomes 'ffr' and 'labor market' becomes 'lab_mkt'.

Finally, we rank the 13,888 remaining tokens in terms of their contribution to discriminating among documents. The ranking punishes words both when they are infrequent and when they appear in many documents.[15] Plotting the ranking indicates a natural cutoff we use to select $V = 8,615$ words for the topic model. Words below this cutoff are removed from the dataset.

After this processing, 2,715,586 total tokens remain. Some statements are empty, so we remove them from the dataset, leaving $D = 46,169$ total documents for input into the Gibbs sampler.

6.2.2 Model selection

There are three parameters of the model that we fix in estimation: the hyperparameters for the Dirichlet priors α and η and the number of topics K. For values of the hyperparameters, we follow the general advice of Griffiths and Steyvers (2004) and Steyvers and Griffiths (2006) and set $\alpha = 50/K$ and $\eta = 0.025$. The low value of η promotes sparse word distributions so that topics tend to feature just a handful of prominent words.

The most common approach to choosing the number of topics is to estimate the model for different values of K on randomly selected subsets of the data (*training documents*), and then to determine the extent to which the estimated topics explain the omitted data (*test documents*). This approach shows that a model with several hundred topics best explains our data. Since our goal is to organize text into easily interpretable categories rather than to predict the content of FOMC meetings *per se*, we consider this number too high.[16] We instead estimate models with $K = 30$, $K = 50$, and $K = 70$, which both

[15] More specifically, we computed each word's term-frequency, inverse-document-frequency—or *tf-idf*—score. The formula for token v is

$$\textit{tf-idf}_v = [1 + \log(N_v)] \times \log\left(\frac{D}{D_v}\right)$$

where N_v is the number of times v appears in the dataset and D_v is the number of documents in which v appears.

[16] According to Blei (2012), interpretability is a legitimate reason for choosing a K different from the

allows us to relatively easily interpret the topics and to explore the effect of altering the number of topics on the results.[17] In the main body of the text, we report results for $K = 50$.

6.3 Communication measures based on LDA estimates

Primary objects of interest for empirical analysis are the proportion of time different FOMC members spend on different topics and, to a lesser extent, the proportion of time the committee as a group devotes to different topics. The Gibbs sampler delivers the estimate $\widehat{\theta}_{d,j}$ at the jth iteration for the topic proportions in document d along with estimated topics $\widehat{\beta}_{k,j}$ for $k = 1, \ldots, K$. We are interested in estimates of $\widehat{\theta}_{i,t.s,j}$ where i indexes an FOMC member, t indexes a meeting, and s indexed a meeting section (as well as estimates of $\widehat{\theta}_{t.s,j}$). The sections of interest are the entire meeting; FOMC1; FOMC2; or FOMC1 and FOMC2 combined. We detail how we obtain estimates for aggregate documents in appendix A. The basic idea is to hold fixed the estimated topics and re-sample—or *query*—the aggregate documents.

We obtain the final estimate $\widehat{\theta}_{i,t.s}$ by averaging $\widehat{\theta}_{i,t.s,j}$ over $j \in \{2100, 2200, \ldots, 3000\}$. Hence, in the language of MCMC estimation, we run 1,000 iterations using a burn-in of 2,000 iterations and applying a thinning interval of 100. Based on perplexity scores, all the chains we estimate converge at or before the 2,000th iteration.

There are several further measures of communication one can build from $\widehat{\theta}_{i,t.s,j}$, as we describe in the rest of this section. In each case, the overall estimate is the average over $j \in \{2100, 2200, \ldots, 3000\}$.

6.3.1 Content

As an unsupervised learning algorithm, LDA produces topics with no meaningful labels. Any attribution of meaning to topics involves a subjective judgement on the part of the researcher. The only such judgement we make in the analysis is to classify topics into "economics" topics Θ_E and "discussion" topics $\{k\}_{k=1}^{K} \setminus \Theta_E$. The following section provides examples of these. Economics topics are those that reflect substantive themes in economics like foreign trade or monetary policy, while discussion topics are formed of words used in general conversation not tied to any obvious economic topic.

one that performs best in out-of-sample prediction. He notes a "disconnect between how topic models are evaluated and why we expect topic models to be useful." In our setting, as the number of topics increases, the identified topics become increasingly specific. As we show in the next section, a 50 topic model produces a single topic relating to risk. By contrast, a 200 topic model produces topics on upside risks, downside risks, risks to growth, financial market risk, etc. Also, as our database is to some extent conversational, a model with a large K picks out very specific conversational patterns as topics, such as addressing Chairman Greenspan prior to discussing one's views on the economy.

[17] The highest marginal effects of changing the number of topics on out-of-sample prediction values arise for $K < 100$.

Two measure arise from this classification. First $E_{i,t.s,j} = \sum_{k \in \Theta_E} \widehat{\theta}^k_{i,t.s,j}$ is the fraction of time spent on economics topics. Second, $\widehat{\chi}_{i,t.s,j}$ is the conditional distribution over economics topics.[18]

6.3.2 Focus

Another quantity of interest is the degree to which members focus their attention on particular topics. To measure focus we compute the Herfindahl index $HC_{i,t.s,j} = \sum_{k \in \Theta_E} \left(\widehat{\chi}^k_{i,t.s,j}\right)^2$ so that higher values are associated with more focus. $HU_{i,t.s,j}$ is the same index computed on the unconditional topic distributions.

6.3.3 Distance from Chairman

Alan Greenspan is clearly a focal member of the FOMC during our sample, so measuring the extent to which other members speak about topics he does is of interest. To measure this we compute the Kullback-Leibler (KL) divergence, which is standard in the information theory literature:

$$DC_{i,t.s,j} = \sum_{k \in \Theta_E} \log \left(\frac{\widehat{\chi}^k_{i,t,s,j}}{\widehat{\chi}^k_{G,t,s,j}} \right) \widehat{\chi}^k_{i,t,s,j}, \tag{4}$$

where G is the index of Greenspan. We also compute the measure $DU_{i,t.s,j}$ on the unconditional topic distributions.

