TRANSPARENCY AND DELIBERATION WITHIN THE FOMC: A COMPUTATIONAL LINGUISTICS APPROACH*

Stephen Hansen
Michael McMahon
Andrea Prat

How does transparency, a key feature of central bank design, affect monetary policy makers' deliberations? Theory predicts a positive discipline effect and negative conformity effect. We empirically explore these effects using a natural experiment in the Federal Open Market Committee in 1993 and computational linguistics algorithms. We first find large changes in communication patterns after transparency. We then propose a difference-in-differences approach inspired by the career concerns literature, and find evidence for both effects. Finally, we construct an influence measure that suggests the discipline effect dominates. *JEL Codes*: E52, E58, D78.

I. Introduction

In this article we study how transparency, a key feature of central bank design, affects the deliberation of monetary policy makers on the Federal Open Market Committee (FOMC). We ask: what are the effects on internal deliberation of greater external communication about those deliberations? Deliberation takes up the vast majority of the FOMC's meeting time and is seen by former members as important for the committee's decisions (see Meyer 2004, for example), yet it remains little studied beyond anecdotal accounts. Determining how monetary policy committees

*We owe special thanks to Omiros Papaspiliopoulos for numerous helpful discussions on MCMC estimation. We are extremely grateful to the editor, Robert Barro, and three anonymous referees for their insightful suggestions. We thank Francesco Amodio, Andreu Arenas Jal, Andrew Bailey, Guillermo Caruana, Francesco Caselli, Diego Garcia, Refet Gürkaynak, Gilat Levy, Seamus MacGorain, Rick Mishkin, Emi Nakamura, Tommaso Nannicini, Bryan Pardo, Lukas Püttmann, Amar Radia, David Romer, Glenn Rudebusch, Cheryl Schonhardt-Bailey, Jón Steinsson, Dave Stockton, Thomas Wood, and Janet Yellen for helpful discussions. We have also benefited from comments and suggestions by seminar attendees at numerous central banks and academic departments as well as at conferences. We thank Eric Hardy, Marcel Schlepper, and Paul Soto for excellent research assistance. We benefited from financial support from Amazon Web Services, the Bank of England's Research Donations Committee, and a British Academy small grant. Any errors remain ours alone.

The Quarterly Journal of Economics (2018), 801–870. doi:10.1093/qje/qjx045.
Advance Access publication on October 31, 2017.

TABLE I

INFORMATION MADE AVAILABLE BY DIFFERENT CENTRAL BANKS AS OF 2014

	Federal Reserve	Bank of England	European Central Bank
Release minutes?	Yes	Yes	No
Release transcripts?	Yes	No	No

deliberate, and how this depends on central bank design, is 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 I shows, as of 2014 there was heterogeneity across three major central banks in terms of how detailed were the descriptions of policy meetings put on the public record, 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 since 2014. 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 implemented major reforms to its disclosure policy that make it more transparent, including the partial publication of transcripts.

In spite of this increase in transparency, whether more transparency is always beneficial is an open question. In fact, policy makers and scholars have identified potential negative and positive effects of an increase in how much information about the

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; Hansen and McMahon 2016) and uses them to address central bank design questions.

2. Minutes of the ECB’s governing council meetings before 2015 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 due to be released after a 30-year lag.

internal workings of a central bank is revealed to the public. On the negative side, a large body of literature on career concerns emphasizes that transparency leads agents—and monetary policy makers specifically—to distort their decisions either by engaging in herding and conformism (Prat 2005; Visser and Swank 2007) or in antiherding and exaggeration (Prendergast and Stole 1996; Levy 2004, 2007). The empirical literature examining transparency has tended to emphasize this negative effect, in particular conformity. For example, Meade and Stasavage (2008) show that the tendency to dissent from the chair on the FOMC decreases with transparency, while Fehrler and Hughes (forthcoming) provide experimental evidence of conformity. Finally, policy makers appear to worry about the potential for transparency to stifle discussion. Before the Fed had released transcripts, Alan Greenspan expressed his views to the House Banking Committee as follows:

> *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 policy making.* (Greenspan 1993, as reported in Meade and Stasavage 2008; emphasis added)

On the positive side, there is a broad argument that transparency increases the accountability of policy makers, and induces them to work harder and behave better. This argument has been explicitly applied to central banking (see Transparency International 2012, for example), and even the ECB, the least transparent 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] At the same time, there is less overall emphasis on this idea in recent empirical work on central bank transparency than the negative, information-distortion effect. Nevertheless, it

3. See http://www.ecb.europa.eu/ecb/orga/transparency/html/index.en.html.

is wholly consistent with the career concerns literature: in the canonical Holmström (1999) model, the more precise the signal the principal observes about the agent, the higher the equilibrium effort of the agent. This is called the discipline effect in agency theory.

Of course, it is possible that both effects—discipline and information distortion—operate simultaneously. Given that previous research indicates that a key advantage of a committee is the aggregation of heterogeneous views on the economy (Blinder and Morgan 2005; Hansen, McMahon, and Velasco 2014b, for example), one should ask whether, on balance, more disclosure improves or worsens information aggregation. The key innovation of this article is to use text data from FOMC transcripts to explore these issues. Since text is inherently high dimensional, we can explore behavioral responses to transparency in a multitude of ways, which allows us to separate out different theoretical effects more clearly than is possible from a unidimensional object like an interest rate preference.

To study transparency, we use the natural experiment, used originally by Meade and Stasavage (2008), that led to the release of the FOMC transcripts. FOMC meetings have been tape-recorded since the 1970s to prepare minutes. Initially, committee members believed that these tapes were erased afterward. In October 1993, following pressure from politicians, Alan Greenspan discovered and revealed that before being erased the tapes had, in fact, been transcribed and stored in archives all along. The Fed quickly agreed to publish all past transcripts and, a short time later, extended that policy to cover all future transcripts with a five-year lag. This gives us access to periods both when policy makers did and did not believe their deliberations would be public.

To quantify text, we use both basic character counts and latent Dirichlet allocation (LDA) (Blei, Ng, and Jordan 2003)—a machine learning algorithm for probabilistic topic modeling that decomposes documents in terms of the fraction of time spent covering a variety of topics. For our empirical analysis, we first identify topics that are informative about policy preferences, then construct various communication measures from them. FOMC meetings have two major parts related to the monetary policy decision: the economic situation discussion (FOMC1) followed by the monetary policy strategy discussion (FOMC2). A novel aspect of our research is to treat these sections separately. We generate counts and communication measures at the meeting-speaker-section level and use them to make three distinct contributions.

First, controlling for person fixed effects, we show large behavioral responses to transparency along many dimensions. The most striking results are that meetings become less interactive, more scripted, and more quantitatively oriented. This in itself is an important finding because it suggests that transparency matters a great deal for deliberation.

Attributing the average effect of transparency to career concerns is problematic in the FOMC context because the macroeconomy (and therefore discussions surrounding it) evolves over time. Trends and cycles may drive average differences as much as or more than reputation concerns. Our second contribution is to conduct a difference-in-differences analysis with time fixed effects. We use members' experience in monetary policy making as a proxy for career concerns, as theoretical models predict career concerns decline with experience. We find that less experienced members speak more quantitatively in FOMC1 while also discussing a more diverse range of topics; in FOMC2 they make fewer interjections, discuss a less diverse and narrower range of topics, and use less dissenting language. This is consistent with discipline operating in FOMC1, for which members prepare in advance, and then engaging in conformity in FOMC2, which is more extemporaneous.

Third, since the discipline and information-distortion effects appear present in the data, we propose an influence score in the spirit of the PageRank algorithm to compare the two effects. After transparency, more inexperienced members become more influential in terms of their colleagues' (particularly Alan Greenspan's) topic coverage, indicating that their statements contain relatively more information after transparency than before.

The ultimate message of the article is that career concerns matter for how policy makers respond to transparency. Moreover, while we present evidence strongly indicating the presence of a negative conformity effect among rookie members, the fact that they nevertheless become more influential in shaping debate suggests that the positive discipline effect is as, if not more, relevant for affecting their underlying information sets. This is notable since, in our view, the discipline effect has received less attention in discussions surrounding transparency in monetary policy.

Our article also makes a methodological contribution by introducing LDA to the economics literature. LDA is a widely used topic model and has been cited over 18,000 times between 2003 and the start of 2017, although we are aware of no applications in economics that predate the original draft of this article (Hansen, McMahon, and Prat 2014).[4] An important distinction in the analysis of text is whether documents come with natural labels. 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. LDA instead uncovers 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.

