Transparency and Deliberation within the FOMC: a Computational Linguistics Approach*

Stephen Hansen† Michael McMahon‡ Andrea Prat§

First Draft: April 29, 2014
This Draft: July 2, 2014

Abstract

How does transparency, a key feature of central bank design, affect the deliberation of monetary policymakers? We exploit a natural experiment in the Federal Open Market Committee in 1993 together with computational linguistic models (particularly Latent Dirichlet Allocation) to measure the effect of increased transparency on debate. Commentators have hypothesized both a beneficial discipline effect and a detrimental conformity effect. A difference-in-differences approach inspired by the career concerns literature uncovers evidence for both effects. However, the net effect of increased transparency appears to be a more informative deliberation process.

Keywords: Monetary policy, deliberation, FOMC, transparency, career concerns

JEL Codes: E52, E58, D78

*We would like to thank Francesco Amodio, Andrew Bailey, Francesco Caselli, Gilat Levy, Rick Mishkin, Emi Nakamura, Tommaso Nannicini, Bryan Pardo, Amar Radia, Glenn Rudebusch, Cheryl Schonhardt-Bailey, Jón Steinsson, Dave Stockton, Thomas Wood, and Janet Yellen for insightful discussions. We are particularly grateful to Omiros Papaspiliopoulos for numerous helpful discussions on MCMC estimation and Refet S Gürkaynak for sharing the monetary policy surprise data. We have also benefited from comments and suggestions by seminar attendees at the San Francisco Federal Reserve, University of Warwick, University of Manchester, INSEAD, Bank of England, LSE, the New York Federal Reserve, Columbia University, the ESRI, Universitat Pompeu Fabra, the CEP conference, ESSIM and the ECB. We thank Eric Hardy for excellent research assistance in gathering biographical data, and the Bank of England's Research Donations Committee for seed corn financial support. 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 key feature of central bank design, affects the deliberation of monetary policymakers on the Federal Open Market Committee (FOMC). In other words, we ask: what are the effects on internal deliberation of greater external communication? Deliberation takes up the vast majority of the FOMC's time and is seen by former members as important for the ultimate decision (see Meyer 2004, for example), but yet it remains little studied beyond anecdotal accounts. Determining how monetary policy committees deliberate, and how this depends on central bank design, is therefore important for understanding monetary policy decision making.[1] 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.

As table 1 shows, there is heterogeneity across three major central banks in terms of how detailed are the descriptions of policy meetings that are put on the public record.[2] 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). It is unclear, though, whether a central bank wishing to increase transparency should move to only release minutes, or whether there would be an additional benefit from disclosing full transcripts, as occurs with FOMC meetings. This is precisely the question currently facing the Bank of England, which has just announced a review of its policy to not release transcript information.

Table 1: Information made available by different central banks

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

What is the optimal disclosure policy? Policymakers and academics have identified potential positive and negative effects of an increase in how much information about the internal workings of a central bank is revealed to the public.

On the positive side, there is a broad argument that transparency increases the accountability of policymakers, and induces them to work harder and behave better. This argument has been explicitly applied to central banking (Transparency International

[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] 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.

2012), and even the ECB, the least open of the large central banks, states that: "Facilitating public scrutiny of monetary policy actions enhances the incentives for the decision-making bodies to fulfill their mandates in the best possible manner."[3] This effect is often labeled as discipline in agency theory and it arises in the Holmström (1999) career concerns model. The more precise the signal the principal observes about the agent, the higher the equilibrium effort of the agent.

On the negative side, many observers argue that too much transparency about deliberation will stifle committee discussion. In fact, before the Fed had released transcripts, Alan Greenspan expressed his views to the Senate Banking Committee (our emphasis):

> "**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.**" Greenspan (1993), as reported in Meade and Stasavage (2008).

The view that more transparency may lead to more conformity and hence less information revelation is formalized in the career concerns literature. Greater disclosure can induce experts who are concerned with their professional reputation to pool on actions that are optimal given available public signals even when their private signals would suggest that other actions are optimal (Prat 2005). In such circumstances, the principal benefits from committing to a policy of limited transparency.[4]

Of course, it is possible that both effects—discipline and conformity—operate simultaneously, in which case one should ask whether on balance more disclosure improves or worsens information aggregation. We are able to explore these issues by exploiting the natural experiment that led to the release of the FOMC transcripts. Since the 1970s, FOMC meetings were tape recorded to help prepare minutes. Unknown to committee members, though, these tapes were transcribed and stored in archives before being

[3]From `http://www.ecb.europa.eu/ecb/orga/transparency/html/index.en.html`.

[4]Conformity arises when agents wish to signal expertise. Another potential cost of transparency is that policymakers may start pandering to their local constituencies in order to signal their preferences. While this may be a concern for the ECB, in the US there is much less regional heterogeneity than in the euro area. In any case, models of preference signalling do not make any clear predictions about the communication measures we study in this paper.

recorded over. They only learned this 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.[5] To avoid accusations of hiding information, and to relieve potential pressure to release information in a more timely fashion, the Fed quickly agreed to publish the past transcripts and all future transcripts with a five-year lag. We thus have a complete record of deliberation both when policymakers did not know that their verbatim discussions were being kept on file let alone that such information would be made public (prior to November 1993), and when they 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. This dissent data, recorded in Meade (2005), is a binary measure based on whether a policymaker voiced disagreement with Chairman Greenspan's policy proposal during the policy debate. Their main finding, which they interpret as conformity, is that the probability that members dissent declines significantly after transparency. We instead generate communication measures based on basic text counts and on *topic models*, a class of machine learning algorithms for natural language processing that estimates what fraction of time each speaker in each section of each meeting spends on a variety of topics.

This approach allows one to construct several measures of communication relating to both discipline and conformity, and also to compare which effect is stronger. The wealth of data also allows us to extend Meade and Stasavage (2008) in another direction. Rather than compare changes before and after transparency, we also use a difference-in-differences approach to pin down the precise effect of career concerns. Since career concerns models predict that reputational concerns decline with labor market experience, we estimate the differential effect of transparency on FOMC members with less experience in the Fed.

[5]The issue came to a head in October 1993, between the September and November scheduled FOMC meetings, when there were two meetings of the Senate Banking Committee to discuss transparency with Greenspan and other FOMC members. In preparation for the second of these meetings, during an FOMC conference call on October 15 1993, most of the FOMC members discovered the issue of the written copies of meeting deliberation. As President Keehn says in the record of this meeting (Federal Open Market Committee 1993): "Until 10 minutes ago I had no awareness that we did have these detailed transcripts." President Boehne, a long-standing member of the committee, added: "...to the very best of my recollection I don't believe that Chairman Burns or his successors ever indicated to the Committee as a group that these written transcripts were being kept. What Chairman Burns did indicate at the time when the Memorandum was discontinued was that the meeting was being recorded and the recording was done for the purpose of preparing what we now call the minutes but that it would be recorded over at subsequent meetings. So there was never any indication that there would be a permanent, written record of a transcript nature." He then added "So I think most people in the subsequent years proceeded on that notion that there was not a written transcript in existence. And I suspect that many people on this conference call may have acquired this knowledge at about the same time that Si Keehn did." Schonhardt-Bailey (2013) contains more contemporary recollections by FOMC members about the release of transcripts.

We find evidence of both discipline and conformity. FOMC meetings have two major parts related to the monetary policy decision, the economic situation discussion (which we label FOMC1) followed by the policy debate (FOMC2). After transparency, more inexperienced members come into the meeting and discuss a broader range of topics during FOMC1 and, while doing so, use significantly more references to quantitative data and staff briefing material. This indicates greater information acquisition between meetings, i.e. discipline. On the other hand, after transparency they disengage more with debate during FOMC2: they are less likely to make interjections, ask less questions, and stick to a narrow range of topics. They also speak more like Chairman Greenspan.

Discipline pushes towards an increase in the informativeness of inexperienced members' statements, while conformity pushes towards a decrease. To gauge the overall effect of transparency, we propose an influence score in the spirit of the PageRank algorithm in order to measure the strength of these two effects. After transparency, more inexperienced members become significantly more influential in terms of their colleagues' topic coverage, indicating that their statements contain relatively more information after transparency than before. Thus, while we confirm Greenspan's worries expressed above, the counteracting force of increased discipline after transparency which he does not mention appears even stronger. The main conclusion of the paper is that central bank designers should take seriously the role of transparency in disciplining policymakers.

The primary algorithm we use is Latent Dirichlet Allocation (LDA) introduced by Blei, Ng, and Jordan (2003). LDA is widely used in linguistics, computer science, and other fields and has been cited over 8,000 times in ten years. While topic modelling approaches are beginning to appear in the social science literature, there use so far is mainly descriptive. For example, Quinn, Monroe, Colaresi, Crespin, and Radev (2010) apply a topic model similar to LDA to congressional speeches to identify which members of Congress speak about which topics. An innovation of our paper is to use communication measures constructed from LDA output as dependent variables in an econometric model explicitly motivated by economic theory (more specifically, career concerns). We believe this illustrates the potential fruitfulness of combining traditional economic tools with those from the increasingly important world of "Big Data" for empirical research in economics more broadly.

Fligstein, Brundage, and Schultz (2014)—developed independently[6] from this paper—also apply LDA to FOMC transcripts focusing on the period 2000-2007. They describe the topics that the meeting as a whole covers rather than the topics of individuals, and verbally argue they are consistent with the sociological theory of "sense-

[6]The first public draft of Fligstein, Brundage, and Schultz (2014) of which we are aware is from February 2014. Our paper was developed in 2012 and 2013, with the main results first presented publicly in September 2013.

making". They claim that the standard models that macroeconomists use led them to fail to connect topics related to housing, financial markets and the macroeconomy. In contrast, this paper uses LDA applied to all data from the Greenspan era (1987-2006) to construct numerous measures of communication patterns at the meeting-section-speaker level and embeds them within a difference-in-differences regression framework to identify how transparency changes individual incentives.

Bailey and Schonhardt-Bailey (2008) and Schonhardt-Bailey (2013) also use text analysis to examine the FOMC transcripts. They emphasize the arguments and persuasive strategies adopted by policymakers (measured using a computer package called "Alceste") during three periods of interest (1979-1981, 1991-1993, and 1997-1999). Of course, many others have analyzed the transcripts without using computer algorithms; for example, Romer and Romer (2004) use the transcripts to derive a narrative-based measure of monetary policy shocks.

The paper proceeds as follows. Section 2 reviews the career concerns literature that motivates the empirical analysis, and section 3 describes the institutional setting of the FOMC. Section 4 lays out the econometric models used to study transparency. Section 5 then describes how we measure communication, while section 6 presents the main results on how transparency changes these measures. Section 7 examines the overall effect of transparency on behavior. Section 8 explores the effect of transparency on policy, and section 9 concludes.

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 the long-term payoffs of FOMC members depend on what people outside the FOMC think of their individual expertise in monetary policy. This is either because a higher perceived expertise leads to better employment prospects or because of a purely psychological benefit of being viewed as an expert in the field. The intended audience may 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 gathering and studying data 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 to increase incentives to acquire information prior to meetings.[7]

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. Ottaviani and Sørensen (2006) show (see their proposition 6) that the bias toward conformity or exaggeration depends on how well the agent knows his own type: experts with no self-knowledge conform to the prior while experts with high self-knowledge may exaggerate their own information in order to appear more confident. (See also Avery and Chevalier (1999) for a related insight.)

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.

Overall, the effect of increased transparency can be positive (through increased discipline) or negative (through increased conformity/non-conformity). In section 7 we return to examining which effect is stronger in the data.

[7]Equilibrium effort in period t in the Holmström model is $g'(a_t^*) = \sum_{s=1}^{\infty} \beta^s \frac{h_\varepsilon}{h_t + s h_\varepsilon}$ where g is the (convex) cost of effort, β is the discount factor, h_t is the precision on the agent's type (increasing in t), and h_ε is the precision of the agent's output. Clearly the cross derivative of a_t^* with respect to h_ε and h_t is decreasing. So, if one interprets transparency as increasing h_ε, the discipline effect will be higher for those earlier in their careers. Gersbach and Hahn (2012) explore this idea specifically for monetary policy committees.

3 The FOMC and its 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 one 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.[8]

The US president nominates members of the Board of Governors who are then subject to approval by the US Senate. A full term as a Governor is 14 years (with an expiry at the end of January every even-numbered year), but the term is actually specific to a seat around the table rather than an individual member so that most Governors join to serve time remaining on a term. 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.[9] 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.[10]

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. Before FOMC meetings, the members receive briefing in advance such as the "Green Book" (staff forecasts), "Blue

[8]Federal Reserve staff also attend the meeting and provide briefings in it.

[9]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.

[10]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.

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). All members participate in both stages regardless of whether they are currently voting members:[11]

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 (FOMC1):**
   - Board of Governors' staff present the economic situation (including forecast).
   - There are a series of questions on the staff presentations.
   - FOMC members present their views of the economic outlook. The Chairman tended to speak reasonably little during this round.
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 (FOMC2):**
   - 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.

The econometric analysis focuses mainly on the part of the meeting relating directly to the economic situation discussion which we call FOMC1, and the part relating to the discussion of the monetary policy decision which we call FOMC2. However, we estimate our topic models for all statements in each meeting in the whole sample.

[11] 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?

One concern may be that formal FOMC meetings might not be where the FOMC actually meets to make policy decisions but rather the committee meets informally to make the main decisions. Thankfully, this is less of a concern on the FOMC than it would potentially be in other central banks. This is because the Government in Sunshine Act, 1976, aims to ensure that Federal bodies make their decisions in view of the public and requires them 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.[12] This position suggests that the Committee takes very seriously the discussion of its business in formal meetings, which accords with what we have been told by staff and former members of the FOMC, as well as parts of the transcripts devoted to discussing how to notify the public that members had chosen to start 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

We now discuss the natural experiment that allows us to identify the effect of transparency, the econometric specification within which we embed it, and the data sources on which we draw.

4.1 Natural experiment

The natural experiment for transparency on the FOMC resulted from both diligent staff archiving and external political pressure. In terms of the former, for many years prior to 1993 Fed staff had recorded meetings to assist with the preparation of the minutes. As highlighted in the FOMC's own discussions (Federal Open Market Committee 1993, quoted in the introduction), members believed that 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—they first typed up and archived a verbatim text of the discussion.

FOMC members, including Chairman Greenspan, were not aware of these archives until political pressure from Henry B. Gonzalez forced the Fed to discuss how it might be more transparent, at which point staff revealed them to Greenspan. Shortly thereafter, in

[12] 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.