$DC_{i,t.s,j}$ returns a number between 0 and ∞, and measures the information lost when Greenspan's conditional topic distribution is used to approximate member i's distribution. A higher number represents a worse approxmation.

6.3.4 Influence

Given the importance of Chairman Greenspan to the ultimate policy decision, we shall measure influence of member i as any time that i talks about affects that Greenspan subsequently talks about. We use a standard measure as in Palacios-Huerta and Volij (2004), which is a modified Page-rank algorithm, and exploit the timing of the discussion to measure influence. We measure two forms of influence on Greenspan. Let x^{ECON}_{it} $\left(x^{\text{POL}}_{it}\right)$ be the estimated distribution of over topics for member i in the economic discussion (policy debate) in meeting t. If the pattern of focus of i in Econ is positively correlated with the pattern of focus of Greenspan in Policy, this is influence of i on Greenspan within the

[18] Formally,

$$\widehat{\chi}^k_{i,t.s,j} = \begin{cases} \widehat{\theta}^k_{i,t.s,j} / E_{i,t.s,j} & \text{if } k \in \Theta_E \\ 0 & \text{otherwise} \end{cases}$$

meeting (v_{it}^G). While if what i talks about in meeting t leads to Greenspan putting more weight on that topic in meeting $t+1$, this is inter-meeting influence of i (d_{it}^G).

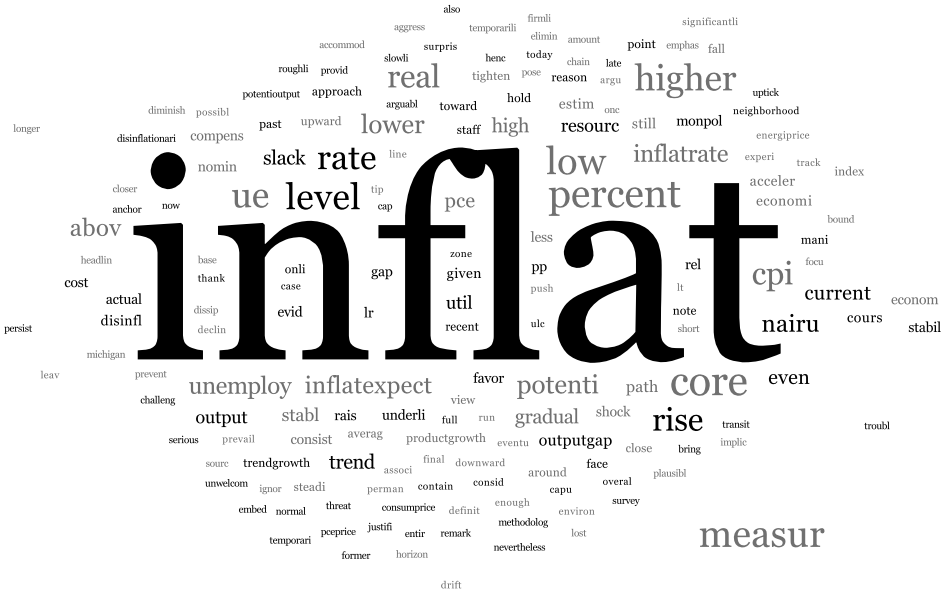
7 The estimated topics and other LDA outputs

While space constraints mean that we cannot show all the estimated topics, we thought that it would be useful to show some of the topic outputs and also to examine, from a purely descriptive point of view, how the FOMC was talking about some these different topics overtime.

7.1 Estimated topics and how intuitively LDA works

As an obvious place to start, we look at the topic which we call "The Inflation Topic". We represent the probability distribution over words (precisely, tokens) as a word cloud in which the size of the word indicates the probability of choosing that word when we talk about that topic. Figure 1 shows that this topic is clearly about inflation; the stemmed token "inflat", capturing "inflation", "inflationary", etc... is highly prominent.

Figure 1: Topic 3 - Inflation



Using the same word cloud approach, Figures 2 and 3 show topics related consumption and investment, and productivity respectively. There is a specific topic covering banking issues (figure 4) which includes words like bank, debt, liquid, mortgage, and loan with high probability.

But not all topics are about economic topics. In order for conversation to work, there are a lot of "filler" or discussion words which are necessary to convey meaning

Figure 2: Topic 12 - Consumption and Investment demand



Figure 3: Topic 13 - Productivity



when talking about any economic topic. We present two examples below. Our analysis suggests that there is a topic dedicated to the names of other members on the committee which is presented as a word cloud in figure 5. This is driven both by members using names to refer to others arguments, but especially by the roll call vote at the end of the meetings. Another topic captures points in the discussion where specific reference is made to materials presented or circulated by the staff. As seen in figure 6, charts, and lines on charts, are referred to in this topic.

Given that these topics seem incredibly close to the topics that economists might generate, a key questions is how the LDA model actually works to generate topics which make sense? The basic intuition is about the co-occurrence of words in documents. As discussed by Blei (2009), LDA places regularly co-occurring words together into topics because it wants to spread words across few topics to maximise the word probabilities for each given topic (the word probabilities are the $p(w_{d,n}|z_{d,n}, \beta_k)$ term above). This is not driven by the driven by sparsity imposed on the distribution (which comes from the Dirichlet hyperparameter) but rather simply reflects the best fitting posterior.

Another reason that this approach does so well is that it is a mixed membership model and so allows words to appear in many documents and words in a given document to be from different topics. This is a less restrictive approach than a mixture model which would constrain each document to have only a single topic. The less restrictive approach actually means we get topic distributions which look like more naturally co-occurring words. For example, imagine an article about statistics in football. A mixture model would need "Gaussian" and "Touchdown" to be in the same topic, but a mixed membership model can split them across two topics meaning that one topic has "Gaussian" together with other statistics words, while the other has "Touchdown" and other football words.

7.2 Allocation of attention over topics

Along with the topics, we also have the allocation of attention across topics by each member for each of their interjections in the discussion. We can aggregate these document distributions (over topics) at any level, but to illustrate the basic analysis, we present some illustrative time-series of the time the FOMC (as an aggregate) spent on different topics.

