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

Stephen Hansen[‡] Michael McMahon[§] Andrea Prat[¶]

First Draft: April 29, 2014
This Draft: February 2, 2015

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, Guillermo Caruana, Francesco Caselli, Diego Garcia, Refet Gürkaynak, Gilat Levy, Seamus MacGorain, Rick Mishkin, Emi Nakamura, Tommaso Nannicini, Bryan Pardo, Amar Radia, David Romer, 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. 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 annual conference; ESSIM; the NBER Summer Institute; the ECB; the CEPR Workshop on Incentives, Management, and Organization; the BIS; CEMFI; Torino; University of Hamburg; London Business School; CAGE; Norwegian Business School BI; and the OFCE EME workshop. We thank Eric Hardy for excellent research assistance in gathering biographical data. We also benefited from seed corn financial support from the Bank of England's Research Donations Committee, and a British Academy small grant. Any errors remain ours alone.

[†]All source code for estimation is available from `https://github.com/sekhansen/text-mining-tutorial` and a worked example from `http://nbviewer.ipython.org/url/www.econ.upf.edu/~shansen/tutorial_notebook.ipynb`.

[‡]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 were the descriptions of policy meetings put on the public record as of 2014, a major aspect of procedural transparency (Geraats 2002). At the same time, Geraats (2009) notes a general rise in procedural transparency across central banks. This tendency is also evident in the ECB and the Bank of England. 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), and in this spirit the ECB has committed to release more detailed accounts of its meetings (but not full transcripts) in the future.[2] Moreover, the Bank of England has recently announced major reforms to its disclosure policy that will make it more transparent, including the partial publishing of transcripts. In spite of this increase in transparency, whether more transparency is always beneficial is an open question.

Table 1: Information made available by different central banks as of 2014

	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.

[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 2014) 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.

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 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

[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.

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 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

[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.

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.

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. 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 more influential in terms of their colleagues' (and particularly Alan Greenspan's) 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, and seek to design disclosure policies that maximize this effect while minimizing the conformity effect. As we elaborate in the discussion below, this insight has already directly informed real-world disclosure policies.

Our paper also makes a methodological contribution. An important distinction in the analysis of text is whether documents come with natural labels or not. When they do, an important task is to use text features to predict them. For example, Gentzkow and Shapiro (2010) present a way of determining which phrases best predict party affiliation in congressional speeches. We instead present a way of uncovering hidden themes in unlabeled text data without linking themes to particular word lists prior to estimation, which is currently the de facto standard approach in economics. This approach should be fruitful in many areas of research beyond our particular application.

The machine learning algorithm we introduce to the economics literature is Latent Dirichlet Allocation (LDA) by Blei, Ng, and Jordan (2003). LDA is a widely used topic model and has been cited over 10,000 times since 2003. Topic modeling approaches are beginning to appear in the social science literature—in particular political science (see Grimmer 2010, Quinn, Monroe, Colaresi, Crespin, and Radev 2010)—but to our knowledge ours is the first paper to use it in economics. Moreover, we do not use LDA to just describe which documents cover which topics, but also use its output to construct measures of communication that we embed as dependent variables in an econometric model explicitly motivated by economic theory (i.e. career concerns). We believe this illustrates the value 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-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). Acosta (2014) —also developed independently —uses Latent Semantic Analysis, a precursor to LDA, to analyse the effect of changes in Fed transparency on aggregate measures of meeting communication. 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. A narrative approach to text is also used in Chappell, McGregor, and Vermilyea (2005).

Moreover, there is an existing literature that converts central bank text communication to quantitative measures using different methodologies. Chappell, Havrilesky, and McGregor (2000) is an early example of a literature which classifies speeches or minutes as hawkish or dovish and examines individual reaction functions. Bligh and Hess (2006) explore how the content of statements, measured in terms of optimism, pessimism, certainty and others, can help to forecast financial market variables. Boukus and Rosenberg (2006), Hendry and Madeley (2010) and Hendry (2012) use the thematic content of central bank communications to examine market response. Lucca and Trebbi (2009) use an algorithm to score Fed text data as hawkish or dovish. Apel and Blix Grimaldi (2012) carry out a similar exercise for Swedish Riksbank minutes.

The paper proceeds as follows. Section 2 reviews the career concerns literature that

[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.

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, and section 8 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 (or some other material) 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

[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.

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.

Therefore, the overall effect of increased transparency can be positive (through increased discipline) or negative (through increased conformity/non-conformity). However, we can go one step further and examine how transparency interacts with an other observable: the agent's experience level.

In all career concerns models, the effect of transparency depends on how long the agent has been active. When the agent starts, little is known about him. As time passes, the principals gather more information about him. More experienced agents have less of an incentive to distort their behavior in order to signal their type (Holmström 1999). And the effect of transparency is stronger on agents who have more incentive to signal their types.

The differential effect of experience can be used to study career concerns. Hong, Kubik, and Solomon (2000) compared the behavior of inexperienced and experienced equity analysts, the latter being those who have been providing earnings forecast for at least three years. Consistent with a model of conformity, they found that inexperienced analysts deviate less from consensus forecasts.

In our setting, the differential effect of experience on career concerns means that less experienced agents should be more affected by a change in disclosure rules than their more experienced colleagues. In the case of discipline, this means that effort will go up relatively more for the inexperienced agents. In the case of conformity, this means that incentives to conform (or non-conform) will be relatively stronger among the less experienced agents.

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]

[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.

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 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 using the entire meeting in the whole sample under Greenspan with each unique member intervention being treated as a separate statement for the estimation of topics.

[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

As discussed in detail in Lindsey (2003), 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), the few members who knew of the tapes 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, who was angry at Fed opacity with leaks of sensitive information to the market, forced the Fed to discuss how it might

[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.

be more transparent. It was during these discussions, in October 1993, that FOMC members became aware of the transcripts. Initially Greenspan was evasive on the issue with the Senate Banking Committee and he argued that he didn't want to release any verbatim information as it would stifle the discussion. But pressure on the Fed grew, and so it quickly moved to release the existing transcripts (with a five-year lag). While no commitment on publishing transcripts going forward was immediately made, and the Fed had five years to make a decision due to the publication lag, this was considered a highly likely outcome. The commitment become formal after 15 months. Taken altogether, 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 essentially took for granted 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 "diff" regression.[14] While useful as a descriptive account of behavior before and after transparency, "diff" analysis is potentially problematic because the timing of other changes may have coincided with the change in transparency. This means the estimated effect 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.

Our empirical strategy is inspired by the differential effect of experience discussed in the theory review (Section 2) and is similar to Hong, Kubik, and Solomon (2000). 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 committee.[15] The longer a

[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, a few members were aware of their existence 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 past 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.

[14]In the appendix, section B, we present and discuss the results from a simple "diff" regression specification.

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

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 1 we plot the histogram of this variable across all members in our main sample period.:[16]

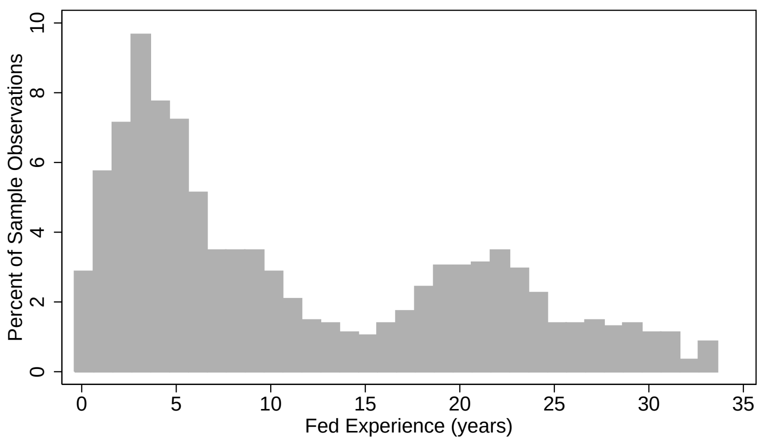


Figure 1: Histogram of Federal Reserve Experience ($FedExp_{i,t}$)

Notes: This figure plots a histogram of the $FedExp_{i,t}$ variable, measured as years of Federal Reserve experience, in our main sample.

The main specification we use in the paper is the following "diff-in-diff" regression:[17]

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

where y_{it} is the communication measure of interest such as a measure of the breadth of topic coverage in meeting t, and $D(Trans)$ is a transparency dummy (1 after November 1993, 0 before).

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. (Below we describe all the different dependent variables that we use in the analysis fully.) 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

[16]In other contexts, one might use age as a good proxy for experience. However, the number of years the member has spent at the Fed is a more appropriate measure in our context. A member who joins the Fed at age 60 will not have established a reputation nearly as much as a member aged 50 with ten years of experience.

[17]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.

analysis presented below, we focus on a sample that uses a window of four years before and four years after the change in transparency (1989-1997). This eight-year window encompasses only Alan Greenspan's tenure as chair of the FOMC. In appendix D, we show that results remain robust to alternative sample selections.[18]

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 of deliberation that are used as dependent variables in (DinD) are constructed using FOMC meeting transcripts.[19]. 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 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. Where staff

[18] In the appendix, we carry out three main robustness tests related to the sample choice. First, we examine the results if we drop all the 1993 observations to control for some of the committee knowing sooner about the transcripts. Second, we drop four members whom internal Fed accounts suggest did know about the written record of the earlier meetings before October 1993. And finally we run a Placebo test using the second half of the Greenspan's tenure as Chairman, and imposing November 2001 as an artificial change in transparency. In all cases, the main results of the analysis are robust.

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

statements were recorded separately in appendices, we re-inserted them into the main transcripts where they were took place in the deliberation. 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

Our text dataset is a collection of D documents, where a document d is a list of tokens $\mathbf{w}_d = (w_{d,1}, \ldots, w_{d,N_d})$. A token is a single element of a string such as a word, number, or punctuation mark. 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 asked a question of staff (the first statement) and a staff member replied (the second statement). Our challenge is to build quantitative communication measures from this unstructured data for the dependent variables in the regressions.

The simplest communication measures rely on counting tokens. For each statement, we count the

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

We also count the total number of statements that FOMC members make in FOMC1 and FOMC2 as a fourth count-based measure of communication.

More abstractly, one can represent each document in terms of a frequency count of the V unique tokens in the data. This is called the bag-of-words model, and its most important simplifying feature is to ignore word order entirely.[20] The bag-of-words transformation converts each document into a highly sparse V-dimensional histogram: while individual statements contain a few hundred words at most, V is on the order of 10,000-20,000 depending on how one selects vocabulary. Dimensionality reduction is therefore key.