October 1993, Greenspan acknowledged the transcripts' existence to the Senate Banking Committee. Initially Greenspan argued that he didn't want to release any verbatim information as it would stifle the discussion. But pressure on the Fed was growing, and so it 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.[13]

4.2 Econometric specification

Since the decision to change transparency was not driven by FOMC concerns about the nature or style of deliberation, and the change came as a surprise, the most straightforward empirical strategy to identify the effects of transparency on deliberation is to estimate a baseline "diff" regression of the following form:

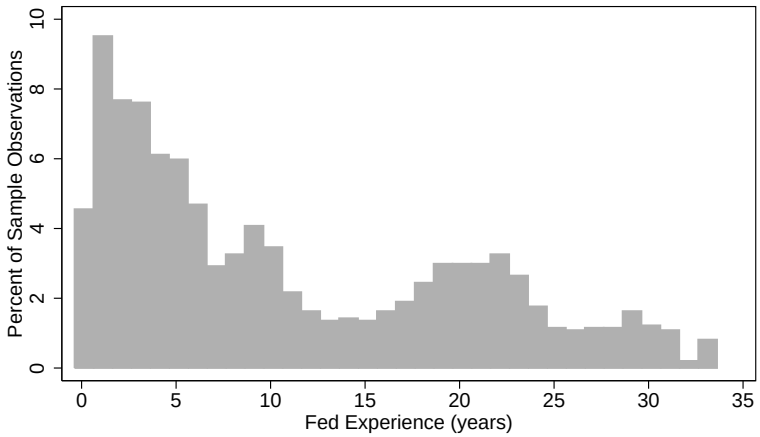
$$y_t = \alpha + \beta D(Trans) + \lambda X_t + \varepsilon_t, \tag{DIFF}$$

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.

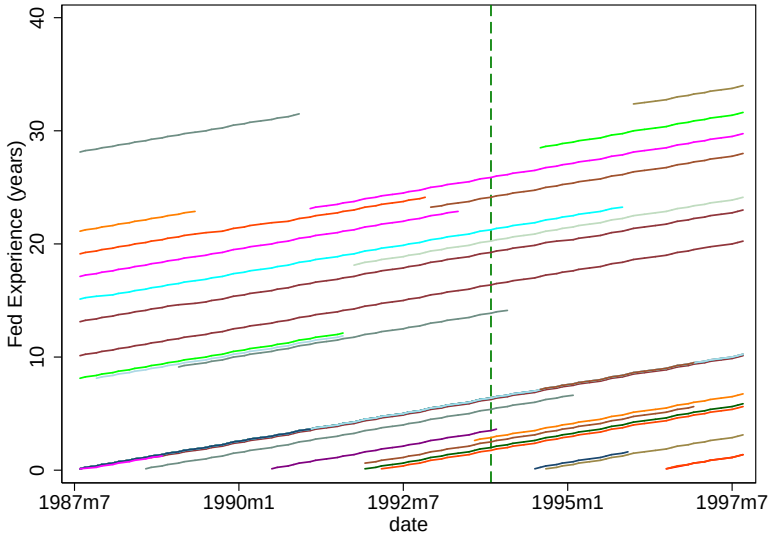
While useful as a descriptive account of behavior before and after transparency, the "diff" analysis is potentially problematic because the timing of other changes may have coincided with the change in transparency. As such, the β estimated in (DIFF) may capture the effects of these other changes, making it impossible to disentangle the different effects. In order to more clearly attribute the changes one observes to 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 expert's career concerns increases in the uncertainty of the principal's belief about the expert's ability. To capture uncertainty in FOMC members' ability, we define a variable $FedExp_{i,t}$ that measures the number of years member i has spent working in the Fed system through meeting t. This includes both years spent in the Fed before appointment to the FOMC, and years spent on the

[13]While the majority of members only found out about the existence of the transcripts in October 1993 as a result of the Senate hearings and a series of conference calls by FOMC members related to this process, some members were aware a bit earlier. Nonetheless, we choose November 1993 as the point at which the main transparency effects occur; this is the first meeting at which all members were aware of the transcripts and a decision to release the transcripts with a five-year lag had been put forward. If the few members that knew of the transcripts before October 1993 started to react to the possibility of the transcripts becoming public, this would tend to bias our estimates away from finding a change after November 1993.



(a) Histogram of $FedExp_{i,t}$



(b) Time-series of $FedExp_{i,t}$

Figure 1: Federal Reserve Experience ($FedExp_{i,t}$)

Notes: This figure plots a histogram (upper) and the individual time-series (lower) of the $FedExp_{i,t}$ variable, measured as years of Federal Reserve experience, in our main sample. The dashed line in the lower figure indicates November 1993 and the change in transparency on the FOMC.

committee.[14] The longer a member has served in the Fed, the more time the policymaking community has observed them, and so the less uncertainty there should be about their expertise in monetary policy. In other words, we expect career concerns to decline in $FedExp_{i,t}$. In figure 1a we plot the histogram of this variable across all members in our main sample period and in figure 1b we plot the individual evolution of this variable across each member.

The main specification we use in the paper is the following "diff-in-diff" regression:[15]

$$y_{it} = \alpha_i + \delta_t + \beta D(Trans)_t + \eta FedExp_{i,t} + \phi D(Trans)_t \times FedExp_{i,t} + \epsilon_{it} \qquad \text{(DinD)}$$

The main coefficient of interest is the ϕ coefficient on the interaction term. Since career concerns decline with $FedExp_{i,t}$, a positive (negative) ϕ indicates that members with greater career concerns do less (more) of whatever $y_{i,t}$ is measuring. Given the inclusion of time and member fixed effects, the identification comes mostly off those members who served both before and after the change in transparency. For the baseline analysis presented below, we will focus on a sample that uses the first ten years of Alan Greenspan's tenure as chair of the FOMC (1987-1997). In appendix C, we show that results remain robust to alternative sample selections.

Testing the statistical significance of the ϕ coefficient requires us to have a well-estimated variance-covariance matrix. This is particularly a challenge with a fixed-effects panel data model because the data can be autocorrelated, there may be heteroskedasticity by member, and there may be cross-sectional dependence. All of these reduce the actual information content of the analysis and may lead us to overstate the significance of estimated relationships. We use the nonparametric covariance matrix estimator proposed by Driscoll and Kraay (1998). This helps to make our standard errors robust to general forms of spatial and temporal dependence, as well as being heteroskedasticity- and autocorrelation-consistent.

4.3 FOMC transcript data

The y_{it} measures in (DinD) are constructed using FOMC meeting transcripts.[16]. 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. In this paper, we use the set of transcripts from the tenure of Alan Greenspan—August 1987 through January 2006, inclusive, a total of

[14]This information came from online sources and the *Who's Who* reference guides.

[15]For the purposes of the analysis, we treat all staff members as a single homogenous group. So, in meeting t, i indexes all FOMC members plus a single "individual" called staff.

[16]These are available for download from `http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm`

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 directly about about monetary policy we do not use them in our analysis.

The transcripts available from the Fed website need to be cleaned and processed before they can be used for empirical work. We have ensured the text is appropriately read in from the pdf files, and have removed non-spoken text such as footnotes, page headers, and participant lists. 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 took the separate staff statements from appendices and then matched them into the main transcripts. The final dataset contains 46,502 unique interjections along with the associated speaker.

While we estimate topic models on the whole meeting, we focus our analysis on the statements in each meeting that corresponded to the economic situation discussion (FOMC1) and the policy discussion (FOMC2), as described in section 3. To do this, we manually coded the different parts of each meeting in the transcript; FOMC1 and FOMC2 make up around 31% and 26% of the total number of statements.

5 Measuring Communication

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.[17] 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.

[17]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).

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.

5.1 Statistical model

Our text dataset is a collection of D documents, where a document d is a list of words $\mathbf{w}_d = (w_{d,1}, \ldots, w_{d,N_d})$.[18] In our dataset, a document is a single statement, or interjection, by a particular member in a particular meeting. For example, we would have two statements if Alan Greenspan asks a question of staff (the first statement) and a staff member replies (the second statement).

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. In turn, each document is modeled as a distribution over topics. Documents are independently but not identically distributed. Let $\theta_d \in \Delta^K$ be the distribution of topics in document d, where θ_d^k represents the "share" of topic k in document d. In the FOMC context, we imagine θ_d as a choice variable of the policymaker that generates document d.

The statistical process that generates the list of words in document d involves two steps. We dispose of the d subscripts for notational convenience. Imagine a document as composed of N slots corresponding to the N observed words. In the first step, each slot is independently allocated a topic assignment z_n according to the probability vector θ corresponding to the distribution over topics in the document. These topic assignments are unobserved and are therefore latent variables in the model. In the second step, a word is drawn for the nth slot from the topic β_{z_n} that corresponds to the assignment z_n. Given θ and the topics β_k for $k = 1, \ldots, K$, the overall probability of observing the list

[18] Here "word" should more formally be "token" which is not necessarily an English word but rather should be understood as simply an abstract element.

of words corresponding to document d is

$$\prod_{n=1}^{N} \sum_{z_n} \Pr[\, z_n \mid \theta \,] \Pr[\, w_n \mid \beta_{z_n} \,] \tag{1}$$

where the summation is over all possible topic assignments for word w_n.

Computations based on (1) are generally intractable, so direct maximum likelihood approaches are not feasible. Instead, LDA assumes that each θ_d is drawn from a symmetric Dirichlet(α) prior with K dimensions, and that each β_k is drawn from a symmetric Dirichlet(η) prior with V dimensions. Realizations of Dirichlet distributions with M dimensions lie in the M-simplex, and the hyperparameters α and η determine the concentration of the realizations. The higher they are, the more even the probability mass spread across the dimensions. Given these prior probabilities, the probability of document d becomes

$$\int \dots \int \prod_{k=1}^{K} \Pr[\, \beta_k \mid \eta \,] \Pr[\, \theta \mid \alpha \,] \left(\prod_{n=1}^{N} \sum_{z_n} \Pr[\, z_n \mid \theta \,] \Pr[\, w_n \mid \beta_{z_n} \,] \right) d\theta d\beta_1 \dots d\beta_K \tag{2}$$

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 independent. LDA can be extended to model various dependencies across documents.[19] Dynamic topic models allow β_k to evolve over time.[20] 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.

One reason for the popularity of LDA is its ability to consistently estimate topics that appear natural despite having no pre-assigned labels. As we show in section 5.4, it indeed

[19]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).

[20]A distinct issue is whether the distribution over topics in a particular statement is affected by the distribution over topics in previous statements. Rather than explicitly building such dependence directly into the statistical model, we explore it with the influence measure we construct in section 7.

estimates topics that are close to ones economists would generate. The basic intuition for how LDA generates topics relates to the co-occurrence of words in documents. As discussed by Blei (2009), LDA places regularly co-occurring words together into topics because it tends to spread words across few topics to maximise the word probabilities for each given topic—i.e. the $\Pr\left[w_{d,n} \mid \beta_{z_{d,n}} \right]$ term in (1). Another advantage of LDA is that it is a mixed membership model that allows the same word to appear in multiple topics with different probabilities, whereas a standard mixture model would force each word to appear in just one topic. For example, returning to the example above, the word "growth" can appear both in a topic about activity (along with words like "gdp") and in a topic about labor markets (along with words like wage). This flexibility loosens the typical definition of co-occurrence and leads to more accurate descriptions of content.

5.2 Estimation

The parameters of interest of the model are the topics β_k and document-topic distributions θ_d. For estimation, we use the Gibbs 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. The algorithm begins by randomly assigning topics to words, and then updating topic assignments by repeatedly sampling from the appropriate posterior distribution. Full details of the approach are in appendix A.[21]

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 8,000 iterations beginning from 5 different starting values and choose for analysis the chain that achieves the best fit of the data based on its average post-convergence *perplexity*, a common measure of fit in the natural language processing literature.[22] In practice the differences in perplexity across chains are marginal, indicating that the estimates are not especially sensitive to starting values.

5.2.1 Vocabulary selection

Before sampling the model, one must choose which vocabulary will be excluded from the analysis. The reason for this is to ease the computational burden of estimation by

[21] For estimation we adapt the C++ code of Phan and Nguyen (2007) available from `gibbslda.sourceforge.net`.

[22] 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.

removing words which will not be very informative in the analysis. 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 *stop words*, or extremely common words that appear in many documents. Our list is rather conservative, and contains all English pronouns, auxiliary verbs, and articles.[23]

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'. Hence, the words are now most appropriately called "tokens".

We then tabulate the frequencies of all two- and three-token sequences in the data, known as *bigrams* and *trigrams*, respectively. For those that occur most frequently and which have a specific meaning as a sequence, we construct a single token and replace it for the sequence. For example, 'fed fund rate' becomes 'ffr' and 'labor market' becomes 'labmkt'. The former example ensures that our analysis does not mix up "fund" when used as part of the noun describing the main policy instrument and when members discuss commercial banks' funding.

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.[24] 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.

[23]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 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').

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

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

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.

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.

5.2.2 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 a limited number 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.[25] We instead estimate models with $K = 50$ and $K = 70$, which both allows us to relatively easily interpret the topics and to explore the effect of altering the number of topics on the results.[26] In the main body of the text, we report results for $K = 50$.

5.3 Document aggregation

The primary object of interest for the empirical analysis is 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 $k = 1, \ldots, K$. While considering individual statements is useful for estimating the LDA model, for estimating (DinD) we are more interested in measures of the form $\widehat{\theta}_{i,t.s,j}$, where i indexes an FOMC member, t indexes a meeting, and s indexes a meeting section (FOMC1 or FOMC2). We detail how we obtain estimates for aggregate

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

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

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 $j \in \{4050, 4100, \ldots, 8000\}$. Hence, in the language of MCMC estimation, we run 4,000 iterations after a burn-in of 4,000 iterations, and apply a thinning interval of 50. Based on perplexity scores, all the chains we estimate converge at or before the 4,000th iteration. There are several further measures of communication we use in section 6 derived from $\widehat{\theta}_{i,t.s,j}$. In each case, they are also computed at each relevant iteration and then averaged.

5.4 Estimated topics

In appendix B we report the top ten tokens in each topic, but here discuss a handful to give a sense of the kind of content that LDA estimates. LDA is an unsupervised learning algorithm, and so produces no meaningful topic labels. Any attribution of meaning to topics requires a subjective judgement on the part of the researcher. Most of the empirical results depend only on mild such judgements, but it is still important that the topics are reasonable in the context of macroeconomics.