Given the Baker, Bloom, and Davis (2013), BBD hereafter, economic uncertainty index picks up the effects of expiring fiscal measures, as well market measures of risk, we shall compare the time-series of time allocated to the estimated "Fiscal" topic (Figure 7 below) and the "Risk" topic (Figure 8) to the BBD measure. As can be seen in Figure 9, in the late 1980s, and around times of war, the committee discusses the fiscal topic more often and this tends to coincide with increases in BBD uncertainty. However, in the later

Figure 4: Topic 38 - Banking



Figure 5: Topic 4 - Referring to other people



years, there are many spikes in BBD uncertainty missed by the FOMC in terms of the fiscal topic; instead these are captured by the risks topic as shown in Figure 10. What is particularly noteworthy is that these data suggest that up to the end of 2007, the BBD measure of risk had not yet picked up while the FOMC was increasingly discussing it.

Topic 16, which we call "Central Bank Communication", is regularly used in meetings to discuss things like statements and the previous minutes. The words which make it up are presented in figure 11 and concern all aspects of FOMC communication strategy such as statement, press, announcement and transparent. This is useful to help check whether the decision to reveal the transcripts was surprising. As we argue for our natural experiment, FOMC members only learned of the transcripts and the likelihood that they would be published in November 1993. If it were indeed a big surprise, you would expect there to be more than usual discussion of issues of communication. Figure 12 shows that a typical meeting might spend 2% of the meeting on this topic, and in an unusual meeting, perhaps discussing a particularly tricky statement, the FOMC would spend up to 8% of the meeting on the topic; the November 1993 meeting spent around 18% of the meeting discussing the issue of transparency and transcripts being made public.

8 Empirical Results

8.1 Results: Counts

Before even turning to the results of the LDA analysis, we can use the primitive descriptive statistics from the transcripts to get a sense of whether there were changes in the deliberation after transparency. To be begin with, we estimate equations of the form (1) in which simply look at the change between before and after transparency. We control in these regressions for a variety of meeting and time-specific controls, but there are not time fixed effects and as it is the aggregate meeting level, we cannot perform the difference-in-difference analysis. We carry out the analysis separately for the discussion of the economic situation (FOMC1) and the policy discussion (FOMC2).

Table 2 shows the changes that took place in FOMC1. Meetings are characterized by more words, but delivered in fewer statements resulting in simply more words per statement. The large drop in statements perhaps reflects a lack of back-and-forth dialogue which would generate many statements as the debate bounces from member to member. The discussion part is also characterized by less questions. These simple counts already paint a picture of FOMC members coming to the meeting with longer, more prepared views on the economy, but somewhat less likely to question the staff and there colleagues during the discussion.

Table 3 shows that there is less evidence of prepared statements in FOMC2; there is

Figure 6: Topic 41 - Referring to staff materials



Figure 7: Topic 42 - Fiscal

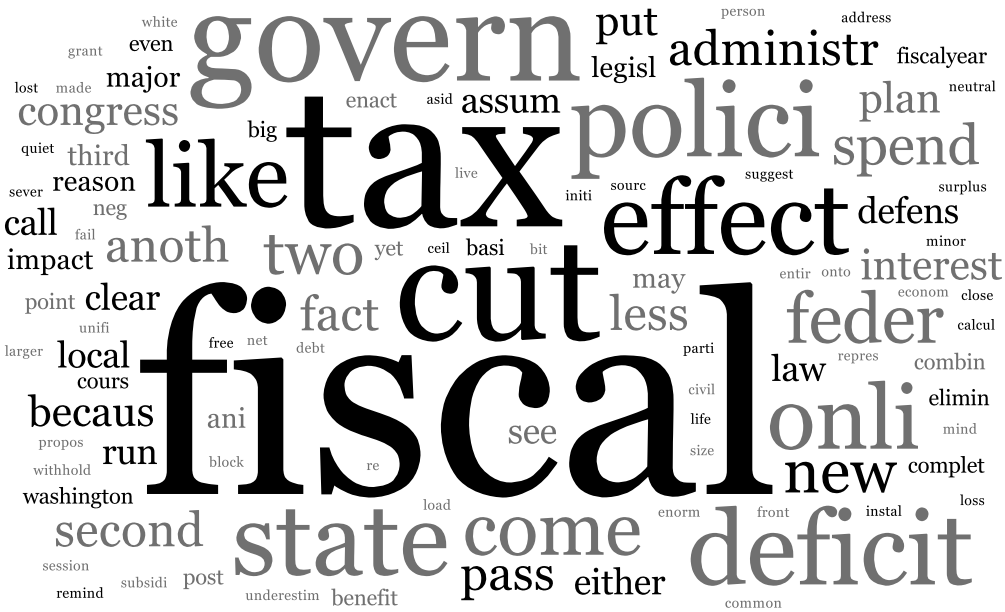


Table 2: Language measures - Overall Meeting level, Economic Situation Part

VARIABLES	(1) Total Words	(2) Statements	(3) Questions	(4) Sentences	(5) Words p.s.
D(transparent)	1,286**	-19.7**	-5.48**	84.6***	47.8***
	[0.013]	[0.016]	[0.048]	[0.003]	[0.001]
D(NBER recession)	684	-12.7	-5.01	11.9	35.1
	[0.253]	[0.253]	[0.362]	[0.714]	[0.146]
D(2 Day)	710	21.0*	9.02**	53.0**	-35.5***
	[0.106]	[0.059]	[0.010]	[0.028]	[0.004]
Uncertainty(t-1)	0.86	-0.055*	-0.0096	0.017	0.085*
	[0.720]	[0.077]	[0.395]	[0.890]	[0.060]
Constant	8,436***	101***	26.5***	468***	94.7***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.253	0.145	0.146	0.298	0.373
Lag Dep. Var?	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	79	79	79	79	79

no increase in the number of words for example. However, similar to the effect in FOMC 1, there seems to have been a large drop in the dynamism of debate after transparency. Questions and interactions both fall by almost 50%.