More broadly, this article contributes to the literature on the impact of transparency on FOMC deliberation initiated by Meade and Stasavage (2008), who showed a tendency for reduced dissent in voice following the natural experiment. These papers include Woolley and Gardner (2017), Schonhardt-Bailey (2013), Acosta (2015), and Egesdal, Gill, and Rotemberg (2015); all use automated approaches to analyze the text of FOMC transcripts.[5] Our article makes three key contributions beyond the current literature. First, we frame the decision to increase transparency as a trade-off between discipline and conformity, whereas existing publications focus on conformity in their empirical analyses and thereby miss an important channel. This is of first-order importance since discipline appears as strong as or stronger than conformity in this setting. Second, we use a difference-in-differences approach to identify the impact of the natural experiment on behavior. The current literature compares the average behavior of FOMC members before and after the experiment, which we

4. Fligstein, Brundage, and Schultz (2014) is a paper in sociology we became aware of afterward that uses LDA on FOMC transcripts to discuss sociological theories of "sense-making." Since 2014, a number of papers in economics make use of LDA, such as Budak et al. (2014), Nimark and Pitschner (2016), Bandiera et al. (2017) and Mueller and Rauh (2017).

5. There is a also a literature that uses text mining techniques to study central bank communication to the public rather than deliberation. Examples include Chappell, Havrilesky, and McGregor (2000), Boukus and Rosenberg (2006), Lucca and Trebbi (2009), Hendry and Madeley (2010), Hendry (2012), Apel and Blix Grimaldi (2012) and Bligh and Hess (2013). Of course, many others have analyzed the transcripts without using computer algorithms (e.g., Romer and Romer 2004; Chappell, McGregor, and Vermilyea 2005; Cieslak and Vissing-Jørgensen 2017).

argue is problematic given the importance of time-varying factors for communication. Third, LDA allows us to construct a more interpretable set of communication measures than other methods, so we can link more clearly to the underlying economic objects of interest: discipline, conformity, and influence. Taken together, these contributions provide a novel view of how central bankers respond to transparency and can inform important debates in central bank design.

The article proceeds as follows. Section II reviews the career concerns literature that motivates the empirical analysis, and Section III describes the institutional setting of the FOMC and the natural experiment we exploit. Section IV describes how we measure communication, and Section V presents the main results on how transparency affects these measures. Section VI examines the overall effect of transparency on behavior using influence. Section VII explores robustness, and Section VIII concludes.

II. 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, and so on. 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.

II.A. Discipline

The canonical reference in the literature is Holmström (1999), who shows that career concerns motivate agents to undertake costly, noncontractible 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 have more information on the economy.

As for transparency, Holmström (1999) predicts that effort exertion increases as the noise in observed output decreases. Interpreting transparency as increasing the precision of observers' information regarding member productivity, one would expect transparency to increase incentives to acquire information prior to meetings.[6]

II.B. Conformity / Nonconformity

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

6. 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.

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. It should be noted that Levy (2007), and especially Visser and Swank (2007), explicitly use transparency of monetary policy discussions to motivate their analyses.

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

In all standard 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 engage in behavior that signals their type (Holmström 1999). 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/nonconformity, this means that incentives to conform (or nonconform) will be relatively stronger among the less experienced agents. To the extent that knowledge of type is less likely for the less experienced, one would expect them to be more likely to conform. This hypothesis is corroborated by anecdotal evidence. Greider (1987) (referenced in Visser and Swank 2007) quotes Lawrence Roos, a former St. Louis Fed president, as saying, “If one is a young, career-oriented President who's got a family to feed, he tends to be more moderate in his opposition to Governors.”

III. FOMC Transcript Data and Natural Experiment

The FOMC meets eight times a year to formulate monetary policy (by law it must meet at least four times) and determine other Federal Reserve policies. It contains 19 members: 7 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 12 presidents of Regional Federal Reserve Banks of whom one—the president of the New York Fed—is vice-chair of the FOMC.[7] Federal Reserve staff also attend the meeting and provide briefings. The main policy variable of the FOMC is a target for the Federal Funds rate. Though all members attend the meetings and take part in the discussion, at any given time only 12 of the FOMC have policy voting rights. All seven governors have a vote; the president of the New York Fed is a permanent voting member; and four of the remaining eleven Fed presidents vote for one year on a rotating basis.[8]

FOMC meeting transcripts are available for download from the Federal Reserve website. Apart from minor redactions relating, for example, to maintaining confidentiality of certain participants in open market operations, they provide a nearly complete account of every FOMC meeting from the mid-1970s onward. In this article, the set of transcripts from the tenure of Alan Greenspan—August 1987 through January 2006 inclusive, covering 149 meetings—form the basis of our deliberation analysis.[9]

7. The U.S. president nominates members of the Board of Governors, who are then subject to approval by the Senate. A full term as a governor is 14 years (with an expiration 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, subject to approval by the Board of Governors, and serve five-year terms.

8. Chicago and Cleveland Fed residents vote one year on and one year off, while the remaining nine presidents vote for one of every three years.

9. The raw transcripts 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 nonspoken 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 statements were recorded separately in appendixes, we reinserted them into the main transcripts where they took place in the deliberation.

During this period, the FOMC also engaged in numerous conference calls. However, because many of these were not directly about monetary policy, the transcripts are either partial or nonexistent, and the calls did not follow specific structures even when about monetary policy, we do not use them in our baseline analysis.

The final data set contains 46,502 unique interjections along with the associated speaker. For example, we would have two interjections if Alan Greenspan asked a question of staff (the first interjection) and a staff member replied (the second interjection). In total there are 5,507,304 words excluding punctuation, numbers, and so on.

III.A. Meeting Structure under Chairman Greenspan

Most FOMC meetings in our sample 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).[10]

During the meeting there are a number of stages, including two core discussion stages relevant to the monetary policy decision. All members participate in both stages regardless of whether they are currently voting members.[11]

- i. A New York Fed official presents financial and foreign exchange market developments, and staff answer questions on these financial conditions.
- ii. Economic situation discussion (FOMC1)
  - a. Board of Governors' staff present the economic situation (including forecast).
  - b. There are a series of questions on the staff presentations.
  - c. FOMC members present their views of the economic outlook. Chairman Greenspan tended to speak reasonably little during this round.

10. In June 2010 the Blue Book and Green Book were merged into the Teal Book.

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

iii. 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. Later in the sample, the two-day meetings were used to discuss special topics in more details.

iv. Monetary policy strategy discussion (FOMC2)
   a. The board's director of monetary affairs then presents a variety of monetary policy alternatives (without a recommendation).
   b. A potential round of staff questions.
   c. The chairman (first) and the other FOMC members discuss their policy preferences.

v. The FOMC votes on the policy decision—FOMC votes are generally unanimous (or close to) but there is more dissent in the discussion (Meade 2005).

vi. Other items, such as discussions of FOMC disclosure policy or other special topics, tend to be irregularly added to the FOMC meeting agenda. However, these discussions can be quite long and can take up significant portions of given meetings.

One of the unique contributions of our article, compared with other papers that look at Fed deliberations, is that we distinguish between these different sections of the meeting. In particular, in our article we limit our attention to FOMC1 and FOMC2, which contain, respectively, a total of 2,748,030 (50% of total) and 1,169,599 (21% of total) words. One important reason to treat these two sections separately is that, as the two core monetary policy sections, they appear consistently across the whole of the Greenspan era. By focusing on these sections, we can be more confident that our findings relate to changes in the deliberation about monetary policy as opposed to other topics.

There is also good reason to examine FOMC1 and FOMC2 separately, as opposed to simply considering discussions of monetary policy jointly. The two sections are structured differently, which means that the likelihood of information distortion and discipline effects vary between sections. For instance, FOMC1 is an information-sharing exercise in which each member shares their reading of the current economic situation and its likely path. The fact that the FOMC members "have prepared for this go-round through weeks of information gathering" (Federal Reserve Bank of Philadelphia 2008) makes FOMC1 the part of the meeting most likely to benefit from a discipline effect encouraging more comprehensive information analysis. The chair speaks very little in FOMC1 but actually sets out his vision for the correct policy at the start of FOMC2. While there can be some preparation for the FOMC2 discussion on policy strategy, having to react to the position laid out by the chair, as well as to other FOMC members, makes the discussion more extemporaneous in nature. Having a clear position to react to means that this section of the meeting would be relatively more likely to reveal any inclination toward conformity or nonconformity.

III.B. 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 that the committee meets informally to make the main decisions. 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, and so on. Although the FOMC is not obliged to operate under the rules of the Sunshine Act, they maintain a position that is 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.