An obvious place to start is to examine discussion of inflation. A single topic—topic 25—gathers together many of the terms macroeconomists associate with inflation. Figure 2 represents the topic with a word cloud in which the size of the token represents the probability of its appearing in the topic.[27] The dominant token is "inflat" which captures words relating to inflation, but there are others like "core", "cpi", etc. Given recent events, also of interest is topic 38 (figure 3), which collects together terms relating to banking and finance more generally. There are also topics on consumption and investment (figure 4) and productivity (5) which, as we show in section 8, predict policy outcomes.

So far the topics we have displayed relate to obvious economic themes, but there are also quite a few topics that do not. We call these topics *discussion* as opposed to *economics* topics, and have classified each topic into one of the two categories. This is the main subjective labeling exercise we use in the analysis. In the 50-topic model we analyze, there are 30 economics topics and 20 discussion topics. Discussion topics comprise both topics made up of words that are used in conversation to convey meaning when talking about economics topics, and some topics which are pure conversational words. For example, there is a topic which just picks up the use of other members' names as well as the voting roll call (figure 6); and the five most likely tokens in topic 49 (figure 7) are 'say', 'know', 'someth', 'all', and 'can' which can be used in general conversation regardless of what specific topic is being discussed. But a few of the other discussion topics may also be informative about the behaviour of FOMC members such as the topic

[27]The use of a word cloud is purely for illustrative purposes and the clouds play no role in the analysis; the precise probability distribution over tokens for each topic is available on our websites.

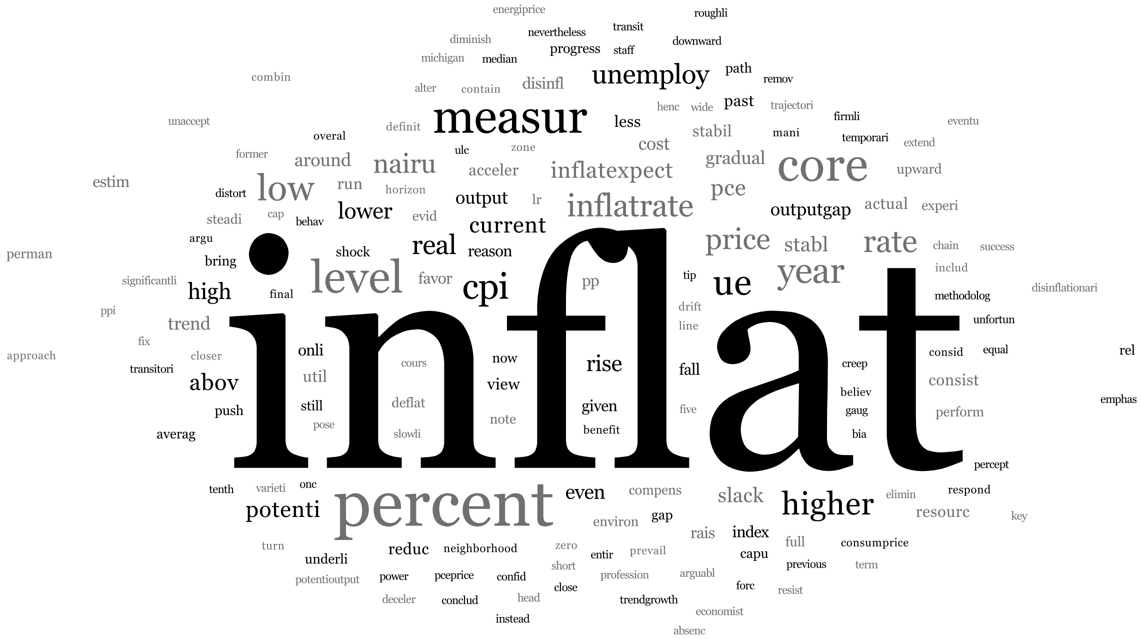


Figure 2: Topic 25—"Inflation"

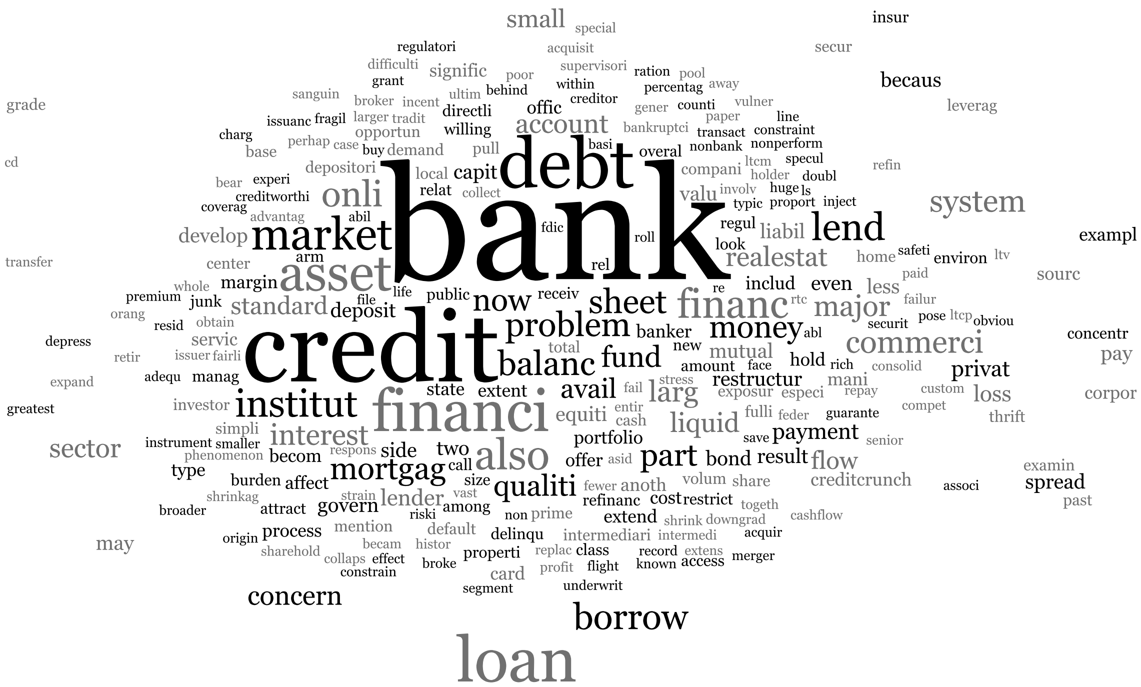


Figure 3: Topic 38—"Banking"

Notes: Each word cloud represents the probability distribution of words within a given topic; the size of the word indicates its probability of occurring within that topic.

Figure 4: Topic 23—"Consumption and Investment"

Figure 5: Topic 29—"Productivity"

Notes: Each word cloud represents the probability distribution of words within a given topic; the size of the word indicates its probability of occurring within that topic.

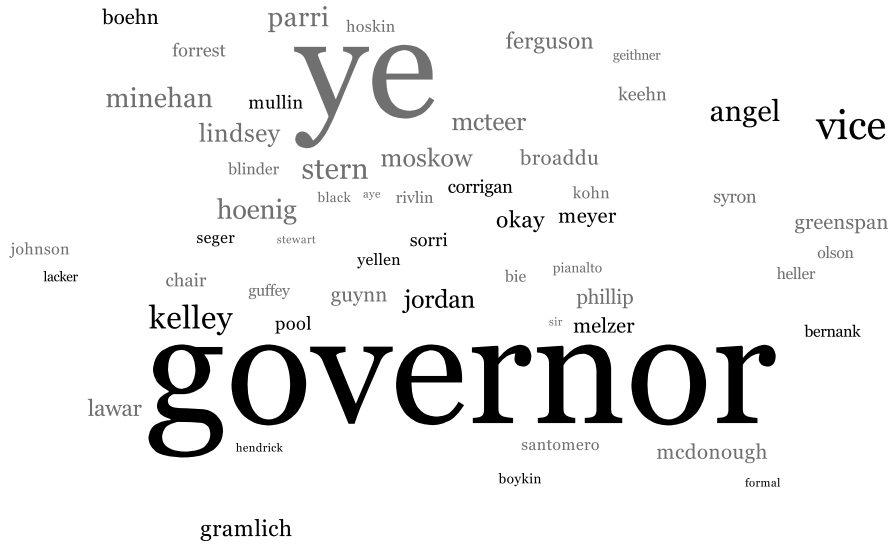


Figure 6: Topic 27—"Discussion topic: FOMC Names"

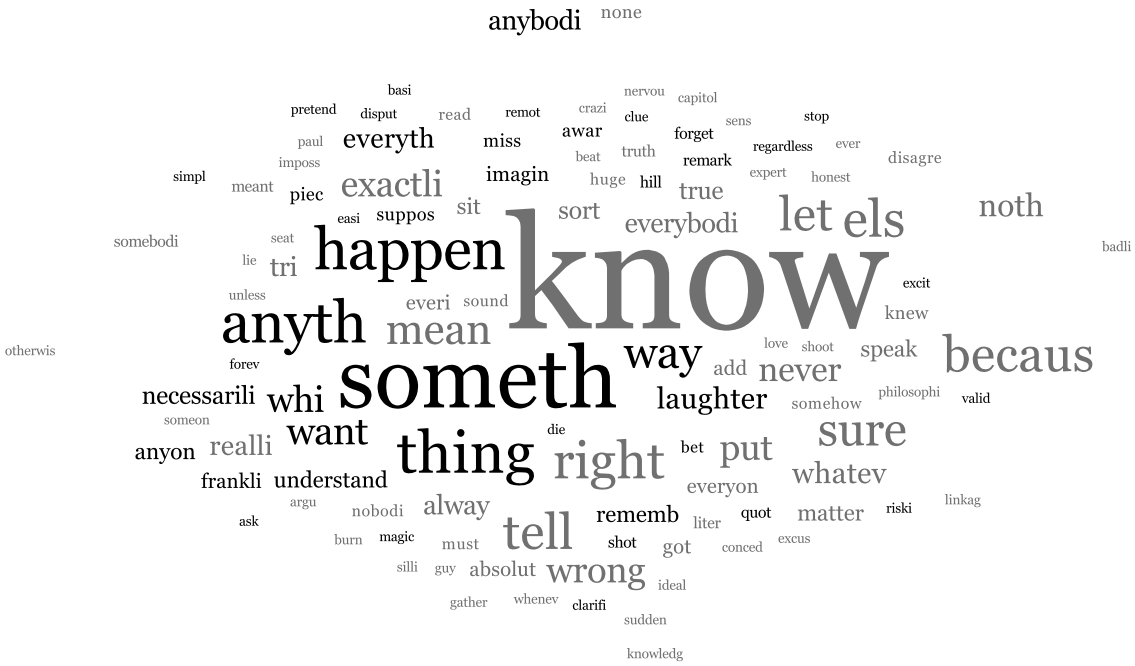


Figure 7: Topic 49—"Discussion topic: General terms"

Notes: Each word cloud represents the probability distribution of words within a given topic; the size of the word indicates its probability of occurring within that topic.

containing terms relating to discussions of data and also one relating to discussions of staff materials; we return to discussing these topics in more detail in section 6.

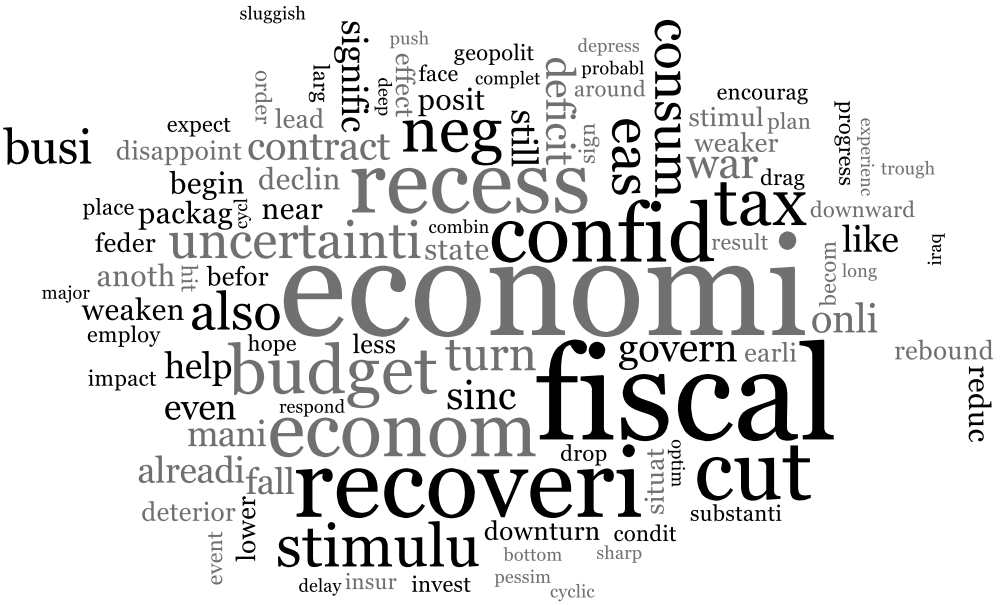
5.5 Connecting topics to external events

A common approach for assessing the quality of the output of machine learning algorithms is to validate them against external data. Since we do not rely heavily on specific topic labels, such an exercise is not crucial for interpreting our results, but for interest we have explored the relationship of the estimated topics to the recently developed uncertainty index of Baker, Bloom, and Davis (2013) (BBD hereafter). This index picks up the public's perceptions of general risk as well as expiring fiscal measures. It is also methodologically related to our data in that the primary input for the index is text data from the media, albeit measured differently (via the number of articles per day that contain a set of terms the authors select).