However, to shed some light on whether these results apply to everyone, or whether they are focused on those less experienced in monetary policy, we proceed to estimate our diff-in-diff equation presented as (2). The key coefficient in these regressions is the ϕ coefficient on the interaction between our "D(Transparency)" variable and the measure of Fed Experience. Because this interaction is decreasing in the extent of career concerns, a positive coefficient suggests that newer members on the FOMC with more reputation to build actually do less of the dependent variable. For example, column (6) of table 4, where for the sake of brevity we present the results for FOMC1 and FOMC2 in a single table but reduced variables, suggests that the reduced questions in the policy discussion is especially driven by newer members. The same is true for the lower statements in FOMC 2. But the results suggest that the greater preparedness in the discussion of the economic situation is not specific to newer members. This seems to suggest that after transparency everyone started to come to meetings with more prepared information on the state of the economy, but then, having provided their information in FOMC 1, newer members took much less part in the debate about the correct policy choice.

8.2 Results: LDA

Table 5 shows that while there is no evidence that the less experienced members spend more time on topics that are clearly economics in the economic discussion part, they certainly cover more topics as picked up by the Herfinahl concentration index across

Table 3: Language measures - Overall Meeting level, Policy Part

VARIABLES	(1) Total Words	(2) Statements	(3) Questions	(4) Sentences	(5) Words p.s.
D(transparent)	648	-50.1***	-15.8***	4.50	59.7***
	[0.365]	[0.001]	[0.000]	[0.900]	[0.000]
D(NBER recession)	-180	-2.81	-4.21	-18.2	2.19
	[0.868]	[0.911]	[0.617]	[0.765]	[0.755]
D(2 Day)	1,739**	8.73	5.98	79.9**	12.5
	[0.017]	[0.505]	[0.178]	[0.030]	[0.159]
Uncertainty(t-1)	-0.18	-0.032	-0.020	-0.0054	0.012
	[0.943]	[0.482]	[0.119]	[0.965]	[0.647]
Constant	6,339***	111***	27.6***	344***	51.2***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.082	0.164	0.159	0.063	0.452
Lag Dep. Var?	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	79	79	79	79	79

their allocated topics. Interestingly, these broad discussants in the first part, flip to become less about economics topics in general and more about just a few, select topics.

Table 6 provides marginal evidence that the newer members tend to become closer to Greenspan in terms of the distance between what Greenspan and they talk about.

We now examine the persistence of topics across time. We are especially interested in whether, as predicted by Prendergast and Stole (1996), there is increased persistence after transparency. Moreover, their prediction is that the greater persistence should be driven more by experienced members.

As shown in table 7, the findings are exactly in line with the theoretical predictions. Column (1) shows that after transparency, the average AR(1) coefficient on a topic increases after transparency. Column (2) shows that this increase is driven mainly by the experienced members.

8.3 Results: Dissent on the FOMC

In order to examine whether transparency induces those with greater career concerns to dissent less often, we follow the earlier literature and use dissents in voice rather than votes. The reason for this is that the FOMC meetings are characterized by a reluctance for too many members to formally disagree with the Chair's proposal. Instead, in the second discussion round, there are often voiced disagreements which are not then reflected in the vote. We measure voiced dissents using the data provided by Meade (2005); she used the FOMC transcripts and coded member i in meeting t as dissenting in voice when, in meeting t, member i formally expresses disagreement with the Chair's favored strategy. These are the same dissent data that were used in Meade and Stasavage (2008) and they

Table 4: Language measures - Member level

VARIABLES	(1) Total Words	(2) Statements	(3) Questions	(4) Total Words	(5) Statements	(6) Questions
D(transparent)	-2,102	-26.4	-2.10	-2,707	39.5	23.4*
	[0.364]	[0.449]	[0.898]	[0.155]	[0.272]	[0.071]
Fed Experience	250	2.61	0.23	456	-6.70	-3.72*
	[0.294]	[0.468]	[0.891]	[0.114]	[0.219]	[0.058]
D(Trans) x Fed Experience	-0.54	0.033	0.0052	-2.38	0.13***	0.048***
	[0.797]	[0.309]	[0.728]	[0.345]	[0.007]	[0.005]
Constant	-862	-11.6	-0.098	-2,279	44.9	22.5**
	[0.532]	[0.576]	[0.992]	[0.168]	[0.150]	[0.045]
R-squared	0.145	0.203	0.129	0.268	0.272	0.228
Number of mem_num	38	38	38	38	38	38
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC2	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1449	1449	1449	1428	1428	1428

are available from 1989 to 1997.

Meade and Stasavage (2008) examined the overall change ("diff") in likelihood of dissent and found that members are less likely to dissent after transparency. As with any "diff" analysis, these findings are vulnerable to the suggestion that something else changed at the same time, such as economic conditions, and so the finding could have nothing to do with transparency. Even if the result is driven by transparency there is no reason to assume that it is evidence of conformism; if discipline leads to the committee having more high quality information, then members might be more likely to come to the same conclusion.

In table 8, we examine the diff-in-diff results on these voiced dissents data. While we find that transparency does indeed reduce the likelihood of dissent on the committee as a whole, there is no evidence of the effect being greater for rookies.

8.4 Results: Influence

Using our measure of influence, we find evidence that newer members become influential over Alan Greenspan across meetings. The results are reported in table 9. While the sign is suggestive of growing influence within the meeting, the effect is not statistically significant.

Table 5: Topics discussed - Member level

VARIABLES	(1) Econ A	(2) HH_U_A	(3) Econ B	(4) HH_U_B
D(transparent)	2.35***	-0.19	0.012	-0.080
	[0.000]	[0.231]	[0.972]	[0.139]
Fed Experience	-0.23***	0.021	0.0039	0.013
	[0.000]	[0.206]	[0.939]	[0.107]
D(Trans) x Fed Experience	0.00043	0.00040***	-0.0014***	-0.00015**
	[0.401]	[0.006]	[0.002]	[0.037]
Constant	1.90***	-0.098	0.52*	-0.060
	[0.000]	[0.301]	[0.073]	[0.200]
R-squared	0.262	0.148	0.204	0.245
Number of mem_num	38	38	38	38
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1449	1449	1428	1428

9 The effect of transparency on the Federal Funds Rate

The key macroeconomic question is whether there are any effects on the Federal Funds Rate that the FOMC decides? This is a very tricky question because overall analysis can only rely on a "diff" analysis. However, subject to that caveat, we show two pieces of evidence which support a committee that makes decisions more quickly. These results could, potentially, be the result of greater certainty which itself results from greater information brought to the discussion by more prepared members.