By far the most common solution in the economics literature is to employ *dictionary methods*. These involve the researcher defining a list of words that she believes captures

[20] 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.

relevant content, and then representing each document as a (normalized) count of terms in this list. For example, to measure economic activity, we might construct a list which includes ‘growth’. But clearly other words are also used to discuss activity, and choosing these involves numerous subjective judgments. More subtly, ‘growth’ is also used in other contexts, such as in describing wage growth as a factor in inflationary pressures, and accounting for context with dictionary methods is practically very difficult.

For these reasons, we instead adopt a machine learning approach to dimensionality reduction that alleviates these concerns by using Latent Dirichlet Allocation (LDA). The rest of the section discussing LDA as a statistical model, then describes how we estimate it and build communication measures from its output. Many details are left out, and are filled in by the accompanying online technical appendix.

5.1 Statistical model

The first important objects in LDA are K *topics*, each of which is a distribution $\beta_k \in \Delta^V$ over the V elements in the vocabulary. The choice of probability distributions is important since it allows the same token to appear in different topics with potentially different weights. Informally, one can think of a topic as a weighted word list that groups together words that all express the same underlying theme. Unlike with dictionary methods, the form of β_k is not imposed on the data ex-ante.

Given topics, the simplest statistical model would be to associate each document with a single topic, yielding a basic mixture model. Instead, LDA is a mixed-membership model in which each document can belong to multiple topics. Formally, this is represented by each document d having its own distribution over topics given by θ_d. Informally, θ_d^k represents the “share” of topic k in document d.

To describe the data generating process, first think of document d as having N_d slots to fill. In the first step, each slot (n, d) is independently allocated a topic assignment $z_{n,d}$ according to the probability vector θ_d. These unobserved topic assignments are latent variables in the model. In the second step, a word is drawn for the nth slot from the topic $\beta_{z_{n,d}}$ that corresponds to the assignment $z_{n,d}$. The probability of observing the data (words and topic assignments) is thus

$$\prod_{d=1}^{D} \prod_{n=1}^{N_d} \sum_{z_{n,d}} \Pr\left[\, w_{n,d} \mid \beta_{z_{n,d}} \,\right] \Pr\left[\, z_{n,d} \mid \theta_d \,\right]. \tag{1}$$

Importantly, LDA reduces the dimensionality of each document substantially. In the bag-of-words model, documents live in a V-dimensional space. After estimating LDA, one obtains a representation of each document in terms of the (estimated) θ_d, which lives in the $K-1$ simplex. In our data, this reduces the dimensionality of each document from many thousands to less than 100. Importantly, though, LDA does not ignore any dimensions of variation in the bag-of-words counts since the underlying topics are free to lie anywhere the $V-1$ simplex.

Due to the high dimensionality of the parameter space and the sparsity of the underlying data, topic models generally have a Bayesian formulation. We assign a symmetric Dirichlet prior with K dimensions and hyperparameter α to each θ_d, and a symmetric Dirichlet prior with V dimensions and hyperparameter η to each β_k. Realizations of Dirichlet distributions with X dimensions lie in the $X-1$ 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.

It is also worth locating LDA in the context of machine learning. Broadly speaking, machine learning algorithms (not just those for text mining) either solve supervised or unsupervised learning problems. Supervised learning is the task of taking labeled observations, and using features of the observations to predict those labels. For example, Gentzkow and Shapiro (2010) propose an algorithm for finding which phrases in congressional speeches (a speech is an observation) best predict party affiliation (the party of the speaker is a label). In unsupervised learning, observations have no labels, and the task is to uncover hidden patterns that allow one to structure the observations in some meaningful way. Clustering and factor analysis are examples of unsupervised learning tasks. LDA is an unsupervised learning algorithm, as its goal is to find K meaningful word groupings in the data and to represent each document in terms of these groupings.

5.2 Vocabulary and model selection

Before estimating the model, we need to select which tokens to keep and how to represent them (vocabulary selection), as well as the number of topics K and the values of the hyperparameters α and η (model selection). For vocabulary selection, we drop all tokens containing non-alphabetic characters, remove both common and rare words, and convert the remaining tokens into a common linguistic root through stemming so that, for example, 'preferences', 'preference', and 'prefers' all become 'prefer'. The outcome of stemming need not be an English word.

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'. After this processing, $V = 8,615$ unique and 2,715,586 total tokens remain. Some statements are empty, so we remove them from the dataset, leaving $D = 46,169$ total documents.

For values of the hyperparameters, we follow 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.

Our goal is to organize text into easily interpretable categories, and this informs our choice of K. If one picks too few topics, they tend to mix together underlying themes and become very general, while if one picks too many, topics become highly specific to particular conversational patterns. We settle on models with $K = 50$, which we report in the main text, and $K = 70$, which we report in the appendix. Another common approach to selecting K is cross-validation on out-of-sample data, but this assesses a model's pure predictive power. Since we are not interested in predicting the content of FOMC meetings *per se*, we do not adopt this approach.[21]

5.3 Estimation

For estimation we use the collapsed Gibbs sampling algorithm of Griffiths and Steyvers (2004) (see also Steyvers and Griffiths 2006). Since the Dirichlet prior is conjugate to the categorical distribution, one can easily analytically integrate out the θ_d and β_k parameters from the probability in (1), and express the probability of the data in terms of just the observed words and unobserved topic assignments. This is why sampling is "collapsed". The remaining challenge is to estimate the topic assignments. To do so, we construct a Gibbs sampler with the following features:

1. Randomly allocate to each token in the corpus a topic assignment drawn uniformly from $\{1, \dots, K\}$.
2. For each token, sequentially draw a new topic assignment via multinomial sampling. The probability that token n in document d is assigned to topic k is increasing in:
   (a) The number of other tokens in document d that are currently assigned to k.
   (b) The number of other occurrences of the token $w_{n,d}$ in the entire corpus that are currently assigned to k.
3. Repeat step 2 4,000 times as a burn in phase.
4. Repeat step 2 4,000 more times, and store every 50th sample.

Steps 2a and 2b mean that tokens that regularly co-occur in documents will be grouped together to form topics. Also, step 2a means that tokens within a document will tend to be grouped together into few topics rather than spread across many separate topics. The burn in phase of sampling allows the chain to converge sufficiently, after which we begin

[21] 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."

drawing the samples we use to construct communication measures. Allowing a thinning interval between samples reduces autocorrelation between them. The online technical appendix reports how we assess convergence and the selection of the chain we use for the analysis.

The basic object of interest for the analysis is the predictive distribution $\widehat{\theta}^k_{i,t,s,j}$ that expresses the probability a word spoken by member i in meeting t in section s belongs to topic k, computed at the jth iteration of the chain. Its construction is detailed in online technical appendix. Below we report numerous communication measures constructed from these distributions. In each case, we compute the measure for each iteration in $j \in \{4050, 4100, \ldots, 8000\}$, and then take an average over the 80 samples.

5.4 Estimated topics

One reason for the popularity of LDA is its ability to consistently estimate topics that appear natural despite having no pre-assigned labels. In appendix A 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.[22] 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 judgments, 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.[23] 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).

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

[22] We report the predictive topic distributions at the 8,000th iteration of the Markov chain. The probability that token v appears in topic k is $\widehat{\beta}^v_k = \frac{\eta + n^v_k}{\eta V + N_k}$, where n^v_k is the number of times that token v is assigned to topic k in the corpus, and N_k is the total number of tokens assigned to topic k.

[23] 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.

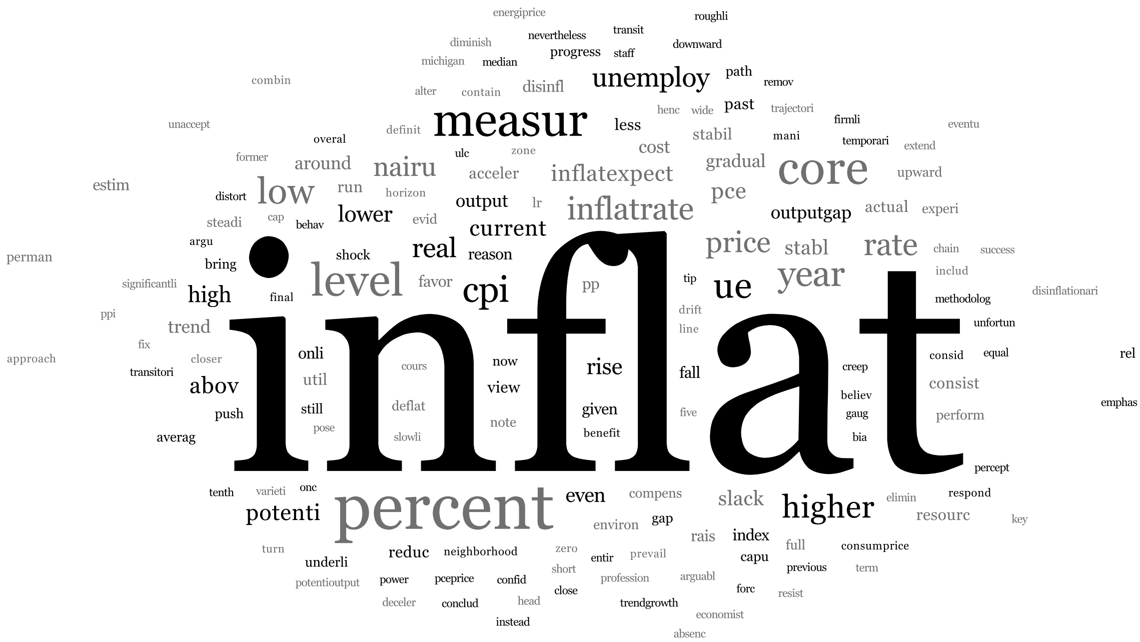


Figure 2: Topic 25—"Inflation"