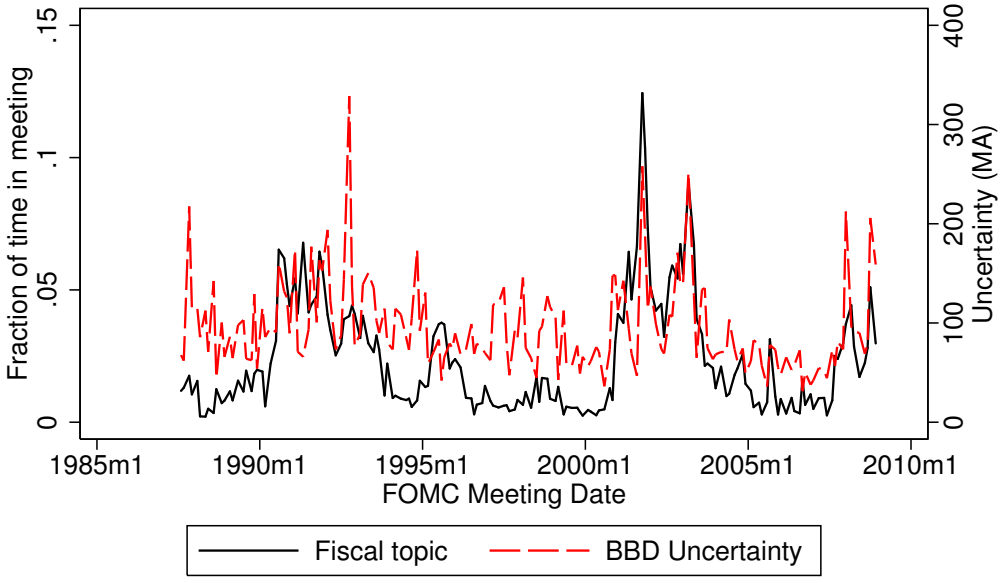
Figure 8 displays the estimated topic most associated with fiscal issues, and plots the amount of time the FOMC as a whole spends on it against the BBD index.[28] The relationship between BBD-measured uncertainty and FOMC attention towards fiscal matters is quite strong, with both notably spiking during times of war and recession. Figure 9 displays the topic most associated with risk and uncertainty and also plots the attention it received during FOMC meetings against the BBD index. While the two series co-move, it is particularly noteworthy that the estimates suggest that in the run-up to the financial crisis in 2007 the market was not yet concerned with risk while the FOMC was increasingly discussing it.

Finally, the estimates pick up a topic related to central bank communication that appears regularly in meetings to capture discussion of statements and previous minutes. Its associated word cloud is in figure 10a. This topic is useful to check whether the decision to reveal the transcripts was surprising. As we argue for our natural experiment, FOMC members only learned of the transcripts in October 1993 and discussed the right policy to deal with their release at the start of the meeting in November 1993. If it were indeed a big surprise, one would expect there to be more than usual discussion of issues of communication. Figure 10b shows that during a typical meeting FOMC members might spend 2% of their time on this topic, and in an unusual meeting—perhaps discussing a particularly tricky statement—up to 8% of their time. By contrast, in November 1993 the FOMC spent over 20% of the meeting discussing the issue of transparency and transcripts being made public. We are therefore comfortable interpreting the publication of transcripts as a genuine surprise.

[28] The distributions for the out-of-sample years coinciding with Ben Bernanke taking over as Chairman are estimated through the querying procedure discussed in appendix A.



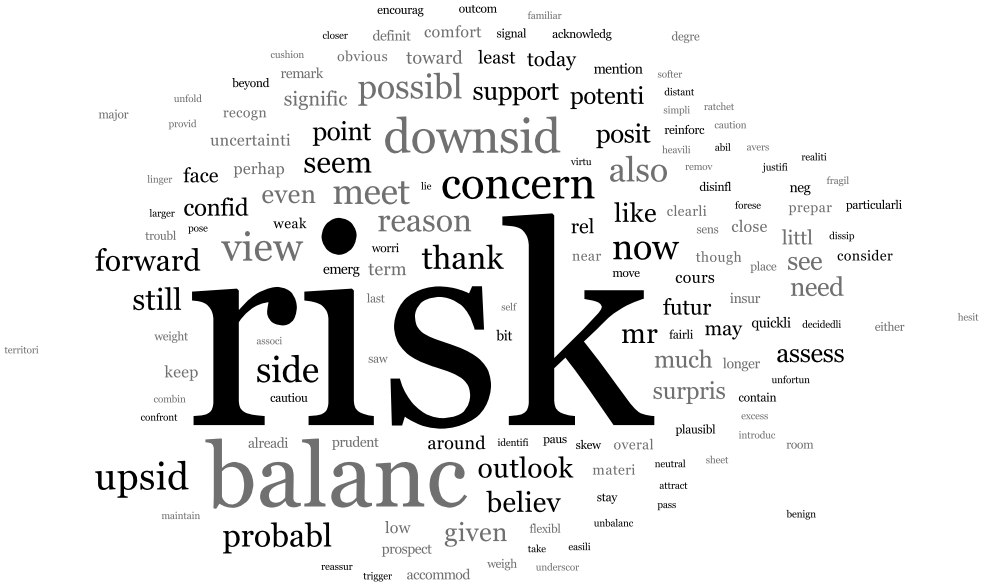
(a) Topic 45—"Fiscal"



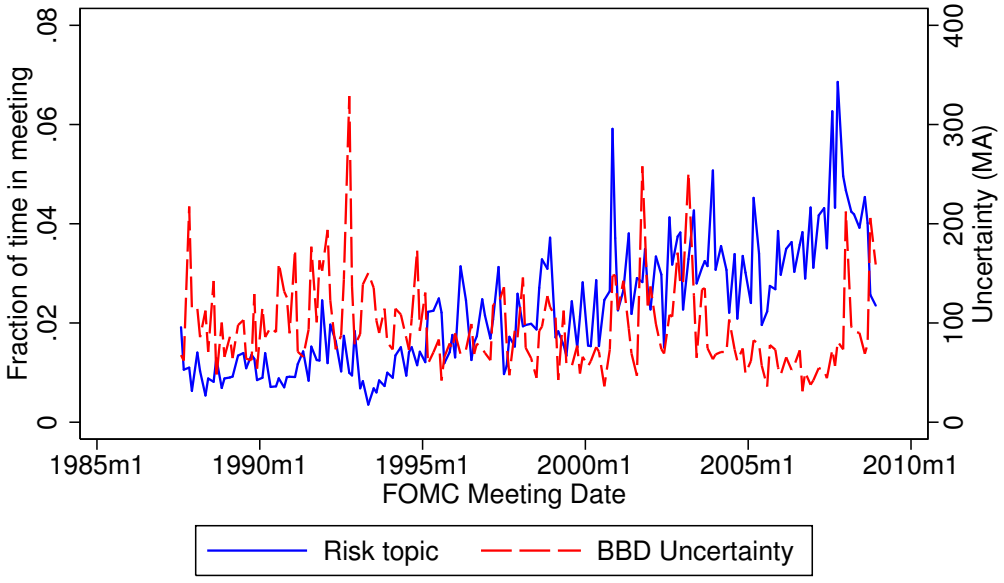
(b) BBD uncertainty and discussion of topic 45

Figure 8: BBD uncertainty measure and FOMC attention to fiscal issues

Notes: The word cloud (top) represents the probability distribution of words within a given topic. The time-series (bottom) captures the time allocated to that topic in each meeting.



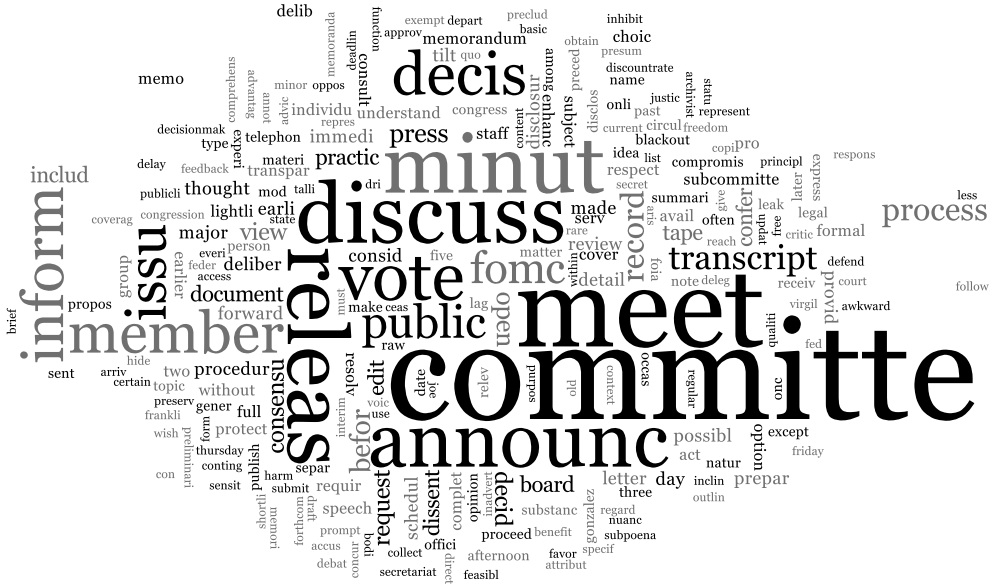
(a) Topic 40—"Risk"



(b) BBD uncertainty and discussion of topic 40

Figure 9: BBD uncertainty measure and FOMC attention to risk

Notes: The word cloud (top) represents the probability distribution of words within a given topic. The time-series (bottom) captures the time allocated to that topic in each meeting.



(a) Topic 6—"Central Bank Communication"

(b) Discussion of topic 6 across meetings

Figure 10: FOMC attention to communication: surprised by transparency revelation?

Notes: The word cloud (top) represents the probability distribution of words within a given topic. The time-series (bottom) captures the time allocated to that topic in each meeting.

6 Empirical Results

We now present the estimates of the econometric models in section 4 using numerous measures of communication. The first are derived from the token counts described in section 5. After documenting how these shifted with transparency, we study how the content of statements changed using various measures constructed from the LDA model.

6.1 Transparency and basic language counts

We first use token counts to judge whether there were substantive changes in deliberation after transparency. To begin, we estimate (DIFF) to compare meeting-level aggregates before and after transparency. The regressions include meeting and other time-specific controls (but, obviously, no time fixed effects), and are estimated separately on the discussion of the economic situation (FOMC1) and of policy (FOMC2).

Table 2a shows the change in FOMC1. After transparency, there are more words delivered in fewer statements, resulting in more words per statement. We interpret the drop in statements as reflecting a reduction in back-and-forth dialogue, since this would generate many statements as the debate bounced from member to member. There are also significantly fewer questions. These simple counts paint a picture of FOMC members coming to the meeting with longer, more scripted views on the economy, and being somewhat less likely to question the staff and their colleagues during the discussion.

Table 2b shows the change in FOMC2. While the change in the number of words and sentences is not statistically significant, there are dramatic effects in the rest of the measures. The average number of questions and statements both drop by around 35% and the number of words per statement increases by nearly 40%. This indicates a stark shift away from a dynamic, flowing discussion towards one in which members share their views on policy in one long statement, and then disengage from their colleagues.

Since the results in tables 2a and 2b are based on the (DIFF) specification, it is unclear whether one can attribute the observed changes to career concerns or to some other factor that shifted near November 1993. To link the results to career concerns, we estimate (DinD) to examine which changes are more pronounced for members with less experience in the Fed. The results are in table 3 (which presents the results for FOMC1 and FOMC2 in a single table but covers a reduced number of variables). The key coefficient is that estimated for the interaction term between the transparency dummy and the Fed experience variable. Recall that since career concerns decline with experience, the direction of the effect of career concerns is opposite in sign to the estimated coefficient. The main result is that in FOMC1 all members make statements of similarly increased length, but that in FOMC2 less experienced members are particularly inclined to opt out of debate in the sense that they make significantly fewer interjections and ask fewer

Table 2: The effect of transparency on count measures of deliberation—meeting level

(a) Economic situation discussion

Main Regressors	(1) Total Words	(2) Statements	(3) Questions	(4) Sentences	(5) Words/Statement
D(Trans)	1,005**	-20.1***	-5.62**	67.7***	42.4***
	[0.038]	[0.007]	[0.044]	[0.009]	[0.001]
Serving FOMC members	375	-0.22	-0.25	21.9*	1.32
	[0.101]	[0.944]	[0.849]	[0.061]	[0.824]
D(NBER recession)	487	-13.9	-5.35	5.89	29.8
	[0.394]	[0.173]	[0.271]	[0.846]	[0.172]
D(2 Day)	720*	20.3**	8.87***	52.4**	-31.7***
	[0.079]	[0.047]	[0.008]	[0.022]	[0.006]
Uncertainty(t-1)	1.01	-0.052*	-0.0086	0.026	0.083**
	[0.659]	[0.095]	[0.438]	[0.825]	[0.035]
Constant	230	97.0	29.2	-4.22	68.5
	[0.955]	[0.102]	[0.243]	[0.984]	[0.540]
R-squared	0.314	0.166	0.167	0.344	0.348
Lag Dep. Var?	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	79	79	79	79	79

(b) Policy discussion

Main Regressors	(1) Total Words	(2) Statements	(3) Questions	(4) Sentences	(5) Words/Statement
D(Trans)	283	-51.6***	-16.4***	-12.5	51.8***
	[0.672]	[0.001]	[0.000]	[0.715]	[0.000]
Serving FOMC members	-184	-1.15	-1.35	-8.75	-4.19
	[0.543]	[0.785]	[0.262]	[0.545]	[0.345]
D(NBER recession)	-401	-5.29	-5.04	-28.9	-1.67
	[0.703]	[0.829]	[0.539]	[0.628]	[0.785]
D(2 Day)	1,632**	8.33	5.77	75.0**	12.7
	[0.013]	[0.495]	[0.165]	[0.023]	[0.121]
Uncertainty(t-1)	-0.27	-0.035	-0.020*	-0.014	0.013
	[0.914]	[0.429]	[0.079]	[0.909]	[0.613]
Constant	9,574*	130	51.5**	498*	125
	[0.093]	[0.114]	[0.027]	[0.072]	[0.114]
R-squared	0.085	0.179	0.177	0.071	0.468
Lag Dep. Var?	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	79	79	79	79	79

Notes: These tables report the results of estimating (DIFF) on variables related to meeting-level counts of measures of the discussion. The upper (lower) table reports the results for FOMC1 (FOMC2). Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Table 3: Count measures—member level

Main Regressors	(1) Total Words	(2) Statements	(3) Questions	(4) Total Words	(5) Statements	(6) Questions
D(Trans)	-1,481	-17.8*	-1.78	241***	3.15	1.62*
	[0.384]	[0.094]	[0.676]	[0.009]	[0.198]	[0.093]
Fed Experience	250	2.61	0.23	464**	-6.55	-3.68**
	[0.331]	[0.106]	[0.721]	[0.016]	[0.137]	[0.050]
D(Trans) x Fed Experience	-0.54	0.033	0.0052	-2.44	0.13***	0.047***
	[0.761]	[0.226]	[0.663]	[0.359]	[0.007]	[0.004]
Constant	-876	-11.8	-0.11	-5,363**	79.9	44.0**
	[0.559]	[0.206]	[0.976]	[0.019]	[0.126]	[0.048]
Number of groups	38	38	38	38	38	38
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Within Meeting	FOMC1	FOMC1	FOMC1	FOMC2	FOMC2	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	1449	1449	1449	1432	1432	1432
% Rookie effect	-	-	-	-	-36.9	-67.5

Notes: This table reports the results of estimating (DinD) on variables related to count measures of the discussion. Where the difference in difference is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

questions.