However, while the Sunshine Act prohibits a premeeting of the whole committee, we cannot rule out bilateral meetings and we know that premeeting communication between individual governors and the chair did take place through less formal engagements.[13] However, such informal communication is much more

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.

13. As Meyer (2004) says: "When I began my term, the Chairman would meet individually with the other governors during the week before FOMC meetings. His assistant would call to make an appointment, and he would then come to the office of each of the governors. He would sit down and explain his views on the outlook and his 'leaning' with respect to the policy decision that would be considered by the Committee at the upcoming meeting."

likely to occur between board members and the chairman, or among board members, as they are all situated in the Federal Reserve Board buildings in Washington, DC. In Section VII, we show that limiting the analysis to presidents actually strengthens the results. As such, we do not believe that our results are driven by premeeting communication.

III.C. 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, since the chairmanship of Arthur Burns in the mid-1970s, Fed staff had recorded meetings to assist with preparing the minutes. To help the minute writers, the tapes were first transcribed into a nearly verbatim text of the discussion. While the staff did record over the older tapes after the release of the minutes, unknown to FOMC members, a copy of the typed-up written record was archived. FOMC members were only made aware of these archives when political pressure from U.S. Representative 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 be more transparent.

The issue came to a head in October 1993, between the September and November scheduled FOMC meetings, when there were two meetings of the House 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. Initially Greenspan was evasive on the issue with the House 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, so it quickly moved to release the existing transcripts (with a five-year lag). Although 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 and finally became formal on February 2, 1995.[14]

14. By July 1994, the FOMC's Disclosure Subcommittee had recommended the lagged release of future transcripts (Lindsey 2003). Although the FOMC had

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 after five years.[15] Since the decision to change transparency was not driven by the FOMC's own concerns about the nature or style of deliberation, and the change came as a surprise to members, we can use this natural experiment to evaluate the effects of transparency on deliberation.

IV. Measuring Communication

Our key empirical challenge is to construct measures of communication from the 26,645 statements in the economic situation (FOMC1) and monetary policy strategy (FOMC2) discussions of FOMC meetings. We propose simple measures that capture the nature of deliberation without needing to determine the linguistic content of statements, but we are also interested in this. At an abstract level, the data set can be represented as a 26,645 by 24,314 document-term matrix, where 24,314 is the number of unique words in the data set. The (d, v)th element of the matrix is the number of times the vth unique word appears in the dth statement. This representation is high dimensional and sparse, so dimensionality reduction is key.

By far the most common approach to automated content analysis in economics relies on so-called dictionary methods in which the researcher defines a set of words of interest and then computes their counts or frequencies across documents. For example, to measure economic activity, we might construct a word list which

deferred the final decision, these recommendations were communicated to the FOMC and coincide with what was formally ratified by the FOMC.

15. While the majority of members only found out about the existence of the transcripts in October 1993 as a result of the House Banking Committee 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.

includes "growth." 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.

We alleviate these concerns by instead using LDA for dimensionality reduction. An important advantage of machine learning over dictionary methods is that it uses variation in all terms to represent statements on a low-dimensional latent space. Also, machine learning approaches determine which words are most important for discriminating between statements rather than imposing this on the data. Finally, a distinguishing feature of LDA compared with other algorithms for dimensionality reduction is that it is fully probabilistic. For example, latent semantic indexing (which has already appeared in the economic literature, see Boukus and Rosenberg 2006; Hendry and Madeley 2010; Hendry 2012; Acosta 2015) is essentially a principal components analysis that performs a singular value decomposition on the document-term matrix and retains the most informative dimensions. In contrast to this linear algebra approach, LDA explicitly estimates a flexible statistical model, which makes interpreting its output easier. More broadly, LDA can also easily serve as a statistical foundation for more complex latent variable models of text, such as dynamic (Blei and Lafferty 2006) or correlated (Blei and Lafferty 2007) topic models.

It is also useful to locate LDA in the broader context of machine learning. Generally speaking, machine learning algorithms (not just those for text mining) solve either 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 represent each document in terms of these groupings.

The rest of this section discusses LDA as a statistical model, then discusses the output it generates on the FOMC transcript data. Finally, it describes how we build communication measures from this output. Many details are left out, and are filled in by the Online Appendix.

IV.A. LDA Statistical Model

LDA is a Bayesian factor model for discrete data. Suppose there are D documents that make up a corpus of texts with V unique terms. The first important objects in LDA are K topics (i.e., factors), each of which is a probability vector $\boldsymbol{\beta}_k \in \Delta^{V-1}$ over the V unique terms in the data. The choice of probability distributions is important because it allows the same term to appear in different topics with potentially different weights. Informally, one can think of a topic as a weighted word list that groups words expressing the same underlying theme.

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 $\boldsymbol{\theta}_d$ (i.e., factor loadings). Informally, θ_d^k represents the "share" of topic k in document d.

The probability that any given word in document d is equal to the vth term is therefore $p_{dv} \equiv \sum_k \beta_k^v \theta_d^k$ and the overall likelihood is $\prod_d \prod_v p_{d,v}^{n_{d,v}}$ where $n_{d,\,v}$ is the number of times terms v appears in document d. Importantly, LDA reduces the dimensionality of each document substantially. In the document-term matrix, 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 raw term counts since the underlying topics are free to lie anywhere in the $V-1$ simplex.

The model described so far is probabilistic latent semantic indexing (Hofmann 1999). A key innovation of LDA is to extend this model by placing Dirichlet priors on the probability vectors for document-topic shares (and, in most formulations, topics) to smooth estimation and specify a generative process for documents. Accordingly, 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.

The inference problem in LDA is to approximate the posterior distributions over $\boldsymbol{\beta}_k$ for every k and over $\boldsymbol{\theta}_d$ for every d given K, α, and η. We adopt a popular Markov chain Monte Carlo algorithm for estimation (Griffiths and Steyvers 2004), which we describe in the Online Appendix. Next we simply describe the output, where all point estimates are constructed by averaging over draws from a particular Markov chain.

IV.B. Vocabulary and Model Selection for LDA

Prior to estimation we preprocess the raw text in several steps. The purpose is to reduce the vocabulary to a set of terms that are most likely to reveal the underlying content of interest, and thereby facilitate the estimation of more semantically meaningful topics.

First, we identify collocations, or sequences of words that have a specific meaning. For example, "labor market" corresponds to a single economic concept but is composed of two separate words. To do this we first use the part-of-speech tagger described in Toutanova et al. (2003) to tag every word in the FOMC transcripts. We then tabulate the frequencies of part-of-speech patterns identified in Justeson and Katz (1995) as likely to correspond to collocations.[16] Finally we create a single term for two-word (three-word) sequences whose frequency is above 100 (50).

The second step of preprocessing is to remove common stopwords like "the" and "of" that appear frequently in all texts. The third step is to convert the remaining terms into their linguistic roots through stemming so that, for example, "preferences", "preference," and "prefers" all become "prefer." The outcome of stemming need not be an English word. Finally, we follow the suggestion of Blei and Lafferty (2009) and rank the remaining words using term frequency-inverse document frequency (tf-idf), a measure of informativeness that punishes both rare and frequent words. Figure I plots the tf-idf values for each word; based on inspection we drop all terms ranked 9,000 or lower. Because a large number of words share the same tf-idf weight, we end up with

16. These are adjective-noun; noun-noun; adjective-adjective-noun; adjective-noun-noun; noun-adjective-noun; noun-noun-noun; and noun-preposition-noun.