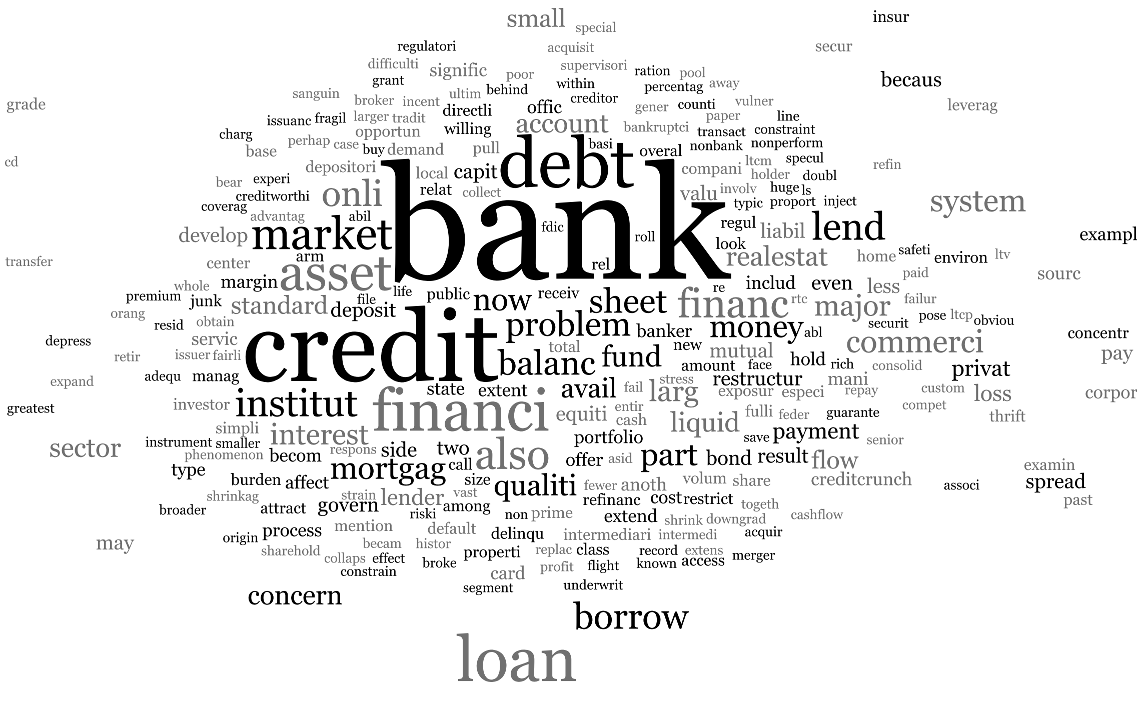


Figure 3: Topic 38—"Banking"

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”

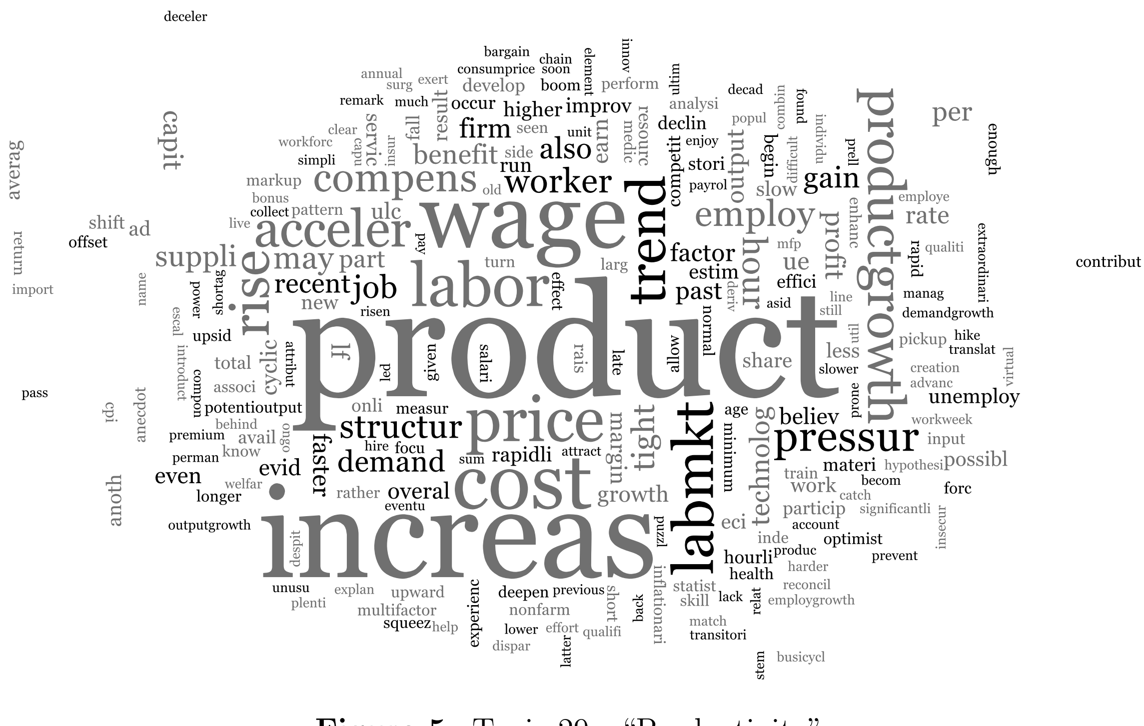


Figure 5: Topic 29—“Productivity”

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.

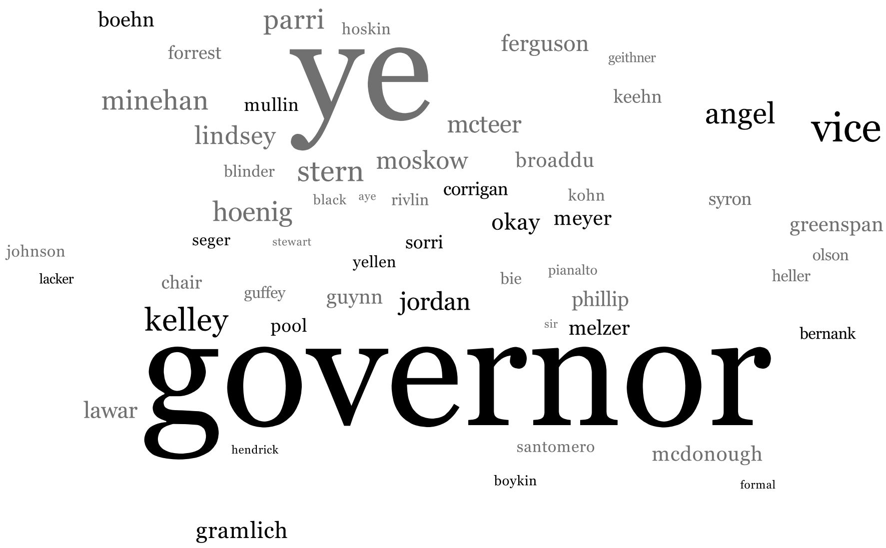


Figure 6: Topic 27—"Discussion topic: FOMC Names"

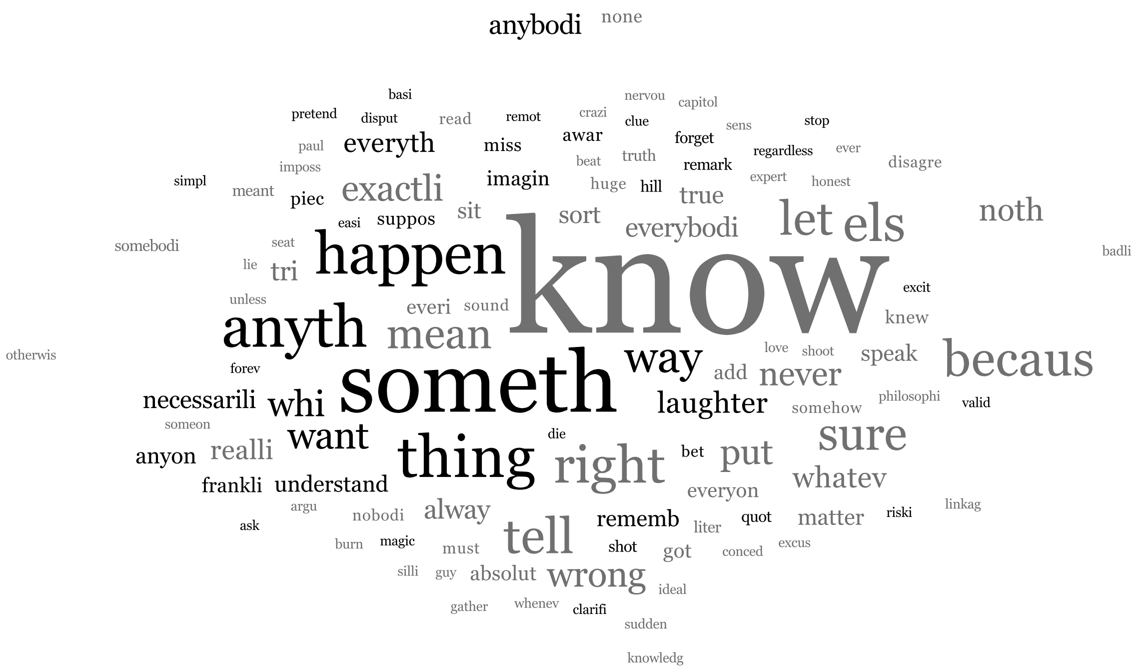


Figure 7: Topic 49—"Discussion topic: General terms"

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.

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 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.

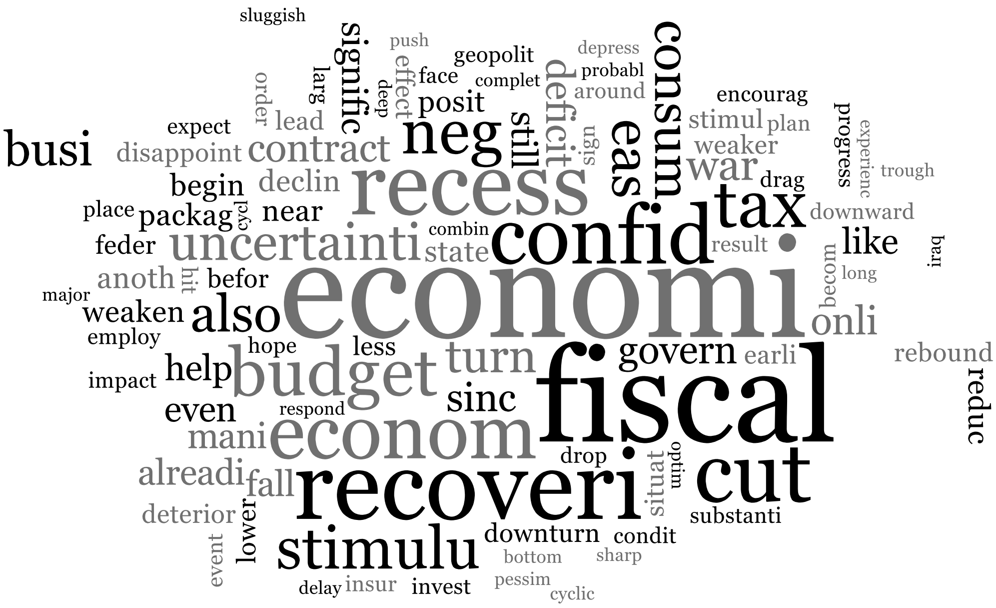
5.5 Connecting topics to external events

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).