In order to quantify the economic importance of the estimated effect of career concerns, whenever the coefficient on the interaction term is significant, we report what we term the *rookie effect*. This measures by how much a member with one year of Fed experience differs in terms of the dependent variable compared to a member with 20 years' experience, where the difference is expressed as a percentage of the pre-transparency average for FOMC members. (The choice of one and 20 years' experience is driven by the distribution of experience as presented in figure 1a.) Take for example the significant coefficient of 0.13 in column (5) of table 3. This indicates that each additional year of experience leads members to make an additional 0.13 statements per meeting during FOMC2. Then a member with 20 years' experience will make $19 \times 0.13 = 2.47$ more statements than a member in the first year. Moreover, 2.47 represents 36.9% of the pre-transparency mean number of statements made in FOMC2 by all members. We report the rookie effect as -36.9 in this case. Similarly, the fall in questions represents a 67.5% decline relative to the pre-transparency average of questions by members in FOMC2.

The view that emerges is that all members come into the meeting after transparency with long, scripted statements that they share with their colleagues in FOMC1. Then, when the time comes to debate policy in FOMC2, inexperienced members stay relatively

silent and let their more senior colleagues drive the debate.

6.2 Transparency and statement content

As explained in section 5.4, we label each estimated topic as economic or discussion. The first measure of statement content we construct from the LDA output is the fraction of time devoted to economics topics. This labeling also allows one to define a conditional probability distribution over economics topics for FOMC statements. The second measure of statement content is a Herfindahl concentration index applied to this conditional distribution. This index measures the scope of the discussion: higher values indicate a narrow discussion, while lower values indicate a broader discussion.

We again begin the analysis by looking at the meeting-level response of these two measure to transparency by way of estimating (DIFF). Table 4a contains the results. There is a marked shift in both FOMC1 and FOMC2 towards more attention towards economic topics, with a 5.5 and 3.4 percentage point effect in the respective sections. There is not a measure change in concentration at the aggregate level. Table 4b shows that these meeting-level results mask heterogeneity at the individual level. First, the increase in attention to economics topics during the policy discussion is driven more by inexperienced members. The could either be driven by their engaging less in back and forth debate—and therefore using less conversational speech patterns—or staying more focused on substantive issues. The pattern for the topic concentration index moves in opposite directions during FOMC1 and FOMC2. Inexperienced members come into the meeting and discuss more topics on average compared to experienced members when analyzing the economic situation, but when the meeting moves to policy debate inexperienced members limit their attention to fewer topics. This is consistent with inexperienced members bringing additional information into the meeting in the form of a more diverse statement in FOMC1, but then not engaging with their colleagues in FOMC2 since such engagement would force them to touch on viewpoints other than their own.

In order to push the idea that inexperienced members bring additional information to meetings after transparency, we explore the behavior of the topic that relates to discussion of quantitative data (topic 7, figure 11a), and also the one that relates to discussion of staff materials (topic 22, figure 11b). One of the primary ways in which FOMC members can prepare for meetings is to gather and study data to provide evidence for their views. Both of these topics indicate introducing such evidence into the discussion—topic 7 is made up of words that one discussing quantitative data would use, while topic 22 is made up of words that one engaging with staff briefings and presentations would use. A member without career concerns who spent little time preparing for meetings (nor paying attention to colleagues during them) would most likely not discuss their views

Table 4: Economics focus and concentration of topics discussed

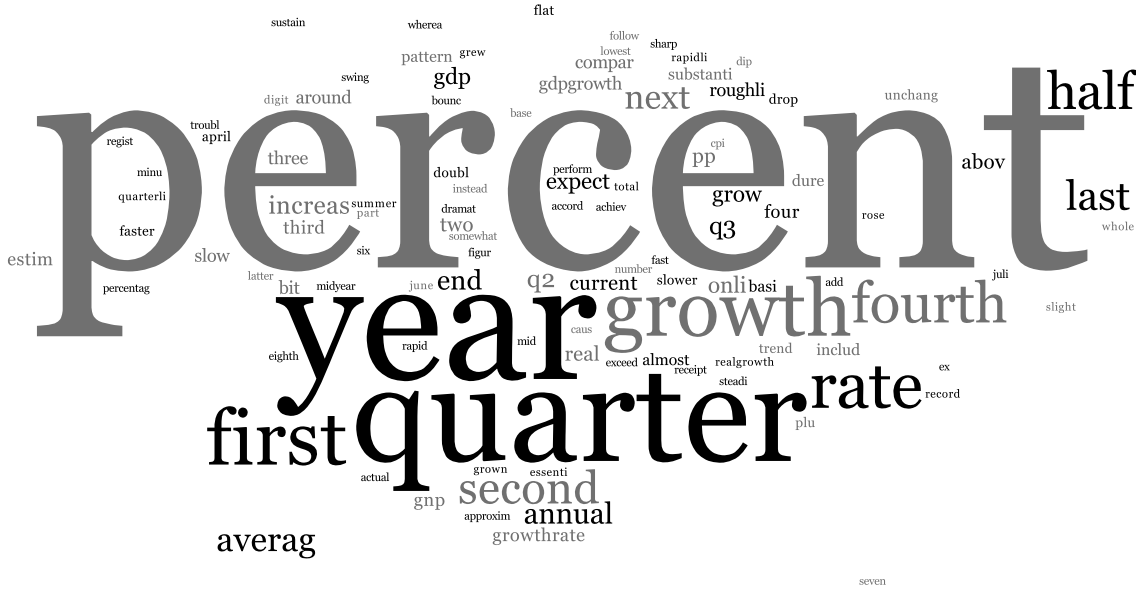
(a) Meeting level

Main Regressors	(1) Economics	(2) Economics	(3) Herfindahl	(4) Herfindahl
D(Trans)	0.055***	0.034***	0.00042	-0.0037
	[0.000]	[0.001]	[0.834]	[0.101]
Serving FOMC members	-0.0018	-0.0055	-0.0012	-0.0016
	[0.498]	[0.236]	[0.153]	[0.158]
D(NBER recession)	0.0068	0.014	0.0016	-0.0015
	[0.530]	[0.461]	[0.386]	[0.667]
D(2 Day)	-0.016**	-0.034***	-0.0053***	-0.0054**
	[0.013]	[0.001]	[0.009]	[0.017]
Uncertainty(t-1)	9.5e-06	0.000024	-6.4e-06	-0.000017
	[0.654]	[0.679]	[0.478]	[0.131]
Constant	0.44***	0.62***	0.080***	0.10***
	[0.000]	[0.000]	[0.000]	[0.000]
R-squared	0.764	0.281	0.141	0.125
Lag Dep. Var?	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC2	FOMC1	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	79	79	79	79

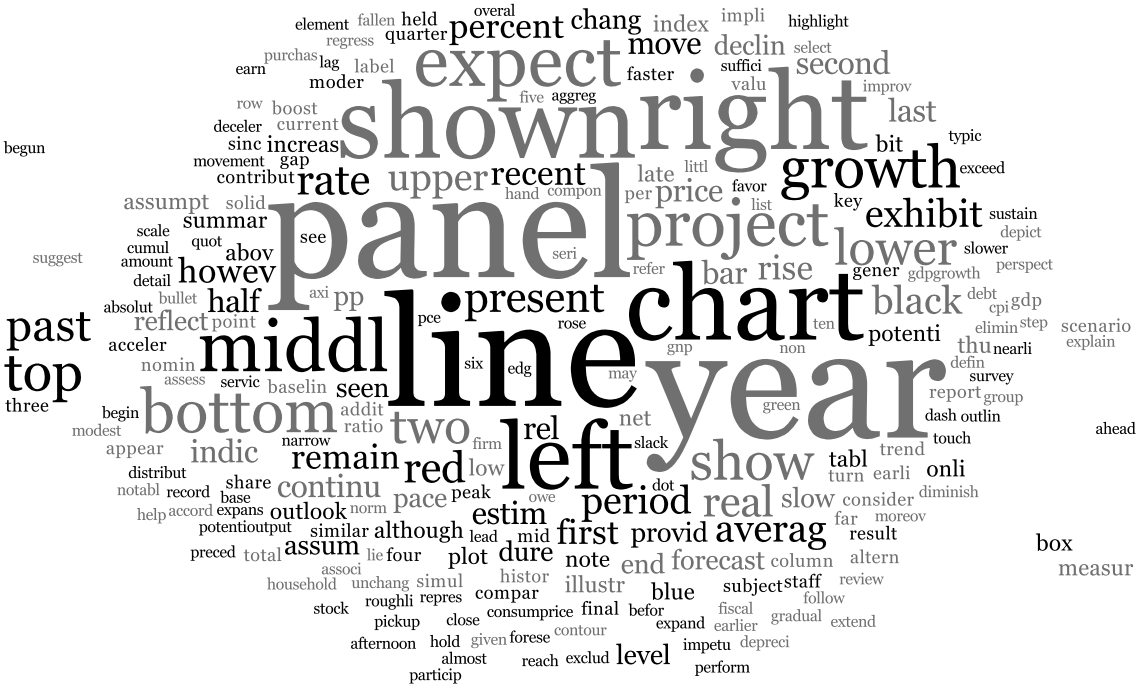
(b) Member level

Main Regressors	(1) Economics	(2) Economics	(3) Herfindahl	(4) Herfindahl
D(Trans)	1.43***	-0.041	-0.21***	-0.10
	[0.003]	[0.908]	[0.002]	[0.323]
Fed Experience	-0.22***	0.0081	0.035***	0.017
	[0.003]	[0.878]	[0.001]	[0.267]
D(Trans) x Fed Experience	0.00021	-0.0014**	0.00071**	-0.00037***
	[0.622]	[0.015]	[0.035]	[0.000]
Constant	1.86***	0.56*	-0.12*	-0.030
	[0.000]	[0.073]	[0.050]	[0.737]
Number of groups	38	38	38	38
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Within Meeting	FOMC1	FOMC2	FOMC1	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	1449	1431	1449	1431
% Rookie effect	-	4.7	-13.0	10.4

Notes: The upper table reports the results of estimating (DIFF) and the lower table reports the results of estimating (DinD) on variables related to LDA measures of the discussion. Where the difference in difference is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.



(a) Topic 7—"Data Indicators"



(b) Topic 22—"Staff Material Discussion"

Figure 11: Topics relating to the extent of preparation

Notes: Each word cloud represents the probability distribution of words within a given topic; the size of the word indicates its probability of occurring within that topic.

using specific references to relevant data, while one who had done their homework would likely bring into the meetings a dossier of background information on which to draw. FOMC1 is the important section for determining the relevance of this effect, since during it members mainly read from prepared texts.

Tables 5a and 5b present meeting-level and member-level evidence, respectively, on the allocation of attention to these evidence topics in FOMC1. While there is no meeting-level change, at the member level we indeed find evidence that rookie members are more likely to discuss these topics, and especially topic 7, during their analysis of the economy. Members with one year of experience increase their allocation to topics 7 and 22 relative to those with 20 years' experience by an amount equivalent to around 44% and 11%, respectively, of the pre-transparency average allocation. There are also changes in FOMC2, with less overall discussion of topic 7, and relatively more coverage of topic 22 by older members. Since FOMC2 is more extemporaneous than FOMC1, a theory of greater preparation has no clear prediction on its content, so we simply note these further empirical results associated with transparency.

Our final measure of content compares the statements of each FOMC member to those of Alan Greenspan, who is clearly a focal member during the sample. One obvious way that FOMC members might engage in herding is to mimic the Chair's views and bring up similar topics; anti-herding would involve the opposite behavior. Let $\boldsymbol{\chi}_{i,t,s}$ denote i's conditional probability distribution over economics topics in section s of meeting t; we are interested in comparing the similarity of $\boldsymbol{\chi}_{i,t,s}$ with $\boldsymbol{\chi}_{G,t,s}$, where G is Greenspan's speaker index. Although $\boldsymbol{\chi}_{i,t,s}$ has thirty dimensions, members almost certainly discuss far fewer topics in each section of each meeting. Hazen (2010) compares several ways of computing the similarity of documents estimated by LDA, and concludes that the dot product substantially outperforms other standard measures like cosine similarity and Kullback-Leibler divergence in conversational speech data when each statement is composed of a limited number of topics relative to K. Accordingly we define

$$SG_{i,t.s} = \boldsymbol{\chi}_{i,t,s} \cdot \boldsymbol{\chi}_{G,t,s} \tag{3}$$

as the similarity between member i and Greenspan. If i and G each discuss a single topic, $SG_{i,t.s}$ is the probability they discuss the same topic. More generally, a higher value of $SG_{i,t.s}$ indicates a greater overlap in topic coverage.