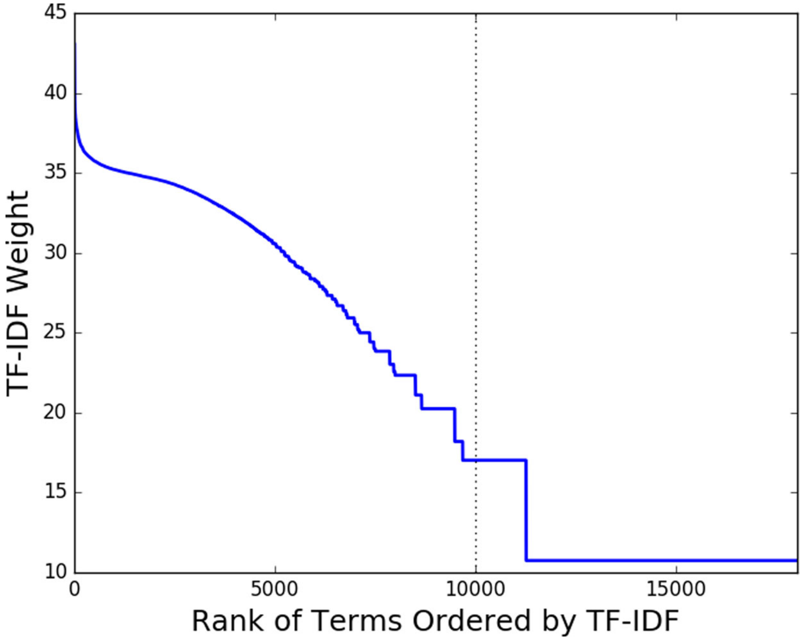


FIGURE I

Ranking of Stems with Term Frequency-Inverse Document Frequency

Let n_v be the count of term v in the data set. The term frequency tf_v is $1 + \log(n_v)$. The document frequency is $df_v = \log\left(\frac{D}{D_v}\right)$ where D is the number of documents and D_v is the number of documents in which term v appears. The tf-idf weight of term v is $tf_v \times df_v$. This figure plots the tf-idf weight of each stem in FOMC1 and FOMC2, and the dotted line indicates that the threshold we choose to drop stems from the data for the analysis.

8,206 unique stems. The set of words we remove are made up of all words that appear in two or fewer statements (these have a low term frequency) and the term "think" (this has a low inverse document frequency, i.e., it appears in many individual statements in the transcripts). Table II shows the effect of preprocessing on the size of the data. Although reductions are substantial, we still face an inherently high-dimensional problem.

For values of the hyperparameters, we follow Griffiths and Steyvers (2004) and set $\alpha = \frac{50}{K}$ 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.

TABLE II
DATA DIMENSIONALITY REDUCTION OF EACH PREPROCESSING STEP

	Raw text	Identify collocations	Remove stopwords	Stemming	TF-IDF adjustment
Total words	3,917,629	3,814,074	1,732,323	1,732,323	1,672,869
Unique words	24,314	25,019	24,822	15,394	8,206

Notes. Our raw text contains 3,917,629 words, 24,314 of which are unique. This table shows how these numbers evolve through preprocessing. The stopword list we use is from http://snowball.tartarus.org/algorithms/english/stop.txt. The stemming algorithm is the Porter stemmer implemented in Python's Natural Language Toolkit (Bird, Klein, and Loper 2009). TF-IDF weighting is as described in the main text.

A persistent challenge in unsupervised learning is to choose the appropriate dimensionality of the latent space, in our case the number of topics K. In probabilistic topic modeling, there is typically a trade-off between the interpretability of the model's output—which favors a lower K—and its statistical goodness-of-fit—which favors a higher K (see Chang et al. 2009). For our baseline analysis, we favor the former and settle on $K = 40$ after experimenting with different values.[17] (If one picks too few topics, they tend to mix underlying themes and become very general, whereas if one picks too many, topics become highly specific to particular conversational patterns.) However in Appendix A we conduct a formal model selection exercise and find a model with $K = 70$ best fits the data. In Section VII we report results based on this number of topics and find a general concordance with those from the baseline.

IV.C. LDA Output

We estimate LDA on the set of individual statements in FOMC1 and FOMC2, which form the topics we anlayze next. The estimation also produces a distribution of topics within individual statements. However, we are interested in the distribution of topics within more aggregated documents, for example, how individual speakers allocate attention within a meeting. To estimate these, we keep topics fixed at their originally estimated values but reestimate document-topic distributions for more aggregated documents. For more details, see the Online Appendix.

17. 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."

													Pro-cyclicality
Topic0[1]	product	increas	wage	price	cost	labor	rise	acceler	inflat	pressur	trend	compens	0.024
Topic1[1,2]	growth	slow	economi	continu	expans	strong	trend	inflat	will	recent	slowdown	moder	0.023
Topic2[2]	inflat	expect	core	measur	higher	path	slack	gradual	continu	remain	view	suggest	0.017
Topic3[1]	percent	year	quarter	growth	month	rate	last	next	state	averag	california	employ	0.007
Topic4	number	data	look	chang	measur	use	point	show	revis	estim	gdp	actual	0.007
Topic5[1,2]	polici	inflat	monetarpol	need	time	can	monetari	move	tighten	view	action	believ	0.005
Topic6[2]	rate	term	expect	real	lower	increas	rise	level	declin	short	nomin	year	0.005
Topic7	statement	word	chang	meet	languag	discuss	issu	want	read	sentenc	view	use	0.005
Topic8[2]	chairman	support	mr	direct	recommend	agre	asymmetr	prefer	symmetr	move	toward	favor	0.004
Topic9[1]	employ	continu	growth	job	nation	region	seem	state	manufactur	greenbook	busi	bit	0.004
Topic10	dollar	unitedstates	export	countri	import	foreign	japan	growth	abroad	trade	develop	currenc	0.003
Topic11	model	use	simul	shock	effect	scenario	nairu	differ	rule	chang	baselin	altern	0.003
Topic12[2]	risk	may	balanc	seem	side	uncertainti	possibl	economi	probabl	reason	upsid	much	0.003
Topic13	forecast	greenbook	staff	project	differ	assumpt	littl	assum	somewhat	lower	end	period	0.002
Topic14	period	committe	consist	econom	run	maintain	futur	read	slightli	stabil	expect	develop	0.002
Topic15	invest	incom	spend	capit	household	consum	busi	hous	consumpt	sector	stock	stockmarket	0.002
Topic16[1]	month	report	increas	survey	expect	indic	remain	continu	last	recent	data	activ	0.002
Topic17[1]	project	forecast	year	quarter	expect	will	percent	revis	anticip	growth	next	recent	0.002
Topic18	question	ask	issu	let	want	answer	rais	discuss	don	start	without	okay	0.001
Topic19	peopl	talk	lot	much	comment	around	differ	number	realli	look	thing	hear	0.001
Topic20	presid	ye	governor	parri	stern	vice	hoenig	minehan	kelley	jordan	moskow	mcteer	0.001
Topic21	move	can	evid	signific	stage	inde	will	issu	economi	may	quit	clearli	0.001
Topic22[2]	chairman	thank	mr	time	meet	laughter	comment	let	will	point	call	may	0.0
Topic23[1]	year	panel	line	shown	right	chart	expect	project	percent	middl	left	next	0.0
Topic24	district	nation	area	continu	sector	construct	manufactur	report	activ	region	economi	remain	0.0
Topic25	know	someth	happen	right	thing	want	look	sure	can	realli	anyth	els	0.0
Topic26[1,2]	polici	might	committe	market	may	tighten	eas	risk	action	staff	possibl	potenti	-0.001
Topic27	year	continu	product	price	level	industri	will	sale	increas	auto	last	district	-0.001
Topic28[1]	inventori	product	sale	level	order	will	sector	come	good	quarter	much	adjust	-0.001
Topic29	price	oil	increas	energi	effect	import	suppli	product	demand	will	market	oilprices	-0.002
Topic30	term	might	point	can	sens	run	short	probabl	time	longer	tri	someth	-0.002
Topic31	seem	may	time	certainli	bit	littl	quit	much	far	perhap	better	might	-0.003
Topic32	money	aggreg	borrow	seem	rang	reserv	rate	target	time	altern	suggest	million	-0.003
Topic33[2]	move	market	point	will	fundsrate	rate	basispoints	need	fed	today	basi	time	-0.004
Topic34[1]	report	busi	compani	year	contact	firm	sale	worker	expect	plan	director	industri	-0.004
Topic35	will	fiscal	ta	budget	cut	govern	effect	billion	state	spend	deficit	year	-0.005
Topic36	will	economi	world	rather	problem	believ	can	situat	much	seem	view	good	-0.008
Topic37	realli	look	side	thing	lot	problem	concern	littl	pretti	situat	kind	much	-0.012
Topic38	bank	credit	market	loan	financi	debt	lend	fund	concern	financ	problem	spread	-0.018
Topic39[1,2]	economi	weak	recoveri	recess	confid	eas	neg	econom	will	turn	declin	period	-0.059