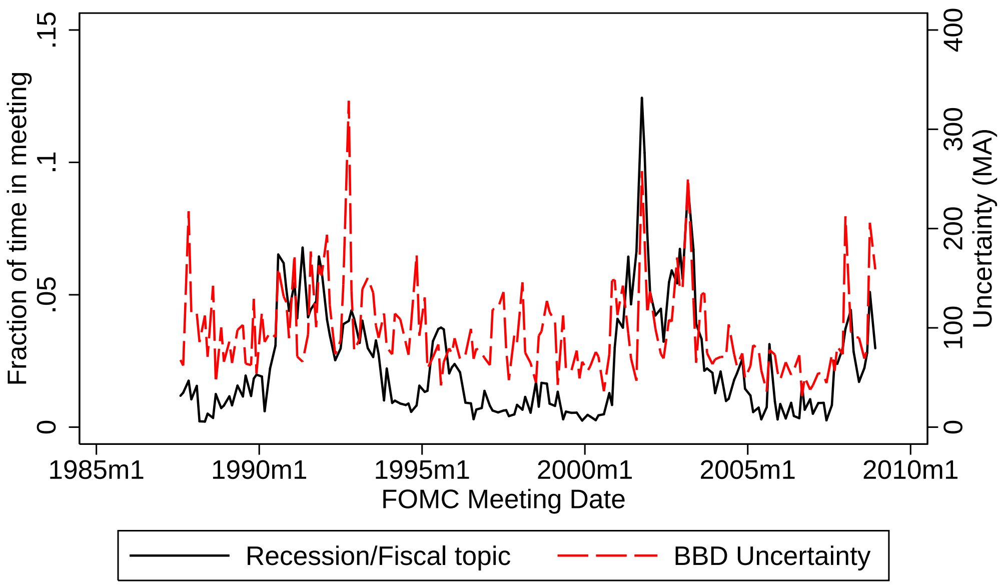
Figure 8 displays the estimated topic most associated with issues of recession and fiscal policy, and plots the amount of time the FOMC as a whole spends on it against the BBD index.[24] The relationship between BBD-measured uncertainty and FOMC attention towards recession/fiscal matters is quite strong, 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

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



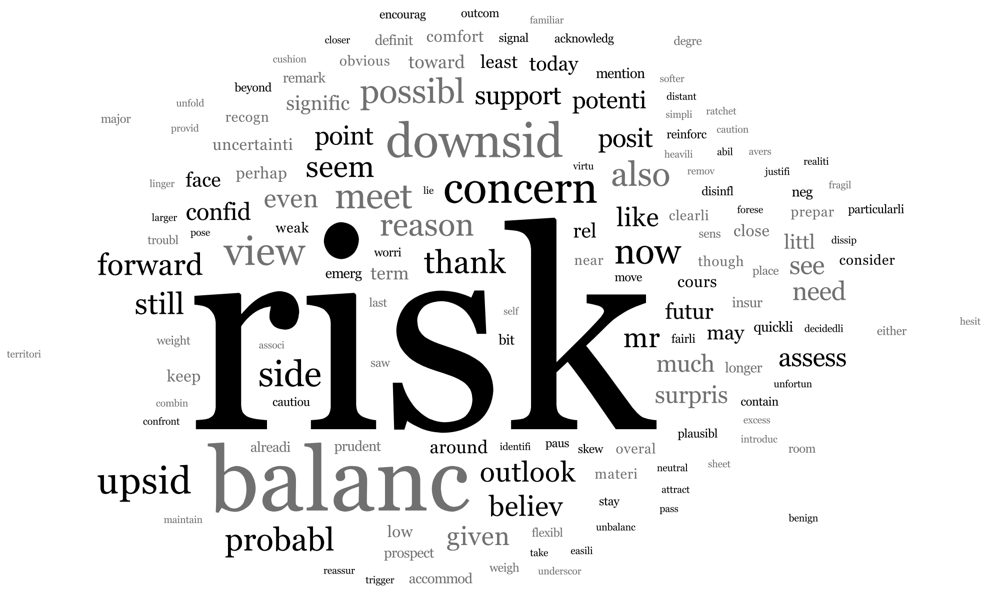
(a) Topic 45—“Recession/Fiscal”



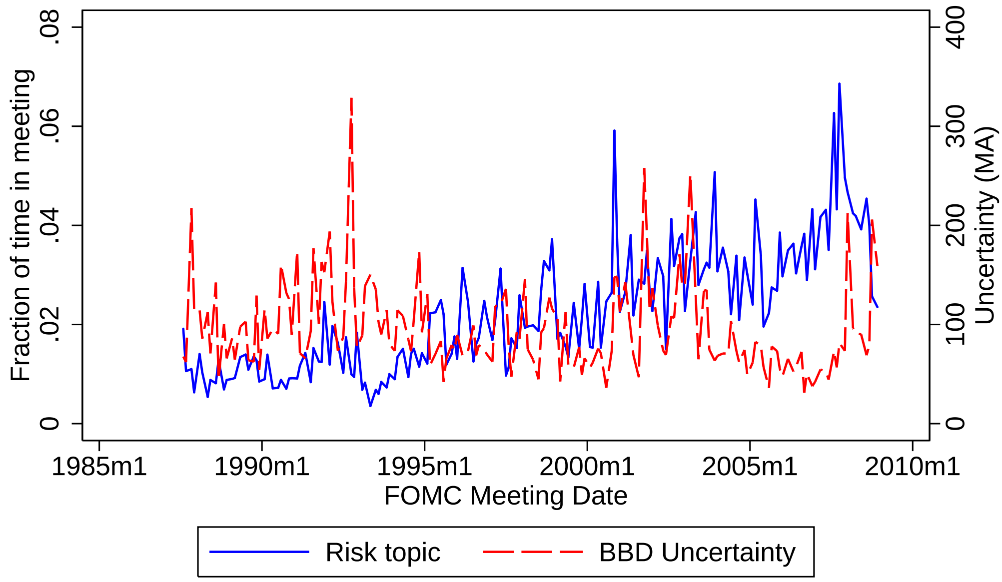
(b) BBD uncertainty and discussion of topic 45

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.



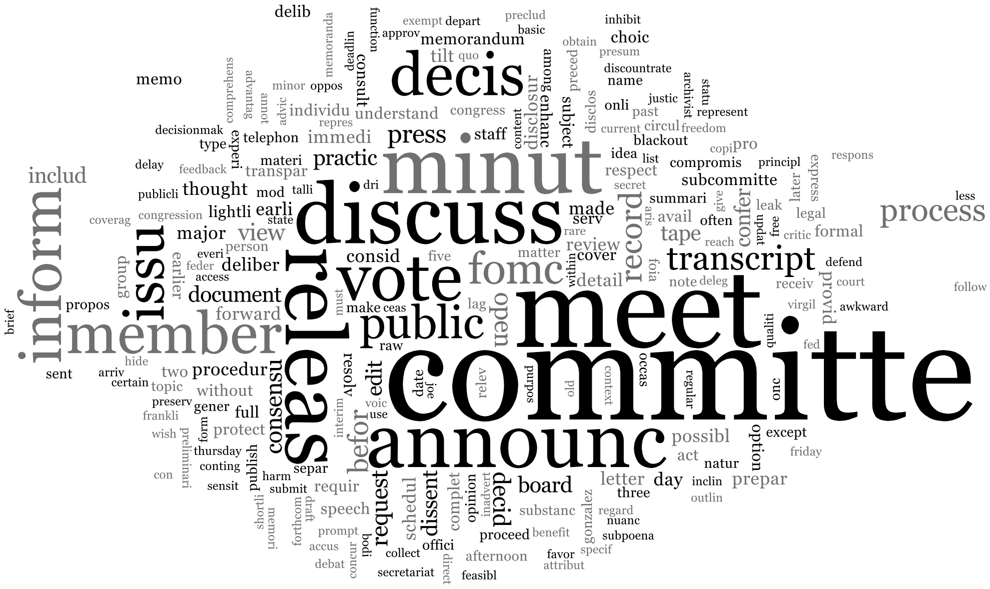
(a) Topic 40—“Risk”



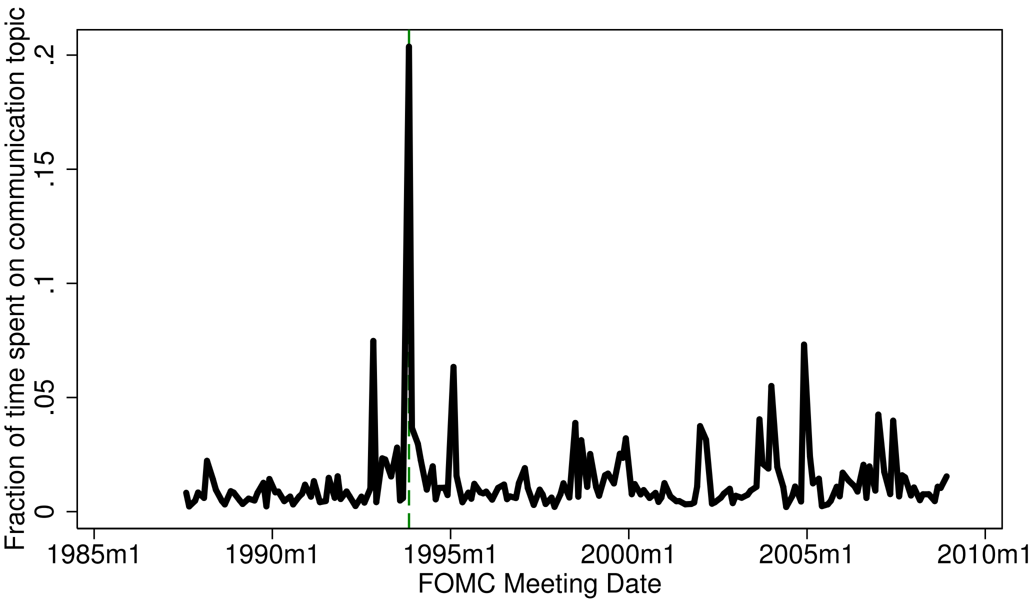
(b) BBD uncertainty and discussion of topic 40

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.