9.1 Kuttner Monetary Policy Shocks

As a first test, we examine whether the FOMC was more or less likely to surprise the markets with the interest rate decisions. To this, we use the surprise data developed in Kuttner (2001) and subsequently used in Bernanke and Kuttner (2005). These data are derived from traded futures securities which allow him to decompose movements in the Fed Funds target rate into expected and unexpected moves. His approach corrects for many of the concerns that arise from just doing these calculations raw, and these data are available from July 1989 to 2008.

In table 10 we report the results of estimating (1) on these data. The results suggest that there was more (absolute) surprise following transparency. This is true even after we control also for measures of the extent of the discussion such as how many questions there were, how much discussion there was, etc...

Table 6: Distance - Member level

VARIABLES	(1) Cdist_chairman_A	(2) Udist_chairman_A	(3) Cdist_chairman_B	(4) Udist_chairman_B
D(transparent)	0.44	0.67	-1.05	-0.93
	[0.822]	[0.671]	[0.294]	[0.276]
Fed Experience	-0.046	-0.064	0.15	0.14
	[0.822]	[0.692]	[0.339]	[0.294]
D(Trans) x Fed Experience	0.0024	0.0028*	0.0025*	0.00074
	[0.177]	[0.057]	[0.063]	[0.510]
Constant	1.03	1.27	-0.15	-0.034
	[0.379]	[0.179]	[0.864]	[0.964]
R-squared	0.377	0.314	0.443	0.446
Number of mem_num	38	38	38	38
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1449	1449	1428	1428

9.2 Gradualism changes

If it were the case the FOMC meetings were characterized by more information, one possible effect would be reduced policy inertia, or gradualism. This is because it is sometimes argued that gradualism is driven by policy makers reluctance to change policy when uncertain and then to have to quickly reverse their change as it would reveal the extent of their uncertainty. This helps explain why interest rates can change very quickly in times of crisis (when it is clear that interest rates need to change). One potential way to see whether there is a change in gradualism is to estimate an interest rate reaction function and see if there are changes with transparency.

To do this, we take the state of the art approach in Coibion and Gorodnichenko (2012). Their specification allows for flexible time series structures which allows for persistent shocks. We also employ their real time data and forecasts. In table 11, we present the estimates for pre-1993 (column (1)), post-1993 (column (2)) and a specification where we include data but allow the post-transparency coefficients to vary using interaction terms. In all these specifications, the specification allows that shocks to have become less persistent over time. Consistent with the committee acting with greater certainty, the results clearly point to reduced gradualism.

10 Conclusions

Overall, we find results that are in line with the predictions of the career concerns literature. Our difference-in-difference results show transparency induces the inexperienced to engage less in debate; prepare more; and be more willing to change positions between

Table 7: Persistence of topic weights - Member-Topic level

VARIABLES	(1) Topic(ik)	(2) Topic(ik)
L.Topic(i,k)	0.18***	0.18***
	[0.000]	[0.000]
D(Transparent) x L.Topic(i,k)	0.025***	0.0018
	[0.002]	[0.892]
D(Transparent) x L.Topic(i,k) x Fed Experience		0.0020***
		[0.005]
D(Transparent)	-0.00093	-0.000065
	[0.644]	[0.999]
Fed Experience		0
		[1.000]
D(Transparent) x Fed Experience		-0.000076
		[0.212]
Constant	0.030***	0.030
	[0.000]	[0.719]
R-squared	0.037	0.037
Number of mem_topic	918	891
Member-Topic Effects	Yes	Yes
Time Effects	Yes	Yes
Meeting Section	FOMC1 & FOMC2	FOMC1 & FOMC2
Sample	89:11-97:09	89:11-97:09
Obs	32103	30375

meetings. The net effect makes them more influential in shaping discussion. We therefore believe that central bank designers should consider the role of transparency in disciplining policymakers.

Table 8: Voiced Dissent - Member level

VARIABLES	(1) D(Voice Dissent)
D(transparent)	-5.71**
	[0.017]
Fed Experience	1.10**
	[0.020]
D(Trans) x Fed Experience	0.0028
	[0.349]
Constant	-7.98**
	[0.023]
Number of mem_num	33
R-squared	0.230
Member FE	Yes
Time FE	Yes
Meeting Section	FOMC2
Sample	89:11-97:09
Obs	1164

Table 9: Influence - Member level

VARIABLES	(1) influenceG	(2) influenceG
D(transparent)	0.0059	-0.0065
	[0.881]	[0.861]
Fed Experience	-0.00076	0.00097
	[0.897]	[0.801]
D(Trans) x Fed Experience	-0.000070	-0.00011***
	[0.174]	[0.001]
Constant	0.0082	-0.0013
	[0.809]	[0.953]
R-squared	0.254	0.063
Number of mem_num	38	35
Member FE	Yes	Yes
Time FE	Yes	Yes
Within Meeting	Intra	Inter
Sample	89:11-97:09	89:11-97:09
Obs	1424	1422

References

Bailey, A., and C. Schonhardt-Bailey (2008): "Does Deliberation Matter in FOMC Monetary Policymaking? The Volcker Revolution of 1979," *Political Analysis*, 16, 404–427.

——— (2013): *Deliberating Monetary Policy*. MIT Press, Cambridge.

Baker, S. R., N. Bloom, and S. J. Davis (2013): "Measuring Economic Policy Uncertainty," Mimeograph, Stanford University.

Bernanke, B. S., and K. N. Kuttner (2005): "What Explains the Stock Market's Reaction to Federal Reserve Policy?," *Journal of Finance*, 60(3), 1221–1257.