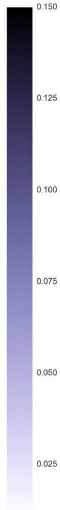


FIGURE II

Topics Ranked by Procyclicality; Terms within Topics Ranked by Probability

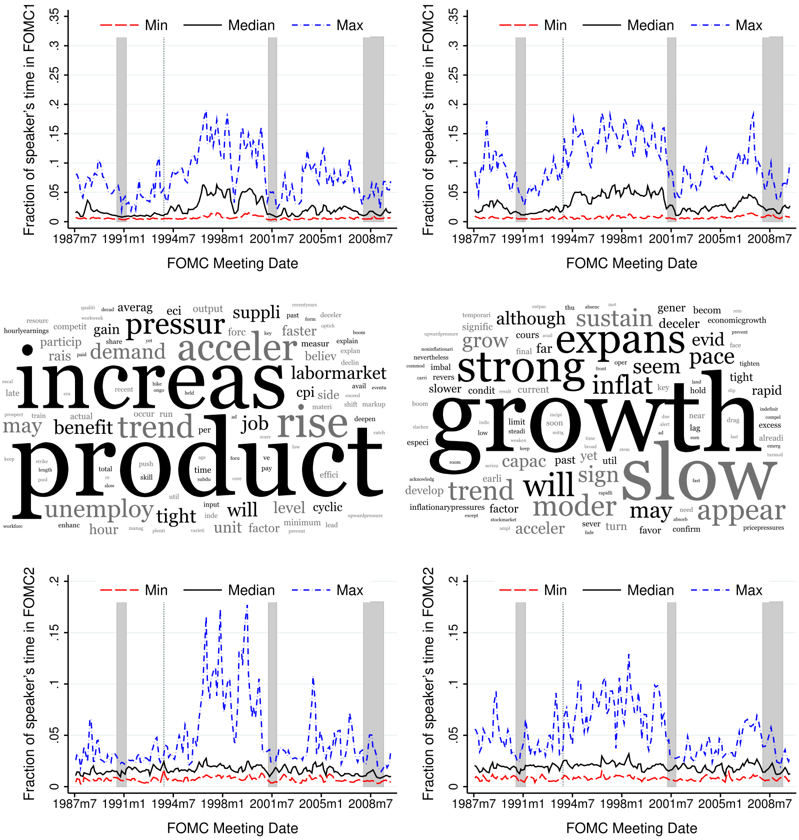
This figure summarizes the 40 separate distributions over vocabulary terms that LDA estimates to represent topics. We order these distributions from 0 to 39 based on a procyclicality index that computes the difference in average time the FOMC as a whole spends discussing the corresponding topic in expansions versus contractions, where we use the standard NBER definition of recessions. Within each row, terms are ordered left to right by the probability they appear in each topic, with differential shading indicating approximate probability values. The raw values for this table are available in the Online Appendix. For an explanation of the superscripts on topics, see Section IV.D.

1. Estimated Topics. The first LDA output of interest is the topics themselves. Topics are probability vectors over the 8,206 unique terms in the vocabulary that remain after preprocessing. Figure II represents each in a heatmap in which darker shades on terms indicate higher probabilities. As expected given our choice of hyperparameter, topics have a limited number of words with relatively high probability and a much larger number with relatively low probability. Overall, topics also form natural groupings of words, as can be seen by the interpretable output. Although nothing in the estimation procedure guarantees this, topics appear to have natural labels: topic 2 is "inflation"; topic 10 is "trade"; topic 29 is "commodity prices"; and so on. (An important caveat is that these interpretations are subjective insofar as they rely on judgments of the researcher and are outside of the statistical model, but these labels play no role in the formal analysis: they are just used for mnemonics.) The ability of LDA to generate easy-to-interpret topics is one reason for its popularity.

Since topics have no natural ordering, we define our own based on a procyclicality index. We pool all FOMC meetings in our sample that correspond to recessions in one group and to expansions in another (as defined by the National Bureau of Economic Research). For each topic, we compute the difference in its share for the FOMC as a whole during expansions versus recessions, and rank topics according to this number. A positive (negative) number indicates more attention during expansions (recessions). This is another dimension on which estimated topics are intuitive. The most procyclical topics include those relating to productivity (0), growth (1), and inflation (2), and the most countercyclical topics include those relating to economic weakness (39), the financial sector (38), and fiscal issues (35). Topics with neglible relationship to the business cycle include those relating to engagement with Alan Greenspan (22), discussion of staff material (23), and reports on regional economic activity (24). These make sense because these topics occur in each meeting regardless of the economic cycle.

2. Estimated Content. Because our main focus is at the meeting-section-speaker level, we compute the distribution over topics for each FOMC member in FOMC1 and FOMC2 separately for every meeting in the sample. For illustrative purposes, in this section we also extend this analysis to transcripts through 2009 using the topics reported above. In Figure III, we plot the minimum, median, and maximum shares for FOMC members in each meeting section (using a three-meeting moving average) for the two most procyclical topics. To further illustrate these topics' key words, we provide an alternative visualization with word clouds, where the size of the word in the cloud is approximately proportional to its probability in the topic. Figure IV does the same but for the two most countercyclical topics. In both figures, recessions are indicated with shaded windows. We also indicate the revelation of the transcripts' existence in October 1993 with a dashed vertical line.

Several interesting points emerge. First, one observes large movements in some of the time series near turning points in the business cycle. Prior to each of the three observed recessions, the maximum attention devoted to topic 1 drops significantly.

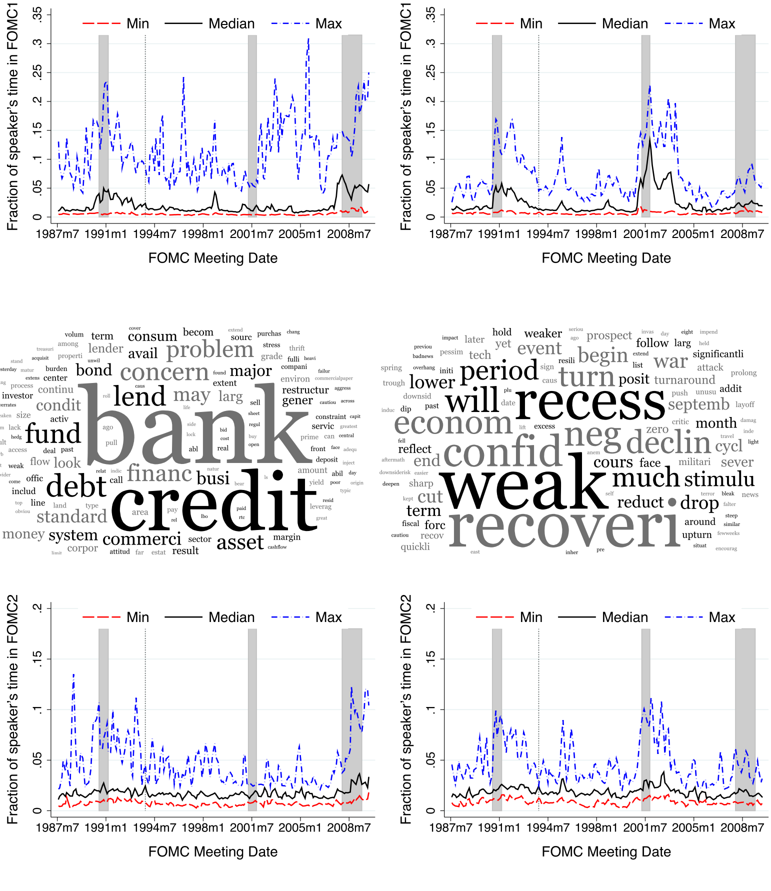


(A) Topic 0 'Productivity' (B) Topic 1 'Growth'

Figure III

Procyclical Topics by Meeting Section (Recessions Shaded)

These figures present information on the prevalence of the two most procyclical topics at the meeting-section-speaker level. For each FOMC meeeting between August 1987 and December 2009, we record the maximum, median, and minimum shares among FOMC members in both the discussion of the economic situation (FOMC1) and the monetary policy strategy (FOMC2). Recession periods are shaded in gray, and the vertical dashed line represents the November 1993 meeting, the first for which the existence of the transcripts was common knowledge. The distributions over terms that each topic induces are represented as word clouds, where the size of term is approximately proportional to its probability.