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.

6 Empirical Results

We now present the estimates of the econometric models in section 4 using various measures of communication. We begin by examining broad changes in the nature of deliberation after transparency before turning to statements' quantitative content. We then compare each member's topic distribution with Greenspan's.

6.1 Transparency and deliberation

To begin the exploration of whether and how deliberation changed with transparency, we first examine the evolution of word, statement, and question counts. Table 2 presents estimates of (DinD) using these measures. The key coefficient is on 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 FOMC2 less experienced members make significantly fewer interjections and ask fewer questions. We interpret the drop in statements as reflecting a reduction in back-and-forth dialogue, since open debate would generate many statements as arguments bounced from member to member. Similarly, the reduction in questions reflects a lower willingness to engage with colleagues and staff.

To quantify the economic importance of the statistically significant coefficients in columns (5) and (6), we report what we term the *rookie effect*. The first step in constructing this is to compute the estimated difference between how a member with one year of Fed experience (a rookie) and one with 20 (a veteran) react to transparency. These numbers roughly correspond to modes of the distribution of experience presented in figure 1. For example, the estimated coefficient of 0.11 in column (5) implies that the difference between the number of statements a rookie and a veteran make drops by $19 \times 0.11 = 2.09$ after transparency. The second step is to report this difference in terms of a percentile change from the median of the pre-transparency distribution of the dependent variable, which in the case of statements is 2. So, the rookie effect takes one into the first percentile of the distribution, implying a change of 49 percentiles relative to the pre-transparency median. This effect and that for questions are thus particularly dramatic. Throughout the paper, we continue to report the rookie effect for all statistically

Table 2: Diff-in-Diff Results: Count measures of deliberation

Main Regressors	(1) Total Words	(2) Statements	(3) Questions	(4) Total Words	(5) Statements	(6) Questions
D(Trans)	-486***	-5.85**	-2.72	-2,940	82.3	38.9
	[0.000]	[0.010]	[0.139]	[0.268]	[0.293]	[0.117]
Fed Experience	973***	6.38	5.04	232	-5.49	-2.62
	[0.000]	[0.142]	[0.163]	[0.200]	[0.305]	[0.124]
D(Trans) x Fed Experience	0.42	0.026	0.0047	-0.68	0.11***	0.037***
	[0.798]	[0.298]	[0.667]	[0.738]	[0.010]	[0.007]
Constant	-10,240***	-66.7	-55.7	0	0	0
	[0.000]	[0.175]	[0.172]	[.]	[.]	[.]
Unique Members	36	36	36	36	36	36
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	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1148	1148	1138	1138	1138
Rookie effect	-	-	-	-	-49	-49

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 how many percentile points the pre-transparency median member would move if their behaviour changed by 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.

significant interaction terms.

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.

Table 3 shows the results. First, during the policy discussion, rookies devote relatively more attention to economics topics. 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

Table 3: Diff-in-Diff Results: Economics focus and concentration of topics discussed

Main Regressors	(1) Economics	(2) Economics	(3) Herfindahl	(4) Herfindahl
D(Trans)	0.16***	0.064	-0.0038	0.34*
	[0.002]	[0.933]	[0.674]	[0.055]
Fed Experience	-0.28***	0.036	-0.0035	-0.018
	[0.008]	[0.484]	[0.852]	[0.134]
D(Trans) x Fed Experience	0.00019	-0.0014**	0.00061*	-0.00028***
	[0.655]	[0.018]	[0.060]	[0.003]
Constant	3.70***	0	0.15	0
	[0.002]	[.]	[0.483]	[.]
Unique Members	36	36	36	36
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Within Meeting	FOMC1	FOMC2	FOMC1	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1138	1148	1138
Rookie effect	-	21	-20	12

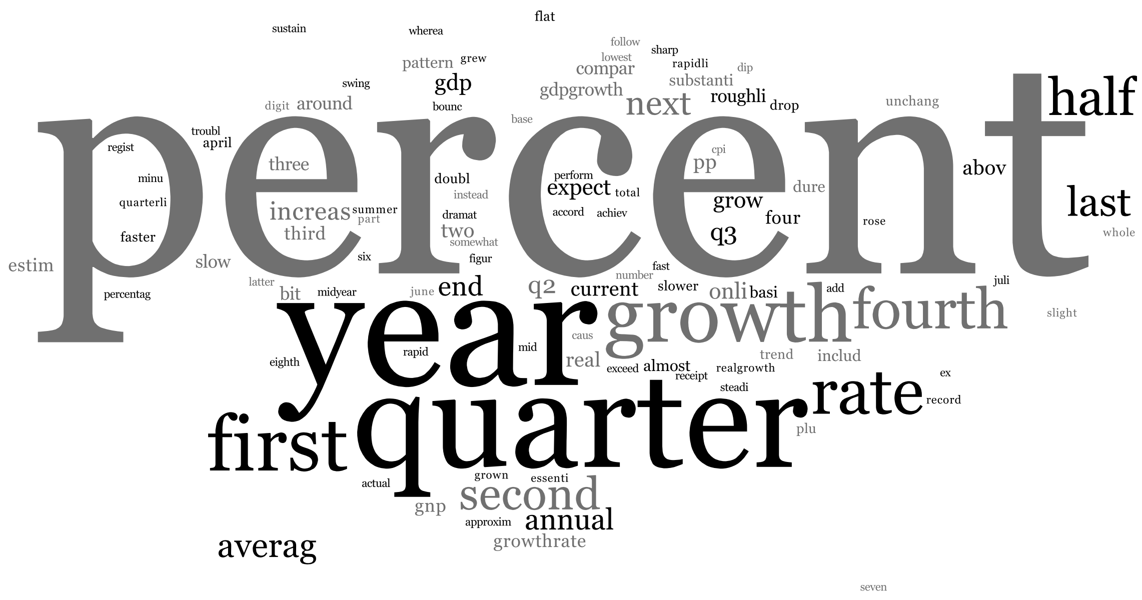
Notes: This table reports the results of estimating (DinD) on variables related to count measures of the discussion. Where the difference in difference is statistically significant, the rookie effect reports how many percentile points the pre-transparency median member would move if their behaviour changed by 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.

FOMC2 since such engagement would force them to touch on viewpoints other than their own.

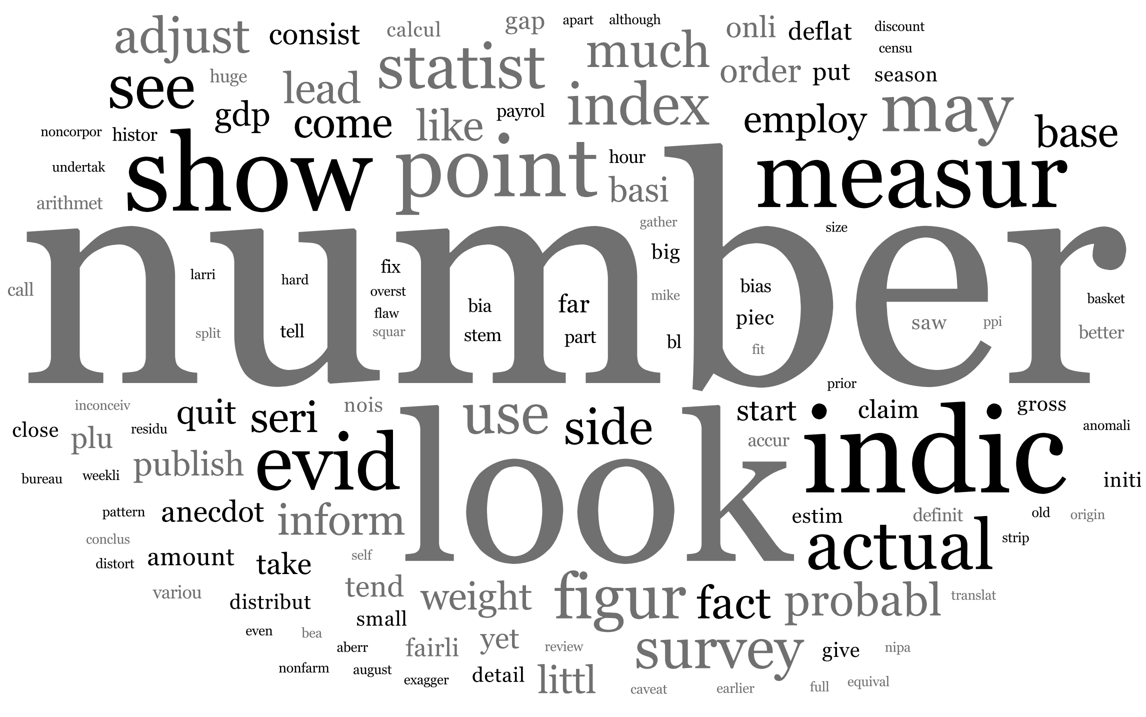
6.2 Transparency and quantitative discussion

As discussed in section 2, a primary channel through which we expect discipline to operate on the FOMC is to encourage especially rookie members to gather additional data between meetings. One would expect such efforts to subsequently appear in the text data in the form of greater reference to numbers and quantitative indicators. 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 using specific references to relevant data, while one who had done their homework would likely bring into the meetings a dossier of evidence on which to draw.

To measure preparation we use two strategies. First, we count the number of tokens in each statement that are numbers (strings that consist solely of numeric characters like '99' and '1' but not tokens like 'one'). Second, we identify two topics from the topic model output that appear to reflect quantitative discussion. These are topics 7 and 11, whose word clouds appear in figures 11a and 11b. The most likely terms in these topics are clearly those would use when discussing data.



(a) Topic 7—"Data Indicators 1"



(b) Topic 11—"Data Indicators 2"

Figure 11: Topics relating to data

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.