Bird, S., E. Klein, and E. Loper (2009): *Natural Language Processing with Python*. O'Reilly Media.

Table 10: The transparency effect on meeting surprises

VARIABLES	(1) |Surprise|	(2) |Surprise|	(3) |Surprise|	(4) |Surprise|
|Surprise|(−1)	0.060	0.060	0.063	0.047
	[0.535]	[0.535]	[0.524]	[0.657]
Words p.m.	0.0053**			
	[0.024]			
D(transparent)	5.94**	5.94**	5.96***	6.06***
	[0.011]	[0.011]	[0.009]	[0.007]
Sentences p.m.		0.100**		
		[0.027]		
Questions p.m.			0.68	
			[0.409]	
Exclamations p.m.				2.92
				[0.479]
Constant	-20.2***	-20.0***	-17.3***	-17.9***
	[0.001]	[0.001]	[0.004]	[0.005]
R-squared	0.139	0.131	0.117	0.117
Macro controls	Yes	Yes	Yes	Yes
Meeting Section	FOMC2	FOMC2	FOMC2	FOMC2
Sample	90:1-06:12	90:1-06:12	90:1-06:12	90:1-06:12
Obs	134	134	134	134

Blei, D. (2009): "Topic Models: Lectures at the Machine Learning Summer School (MLSS), Cambridge 2009," http://videolectures.net/mlss09uk_blei_tm/.

——— (2012): "Probabilitic Topic Models," *Communcations of the ACM*, 55(4).

Blei, D., and J. Lafferty (2006): "Dynamic Topic Models," in *Proceedings of the 23rd Internaional Conference on Machine Learning*, pp. 113–120.

——— (2009): "Topic models," in *Text Mining: Classification, Clustering, and Applications*, ed. by A. Srivastava, and M. Sahami. Taylor & Francis, London, England.

Blei, D. M., A. Y. Ng, and M. I. Jordan (2003): "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 3, 993–1022.

Blinder, A. S., and J. Morgan (2005): "Are Two Heads Better than One? Monetary Policy by Committee," *Journal of Money, Credit and Banking*, 37(5), 789–811.

Chappell, H. W., R. R. McGregor, and T. A. Vermilyea (2005): *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee*, vol. 1. The MIT Press, 1 edn.

Coibion, O., and Y. Gorodnichenko (2012): "Why Are Target Interest Rate Changes So Persistent?," *American Economic Journal: Macroeconomics*, 4(4), 126–62.

Financial Times (2013): "ECB heads towards more transparency over minutes," http://www.ft.com/cms/s/0/a678451e-fa99-11e2-87b9-00144feabdc0.html#axzz2fou0Xcm5, 01 August 2013, last accessed 24 September 2013.

Fry, M., D. Julius, L. Mahadeva, S. Roger, and G. Sterne (2000): "Key issues in the choice of monetary policy framework," in *Monetary Frameworks in a Global Context*, ed. by L. Mahadeva, and G. Sterne. Routledge.

Table 11: Gradualism before and after transparency

VARIABLES	(1) FFR_t	(2) FFR_t	(3) FFR_t
Expected inflation	-0.15	0.26***	-0.053
	[0.432]	[0.000]	[0.634]
Expected output gap	0.067	0.069*	0.024
	[0.312]	[0.064]	[0.489]
Expected output growth	0.078*	0.085***	0.100***
	[0.054]	[0.000]	[0.000]
FFR(-1)	1.02***	0.90***	1.03***
	[0.000]	[0.000]	[0.000]
D(Transparent) x Expected Inflation			0.32**
			[0.029]
D(Transparent) x Expected output gap			0.062
			[0.220]
D(Transparent) x Expected real GDP growth			-0.024
			[0.419]
D(Transparent) x FFR(-1)			-0.16**
			[0.035]
Constant	0.32	-0.29**	-0.20
	[0.594]	[0.039]	[0.239]
R-squared	0.986	0.992	0.991
Frequency	Quarterly	Quarterly	Quarterly
Sample	1987:1-1993:4	1994:1-2006:4	1987:1-2006:4
Obs	28	52	80
Method	Prais-Winsten	Prais-Winsten	Prais-Winsten

Gersbach, H., and V. Hahn (2012): "Information acquisition and transparency in committees," *International Journal of Game Theory*, 41(2), 427–453.

Greenspan, A. (1993): "Comments on FOMC communication," in *Hearing before the Committee on Banking, Finance and Urban Affairs, 103rd Congress, October 19*, ed. by U. H. of Representatives.

Griffiths, T. (2002): "Gibbs sampling in the generative model of Latent Dirichlet Allocation," Discussion paper, Stanford University.

Griffiths, T. L., and M. Steyvers (2004): "Finding scientific topics," *Proceedings of the National Academy of Sciences*, 101(Suppl. 1), 5228–5235.

Hansen, S., M. McMahon, and C. Velasco (2012): "How Experts Decide: Preferences or Private Assessments on a Monetary Policy Committee?," manuscript.

Heinrich, G. (2009): "Parameter estimation for text analysis," Discussion paper.

Holmström, B. (1999): "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66(1), 169–82.

Kuttner, K. N. (2001): "Monetary policy surprises and interest rates: Evidence from the Fed funds futures market," *Journal of Monetary Economics*, 47(3), 523–544.

Levy, G. (2004): "Anti-herding and strategic consultation," *European Economic Review*, 48(3), 503–525.

——— (2007): "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, 97(1), 150–168.

Lombardelli, C., J. Proudman, and J. Talbot (2005): "Committees Versus Individuals: An Experimental Analysis of Monetary Policy Decision-Making," *International Journal of Central Banking*, 1(1).

Meade, E. (2005): "The FOMC: preferences, voting, and consensus," *Review*, (March), 93–101.

Meade, E., and D. Stasavage (2008): "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve," *Economic Journal*, 118(528), 695–717.

Meyer, L. (2004): *A Term at the Fed : An Insider's View*. Collins.