Table 6 presents the results of estimating DinD with $SG_{i,t.s}$ as the dependent variable. The main result is that after transparency, inexperienced members speak more like Greenspan in FOMC2—the negative and significant coefficient on the interaction indicates that the conditional topic distributions of experienced members diverge more from Greenspan's; members with a single year of experience become relatively closer to

Table 5: Discussion topics relating to data indicators

(a) Meeting level

Main Regressors	(1) Data Topic (7)	(2) Data Topic (7)	(3) Figures Topic (22)	(4) Figures Topic (22)
D(Trans)	0.0022	-0.0072***	-0.0044	-0.0031
	[0.388]	[0.005]	[0.255]	[0.280]
Serving FOMC members	0.00096	0.00079	-0.0017	-0.00064
	[0.416]	[0.484]	[0.271]	[0.558]
D(NBER recession)	-0.0024	-0.0023	-0.00093	-0.0028
	[0.500]	[0.509]	[0.747]	[0.514]
D(2 Day)	-0.0015	0.0026	0.074***	-0.0030
	[0.549]	[0.194]	[0.000]	[0.175]
Uncertainty(t-1)	-0.000022**	-0.000021**	-0.000012	1.5e-07
	[0.042]	[0.012]	[0.577]	[0.989]
Constant	0.027	0.0043	0.041	0.023
	[0.221]	[0.826]	[0.183]	[0.291]
R-squared	0.088	0.262	0.910	0.058
Lag Dep. Var?	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC2	FOMC1	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	79	79	79	79

(b) Member level

Main Regressors	(1) Data Topic (7)	(2) Data Topic (7)	(3) Figures Topic (22)	(4) Figures Topic (22)
D(Trans)	-0.12	0.020	0.0090	-0.035
	[0.239]	[0.729]	[0.743]	[0.267]
Fed Experience	0.019	-0.0038	-0.0011	0.0048
	[0.205]	[0.675]	[0.797]	[0.310]
D(Trans) x Fed Experience	-0.00065***	-0.000098	-0.000088**	0.00011*
	[0.002]	[0.103]	[0.035]	[0.078]
Constant	-0.089	0.041	0.017	-0.015
	[0.325]	[0.431]	[0.488]	[0.594]
Number of groups	38	38	38	38
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Within Meeting	FOMC1	FOMC2	FOMC1	FOMC2
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	1449	1431	1449	1431
% Rookie effect	43.9	-	11.7	-14.2

Notes: The upper table reports the results of estimating (DIFF) and the lower table reports the results of estimating (DinD) on variables related to LDA measures of the discussion. Where the difference in difference is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Table 6: Similarity—member level

Main Regressors	(1) SG(it)	(2) SG(it)
D(Trans)	0.00085	-0.054*
	[0.987]	[0.089]
Fed Experience	-0.0020	0.0090*
	[0.805]	[0.067]
D(Trans) x Fed Experience	-0.00015	-0.00023***
	[0.452]	[0.000]
Constant	0.056	-0.0084
	[0.223]	[0.767]
Number of groups	38	38
Member FE	Yes	Yes
Time FE	Yes	Yes
Within Meeting	FOMC1	FOMC2
Sample	87:08-97:09	87:08-97:09
Obs	1449	1431
% Rookie effect	-	9.9

Notes: This table reports the results of estimating (DinD) on LDA derived measures of the discussion. Where the difference in difference is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Greenspan than members with 20 years of experience and the size of the effect is about 10% of the pre-transparency average distance form Greenspan. The lack of effect in FOMC1 is perhaps to be expected. Greenspan's main statement in most meetings in the sample occurred at the beginning of FOMC2. He often spoke very little in FOMC1, whereas members would have only known his stance unambiguously once they began speaking in FOMC2. As discussed in section 2, the theoretical predictions of career concerns models are consistent with both herding and anti-herding. So from a model testing viewpoint, it is notable that herding appears to be the more relevant effect more inexperienced members.[29]

7 Overall Effect of Career Concerns

Ultimately we are interested in linking the effects of transparency to the theoretical framework provided by career concerns models. As discussed in section 2, two key effects of transparency are an increase in discipline and an increase in conformity/non-conformity. The results on inexperienced members' distance from Greenspan together with their disengaging more during debate in FOMC2 clearly point towards fitting in being more

[29]Chevalier and Ellison (1999) have drawn the same conclusion from mutual fund managers.

important than standing out, so we focus on conformity rather than non-conformity.

Table 7: Evidence for career concerns

Discipline	**Conformity**
↑ economics topic coverage in FOMC1	↓ statements in FOMC2
↑ references to data topics in FOMC1	↓ questions in FOMC2
	↓ distance from Greenspan in FOMC2
	↓ economics topic coverage in FOMC2
↑ economics topic percentage in FOMC2	

Table 7 categorizes the main difference-in-differences results from the previous section in terms of their support for discipline or conformity.[30] On the one hand, inexperienced members use the opening part of the meeting (FOMC1) to discuss more economics topics, and when they do so they refer to quantitative evidence more often. Then in FOMC2 they spend more time discussing economics as opposed to discussion topics. This effect is ambiguous to classify since it might reflect their talking about less fluff, but also might reflect less engagement in the discussion. So, we assign this finding to both columns. On the other hand, support for conformity comes from fewer statements and questions in FOMC2; sticking to a narrow agenda of economics topics in FOMC2; and increased mimicry of Greenspan in FOMC2. Of course, ours is not a structural exercise and for each individual result other interpretations might be possible. Taken as a whole, though, we argue that the set of facts we have uncovered can be interpreted plausibly and cleanly through the lens of career concerns.

The effects of discipline and conformity on the informativeness of FOMC members' expressed views go in opposite directions. With discipline, members spend additional time gathering information before meetings, which should tend to increase informativeness. With conformity, members are more likely to avoid expressing their true views, which should tend to decrease informativeness. The rest of this section seeks to determine what the overall effect on informativeness is after the shift to transparency by measuring changes in influence.

7.1 Influence

The basic motivation behind our measurement of influence is the following: as what member i communicates in his or speech becomes more informative, i's colleagues should incorporate i's topics more in their own speech.

This idea is analogous to the measurement of academic impact. A paper is influential if it is cited by other influential papers. The potential circularity of this definition is

[30] In appendix section C we show that the main results are robust to various alternative sample selections. We also show that the main results do not differ by President / Governor splits. And we carry out a placebo tests on the transparency change.

handled by using recursive centrality measures, the most common of which is eigenvector centrality, which is used in a large number of domains (see Palacios-Huerta and Volij (2004) for a discussion and an axiomatic foundation). For instance, PageRank, the algorithm for ranking, builds on eigenvector centrality. Recursive impact factor measures are increasingly common in academia.

In our set-up, the influence measure is built in two steps. First, we construct a matrix of binary directed measures (how i's statements relate to j's future statements). Second, we use this matrix to compute eigenvector centrality.

For the first step, we use the same similarity measure introduced in section 6 for measuring similarity to Greenspan. Let $\mathbf{W}_t$ be a within-meeting influence matrix with elements $\mathbf{W}_t(i,j) = \boldsymbol{\chi}_{i,t,FOMC1} \cdot \boldsymbol{\chi}_{j,t,FOMC2}$. In words, we say member i influences j within a meeting when i's speaking about a topic in FOMC1 leads to j's being more likely to speak about it in FOMC2.

For the second step, use $\mathbf{W}_t$ to obtain a Markov matrix $\mathbf{W}'_t$ by way of the column normalization $\mathbf{W}'_t(i,j) = \frac{\mathbf{W}_t}{\sum_j \mathbf{W}_t(i,j)}$. From there, we measure the within-meeting influence of member i in meeting t as the ith element of the (normalized) eigenvector associated with the unit eigenvalue of $\mathbf{W}'_t$. Denote this value by w_{it}. Loosely speaking, w_{it} measures the relative contribution of member i's FOMC1 topics in shaping the topics of all members in FOMC2. Since Alan Greenspan's views are potentially dominant for shaping policy, another quantity of interest is i's influence just on Greenspan $w^G_{it} \equiv w_{it} \times \mathbf{W}'_t(i,G)$, where G is Greenspan's speaker index.

Some observers—notably Meyer (2004)—have argued that in fact influence *across* meetings is more important than influence within meetings.[31] We therefore define an across-meeting influence matrix $\mathbf{A}_t$ where $\mathbf{A}_t(i,j) = \boldsymbol{\chi}_{i,t,FOMC2} \cdot \boldsymbol{\chi}_{j,t+1,FOMC2}$ and arrive at an overall influence measure a_{it} and a Greenspan-specific influence measure a^G_{it} in a manner identical to that described for the within-meeting measures. We focus on the effect of FOMC2 in meeting t on FOMC2 in meeting $t+1$ since influence on policy is the main quantity of interest.

Before turning to the diff-in-diff analysis, we provide some statistics on the inter-meeting influence measures that we calculate. In table 8, we present a ranking of members by their overall influence (left panel) and their influence on Greenspan (right panel).

[31] 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.

While the table presents the average value of influence for each member, this can be misleading because the influence measures are relative and so the average depends on the period during which the member served. We try to control for the meeting-specific time variation by running a regression of each influence measure in the table on time and member fixed effects ($a_{it}/a_{it}^G = \alpha_{it} + \delta_t + \epsilon_{it}$). We report, and base the ranking on, the member-fixed effects from this regression.

This table shows the cross-sectional variation in the influence measures over the entire tenure of Chairman Greenspan. Members who are highly influential overall tend to exhibit influence over Chairman Greenspan; the Spearman rank correlation between the two rankings is 0.76. However, there are some members who exhibit greater influence over the committee overall than they do over Chairman Greenspan (and vice versa). Interestingly, while Chairman Greenspan is a very good predictor of what Chairman Greenspan will subsequently talk about, his successor Ben Bernanke is also measured to have been influential over what Chairman Greenspan would talk about in the future. Perhaps surprisingly Chairman Greenspan is found to exhibit little influence over the overall FOMC. While we leave a deeper investigation of the reasons that some members are more influential than others for future work, one potential reason for this might be that members tend to use their statements in FOMC2 to reinforce or dispute the proposed policy strategy of Chairman Greenspan by talking about different topics to those which he brought up; because of persistence in what is discussed, this is reflected even in the inter-meeting influence measures. Moreover, in his role as Chairman, Governor Greenspan may discuss some topics every meeting which, in many meetings, are not discussed by others and this would negatively affect his overall influence.

We now turn to the diff-in-diff analysis of influence. Table 9 reports member-level results on the change in influence associated with transparency.[32] Within meetings there is no overall effect, but the inexperienced have a marginally higher influence over Greenspan. The results across meetings show a highly significant increase in the influence of the inexperienced both overall and on Greenspan.

For the ultimate question of interest, the influence results show that what inexperienced members speak about after transparency has a bigger impact on what others (and specifically the Chairman) speak about in future meetings. The natural explanation is that what inexperienced members say after transparency is more worth listening to than before. Another related explanation is that inexperienced members are more likely to identify important topics before the rest of the committee after transparency. In either case, the evidence points towards inexperienced members bringing additional information into deliberation after transparency, even if during that deliberation there is a tendency to disengage from the ebb and flow of debate that occurred before.

[32]The results are checked for robustness in appendix section C.

Table 8: Influence Measures by Member

Speaker	Meetings under Greenspan	Overall Influence Fixed Effect	Overall Influence Average
Keehn	56	0.0052	0.0641
Forrestal	66	0.0047	0.0599
Guffey	33	0.0041	0.0612
Meyer	43	0.0039	0.0592
Fisher	7	0.0034	0.0583
Mcdonough	79	0.0033	0.0583
Lacker	13	0.0032	0.0576
Minehan	93	0.0029	0.0596
Bernanke	22	0.0028	0.0699
Rivlin	24	0.0027	0.0565
Guynn	80	0.0025	0.0572
Yellen	34	0.0020	0.0559
Broaddus	91	0.0017	0.0564
Johnson	23	0.0017	0.0643
Blinder	13	0.0015	0.0551
Moskow	92	0.0015	0.0562
Poole	64	0.0014	0.0591
Laware	53	0.0013	0.0552
Geithner	18	0.0012	0.0555
Ferguson	67	0.0010	0.0558
Phillips	52	0.0010	0.0548
Hoenig	116	0.0007	0.0552
Syron	41	0.0007	0.0562
Parry	134	0.0002	0.0562
Kelley	114	-0.0001	0.0551
Angell	52	-0.0003	0.0587
Santomero	45	-0.0004	0.0646
Mullins	29	-0.0004	0.0532
Gramlich	62	-0.0004	0.0573
Boykin	28	-0.0007	0.0539
Boehne	100	-0.0007	0.0537
Black	41	-0.0009	0.0555
Pianalto	25	-0.0011	0.0648
Lindsey	41	-0.0015	0.0518
Melzer	83	-0.0017	0.0536
Stern	146	-0.0019	0.0525
Kohn	29	-0.0021	0.0586
Bies	34	-0.0023	0.0612
Hoskins	31	-0.0024	0.0519
Heller	15	-0.0025	0.0658
Mcteer	110	-0.0028	0.0529
Morris	10	-0.0031	0.0514
Olson	34	-0.0031	0.0507
Corrigan	47	-0.0031	0.0527
Seger	29	-0.0034	0.0512
Jordan	86	-0.0036	0.0507
Greenspan	149	-0.0053	0.0539
Stewart	4	-0.0072	0.0433

Speaker	Meetings under Greenspan	Greenspan Influence Fixed Effect	Greenspan Influence Average
Bernanke	22	0.00228	0.00854
Greenspan	149	0.00116	0.00515
Keehn	56	0.00077	0.00455
Guffey	33	0.00049	0.00381
Forrestal	66	0.00047	0.00356
Fisher	7	0.00043	0.00354
Minehan	93	0.00036	0.00379
Johnson	23	0.00027	0.00487
Lacker	13	0.00024	0.00328
Laware	53	0.00016	0.00305
Syron	41	0.00009	0.00322
Poole	64	0.00006	0.00369
Rivlin	24	0.00004	0.00299
Blinder	13	0.00004	0.00293
Geithner	18	0.00001	0.00302
Mcdonough	79	0.00001	0.00308
Moskow	92	-0.00001	0.00304
Lindsey	41	-0.00003	0.00282
Yellen	34	-0.00005	0.00290
Meyer	43	-0.00005	0.00305
Santomero	45	-0.00006	0.00526
Mullins	29	-0.00007	0.00280
Guynn	80	-0.00008	0.00296
Black	41	-0.00009	0.00315
Broaddus	91	-0.00010	0.00293
Phillips	52	-0.00011	0.00282
Parry	134	-0.00014	0.00314
Gramlich	62	-0.00015	0.00347
Boykin	28	-0.00016	0.00279
Seger	29	-0.00016	0.00279
Hoenig	116	-0.00016	0.00285
Ferguson	67	-0.00016	0.00289
Kelley	114	-0.00018	0.00291
Angell	52	-0.00018	0.00363
Boehne	100	-0.00019	0.00278
Pianalto	25	-0.00022	0.00570
Stern	146	-0.00025	0.00274
Morris	10	-0.00028	0.00266
Melzer	83	-0.00028	0.00282
Jordan	86	-0.00028	0.00271
Hoskins	31	-0.00030	0.00262
Corrigan	47	-0.00031	0.00284
Olson	34	-0.00038	0.00257
Heller	15	-0.00040	0.00529
Mcteer	110	-0.00041	0.00288
Stewart	4	-0.00049	0.00203
Kohn	29	-0.00064	0.00364
Bies	34	-0.00093	0.00433

Notes: This table reports, for overall FOMC influence (left panel) and influence on Chairman Greenspan (right panel), some statistics on the inter-meeting influence measures. The table presents the average value of influence for each member although the ranking is based the member-fixed effects from a regression of the influence measure of time and member fixed effects ($a_{it}/a_{it}^G = \alpha_{it} + \delta_t + \epsilon_{it}$).