Table 4: Discussion of numbers and data indicators

Main Regressors	(1) Numbers	(2) Numbers	(3) Data Topics (7&11)	(4) Data Topics (7&11)
D(Trans)	-8.43***	-20.2	-0.032**	-0.10
	[0.002]	[0.706]	[0.019]	[0.631]
Fed Experience	21.6***	1.83	0.066**	0.010
	[0.001]	[0.618]	[0.032]	[0.485]
D(Trans) x Fed Experience	-0.21***	-0.078***	-0.00071***	-0.00027**
	[0.004]	[0.005]	[0.003]	[0.035]
Constant	-234***	0	-0.69**	0
	[0.002]	[.]	[0.045]	[.]
Unique Members	36	36	36	36
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Within Meeting	FOMC1	FOMC1	FOMC1	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1138	1148	1138
Rookie effect	14	14	17	16

Notes: This table reports the results of estimating (DinD) on variables related to numbers and data indicators. Where the difference in difference is statistically significant, the rookie effect reports how many percentile points the pre-transparency median member would move if their behaviour changed by 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 4 presents the results. Ex ante, we might expect the discipline effect on text to be strongest in FOMC1 since during this section members generally read a prepared statement, while FOMC2 is more extemporaneous. Consistent with this, we find highly significant effects of transparency on the quantitative discussion of rookies in FOMC1, both on the count and topic measures. But we also find that there is a significant increase in quantitative discussion among rookies in FOMC2 as well relative to veterans. This indicates that rookies indeed prepare more between meetings, and not only show this in their scripted statements on the economy, but also in justifying their policy views.

6.3 Transparency and distance to Greenspan

Our final measures of content compare the statements of each FOMC member to those of Alan Greenspan, who is clearly a focal member during the sample. As explained in section 3, Greenspan tended to speak very little during FOMC1 in our sample, so here we limit attention to statements in FOMC2. Recall also that Greenspan speaks first in FOMC2, with the rest of the members following.

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 $\boldsymbol{\chi}_{i,t}$ denote i's conditional probability distribution over economics topics in meeting t

during FOMC2 (30 of the 50 estimated topics are classified economics topics); we are interested in comparing the overlap of $\boldsymbol{\chi}_{i,t}$ with $\boldsymbol{\chi}_{G,t}$, where G is Greenspan's speaker index. There are many ways in the literature to do so, but we focus on three different measures:[25]

1. *Dot product similarity*: $DP_{it} = \sum_k \chi^k_{G,t}\chi^k_{i,t}$. Although $\boldsymbol{\chi}_{i,t}$ 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 performs well in conversational speech data when each statement is composed of a limited number of topics relative to K. The statistical interpretation is the probability that member i and Greenspan talk about the same topic given if they each discuss just one topic.

2. *Bhattacharyya coefficient*: $BH_{it} = \sum_k \sqrt{\chi^k_{G,t}\chi^k_{i,t}}$. This measures the extent to which two probability distributions overlap, and is widely used in the machine learning literature.

3. *Kullback-Leibler divergence*: $KL_{it} = \sum_k \chi^k_{G,t} \ln\left(\frac{\chi^k_{G,t}}{\chi^k_{i,t}}\right)$. This has strong roots in the information theory literature, and can be interpreted as the amount of information lost when $\chi^k_{i,t}$ is used to approximate Greenspan's distribution.

Before presenting results, we first establish the relationship between topic overlap and policy preferences. One potential criticism of interpreting closeness in topic space as herding is that, because we do not measure sentiment, talking about the same topics as the Chair is not the same as agreeing with the Chair's views. For example, the Chair may spend a lot of time talking about inflation being under control, while a subsequent rookie spends a lot of time talking about inflation being a major risk. Both talk about similar topics, but are not in agreement. To examine this possibility, we correlate our measures of topic overlap with the voiced dissent data coded by Meade (2005), and present results in table 5. Columns (1) and (2) show that increased similarity between i and the Chair lowers the probability that i dissents in voice, while column (3) shows that increased distance from the Chair positively predicts dissent. We are therefore reassured that our measures capture agreement and disagreement about the important dimensions of the policy decision.

Table 6 presents the results of estimating DinD with the overlap measures as dependent variables. The main result is that after transparency, inexperienced members speak

[25] One complication is that some members in some meetings have very short statements. In these cases, using their predictive topic distributions derived from LDA to measure content is problematic since they are essentially uniform. Whenever a speaker has less than five words allocated to economics topics, we replace his predictive distribution with Greenspan's since the implication of a short statement is that he does not disagree with the Chairman's policy view. Moreover, as we show below, distance in topic space correlates with distance in policy space. In all distance regressions we control for very short statements.

Table 5: Relationship between distance and voiced dissent

Main Regressors	(1) D(Voice Dissent)	(2) D(Voice Dissent)	(3) D(Voice Dissent)
D(Non-Voter)	0.0060	0.0072	0.0083
	[0.802]	[0.764]	[0.727]
DP	-1.03*		
	[0.059]		
BH		-0.46***	
		[0.005]	
KL			0.11***
			[0.003]
Constant	0.49***	0.81***	0.34***
	[0.000]	[0.000]	[0.000]
R-squared	0.226	0.229	0.229
Unique Members	35	35	35
Member FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Obs	1194	1194	1194
Type of measure	Similarity	Similarity	Distance

Notes: The table reports the correlation between our three measures of distance, and the voiced dissent variable coded by Meade (2005). Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

Table 6: Overlap of member and Chairman topics

Main Regressors	(1) DP	(2) BH	(3) KL
D(Trans)	0.013***	0.78	1.70
	[0.000]	[0.253]	[0.576]
Fed Experience	-0.011	0.0035	-0.065
	[0.453]	[0.940]	[0.754]
D(Trans) x Fed Experience	-0.00021**	-0.00058**	0.0023*
	[0.037]	[0.033]	[0.055]
Constant	0.17	0	0
	[0.307]	[.]	[.]
Unique Members	35	36	36
Member FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Within Meeting	FOMC2	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1074	1138	1138
Type of measure	Similarity	Similarity	Distance
Rookie effect	12	11	-9

Notes: This table reports the results of estimating (DinD) on variables measuring similarity to or distance from Chairman Greenspan. Where the difference in difference is statistically significant, the rookie effect reports how many percentile points the pre-transparency median member would move if their behaviour changed by 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.

more like Greenspan in FOMC2—for each measure, the experienced either become less similar or more distance from Greenspan after transparency. This shows a systematic difference between rookies and veterans regarding their willingness to deviate from the topics Greenspan discusses initially.

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, which is consistent with the previous empirical literature on career concerns.

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 expected effects of transparency are an increase in discipline and a change in conformity. Though theoretical models do not uniquely predict whether conformity or non-conformity will increase, our results on inexperienced members' distance from Greenspan together with their disengaging more during debate in FOMC2 clearly point towards fitting in being more important than standing out. Therefore we focus on conformity increasing (rather than non-conformity increasing).

Table 7: Evidence for career concerns

<table>
<tr><th>Discipline</th><th>Conformity</th></tr>
<tr><td>↑ economics topic coverage in FOMC1</td><td>↓ statements in FOMC2</td></tr>
<tr><td>↑ numbers in FOMC1</td><td>↓ questions in FOMC2</td></tr>
<tr><td>↑ references to data topics in FOMC1</td><td>↓ economics topic coverage in FOMC2</td></tr>
<tr><td></td><td>↓ distance from Greenspan in FOMC2</td></tr>
<tr><td colspan="2">↑ economics topic percentage in FOMC2</td></tr>
</table>

Table 7 categorizes the main difference-in-differences results from the previous section in terms of their support for discipline or conformity.[26] 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

[26] In appendix section D 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.

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. This section explores the overall effect on informativeness after the shift to transparency by measuring changes in influence.[27]

7.1 Influence

The basic motivation behind our measurement of influence is the following: as i's 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 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 web pages, 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 measures introduced in section 6 for measuring proximity to Greenspan.[28] For concreteness, begin by considering the dot product—the construction using the Bhattacharyya coefficient is identical. 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

[27] An earlier version of this paper,Hansen, McMahon, and Prat (2014), also explored changes in other aspects of policy that coincided with the change in transparency and are consistent with our findings.

[28] We do not use the Kullback-Leibler divergence because its interpretation as an influence measure is unclear. For example, if member i is distant from member j, and member j is distant from member k, it does not follow that i is distant from k.

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 $W_{it}^{G} \equiv W_{it} \times \mathbf{W}_t'(i, G)$, 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.[29] 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_{it}^{G} 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.[30]

Table 8 displays the results for influence. All measures show that rookies become more influential on debate after transparency, although there is some variation in significance depending on the proximity measure we choose. The dot product picks up a highly significant increase in overall influence across meetings, while the Bhattacharyya coefficient picks up a significant increase within meetings. The results on influence on Greenspan are similar, with more statistical and economic significance. During our sample, the FOMC operated rather like an advisory committee with Greenspan as a single decision maker. Other FOMC members offered opinions and disagreement, but rarely if ever could implement a policy that Greenspan did not favor. In this sense, our results on increased influence on Greenspan is particularly important, since they indicate that rookies had increased influence over policy.

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 the future. One natural explanation is that what inexperienced members say after transparency is more worth listening to than before. Another 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.

[29] 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.

[30] Table C.1 in the appendix presents a ranking of members by their overall inter-meeting influence (left panel) and their inter-meeting influence on Greenspan (right panel).

Table 8: Influence

(a) Overall Influence

Main Regressors	(1) W_D	(2) A_D	(3) W_{BH}	(4) A_{BH}
D(Trans)	0.0028***	0.0019**	0.0030***	-0.000057
	[0.000]	[0.036]	[0.000]	[0.810]
Fed Experience	-0.0039	-0.010*	-0.0012	0.00039
	[0.380]	[0.057]	[0.534]	[0.830]
D(Trans) x Fed Experience	-0.000015	-0.00019***	-0.000042**	-0.000012
	[0.732]	[0.009]	[0.041]	[0.439]
Constant	0.096*	0.17***	0.064***	0.051**
	[0.068]	[0.006]	[0.005]	[0.018]
Unique Members	35	32	35	32
Within Meeting	Intra	Inter	Intra	Inter
Obs	1074	1039	1074	1039
Rookie effect	-	17	7	-

(b) Influence on Greenspan

Main Regressors	(1) W_D^G	(2) A_D^G	(3) W_{BH}^G	(4) A_{BH}^G
D(Trans)	0.00031***	0.000032	0.00030***	-0.000012
	[0.009]	[0.814]	[0.000]	[0.711]
Fed Experience	0.00023	-0.0017*	-0.000079	-0.000092
	[0.748]	[0.054]	[0.785]	[0.637]
D(Trans) x Fed Experience	-6.6e-06	-0.000022***	-5.3e-06***	-1.4e-06
	[0.198]	[0.004]	[0.000]	[0.319]
Constant	0.00012	0.022**	0.0034	0.0040*
	[0.988]	[0.030]	[0.309]	[0.080]
Unique Members	35	32	35	32
Within Meeting	Intra	Inter	Intra	Inter
Obs	1074	1039	1074	1039
Rookie effect	-	18	8	-

Notes: This table reports the results of estimating (DinD) on measures of member influence derived from our LDA estimation. The upper table reports the results of estimating influence on the average of the whole committee and the lower table reports the results on influence on Chairman Greenspan. As with earlier tables, all regressions contain member and time fixed effects (rows reporting their inclusion are omitted to save space). Where the difference in difference is statistically significant, the rookie effect reports how many percentile points the pre-transparency median member would move if their behaviour changed by 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 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: even though they are less engaged in the debates, rookies become more influential in shaping discussion. This finding alone does not imply that US monetary policymaking improved after 1993 as a result of transparency, but does suggest that transparency was responsible for changing policymakers' information sets in a meaningful way.