Palacios-Huerta, I., and O. Volij (2004): "The Measurement of Intellectual Influence," *Econometrica*, 72(3), 963–977.

Phan, X.-H., and C.-T. Nguyen (2007): "A C/C++ implementation of Latent Dirichlet Allocation (LDA)," .

Prat, A. (2005): "The Wrong Kind of Transparency," *American Economic Review*, 95(3), 862–877.

Prendergast, C., and L. Stole (1996): "Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning," *Journal of Political Economy*, 104(6), 1105–34.

Quinn, K. M., B. L. Monroe, M. Colaresi, M. H. Crespin, and D. R. Radev (2010): "How to Analyze Political Attention with Minimal Assumptions and Costs," *American Journal of Political Science*, 54(1), 209–228.

Romer, C. D., and D. H. Romer (2004): "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, 94(4), 1055–1084.

Rosen-Zvi, M., C. Chemudugunta, T. Griffiths, P. Smyth, and M. Steyvers (2010): "Learning Author-Topic Models from Text Corpora," in *ACM Transactions on Information Systems*, vol. 28.

Scharfstein, D. S., and J. C. Stein (1990): "Herd Behavior and Investment," *American Economic Review*, 80(3), 465–79.

Steyvers, M., and T. L. Griffiths (2006): "Probabilitic Topic Models," in *Latent Semantic Analysis: A Road to Meaning*, ed. by T. Landauer, D. McNamara, S. Dennis, and W. Kintsch. Laurence Erlbaum.

Thornton, D., and D. C. Wheelock (2000): "A history of the asymmetric policy directive," *Review*, (Sep), 1–16.

Visser, B., and O. H. Swank (2007): "On Committees of Experts," *The Quarterly Journal of Economics*, 122(1), 337–372.

A Sampling Algorithms

A.1 Estimation using Collapsed Gibbs sampling

Table 12: Notation

Notation	Quantity
Basic Notation	
N_d	Number of words in document d
D	Total number of documents
d	Indexes a document
V	Total number of unique tokens (Vocabulary)
v	Indexes a unique tokens
K	Total number of topics
k	Indexes a topic
$w_{d,n}$	Word n in document d
$z_{d,n}$	Topic allocation of word n in document d
Dirichlet Distributions	
β_k	Term distribution for each topic k
η	Dirichlet hyperparameter associated with term distributions
θ_d	Topic distribution for each document d
α	Dirichlet hyperparameter associated with topic distributions
Counts	
n_k^d	Count of words in document d allocated to topic k (K of these)
n_v^k	Count of times unique token v is allocated to topic k (V of these)
$n_{k,-i}^d$	Excluding token i, count of words in document d allocated to topic k (K of these)
$n_{v,-i}^k$	Excluding token i, count of times unique token v is allocated to topic k (V of these)

In practice, the inference problem is made relatively straightforward by assuming a symmetric Dirichlet prior for both topic distributions for each document and for the vocabulary distribution for each topic. Then we use Markov chain Monte Carlo (MCMC) for inference. Following Griffiths (2002), collapsed Gibbs sampling is used. The basic idea of Gibbs sampling is to sampling all variables from their distributions conditional on the current values of all other variable and the data. In the LDA model, the data are the words, and the key quantity is the allocation of each word to a topic which is given by $z_{d,n}$; from the word allocations, we can infer the implied topic and term distributions. As explained in Griffiths (2002) and Heinrich (2009),[19], given the assumptions of the model (especially the assumption of a Dirichlet distribution) the conditional distribution of a word being assigned to a topic, given all other word-topic assignments and the data (vector of words) is given by:

$$Pr(z_i = k|z_{-i}, \vec{w}) \propto \frac{n_{v,-i}^k + \eta}{\sum_{v=1}^V n_{v,-i}^k + \eta} \left(n_{k,-i}^d + \alpha\right) \tag{A.1}$$

[19]These two are the key references which the following description draw upon.

where $\vec{w}$ is the collection of all words, $n^d_{k,-i}$ is the count of words in document d allocated to topic k (K of these) excluding the current token $i = (d, n)$ and $n^k_{v,-i}$ is, again excluding token i, the count of times unique token v is allocated to topic k (V of these).

In terms of the specific approach to the collapsed Gibbs sampling for the LDA model, the approach is the following:

1. We start with a series of words in statements ($w_{d,n}$)
2. We need to initialise the process with an allocation of topics to words (can be completely random) to generate a first iteration of $z_{d,n}$ for each word in each document.
3. Using this allocation of words to topics, we can calculate the necessary word counts n^k_v and n^d_k.
4. Then, iterating through each word in each document (each i), we can:
   (a) Drop the word i from the sample (reducing all the relevant counts) - this gives us the $-i$ sample;
   (b) Sample a new topic allocation for word i using (A.1);
   (c) Update all relevant word counts;
   (d) Move on to the next word $i + 1$
5. After a number of iterations through all words, we stop.
6. Each element in the term distribution matrix (K by V) is given by:

$$\widehat{\beta}^v_k = \frac{n^k_{v,-i} + \eta}{\sum_{v=1}^{V} \left(n^k_{v,-i} + \eta\right)} \tag{A.2}$$

   and each element in the topic distribution matrix (D by K, or is it K by D - CHECK) is given by:

$$\widehat{\theta}^k_d = \frac{n^d_{k,-i} + \alpha}{\sum_{k=1}^{K} \left(n^d_{k,-i} + \alpha\right)} \tag{A.3}$$

The Dirichlet hyperparameters are an important input to the procedure. As mentioned earlier, they determine the amount of sparsity we impose on the Dirichlet distribution; as the value of the hyperparameter falls, we get more a more sparse β_k and θ_d matrices (and so finer term distributions for each topic and more distinct topic distributions in each document). While some earlier forms of topic model estimated the hyperparameters, collapsed Gibbs sampling LDA actually imposes a value on the inference problem. As first shown in Griffiths and Steyvers (2004), the resuts are found to be of highest quality if we set $\eta = 0.01$ and $\alpha = 50/K$.