Table 9: Influence—member level

Main Regressors	(1) w_{it}	(2) a_{it}	(3) w_{it}^G	(4) a_{it}^G
D(Trans)	0.093**	-0.040	0.0039	-0.018***
	[0.041]	[0.273]	[0.507]	[0.000]
Fed Experience	-0.0093**	-0.0041	-0.00036	0.00031
	[0.045]	[0.284]	[0.551]	[0.547]
D(Trans) x Fed Experience	-0.000023	-0.00010***	-7.2e-06*	-0.000027***
	[0.593]	[0.000]	[0.062]	[0.000]
Constant	0.11***	0.16***	0.0053	0.018***
	[0.000]	[0.000]	[0.133]	[0.000]
Number of groups	38	35	38	35
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Within Meeting	Intra	Inter	Intra	Inter
Sample	87:08-97:09	87:08-97:09	87:08-97:09	87:08-97:09
Obs	1427	1377	1427	1377
% Rookie effect	-	3.6	4.5	15.7

Notes: This table reports measures of member influence derived from our LDA estimation. Where the difference in difference is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

8 Effect of Transparency on Policy Outcomes

Ultimately, the key macroeconomic question regarding the shift to transparency is what effect it had on the Federal Funds Rate that the FOMC decides. This section first shows two pieces of evidence that support the idea that the committee had additional information after transparency. Of course, regressions of policy outcomes on transparency can only be done with the (DIFF) specification. We therefore propose an indirect way of measuring individual influence on policy, and show within the (DinD) specification that after transparency rookies' policy influence increases.

8.1 Gradualism changes

One explanation of gradualism, or policy inertia, is that policymakers face uncertainty and only wish to change policy when they are sufficiently sure they will not have to quickly reverse their change (Coibion and Gorodnichenko 2012). This helps explain why interest rates can change very quickly in times of crisis (when it is clear that interest rates need to change). Hence if the FOMC as a whole had more information after transparency, one possible effect would be reduced gradualism.

To see whether gradualism changed after 1993, we estimate an interest rate reaction

Table 10: Gradualism before and after transparency

Main Regressors	(1) FFR_t	(2) FFR_t	(3) FFR_t
Expected inflation	-0.15	0.26***	-0.053
	[0.496]	[0.000]	[0.726]
Expected output gap	0.067	0.069**	0.024
	[0.193]	[0.019]	[0.542]
Expected output growth	0.078**	0.085***	0.100***
	[0.044]	[0.001]	[0.000]
FFR(-1)	1.02***	0.90***	1.03***
	[0.000]	[0.000]	[0.000]
D(Transparent) x Expected Inflation			0.32*
			[0.056]
D(Transparent) x Expected output gap			0.062
			[0.205]
D(Transparent) x Expected real GDP growth			-0.024
			[0.377]
D(Transparent) x FFR(-1)			-0.16*
			[0.090]
Constant	0.32	-0.29*	-0.20
	[0.563]	[0.090]	[0.238]
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

Notes: This table reports the results of estimating a gradualism equation using the specification and data used in Coibion and Gorodnichenko (2012). We estimate the specification separately for the pre-transparency period (column (1)), post-transparency period (column (2)) and a specification which nests both samples allowing for different coefficients pre- and post- transparency.

function along the lines of Coibion and Gorodnichenko (2012). Their specification allows for flexible time series structures which allow for persistent shocks. We also employ their real time data and forecasts from the Greenbook. In table 10, we present the estimates for pre-transparency (column (1)), post-transparency (column (2)) and a specification in which we include all data but allow the post-transparency coefficients to vary using interaction terms. All specifications allow shocks to have become less persistent over time. Consistent with the committee acting with greater certainty, the results clearly point to reduced gradualism.

8.2 Monetary Policy Surprises

As a second test, we examine whether the FOMC was more or less likely to surprise the markets with their interest rate decisions. To do this, we use the surprise data developed in both Gürkaynak, Sack, and Swanson (2005), which measures the effect of both short and longer term movements in the yield curve, and Kuttner (2001) (which was subsequently used in Bernanke and Kuttner (2005)) which captures the shorter end movements

only. We label these variables "Surprise (GSS)" and "Surprise (K)" respectively. These data are derived from traded futures securities which allow one to decompose movements in the Fed Funds target rate into expected and unexpected moves.

Table 11: Effect on market surprises

	(1)	(2)
Main Regressors	|Surprise (GSS)|	|Surprise (K)|
D(Trans)	7.83***	5.36*
	[0.002]	[0.053]
D(NBER recession)	-2.43	-2.08
	[0.351]	[0.505]
D(2 Day)	0.70	0.76
	[0.601]	[0.777]
Uncertainty(t-1)	0.0050	0.0052
	[0.401]	[0.455]
Expected output growth	-0.027	0.42
	[0.971]	[0.616]
Expected inflation	4.00***	3.23*
	[0.006]	[0.058]
Expected output gap	-0.74	-0.83
	[0.139]	[0.256]
Constant	-11.8**	-9.04
	[0.034]	[0.158]
R-squared	0.156	0.063
Lag Dep. Var?	Yes	Yes
Sample	89:11-97:09	89:11-97:09
Obs	71	66

Notes: This table reports the results of estimating (DIFF) on the absolute value of FOMC market surprises measured by Gürkaynak, Sack, and Swanson (2005), Surprise (GSS), and Kuttner (2001), Surprise (K). Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

In table 11 we report the results of estimating (DIFF) on the absolute value of the surprise data. The results indicate that there was more (absolute) surprise following transparency even after controlling for the economic forecast variables such as expected inflation and the expected output gap. This is especially the case for the broader measures of surprises (Surprise (GSS)).

8.3 Do rookie members have influence on important topics?

Our final approach to measure an impact on policy builds on what we have already found and is the only approach where we can analyse the effects as a diff-in-diff analysis. This approach consists of two stages:

1. We first take the level of the GSS surprise measures used above and identify two topics that are consistently correlated with surprises during Chairman Greenspan's tenure. We report the results in table 12a; controlling for uncertainty and whether

Table 12: Rookie Policy Influence

(a) Greenspan's topics and market surprise

Main Regressors	(1) Surprise (GSS)	(2) Surprise (GSS)
Demand topic	0.11	0.13*
	[0.104]	[0.066]
Productivity topic	-0.12**	-0.14**
	[0.019]	[0.042]
D(NBER Recession)	-0.023	0.0065
	[0.331]	[0.770]
BBD uncertainty	-0.000071	-0.000063
	[0.336]	[0.363]
Expected output growth		0.0089*
		[0.073]
Expected inflation		-0.0057
		[0.416]
Expected output gap		-0.0023
		[0.391]
Constant	0.0036	-0.0096
	[0.766]	[0.755]
R-squared	0.054	0.110
Sample	87:08-06:01	87:08-06:01
Obs	139	139

(b) Influence on Greenspan's policy topics

Main Regressors	(1) w_{it}^{G*}	(2) a_{it}^{G*}
D(Trans)	-0.045	-0.17***
	[0.154]	[0.000]
Fed Experience	0.0050	0.015***
	[0.132]	[0.000]
D(Trans) x Fed Experience	-0.00011*	-0.00017**
	[0.065]	[0.014]
Constant	-0.025	-0.054**
	[0.198]	[0.027]
Number of groups	38	35
Member FE	Yes	Yes
Time FE	Yes	Yes
Within Meeting	Intra	Inter
Sample	87:08-97:09	89:08-97:09
Obs	1427	1377
% Rookie effect	46.8	76.3

Notes: The upper table reports the results of estimating (DIFF) on the absolute value of FOMC market surprises measured by Gürkaynak, Sack, and Swanson (2005), Surprise (GSS), and Kuttner (2001), Surprise (K). The lower table reports the results of estimating (DinD) on the measures of influence computed only over the two topics shown to correlate significantly with FOMC surprises of the market. Where the difference in differences is statistically significant, the rookie effect reports, as a % of pre-transparency mean behaviour, the differential effect of transparency on members with one year of Fed experience compared to a member with 20 years of experience. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

the economy is in a recession, we find that topics 23 and 29, introduced in section 5.4 and capturing productivity and demand discussions respectively, are both significantly related to surprises. Specifically, in column (1) we show that when Chairman Greenspan spends more time discussing demand topics, there tends to be a positive surprise while productivity discussions are associated with negative surprises. Column (2) shows that these correlations remain if we additionally control for Greenbook forecast variables.

2. Next, we ask whether rookies—whose influence on all topics increases after transparency—also increase their influence over Greenspan's discussion of just these two topics. This involves recalculating the within- (w_{it}^G) and across-meeting (a_{it}^G) influence of Greenspan measures, previously calculated for all economic topics, but now only looking at topics 23 and 29. These new versions of the influence measures are w_{it}^{G*} and a_{it}^{G*}.

   Table 12b shows that transparency leads rookie members to increase their influence of Chairman Greenspan on these topics. The size of the rookie effect indicates that the effect is substantial relative to the pre-transparency average influence over Greenspan on these topics.

While this is an indirect test that rookies had higher influence on the federal funds rate after transparency, we believe that, together with the rest of the findings in this section, it is certainly consistent with transparency giving rise to an effect on FOMC decisions via an increased information channel.

9 Conclusions

Overall, we find evidence for the two effects predicted by the career concerns literature: discipline and information distortion (the latter taking the form of a bias toward conformity). The net outcome of these two effects appears to be positive: rookies become more influential in shaping discussion and in inducing surprising decisions. We therefore believe that the evidence available from the 1993 natural experiment points toward an overall positive role of transparency. However, policymakers—and future research—should explore ways to structure the deliberation process in order to maximize the discipline effect and minimize the conformity effect.

References

Avery, C. N., and J. A. Chevalier (1999): "Herding over the career," *Economics Letters*, 63(3), 327–333.

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

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.

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.

Chevalier, J., and G. Ellison (1999): "Career Concerns Of Mutual Fund Managers," *The Quarterly Journal of Economics*, 114(2), 389–432.

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

Driscoll, J. C., and A. C. Kraay (1998): "Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data," *The Review of Economics and Statistics*, 80(4), 549–560.

Federal Open Market Committee (1993): "Transcript Federal Open Market Committee Conference Call, October 15, 1993," http://www.federalreserve.gov/monetarypolicy/files/FOMC19931015confcall.pdf, 15 October 1993, last accessed 11 April 2013.

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.

Fligstein, N., J. S. Brundage, and M. Schultz (2014): "Why the Federal Reserve Failed to See the Financial Crisis of 2008: The Role of "Macroeconomics" as a Sensemaking and Cultural Frame," Mimeograph, University of California Berkeley.

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 US House of Representatives.

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

Gürkaynak, R. S., B. Sack, and E. Swanson (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 1(1).

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

Hazen, T. (2010): "Direct and Latent Modeling Techniques for Computing Spoken Document Similarity," in *IEEE Workshop on Spoken Language Technology*.

Heinrich, G. (2009): "Parameter estimation for text analysis," Technical report, vsonix GmbH and University of Leipzig.

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," *Federal Reserve Bank of St. Louis Review*, 87(2), 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.

Ottaviani, M., and P. N. Sørensen (2006): "Reputational cheap talk," *RAND Journal of Economics*, 37(1), 155–175.

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.

Schonhardt-Bailey, C. (2013): *Deliberating Monetary Policy*. MIT Press, Cambridge.

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," *Federal Reserve Bank of St. Louis Review*, 82(5), 1–16.

Transparency International (2012): "Improving the accountability and transparency of the European Central Bank," Position paper.

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

A Sampling Algorithm

Table A.1: 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
$v_{d,n}$	Token index 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	
m_k^d	Count of words in document d allocated to topic k
m_v^k	Count of times token v is allocated to topic k
$m_{k,-n}^d$	Excluding token n, count of words in document d allocated to topic k
$m_{v,-(d,n)}^k$	Excluding token n in document d, count of times unique token v is allocated to topic k

The basic idea of Gibbs sampling is to sample all variables from their conditional distributions with respect to the current values of all other variables and the data. In the LDA model, the data are the words, and the key quantity is the topic allocation of each word; from the word allocations, one can infer the implied topic and term distributions quite easily given the imposition of a symmetric Dirichlet prior. As explained in Griffiths and Steyvers (2004) and in more detail in Heinrich (2009), the conditional distribution of $z_{d,n}$, given all other word-topic assignments $z_{-(d,n)}$ and the vector of words $\mathbf{w}$ in all documents, is given by:

$$\Pr\left[z_{d,n} = k \mid z_{-(d,n)}, \mathbf{w} \right] \propto \frac{m_{v_{d,n},-(d,n)}^k + \eta}{\sum_{v=1}^{V} (m_{v,-(d,n)}^k + \eta)} \left(m_{k,-n}^d + \alpha \right) \tag{A.1}$$

The implementation of (collapsed) Gibbs sampling for the LDA model is the following:

1. Randomly assign all words in all documents to a topic in $\{1, \ldots, K\}$.
2. Form the counts m_k^d and m_v^k.
3. Iterating through each word in each document:
   (a) Drop $w_{d,n}$ from the sample and form the counts $m_{k,-n}^d$ and $m_{v,-(d,n)}^k$.

(b) Assign a new topic for word $w_{d,n}$ by sampling from (A.1).

(c) Form new counts m_k^d and m_v^k by adding the new assignment of $w_{d,n}$ to $m_{k,-n}^d$ and $m_{v,-(d,n)}^k$.

(d) Move on to the next word in the data

4. Repeat 8,000 times.

The estimate of the term distribution matrix (K by V) after any particular iteration is given by

$$\widehat{\beta}_k^v = \frac{m_v^k + \eta}{\sum_{v=1}^V (m_v^k + \eta)} \tag{A.2}$$

and of the topic distribution matrix (D by K) is given by

$$\widehat{\theta}_d^k = \frac{m_k^d + \alpha}{\sum_{k=1}^K \left(m_k^d + \alpha\right)}. \tag{A.3}$$

A.1 Estimating aggregate document distributions

As explained in the text, we are more interested in the topic distributions at the meeting-speaker-section level rather than at the individual statement level. Denote by $\theta_{i,t,s}$ the topic distribution of the aggregate document. Let $w_{i,t,s,n}$ be the nth word in the document, $z_{i,t,s,n}$ its topic assignment, $v_{i,t,s,n}$ its token index, $m_k^{i,t,s}$ the number of words in the document assigned to topic k, and $m_{k,-n}^{i,t,s}$ the number of words besides the nth word assigned to topic k. To re-sample the distribution $\theta_{i,t,s}$, for each iteration $j \in \{4050, 4100, \ldots, 8000\}$ of the Gibbs sampler, we[33]:

1. Form $m_k^{i,t,s}$ from the topic assignments of all the words that compose the aggregate document (i, t, s) from the Gibbs sampling.

2. Drop $w_{i,t,s,n}$ from the sample and form the count $m_{k,-n}^{i,t,s}$.

3. Assign a new topic for word $w_{i,t,s,n}$ by sampling from

$$\Pr\left[\, z_{i,t,s,n} = k \;\middle|\; z_{-(i,t,s,n)}, \mathbf{w}_{i,t,s} \,\right] \propto \widehat{\beta}_k^{v_{i,t,s,n}} \left(m_{k,-n}^d + \alpha\right) \tag{A.4}$$

   where $z_{-(i,t,s,n)}$ is the vector of topic assignments in document (i, s, t) excluding word n and $\mathbf{w}_{i,t,s}$ is the vector of words in the document.