(A) Topic 38 'Financial sector'

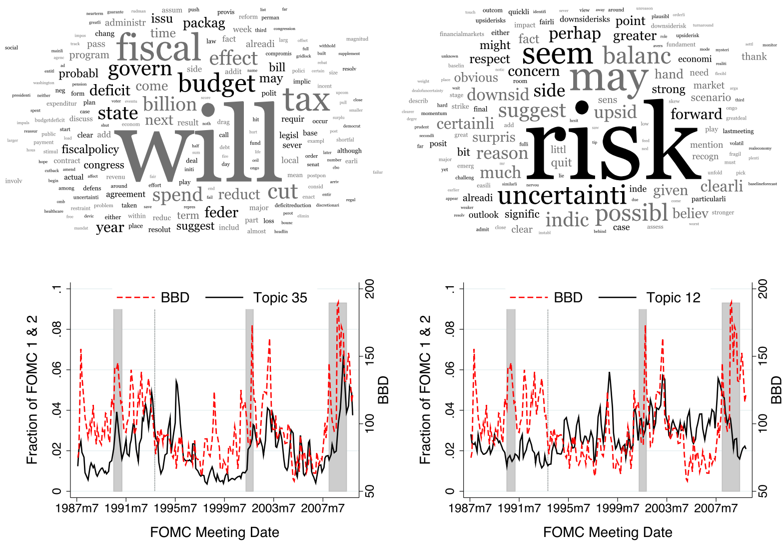
(B) Topic 39 'Economic weakness'

Figure IV

Countercyclical Topics by Meeting Section (Recessions Shaded)

These figures are the equivalent of those presented in Figure III, except for the two most countercyclical topics. See notes for Figure III.

Conversely, prior to the first two recessions, attention to topic 39 surges. This suggests the potential for text to be used in nowcasting exercises (for more on this point see Thorsrud 2016). Second, there is a great deal of speaker heterogeneity in the data. One illustration appears during the build-up of the U.S. housing bubble in the 2000s. The maximum amount of attention on



(A) Topic 35 'Fiscal issues' (B) Topic 12 'Risk'

Figure V

"Fiscal" and "Risk" Topics versus Economic Policy Uncertainty Index

This figure plots the fraction of time the FOMC as a whole spends discussing two topics—one on fiscal issues and on risks—that reflect the same content as the Economic Policy Uncertainty index of Baker, Bloom, and Davis (2016) against the actual EPU index. The comovement in the topic measures with the EPU index provides external validation on the output of LDA.

topic 38—which relates to financial markets—diverges markedly in FOMC1 from the median, reaching nearly 30 percentage points in 2005. More broadly, the figures taken as a whole clearly indicate that policy makers discuss markedly different aspects of the economy in any given meeting. Third, the time series are quite volatile (those computed without the moving average are even more so) and there is no obvious break in levels or ranges around the natural experiment. Instead, attention appears to fluctuate according to phases in the business cycle. We come back to this point when we discuss our econometric approach.

As a final illustration of estimated content, in Figure V we plot the share of attention for the FOMC as a whole on topic 35—which relates to fiscal issues—and topic 12—which relates to economic risk—against the Economic Policy Uncertainty (EPU) index of Baker, Bloom and Davis (2016). We choose these topics because the EPU index captures the public's perceptions of general risk as well as expiring fiscal measures. Both topic series comove with the EPU index, but the relationship is stronger with topic 35. Interestingly, the amount of attention to risk within the FOMC is near its maximum value in the sample at the onset of the Great Recession before falling away; in contrast the EPU index continues to rise substantially after 2007. This is consistent with a view that the FOMC as a whole perceived the buildup of risk better than the public and then shifted attention to other aspects of the economy once this risk was realized.

There is also a methodological link between our measures and the EPU index since the latter is largely constructed based on textual analysis. However Baker, Bloom, and Davis (2016) use a dictionary-method-like approach (they count the frequency of newspaper articles containing a set of predefined keywords) rather than machine learning. It is thus notable that the two methodologies extract comparable content across different texts relating to U.S. economic conditions. Of course, LDA also produces variation in content along many other dimensions of interest.

IV.D. Selecting Topics Relevant for Policy

Ultimately we are interested in how deliberation about monetary policy shifted in response to increased transparency. While the 40 topics represented in Figure II provide a rich description of all the language used on the FOMC, not all of them are necessarily relevant to substantive deliberation. For example, topic 24 about regional economic conditions has the highest overall average share in FOMC1 at 8.3%; however in FOMC2 its share is the lowest at 0.4%. Its high average share in FOMC1 arises because the convention is that each Fed president discusses conditions within his or her region every meeting, which says nothing about beliefs or policy stances. Its low average share in FOMC2 indicates that any information from the regions must be subsumed into more aggregate topics for the strategy discussion. It is unclear whether it should be included in either section's communication measures.

Essentially we face a variable selection problem: which topics are informative about FOMC members' policy preferences? To resolve this, we first obtain the voiced dissent measure from Meade (2005). The FOMC under Greenspan operated under a strong norm for consensus, which means that the voting record has little within-meeting heterogeneity. To obtain a more meaningful preference measure, Meade (2005) records voiced dissents during FOMC2 for the 1989–1997 period by reading the transcripts and finds substantially more dissent in language than in the formal voting record. We use her multinomial variable that records −1 if a member dissents from Greenspan's policy proposal for a lower interest rate; 0 if a member agrees with the proposal; and 1 if a member dissents for a higher rate. Overall there are 1,205 such preferences recorded (for more details see Meade 2005).

We then estimate a multinomial logistic regression with the voiced disagreement variable as a dependent variable and speakers' distributions over topics as independent variables. To select the topics most predictive of voiced dissent, we use the least absolute shrinkage and selection operator (LASSO) of Tibshirani (1996), which has grown rapidly in popularity recently in economics (for example, see Belloni, Chernozhukov, and Hansen 2014). The LASSO adds a weighted ℓ_1 penalty term on the vector of regression coefficients to the standard—in our case, multinomial logistic—regression objective function. The resulting estimated coefficients feature sparsity, and the variables with nonzero coefficients are "selected" as the relevant ones.[18]

We choose the topics selected in the dissent categories −1 and 1 as the ones informative of policy preferences. We call them policy topics. In both categories, policy makers express a clear stance on monetary policy: in the case of −1 (1) they wish rates to be lower (higher) than the baseline stance expressed by Greenspan. In contrast, the topics selected in category 0 have a less clear interpretation since agreement with Greenspan does not indicate a clear dovish or hawkish tilt. We estimate separate models for FOMC1 and FOMC2 and obtain a different set of policy topics for each section. Those in FOMC1 are the topics that a policy maker who will later dissent in FOMC2 uses when analyzing economic conditions. For more details, see the notes to Table III.

Table III displays the policy topics for FOMC1 and FOMC2, which we denote P1 and P2. We also mark the policy topics in Figure II: a 1 (2) superscript indicates the topic is in P1 (P2).

18. Another strategy would be to use the raw word counts in the LASSO instead of topic distributions, but not only are the latter more interpretable, experiments also indicate that using the low-dimensional topic representation as a feature space can actually outperform the high-dimensional word representation (Blei, Ng, and Jordan 2003).

TABLE III

Policy Topics

FOMC1 policy topics (P1)	0	1		3	5			9		16	17		23	26	28		34	39
FOMC2 policy topics (P2)		1	2		5	6	8		12			22		26		33		39

Notes. To select policy topics, we use a penalized multinomial regression with the voiced dissent measure of Meade (2005) as a dependent variable. The independent variables are the distributions over topics for each speaker, as well as real-time contemporaneous CPI and unemployment. We penalize the coefficients on the topic shares with the ℓ_1 norm but not those on CPI nor unemployment. We optimize the resulting LASSO using the glmnet package in R (Friedman, Hastie, and Tibshirani 2010), and select the weight on the penalty using ten-fold cross-validation. Since these folds are generated randomly for each function call, we perform 100 calls and keep as policy topics those selected in at least 50.

There are 12 policy topics in FOMC1, which together account for 31.8% of the situation discussion during the Meade (2005) sample, and 10 in FOMC2, which together account for 33.4% of the strategy discussion. Each section contains policy topics that are among those that vary most significantly over the business cycle as captured by our procyclicality index (such as topics 1 and 39 shared across the sections). This suggests at least some of the policy disagreement on the FOMC arises from different views on the phase of the business cycle that the economy is in. The presence of topics 16 and 17 in P1 further supports this view, as these relate to surveys and forecasts of economic conditions. Also of note is that P1 and P2 share topics 5 and 26, which relate to the committee's policy stance. Finally, P2 contains two topics—8 and 22—that appear to capture the articulation of policy positions vis-à-vis Greenspan's. This is consistent with FOMC2 deliberation as being reactive to the proposed position set out by the chair.

IV.E. Communication Measures

Finally, we describe how we construct empirical measures of communication. We generate all of these at the meeting-section-speaker level, where section corresponds to FOMC1 or FOMC2. Most basically, we first count the total number of words, statements, and questions from the raw text data. These capture broad changes in the nature of deliberation after transparency.