A.2 Querying: New documents and aggregation of documents

Remember that a document for us, in the basic transcript setting, is a single intervention (or statement) by a person at a given FOMC meeting. Sometimes, having "trained" the

topic model on a given set of data, we may wish to see how the model predicts the topics discussed in another new set of documents ($\tilde{d}$. This is called querying.

We also use this to carry out the aggregation of our dataset into the relevant meeting, or sub-meeting sections in order to, for example, work with everything that Janet Yellen discussed in the monetary policy discussion of a particular meeting.

The steps to carry out such a query are to basically hold the term distributions fixed (i.e. what words and in which proportions make up each topic) from the previously sampled topic model M and to resample for the new document's topic distributions given these β_k distributions. The resampling uses:

$$Pr(\tilde{z}_i = k | z_{-i}, \vec{w}; M) \propto \widehat{\beta}_k^v \left(n_{k,-i}^d + \alpha \right) \tag{A.4}$$

This resampling allows us to generate new estimates of $\tilde{\theta}_d^k$ for the new documents.

Figure 8: Topic 20 - Risk



Figure 9: BBD Uncertainty and Discussion of "Fiscal Topic"

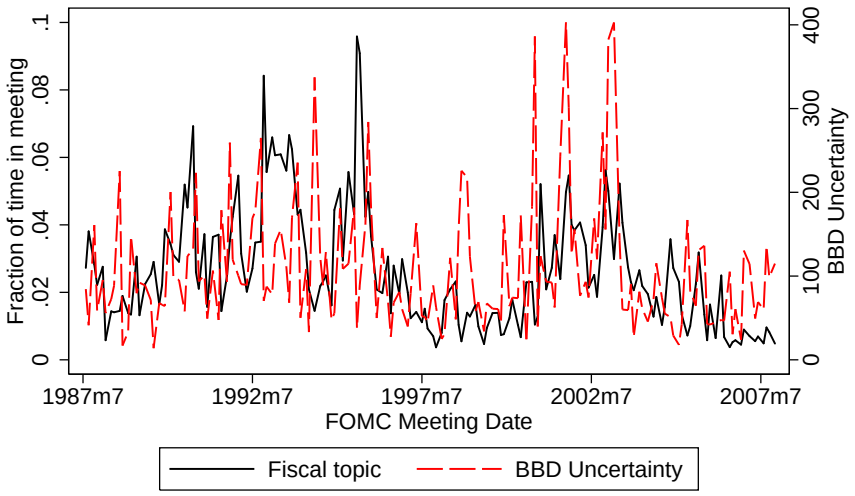


Figure 10: BBD Uncertainty and Discussion of "Risks Topic"

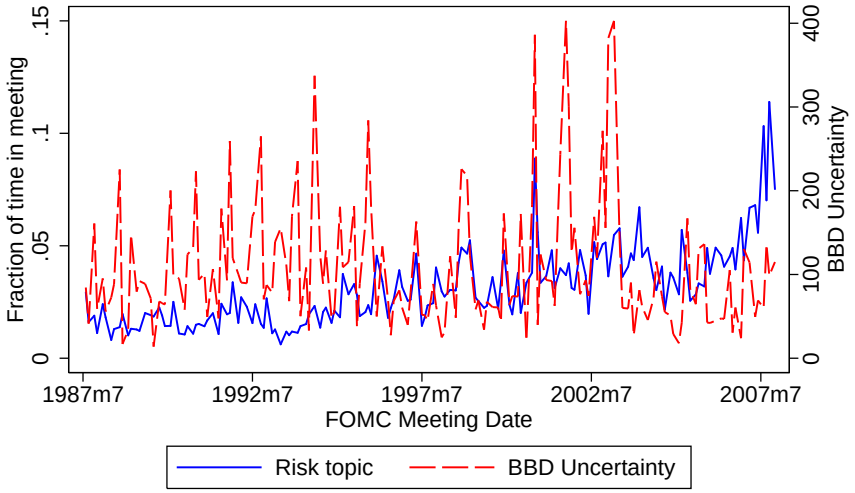


Figure 11: Topic 16 - Central Bank Communication

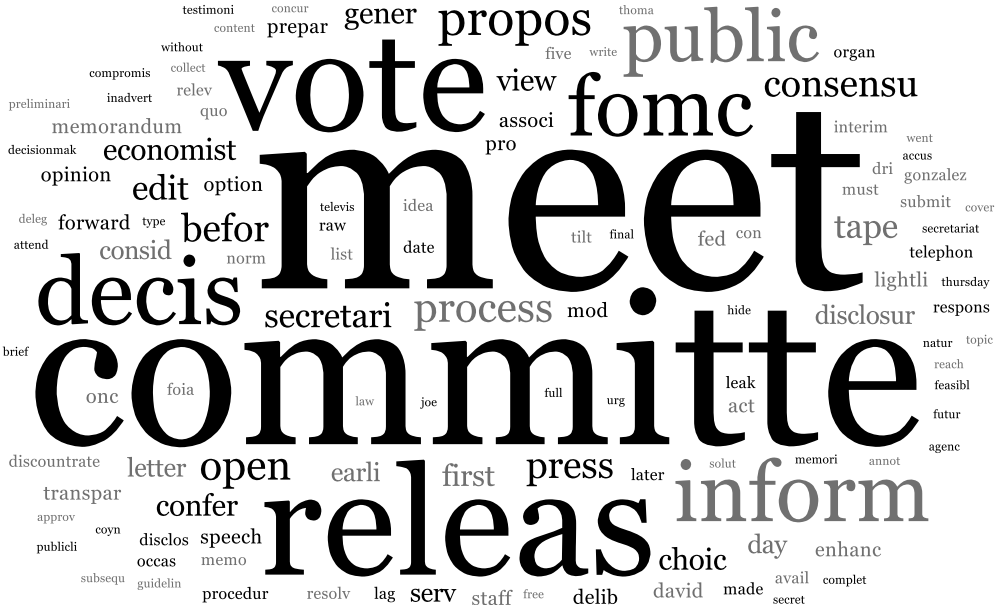


Figure 12: Topic 16 Discussion: Surprised by transparency revelation?