4. Proceed sequentially through all words.

5. Repeat 20 times.

We then obtain the estimate

$$\widehat{\theta}_{i,t,s}^k = \frac{m_k^{i,t,s} + \alpha}{\sum_{k=1}^K \left(m_k^{i,t,s} + \alpha\right)}. \tag{A.5}$$

[33]This procedure is broadly in line with that described in Heinrich (2009) for querying documents outside the set on which LDA is estimated. The key point is that one can estimate out-of-sample document distributions that are formed of the same set of topics as the within-sample documents in the way described. Many fewer iterations are needed since topics do not need to be re-estimated.

B Estimated Topics

Table B.1: Discussion topics

Topic	Top Tokens
T0	side littl see quit better pretti concern good seem much
T4	problem becaus world believ view polit rather make by like
T7	percent year quarter growth first rate fourth half over second
T9	mr without thank laughter let move like peter call object
T14	other may also point first suggest might least indic like
T15	point right want said make agre say comment now realli
T16	now too may all economi seem good much still long
T17	question whether how ask issu rais answer ani know interest
T18	tri can out work way get make how want need
T19	year last month over meet next week two three decemb
T22	year line panel right shown chart by left next middl
T26	up down come out back see off start where look
T27	governor ye vice kelley stern angel parri minehan hoenig no
T32	peopl talk lot say around get thing when all becaus
T33	chang no make reason ani can way other whi becaus
T34	new seem may uncertainti even see much bit by now
T39	look see get seem now when happen realli back regard
T42	get thing problem lot term look realli kind out say
T44	get move can all stage inde signific becaus ani evid
T49	say know someth all can thing anyth happen cannot els

Table B.2: Economics topics

Topic	Top Tokens
T1	price oil increas oilprice effect suppli through up higher demand
T2	target object credibl pricestabil issu goal public achiev strategi lt
T3	direct move support mr recommend prefer asymmetr symmetr favor toward
T5	polici monpol such by action might zero when possibl respons
T6	committe meet releas discuss minut announc vote decis member inform
T8	project expect recent year month data forecast by activ revis
T10	condit committe period reserv futur consist sustain read develop maintain
T11	number data look indic show up measur point evid suggest
T12	statement word languag like use altern sentenc commun refer chang
T13	rate market year spread yield month panel sinc page volatil
T20	model use effect differ rule estim actual result simul relationship
T21	forecast greenbook project assum assumpt staff by baselin scenario path
T23	invest inventori capit incom consum spend busi hous household sector
T24	period reserv market borrow billion day million by treasuri bill
T25	inflat percent core measur level low ue cpi year over
T28	rate market move fund bps ffr polici action point need
T29	product increas wage cost price labor labmkt trend rise acceler
T30	polici might committe may by tighten market eas such seem
T31	district nation manufactur activ region continu area economi employ remain
T35	sale year price industri level continu product auto increas good
T36	rate intrate lt expect real effect lower declin level st
T37	dollar market yen against by intervent mark japanes currenc exrate
T38	bank credit debt loan financi asset by market other also
T40	risk balanc downsid concern view upsid both now side meet
T41	dollar countri export import foreign trade deficit us real other
T43	growth continu economi slow increas strong remain recent expect expans
T45	economi fiscal weak recoveri recess cut confid econom spend budget
T46	treasuri oper secur billion use issu author swap system hold
T47	busi report contact firm compani said up year plan increas
T48	rang money aggreg altern growth nomin monetari veloc year target

C Robustness analysis

In tables C.1-C.4 below, we explore the robustness of the main diff-in-diff results presented in section 6 and 7. In each table we report the main diff-in-diff coefficient (ϕ, on the "D(Trans) $\times$ Fed Experience" regressor) for a number of robustness tests. In particular, we:

1. Try a number of alternative sample periods;
2. Run a placebo test on the change in transparency;
3. Use a 70-topic model, rather than the 50-topic model used in the baseline;
4. Evaluate whether the effects differ between presidents and governors using a triple-diff specification.

To evaluate the robustness to different sample sizes, we first follow Meade and Stasavage (2008) and exclude 1993 from the estimation entirely but proceed otherwise as in the baseline sample. The reason for this is that, despite most members claiming (to each other in a conference call) that they did not know of the transcripts, a few members certainly knew of them prior to October 1993. Therefore we ignore the whole of 1993 as this was a period during which FOMC members may have already known of the transcripts and started to adjust their behavior. The second robustness exercise on sample selection is to remove the first year of Greenspan's tenure; the behavior of the committee may have been different in the first meetings as the new Chairman "settled in". Then we explore a narrower window of four years before and after the change in transparency. Finally we include all years of the Greenspan tenure (1987-2006); in this case the sample is predominantly "post-transparency".

Table C.1 presents the results for the regressions in section 6 for each of these different samples. The table shows that the main coefficient of interest is little changed by the different sample selections. While standard errors do change a bit, the basic messages of the analysis are robust. Table C.2 presents the results for the regressions in section 7; the results are even closer across different sample selections when we consider the influence measures.

We next turn to the other robustness checks. To begin with, we consider a placebo test on the date of the change in transparency. In particular, we take the second half of Alan Greenspan's tenure on the committee, November 1997 to January 2006 (which is not used in the baseline analysis), and we randomly select November 2001 as the meeting at which transparency changed. Of course, since transparency did not actually change at that point, we expect to get zero results on the diff-in-diff with this test and that is exactly what we get in table C.3 and C.4.

As discussed in section 5.2.2, the information criterion favor a larger number of topics but we select 50 in the baseline analysis as it combines both parsimony and interpretability. But have also carried out the analysis using a 70 topic model. The results, rather than the 50 topic model used in the baseline analysis. As shown in the third rows of tables C.3 and C.4, the estimated sign and size of the main coefficients are quite similar using the larger number of topics (though standard errors are wider for some regressions). Only the results on the Herfindahl are markedly changed.

Table C.1: Comparison of results for different sample selections I

D(Trans) × Fed Experience	(1) Economics FOMC1	(2) Economics FOMC2	(3) Herfindahl FOMC1	(4) Herfindahl FOMC2	(5) Data Topic (7) FOMC1	(6) Data Topic (7) FOMC2	(7) Figures Topic (22) FOMC1	(8) Figures Topic (22) FOMC2	(9) SG(it) FOMC1	(10) SG(it) FOMC2
Baseline Coefficient	0.00021	-0.0014**	0.00071**	-0.00037***	-0.00065***	-0.000098	-0.000088**	0.00011*	-0.00015	-0.00023***
	[0.622]	[0.015]	[0.035]	[0.000]	[0.002]	[0.103]	[0.035]	[0.078]	[0.452]	[0.000]
Excluding 1993	0.00067*	-0.0016*	0.00079	-0.00035***	-0.00082***	-0.000096	-0.00013*	0.00012*	-0.00031**	-0.00021***
	[0.098]	[0.053]	[0.136]	[0.005]	[0.001]	[0.191]	[0.055]	[0.099]	[0.020]	[0.001]
Dropping Greenspan's 1st year	0.000079	-0.0015**	0.00066**	-0.00038***	-0.00062***	-0.000086	-0.000088**	0.00011*	-0.00016	-0.00023***
	[0.839]	[0.017]	[0.040]	[0.000]	[0.003]	[0.173]	[0.047]	[0.095]	[0.416]	[0.001]
Narrow window	0.00011	-0.0013**	0.00065**	-0.00037***	-0.00052***	-0.000081	-0.000069	0.00012	-0.00015	-0.00023***
	[0.809]	[0.021]	[0.045]	[0.000]	[0.002]	[0.196]	[0.101]	[0.115]	[0.449]	[0.001]
Full Greenspan Tenure	0.00021	-0.0014**	0.00071**	-0.00037***	-0.00065***	-0.000098	-0.000088**	0.00011*	-0.00015	-0.00023***
	[0.622]	[0.015]	[0.035]	[0.000]	[0.002]	[0.103]	[0.035]	[0.078]	[0.452]	[0.000]

Notes: This table reports, for a variety of robustness tests, the main diff-in-diff coefficient on the $D(Trans) \times FedExp_{i,t}$ regressor. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Table C.2: Comparison of results for different sample selections II

D(Trans) × Fed Experience	(1) w_{it}	(2) a_{it}	(3) w_{it}^G	(4) a_{it}^G
Baseline Coefficient	-0.000023	-0.00010***	-7.2e-06*	-0.000027***
	[0.593]	[0.000]	[0.062]	[0.000]
Excluding 1993	0.000018	-0.000082***	-3.5e-06	-0.000024***
	[0.764]	[0.003]	[0.484]	[0.000]
Dropping Greenspan's 1st year	-0.000023	-0.00011***	-8.7e-06**	-0.000027***
	[0.609]	[0.000]	[0.010]	[0.000]
Narrow window	-0.000020	-0.00011***	-1.0e-05***	-0.000026***
	[0.647]	[0.000]	[0.007]	[0.000]
Full Greenspan Tenure	-0.000023	-0.00010***	-7.2e-06*	-0.000027***
	[0.593]	[0.000]	[0.062]	[0.000]

Notes: This table reports, for a variety of robustness tests, the main diff-in-diff coefficient on the $D(Trans) \times FedExp_{i,t}$ regressor. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Finally, we explore whether the effects that we find, which are averages for FOMC members, are different for further splits of the members into sub-groups. For example, we can examine whether there is a triple-difference effect such that the diff-in-diff results are different for Governors and Bank Presidents. We already capture member fixed effects, so this specification is mostly of interest if we think that rookie presidents and rookie governors respond in very different ways. The downside of this approach is that, with more coefficients to estimate, we lose some power in the estimation. The diff-in-diff-in-diff that we estimate, in the case of Governors versus Presidents, is:[34]

$$\begin{aligned} y_{it} = \alpha_i + \delta_t + \beta_1 D(Trans)_t + \eta_1 FedExp_{i,t} + \eta_2 FedExp_{i,t} \times D(Pres) \dots \\ + \beta_2 D(Trans)_t \times D(Pres) + \phi_1 D(Trans)_t \times FedExp_{i,t} \dots \\ + \phi_2 D(Trans)_t \times FedExp_{i,t} \times D(Pres) + \epsilon_{it} \end{aligned} \tag{DinDinD}$$

From this regression, we can estimate the implied $D(Trans)_t \times FedExp_{i,t}$ coefficients for each group separately as ϕ_1 for Governors and $\phi_1 + \phi_2$ for Presidents, with the difference between these two coefficients given by ϕ_2.

In the final part of tables C.3 and C.4 we report the implied coefficients and the difference between them for Governors versus Presidents. While some of the triple differences are statistically significant, for all of the main coefficients that are significant when we consider the "average" effect, the implied coefficients for each group go in the same direction as the average effect, and there is not a consistent direction to the difference between groups. This convinces us that our baseline approach is not missing important heterogeneity between groups.

[34]To examine other splits of the members, we could simply replace $D(Pres)$ with a variable splitting the sample members along another dimension.

Table C.3: Comparison of results for alternative tests I

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
D(Trans) ×	Economics		Herfindahl		Data Topic (7)		Figures Topic (22)		SG(it)	
Fed Experience	FOMC1	FOMC2	FOMC1	FOMC2	FOMC1	FOMC2	FOMC1	FOMC2	FOMC1	FOMC2
Baseline Coefficient	0.00021	-0.0014**	0.00071**	-0.00037***	-0.00065***	-0.000098	-0.000088**	0.00011*	-0.00015	-0.00023***
	[0.622]	[0.015]	[0.035]	[0.000]	[0.002]	[0.103]	[0.035]	[0.078]	[0.452]	[0.000]
Placebo Coefficient	-0.00038	0.000016	-0.00013	-5.3e-06	0.00013	-7.8e-06	3.2e-06	-0.000050	-0.000035	0.000014
	[0.118]	[0.935]	[0.404]	[0.935]	[0.190]	[0.876]	[0.942]	[0.212]	[0.309]	[0.692]
70-topic Model Coefficient	0.00013	-0.0011**	-0.00022	-0.00030***	-0.00044***	-0.00012**	-0.000048	0.000080	-0.00012	-0.00017***
	[0.724]	[0.028]	[0.594]	[0.000]	[0.009]	[0.023]	[0.190]	[0.180]	[0.428]	[0.005]
Triple-Difference Regressions										
Governor Coefficient	0.00253**	-0.00021	0.00098	-0.00012***	-0.00039	-0.00018	-0.00019	0.00007	-0.00030	-0.00011**
	[.04]	[.828]	[.166]	[.001]	[.144]	[.743]	[.135]	[.324]	[.924]	[.018]
President Coefficient	-0.00062	-0.00199***	0.00080	-0.00042	-0.00128***	-0.00003	-0.00007	0.00009	0.00001	-0.00014
	[.168]	[.006]	[.128]	[.405]	[0]	[.183]	[.155]	[.277]	[.345]	[.378]
Difference	-0.0031**	-0.0018*	-0.00019	-0.00030	-0.00089***	0.00015	0.00012	0.000021	0.00031	-0.000033

Notes: This table reports, for a variety of robustness tests, the main diff-in-diff coefficient on the $D(Trans) \times FedExp_{i,t}$ regressor. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Table C.4: Comparison of results for alternative tests II

D(Trans) × Fed Experience	(1) w_{it}	(2) a_{it}	(3) w_{it}^G	(4) a_{it}^G
Baseline Coefficient	-0.000023	-0.00010***	-7.2e-06*	-0.000027***
	[0.593]	[0.000]	[0.062]	[0.000]
Placebo Coefficient	0.000033	0.000022	3.5e-06	2.4e-06
	[0.151]	[0.518]	[0.505]	[0.657]
70-topic Model Coefficient	-0.000032	-0.000098**	-9.2e-06	-0.000027**
	[0.475]	[0.011]	[0.116]	[0.015]
Triple-Difference Regressions				
Governor Coefficient	0.00007	-0.00002***	0.00000	-0.00002*
	[.339]	[.001]	[.927]	[.065]
President Coefficient	-0.00005	-0.00015	-0.00001	-0.00003***
	[.604]	[.83]	[.209]	[0]
Difference	-0.00013	-0.00014	-7.3e-06	-6.5e-06

Notes: This table reports, for a variety of robustness tests, the main diff-in-diff coefficient on the $D(Trans) \times FedExp_{i,t}$ regressor. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.