Given that macroeconomic policymakers tended to focus on the negative effects on deliberation of too much transparency, our finding of significant discipline effects is important for central bank design. The main policy implication of our results is that central banks designers should seek to maximize the discipline effect and minimize the conformity effect given that both are present in the data and have clear welfare implications. One recent instance of how this insight has already impacted institutional design comes from the Bank of England. As mentioned in the introduction, the Bank of England (as well as the ECB) have recently been reviewing their policies of disclosure about information from their policy discussions. In December 2014, the Bank of England published the independent review authored by former FOMC Governor Kevin Warsh, who writes that our central bank design recommendations "motivate some of the Review's ultimate recommendations" (Warsh 2014, p 34).

In particular, he examined the nature of the discussions at the Bank of England's Monetary Policy Committee's (MPC) monthly two-day meetings. He noted that an informal norm has emerged in which MPC members spend the first day in free-flowing debate about the economy and the second day reading from prepared scripts that explain their policy stances. Thus, publishing transcripts from the second day does not seem to have much downside: the fact that members do all their thinking outside of that day's discussion means that conformity is unlikely to be relevant, while discipline should motivate them to form more coherent, logical and evidence-based arguments in advance. On the other hand, publishing transcripts of the first day runs a real risk of making debate sterile due to conformity, as our results have shown. Due to this reasoning, Warsh ultimately recommended publishing MPC meetings' second day transcripts (with an eight-year delay) but not their first day transcripts, a change the Bank of England has committed to implementing from August 2015. We hope that the findings in this paper can contribute to such improvements to the policymaking process in other contexts in the future, and motivate greater research into topics of deliberation, as well as communication, more generally.

References

Acosta, J. M. (2014): "FOMC Responses to Calls for Transparency: Evidence from the Minutes and Transcripts Using Latent Semantic Analysis," Mimeograph, University of Stanford.

Apel, M., and M. Blix Grimaldi (2012): "The Information Content of Central Bank Minutes," Working Paper Series 261, Sveriges Riksbank (Central Bank of Sweden).

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.

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

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

Bligh, M., and G. D. Hess (2006): "A Quantitive Assessment of the Qualitative Aspects of Chairman Greenspan's Communications," Computing in Economics and Finance 2006 213, Society for Computational Economics.

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.

Boukus, E., and J. V. Rosenberg (2006): "The information content of FOMC minutes," Mimeograph, Federal Reserve Bank of New York.

Chappell, H. W., T. M. Havrilesky, and R. R. McGregor (2000): "Monetary Policy Preferences Of Individual Fomc Members: A Content Analysis Of The Memoranda Of Discussion," *The Review of Economics and Statistics*, 79(3), 454–460.

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 of *MIT Press Books*. The MIT Press.

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.

Gentzkow, M., and J. M. Shapiro (2010): "What Drives Media Slant? Evidence From U.S. Daily Newspapers," *Econometrica*, 78(1), 35–71.

Geraats, P. M. (2002): "Central Bank Transparency," *Economic Journal*, 112(483), 532–565.

——— (2009): "Trends in Monetary Policy Transparency," *International Finance*, 12(2), 235–268.

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.

Grimmer, J. (2010): "A Bayesian hierarchical topic model for political texts: Measuring expressed agendas in Senate press releases," *Political Analysis*, 18(1), 1–35.

Hansen, S., M. McMahon, and A. Prat (2014): "Transparency and Deliberation within the FOMC: a Computational Linguistics Approach," CEPR Discussion Papers 9994, C.E.P.R. Discussion Papers.

Hansen, S., M. McMahon, and C. Velasco (2014): "How Experts Decide: Preferences or Private Assessments on a Monetary Policy Committee?," *Journal of Monetary Economics*, 67(October), 16–32.

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

Hendry, S. (2012): "Central Bank Communication or the Media's Interpretation: What Moves Markets?," Working Papers 12-9, Bank of Canada.

Hendry, S., and A. Madeley (2010): "Text Mining and the Information Content of Bank of Canada Communications," Working Papers 10-31, Bank of Canada.

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

Hong, H., J. D. Kubik, and A. Solomon (2000): "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31(1), 121–144.

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.

Lindsey, D. E. (2003): *A Modern History of FOMC Communication: 1975-2002*. The Federal Reserve Board of Governors.

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).

Lucca, D. O., and F. Trebbi (2009): "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," NBER Working Papers 15367, National Bureau of Economic Research, Inc.

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.

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.

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.

Warsh, K. (2014): "Transparency and the Bank of England's Monetary Policy Committee," Discussion paper, Bank of England.

Appendix For Online Publication

A Estimated Topics

Table A.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 A.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

B Diff results

The most straightforward empirical strategy to examine the effects of the exogenous (to the committee's deliberation) transparency change on deliberation is to estimate a "diff" regression of the following form:

$$y_{it} = \alpha_i + \gamma D(Trans)_t + \lambda X_t + \varepsilon_{it}, \qquad \text{(DIFF)}$$

where y_{it} is the measure of interest (such as a measure of the breadth of topic coverage by member i in meeting t), $D(Trans)$ is the transparency dummy (1 after November 1993, 0 before), and X_t is a vector of macro controls for the meeting at time t. The main difference with this regression specification is that we do not allow for effects of experience, and we do not control for time fixed effects.

In tables B.1 and B.2, we report the results of estimating (DIFF) for each of the variables used in the diff-in-diff except for the influence measures. We do not include estimates for the influence measures because the coefficient γ captures the average effect of the change in transparency and because our measures of influence capture relative influence, it makes no sense for these to change on average.

For most of the estimates, our diff and diff in diff effects go in the same direction. That is, the rookies who are more subject to the career concerns react even more in the direction that everyone else reacts. There are two measures where this is not true. The first is that we find that the average member becomes more narrow in terms of the topics they discuss in FOMC1, while the diff-in-diff highlights that rookies become relatively more broad in FOMC1. The second difference is the overlap with the topics that Chairman Greenspan discusses. On average, we find that FOMC members become on average less similar to Chairman Greenspan even though rookies become relatively closer to him. Moreover, there are some effects where the average response is significant but we find no diff-in-diff, (word spoken in FOMC1 increase on average but not differentially for rookies), or vice versa (on average, there is no change in the discussion of data in FOMC2 even though rookies discuss them relatively more).

However, the problem with the diff analysis remains that the timing of transparency changes may have coincided with other changes that we cannot control for. This means the estimated effect may capture these other effects. This makes it impossible to disentangle the different effects and so it is much harder to interpret the results in the context of a career concerns (or other) framework.

Table B.1: Diff results: The average effect of Transparency I

(a) Count measures of deliberation

Main Regressors	(1) Total Words	(2) Statements	(3) Questions	(4) Total Words	(5) Statements	(6) Questions
D(Trans)	52.3*	-0.36	-0.00093	35.3	-1.98**	-0.67***
	[0.081]	[0.324]	[0.990]	[0.137]	[0.011]	[0.006]
D(Short)	-71.9***	-0.30**	-0.11	-178***	-1.18***	-0.40***
	[0.000]	[0.014]	[0.327]	[0.000]	[0.000]	[0.000]
D(NBER)	14.4	-0.38	-0.075	-30.6	0.46	-0.068
	[0.571]	[0.208]	[0.665]	[0.259]	[0.482]	[0.752]
BBD uncertainty	0.22	-0.0032	0.00021	-0.042	-0.0093**	-0.0042***
	[0.252]	[0.506]	[0.870]	[0.872]	[0.035]	[0.000]
D(2 day)	34.5	1.39**	0.57**	59.5	-0.13	0.071
	[0.115]	[0.048]	[0.020]	[0.219]	[0.841]	[0.733]
Constant	655***	4.59***	1.13***	329***	6.31***	1.69***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Unique Members	36	36	36	36	36	36
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No	No
Within Meeting	FOMC1	FOMC1	FOMC1	FOMC2	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1148	1148	1138	1138	1138
Trans effect	7	-	-	-	-49	-49

(b) Economics focus and concentration of topics discussed

Main Regressors	(1) Economics	(2) Economics	(3) Herfindahl	(4) Herfindahl
D(Trans)	0.069***	0.028***	0.0055**	0.0025**
	[0.000]	[0.000]	[0.042]	[0.026]
D(Short)	-0.0015	0.016***	0.0062**	-0.031***
	[0.830]	[0.007]	[0.014]	[0.000]
D(NBER)	0.0094**	0.0095**	-0.00018	-0.0049***
	[0.015]	[0.017]	[0.864]	[0.000]
BBD uncertainty	-3.5e-06	0.000056	-0.000065***	1.1e-07
	[0.931]	[0.143]	[0.000]	[0.991]
D(2 day)	-0.0042	0.0024	0.000025	0.0038*
	[0.289]	[0.599]	[0.992]	[0.060]
Constant	0.58***	0.56***	0.11***	0.066***
	[0.000]	[0.000]	[0.000]	[0.000]
Unique Members	36	36	36	36
Member FE	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Within Meeting	FOMC1	FOMC2	FOMC1	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1138	1148	1138
Trans effect	25	22	9	6

Notes: These tables report the results of estimating (DIFF). Where the coefficient on D(Trans) is significant, the transparency effect reports how many percentile points the pre-transparency average member would move if their behaviour changed by the average effect of transparency; this is similar to the Rookie effect we report in the main text. Coefficients are labeled according to significance (*** $p<0.01$, ** $p<0.05$, * $p<0.1$) while brackets below coefficients report p-values.