For the topic measures, we mainly rely on the conditional distributions over policy topics generated in the previous subsection. Let $\boldsymbol{\pi}_{i,t,s}$ be the conditional distribution for speaker i in meeting t in section s.

Our first topic measure captures the breadth of the deliberation, which we construct by applying a Herfindahl concentration index to $\boldsymbol{\pi}_{i,t,s}$. Higher values indicate a narrow discussion, while lower values indicate a broader discussion.

As discussed in Section II, 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. A member without career concerns who spent little time preparing for meetings (or paying attention to colleagues during them) would most likely not discuss their views using specific references to relevant data, whereas one who had done their homework would likely bring into the meetings a dossier of evidence on which to draw. Given this, to measure the quantitative content of each members' contribution to the discussion, we first count the number of terms in each statement that are numbers (strings that consist solely of numeric characters like 99 and 1 but not terms like "one"). Second, we identify two topics from the topic model output that appear to reflect quantitative discussion: topics 4 and 23.[19] The most likely terms in these are clearly those that members would use when discussing data.

We next measure the similarity between individual FOMC members and the committee average, which we denote $\overline{\boldsymbol{\pi}}_{t,s}$. This indicates whether FOMC members tend to discuss the same policy topics as their colleagues. Higher values reflect greater congruence in policy topic coverage, and lower values reflect more diversity. There are many ways in the literature of determining the overlap between probability distributions, and we focus on three:[20]

i. Bhattacharyya coefficient: Avg Sim $(\mathrm{B})_{i,t,s} = \sum_k \sqrt{\pi^k_{i,t,s}\overline{\pi}^k_{t,s}}$. This measures the extent to which two probability

19. There is some subjectivity in this choice. For example, topics 11 and 17 also relate to technical discussion. An alternative quantitative measure based on all four topics yields similar results in our baseline regressions (results not reported).

20. One complication is that some members in some meetings have very short statements in FOMC2. LDA estimates their predictive distribution over topics as essentially uniform, as the prior distribution dominates. We have manually examined these statements, and found that in nearly all cases a speaker expresses agreement with Greenspan, such as "I support your recommendation" (Corrigan, November 1988) or "I agree with your recommendation, Mr. Chairman" (Kelley, March 2000). So whenever a speaker has fewer than five stems spoken in FOMC2 (after preprocessing), we replace their predictive distribution with Greenspan's.

distributions overlap, and is widely used in the machine-learning literature.

ii. Dot product similarity: Avg Sim $(\mathrm{D})_{i,t,s} = \sum_k \pi^k_{i,t,s}\overline{\pi}^k_{t,s}$. The policy topics we identify predict voiced dissent on average, but in any particular meeting the debate can be focussed on one or two aspects of the economy. 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.

iii. Kullback-Leibler (KL) divergence: Avg Sim $(\mathrm{KL})_{i,t,s} = \exp\left[-\sum_k \overline{\pi}^k_{t,s} \ln\left(\frac{\overline{\pi}^k_{t,s}}{\pi^k_{i,t,s}}\right)\right]$. The KL divergence is defined to be the argument of the preceding negative exponential function. This has strong roots in the information theory literature, and can be interpreted as the amount of information lost when $\boldsymbol{\pi}_{i,t,s}$ is used to approximate $\overline{\boldsymbol{\pi}}_{t,s}$. We transform the KL divergence into a similarity measure with the negative exponential function for comparability with the other two similarity measures.

Our last communication measure is our most direct measure of conformity. It exploits the estimated multinomial LASSO from Section IV.D. This provides a mapping from speaker-level topic distributions in each section to the probabilities of voicing dovish dissent, no dissent, and hawkish dissent within the Meade (2005) sample. We take these estimated coefficients and construct fitted values for the three dissent categories for the entire sample and thereby obtain a conformity measure from the fitted probability of no dissent. Unlike all of our other measures, we define this measure only for FOMC2, the section in which conformity is most relevant. This measure does not capture the extent to which a member leans towards higher or lower rates, but whether a member is willing to offer dissenting views. One advantage of this measure over extending the Meade (2005) sample via the narrative approach of reading transcripts is that the fitted values are continuous and therefore able to reflect subtle shifts in preferences that a categorical variable constructed manually might miss.

Table IV summarizes the communication measures we use in the empirical analysis in the next section. In the regression tables, we use the shortened names provided in the Name columns

TABLE IV
SUMMARY OF COMMUNICATION MEASURES (MEETING-SECTION-SPEAKER LEVEL)

Count measures		Topic measures	
Name	Description	Name	Description
Words	The count of words spoken	Concentration	The Herfindahl index applied to distribution over policy topics
Statements	The count of statements made	Quant	Percentage of time on data topics
Questions	The count of questions asked	Avg Sim (X) $X \in \{B, D, KL\}$ B = Bhattacharyya D = dot product KL = Kullback − Leibler	The similarity between a speaker's distribution over policy topics and the FOMC average, computed using metric X
Numbers	The count of numbers spoken	Pr (no dissent)	The fitted value for no voiced dissent from the LASSO for policy topic selection (only FOMC2)

to refer to our variables. Including the three different ways of measuring similarity, we have in total four count-based measures and six topic-based measures.

V. EMPIRICAL RESULTS

This section presents the main results of the article on the effect of transparency on deliberation. For all the results, we focus on a sample that uses a window of four years before and four years after the change in transparency (1989–1997). Note that because the FOMC only meets eight times a year, we are constrained in how tightly we can shrink the window while still having enough statistical power to measure the parameters of interest.

The most straightforward empirical approach is to estimate the average effect of transparency on our various communication measures. This is useful to establish whether increased transparency is associated with changes in deliberation. We first present these results as a descriptive exercise, but there are several reasons the analysis is problematic as a test of career concerns. First, the Fed adopted a different editorial stance on the transcripts after 1993 and began to "lightly edit speakers' original words, where necessary, to facilitate the reader's understanding."[21] This might have involved, for example, eliminating small interjections, which would distort our count measures. Second, as discussed in Section IV.C, topic coverage is volatile and appears largely driven by business cycle phases. Any cyclical variation that our control variables do not absorb would be attributed to the effect of transparency. Third, there may be other changes that take effect around November 1993 that affect the nature of deliberation for the FOMC as a whole which the difference analysis would associate with behavioral changes associated with transparency. These may be related to transparency, such as the Fed placing greater emphasis on presenting a united public front, or unrelated to transparency, such as the publication of Taylor (1993) which may have made monetary policy discussions narrower and more technical in all central banks.

All of these criticisms are different variants on a more general concern that many factors beyond individual career concerns drive observed FOMC communication, and that these factors are time varying. We therefore argue for a difference-in-differences analysis that allows the inclusion of time fixed effects to absorb time-varying, unobserved factors affecting the deliberation. We can then isolate behavioral changes of the individuals who should be most affected by the career-concerns channel. This provides a much more reliable test of career concerns than the basic difference regressions.

As pointed out in Mankiw (2001) and Meade and Thornton (2012), the eight-year window for our econometric analysis coincides with a period in which the Clinton administration appointed economists with a more academic background to the Board of Governors. To minimize the impact of the FOMC's changing composition on the results, in the baseline analysis we only include observations for the 19 members who were serving at the moment of the natural experiment. In the window around October 1993 that we examine, this core sample of members represents over 75% of the member-meeting observations; 920 out of 1,220. We return to this in Section VII, where we explore the sensitivity of

21. See https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm.

the analysis to different sample choices and find our main results are robust.

V.A. Difference Results

The basic difference specification we adopt is

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

where the dependent variable y_{it} represents any of the communication measures described in Section IV.E for member i in time t. We run the specification separately for FOMC1 and FOMC2, as explained in Section III.A.

Regarding independent variables, $D(\mathit{Trans})$ is an indicator for being in the transparency regime (1 after November 1993, 0 before), and X_t is a vector of macro controls for the meeting at time t. For these we include whether the economy is in a recession and the EPU index (see Section IV.C for details). We also control for whether the meeting lasted two days or one and the number of FOMC members holding PhDs in a given meeting, since background might affect the choice of language and topics. Finally, we include member fixed effects to account for individual heterogeneity in communication patterns. These fixed effects capture individual members' average level for a communication measure over the eight-year sample window, while the γ coefficient of interest captures the average change observed across all members before and after transparency. This coefficient is identified due to all members in the baseline sample serving on either side of the natural experiment.

For our topic-based communication measures, we also control for the number of stems that form the topic distributions. This determines the weight the observed data gets in forming the estimated distribution over topics relative to the Dirichlet prior. For example, a member who speaks few stems in a meeting section will have an estimated distribution over topics that is close to uniform, which may induce artificial distance from the committee average.