Table B.2: Diff results: The average effect of Transparency II

(a) Discussion of numbers and data indicators

Main Regressors	(1) Numbers	(2) Numbers	(3) Data Topics (7&11)	(4) Data Topics (7&11)
D(Trans)	3.56***	1.60***	0.0088***	-0.000060
	[0.004]	[0.001]	[0.004]	[0.962]
D(Short)	-1.74**	-1.83***	0.0013	0.0048***
	[0.049]	[0.000]	[0.694]	[0.000]
D(NBER)	-1.00	-0.64	-0.00059	-0.0012
	[0.197]	[0.175]	[0.715]	[0.185]
BBD uncertainty	0.0033	0.00018	-7.1e-06	-0.000040***
	[0.527]	[0.969]	[0.725]	[0.001]
D(2 day)	1.44**	1.08*	-0.00042	0.0020
	[0.044]	[0.079]	[0.866]	[0.184]
Constant	7.93***	2.20***	0.045***	0.040***
	[0.000]	[0.004]	[0.000]	[0.000]
Unique Members	36	36	36	36
Member FE	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Within Meeting	FOMC1	FOMC2	FOMC1	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1148	1138	1148	1138
Trans effect	14	14	11	-

(b) Overlap of member and Chairman topics

Main Regressors	(1) DP	(2) BH	(3) KL
D(Trans)	-0.00082	-0.025***	0.12***
	[0.569]	[0.001]	[0.002]
D(Short)	0.045***	0.15***	-0.62***
	[0.000]	[0.000]	[0.000]
D(NBER)	-0.0023**	0.016***	-0.070***
	[0.037]	[0.001]	[0.001]
BBD uncertainty	3.4e-06	0.000046	-0.00023
	[0.860]	[0.418]	[0.401]
D(2 day)	-0.0022*	-0.015*	0.084**
	[0.062]	[0.073]	[0.038]
Constant	0.046***	0.86***	0.58***
	[0.000]	[0.000]	[0.000]
Unique Members	36	36	36
Member FE	Yes	Yes	Yes
Time FE	No	No	No
Within Meeting	FOMC2	FOMC2	FOMC2
Sample	89:11-97:09	89:11-97:09	89:11-97:09
Obs	1138	1138	1138
Type of measure	Similarity	Similarity	Distance
Trans effect	-	-21	22

Notes: These tables report the results of estimating (DIFF). Where the coefficient on D(Trans) is significant, the transparency effect reports how many percentile points the pre-transparency average member would move if their behaviour changed by the average effect of transparency; this is similar to the Rookie effect we report in the main text. Coefficients are labeled according to significance (*** p<0.01, ** p<0.05, * p<0.1) while brackets below coefficients report p-values.

C Influence Ranking

In table C.1, we present a ranking of members by their overall influence (left panel) and their influence on Greenspan (right panel). 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 ($A_{BH}/A_{BH}^G = \alpha_{it}+\delta_t+\epsilon_{it}$). We report, and base the ranking on, the member-fixed effects from this regression.

This table shows that members who are highly influential overall tend to exhibit influence over Chairman Greenspan. However, there are some members who exhibit greater influence over the committee overall than they do over Chairman Greenspan (such as Governor Larry Meyer). Interestingly, while Chairman Greenspan is a good predictor of what Chairman Greenspan will subsequently talk about, other FOMC members seem to influence future Chairman Greenspan even more. Perhaps surprisingly Chairman Greenspan is found to exhibit relatively 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.

Table C.1: Inter-Meeting Influence Measures by Member - $A_{BH}\&A^G_{BH}$

Speaker	Meetings in Sample	Overall Influence Fixed Effect	Overall Influence Average	Speaker	Meetings in Sample	Greenspan Influence Fixed Effect	Greenspan Influence Average
GUFFEY	15	0.0013	0.0612	GUFFEY	15	0.00134	0.00393
BLACK	24	0.0006	0.0582	BLACK	24	0.00059	0.00348
RIVLIN	11	0.0023	0.0576	CORRIGAN	29	0.00041	0.00334
KEEHN	38	0.0012	0.0572	KEEHN	38	0.00125	0.00333
CORRIGAN	29	0.0004	0.0567	RIVLIN	11	0.00228	0.00324
GUYNN	14	0.0009	0.0565	SYRON	35	0.00041	0.00324
MEYER	9	-0.0004	0.0564	KELLEY	64	-0.00002	0.00321
FORRESTAL	49	0.0009	0.0563	FORRESTAL	49	0.00092	0.00320
MOSKOW	25	0.0014	0.0562	GREENSPAN	64	-0.00174	0.00318
SYRON	35	0.0004	0.0561	GUYNN	14	0.00085	0.00317
HOENIG	49	0.0016	0.0557	MOSKOW	25	0.00140	0.00313
BOYKIN	10	0.0013	0.0557	BOYKIN	10	0.00135	0.00309
MELZER	64	0.0002	0.0556	SEGER	11	0.00094	0.00309
MINEHAN	26	0.0009	0.0556	MELZER	64	0.00017	0.00309
KELLEY	64	0.0000	0.0555	HOENIG	49	0.00160	0.00309
SEGER	11	0.0009	0.0554	ANGELL	34	-0.00047	0.00308
MCDONOUGH	34	0.0002	0.0553	MINEHAN	26	0.00086	0.00305
BROADDUS	39	0.0008	0.0553	PHILLIPS	47	0.00127	0.00304
ANGELL	34	-0.0005	0.0552	BOEHNE	63	-0.00005	0.00303
PARRY	64	0.0008	0.0552	MCTEER	54	-0.00020	0.00302
PHILLIPS	47	0.0013	0.0551	PARRY	64	0.00084	0.00302
STERN	62	0.0007	0.0547	MCDONOUGH	34	0.00023	0.00302
YELLEN	20	0.0011	0.0547	MEYER	9	-0.00044	0.00300
LAWARE	43	0.0006	0.0544	LAWARE	43	0.00057	0.00300
BOEHNE	63	-0.0001	0.0543	STERN	62	0.00066	0.00300
MCTEER	54	-0.0002	0.0541	BROADDUS	39	0.00081	0.00299
MULLINS	29	0.0005	0.0539	MULLINS	29	0.00051	0.00290
GREENSPAN	64	-0.0017	0.0537	YELLEN	20	0.00113	0.00289
HOSKINS	15	-0.0010	0.0536	LINDSEY	41	0.00011	0.00288
BLINDER	13	-0.0003	0.0534	HOSKINS	15	-0.00100	0.00283
LINDSEY	41	0.0001	0.0533	JORDAN	45	-0.00086	0.00282
JORDAN	45	-0.0009	0.0531	BLINDER	13	-0.00027	0.00279
JOHNSON	5	-0.0023	0.0504	JOHNSON	5	-0.00234	0.00252

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 ($a_{it}/a^G_{it} = \alpha_{it} + \delta_t + \epsilon_{it}$).

D Robustness analysis

In tables D.1-D.3 below, we explore the robustness of the main diff-in-diff results presented in the main text. In each table we report the estimated rookie effect. The first line of each table replicates the baseline results from the main text for comparison. Where the main diff-in-diff coefficient (ϕ, on the "D(Trans) $\times$ Fed Experience" regressor) is not significant, we do not report a rookie effect. In the tables we :

1. Remove 1993 observations from the baseline sample;
2. Dropping four FOMC members from the analysis;
3. Run a placebo test on the change in transparency;
4. Use a 70-topic model, rather than the 50-topic model used in the baseline.

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 some FOMC members may have already known of the transcripts and started to adjust their behavior. The estimated rookie effect, where significantly different from zero at 10% level, are shown in the second row of results tables (D.1-D.3). Most of the results are virtually unchanged. Where there are some differences, these are driven mainly from the change in the precision of the estimates; the estimated diff-in-diff coefficient on the $D(Trans) \times FedExp_{i,t}$ regressor is very similar. For example, although the rookie effect on Bhattacharyya similarity is not statistically significant when we drop 1993, the ϕ coefficient estimate is relatively similar (-0.00051 compared to -0.00058 in the baseline estimates).

Table D.1: Comparison of results for different robustness checks I

Rookie effect	(1) FOMC1 Total Words	(2) FOMC1 Statements	(3) FOMC1 Questions	(4) FOMC2 Total Words	(5) FOMC2 Statements	(6) FOMC2 Questions
Baseline	-	-	-	-	-49	-49
Excluding 1993	-	-	2	-	-49	-49
Dropping some members	-	-	-	-	-49	-49
Placebo Estimates	-	6	4	-	-	-
70-topic model	-	-	-	-	-49	-49

Notes: This table reports, for a variety of robustness tests, the rookie effect as reported in the main text.

The second robustness exercise is to keep the baseline sample window, but remove four FOMC members who knew of the written record in advance of October 1993. The members that we drop are Presidents Boehne and Melzer, and Governors Mullins and

Table D.2: Comparison of results for different robustness checks II

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Rookie effect	Numbers	Numbers	Data Topics (7&11)	Data Topics (7&11)	Economics	Economics	Herfindahl	Herfindahl
Baseline	13	13	17	16	-	24	-19	10
Excluding 1993	20	13	23	13	-	28	-	-
Dropping some members	-	13	13	29	-	22	-21	10
Placebo Estimates	-	-	-5	-	-	-	-	-
70-topic model	13	13	20	20	-	20	-	10

Notes: This table reports, for a variety of robustness tests, the rookie effect as reported in the main text.

Table D.3: Comparison of results for different robustness checks III

Rookie effect	(1) DP	(2) BH	(3) KL	(4) W_D	(5) A_D	(6) W_{BH}	(7) A_{BH}	(8) W_D^G	(9) A_D^G	(10) W_{BH}^G	(11) A_{BH}^G
Baseline	14	11	-10	-	16	7	-	7	21	8	-
Excluding 1993	12	-	-	-	8	-	-	-	14	7	-
Dropping some members	19	17	-17	-	16	8	-	12	25	10	6
Placebo Estimates	-	-	-	-	-	-	-	-	-	-	7
70-topic model	12	9	-9	-	10	-	-	17	17	-	-

Notes: This table reports, for a variety of robustness tests, the rookie effect as reported in the main text.

Angell. According to the account in Lindsey (2003), they all found out early about the existence of the transcripts. While none of these members necessarily expected the existence of these records to ever be revealed (let alone that the records would be made public), we believe that showing the results are not driven by their behavior is an important robustness check. These results are reported in the third row of the tables. The main results of the paper remain and, in fact, the estimated effects tend to get larger.

We next 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 impose 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. The results in the fourth row of the tables show that there is no systematic evidence to suggest that the results we find are, for example, driven by trends related to Greenspan's growing power over the tenure of his time on the committee.

Finally, as discussed in the main text, we selected 50 topics in the baseline analysis for interpretability. We have also carried out the analysis using a 70 topic model. The results of the analysis of the 70 topic model are shown in the final row of the tables. 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). Overall, the results are very similar.