Testing the statistical significance of the γ 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

TABLE V
Difference Results for Economic Situation Discussion (FOMC1): Count Measures

Main regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Trans)	56.7*	−0.52	−0.039	3.71***
	[.076]	[.162]	[.659]	[.003]
D(Recession)	−1.95	−0.69	−0.19	−0.71
	[.952]	[.159]	[.314]	[.488]
EPU index	0.30	−0.00094	0.00088	0.0040
	[.186]	[.876]	[.586]	[.520]
D(2 day)	27.1	1.36*	0.56*	1.28
	[.256]	[.085]	[.051]	[.188]
# of PhDs	6.68	−0.45***	−0.11***	0.51
	[.561]	[.005]	[.009]	[.109]
Constant	528***	10.0***	2.44***	1.50
	[.002]	[.000]	[.000]	[.740]
Unique members	19	19	19	19
Observations	903	903	903	903
Member FE	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Meeting section	FOMC1	FOMC1	FOMC1	FOMC1
Transparency effect	9.5*	−10	−2.5	53.2***

Notes. This table reports the results of estimating (DIFF) on FOMC member statements from the economic situation discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p < .05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The transparency effect reports the estimated coefficient on $D(Trans)$ as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)$.

and may lead us to overstate the significance of estimated relationships. We use the nonparametric covariance matrix estimator proposed by Driscoll and Kraay (1998), allowing for up to eight meetings (approximately one year) of autocorrelation. This helps make our standard errors robust to general forms of spatial and temporal dependence, as well as being heteroskedasticity- and autocorrelation-consistent.

Tables V and VI show the estimates for FOMC1. For the count measures, there are significant increases in words and the use of quantitative language after transparency. For topics, there is an increase in similarity for two measures. Tables VII and VIII show the estimates for FOMC2. We see particularly strong average effects for the count measures, with the number of words increasing; the number of statements and questions decreasing;

TABLE VI

DIFFERENCE RESULTS FOR ECONOMIC SITUATION DISCUSSION (FOMC1): TOPIC MEASURES

Main regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)
D(Trans)	0.0041	−0.00027	0.0082***	0.0012	0.032***
	[.205]	[.831]	[.001]	[.692]	[.000]
D(Recession)	0.0061**	−0.000056	0.0020	0.015***	−0.0017
	[.028]	[.968]	[.385]	[.000]	[.758]
EPU index	3.7e-06	−9.6e-06	0.000050*	0.000029	0.00015
	[.890]	[.541]	[.077]	[.300]	[.109]
D(2 day)	−0.0040*	0.0042**	0.00044	−0.0037***	0.00051
	[.093]	[.024]	[.802]	[.001]	[.914]
# of PhDs	0.0017	−0.00063	0.000097	0.00079	0.00018
	[.255]	[.292]	[.885]	[.671]	[.928]
# Stems	0.000075***	8.8e-06**	−3.5e-06	0.000030***	0.000049
	[.000]	[.049]	[.837]	[.001]	[.284]
Constant	0.13***	0.037***	0.89***	0.084***	0.62***
	[.000]	[.000]	[.000]	[.001]	[.000]
Unique members	19	19	19	19	19
Observation	903	903	903	903	903
Member FE	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No
Meeting section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Topics	P1	T4 & T23	P1	P1	P1
Similarity measure	—	—	Bhattacharyya	Dot product	Kullback-Leibler
Transparency effect	2.5	−0.7	0.9***	1.1	4.9***

Notes. This table reports the results of estimating (DIFF) on FOMC member statements from the economic situation discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p < .05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The transparency effect reports the estimated coefficient on $D(Trans)$ as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)$.

and an increase in quantitative language. Overall, the picture is consistent with a move toward longer, scripted, more technical language after transparency and, at the same time, a reduction in back-and-forth dialogue during FOMC2, since more open and dynamic debate would generate many statements as arguments bounced from member to member.[22]

22. This finding is similar to that in Woolley and Gardner (2017), who note a decrease in the average number of speakers per 100 words of transcript during our sample period.

TABLE VII
DIFFERENCE RESULTS FOR POLICY STRATEGY DISCUSSION (FOMC2): COUNT MEASURES

Main regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Trans)	92.1**	−0.99***	−0.41**	1.86***
	[.019]	[.007]	[.012]	[.000]
D(Recession)	23.4	1.58***	0.17	−0.34
	[.560]	[.004]	[.356]	[.692]
EPU index	0.34	−0.0025	−0.0027***	0.0031
	[.134]	[.341]	[.004]	[.468]
D(2 day)	48.9	0.45	0.19	0.92
	[.226]	[.251]	[.133]	[.153]
# of PhDs	7.26	0.16	0.039	−0.37
	[.766]	[.560]	[.587]	[.489]
Constant	143	2.76	0.81	5.78
	[.638]	[.416]	[.312]	[.376]
Unique members	19	19	19	19
Observation	895	895	895	895
Member FE	Yes	Yes	Yes	Yes
Time FE	No	No	No	No
Meeting section	FOMC2	FOMC2	FOMC2	FOMC2
Transparency effect	29.9**	−15.7***	−29.4**	44.6***

Notes. This table reports the results of estimating (DIFF) on FOMC member statements from the monetary policy strategy discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance ($^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The transparency effect reports the estimated coefficient on $D(Trans)$ as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)$.

To interpret the economic significance of the estimated coefficients, we report the "transparency effect" in the tables as the value of the estimated γ coefficient as a percentage of the pretransparency average (and stars indicate the statistical significance of the estimated effect). For example, the estimated coefficient in Table VII, column (2) is −0.99, meaning that on average FOMC members made one fewer statement after transparency. The pretransparency average number of statements in FOMC2 is 6.31, so the transparency effect is $100 \times \left(\frac{-0.99}{6.31}\right) = -15.7$. This indicates that the average effect of transparency is equivalent to a nearly 16% reduction of statements in the pretransparency period. Judged on this metric, the largest observed average change after transparency is the increase in quantitative language.

TABLE VIII

DIFFERENCE RESULTS FOR POLICY STRATEGY DISCUSSION (FOMC2): TOPIC MEASURES

Main regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)	Pr (No dissent) (6)
D(Trans)	0.0048*	−0.00045	−0.00079	−0.013***	0.0074	−0.010
	[.097]	[.681]	[.805]	[.000]	[.473]	[.613]
D(Recession)	−0.0055*	0.00016	0.0022	−0.0080**	0.0032	−0.0028
	[.090]	[.908]	[.323]	[.049]	[.636]	[.750]
EPU index	0.000068	−0.000033**	0.000018	−0.000015	0.000097	0.00026**
	[.107]	[.016]	[.605]	[.741]	[.371]	[.012]
D(2 day)	0.0083**	0.00031	−0.0013	0.0017	−0.0032	0.0025
	[.016]	[.701]	[.690]	[.721]	[.786]	[.742]
# of PhDs	−0.0042**	0.0013***	−0.0017	−0.0054***	−0.0058	0.00044
	[.022]	[.007]	[.127]	[.000]	[.113]	[.896]
# Stems	0.000058***	3.3e-06	0.000028**	8.6e-06	0.00012***	−0.00015***
	[.000]	[.805]	[.013]	[.335]	[.001]	[.000]
Constant	0.21***	0.028***	0.94***	0.21***	0.77***	0.82***
	[.000]	[.000]	[.000]	[.000]	[.000]	[.000]
Unique members	19	19	19	19	19	19
Observation	893	893	893	893	893	893
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	No	No	No	No	No	No
Meeting section	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2
Topics	P2	T4 & T23	P2	P2	P2	P2
Similarity measure	—	—	Bhatta-charyya	Dot product	Kullback-Leibler	—
Transparency effect	2.6*	−1.2	−0.1	−8.8***	1	−1.3

Notes. This table reports the results of estimating (DIFF) on FOMC member statements from the monetary policy strategy discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p < .05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The transparency effect reports the estimated coefficient on $D(Trans)$ as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)$.

V.B. Difference-in-Differences Results

To more clearly attribute the changes associated with transparency to career concerns, we now move to a difference-in-differences analysis. To do so requires defining a proxy for the strength of reputational concerns, and then identifying whether there is a differential response to transparency in this proxy. As discussed in Section II, a natural proxy is a member's experience in monetary policy making. The idea of using experience to empirically test career concerns has also been previously used in Hong, Kubik, and Solomon (2000).

Our specific measure of experience is $FedExp_{it}$, or the number of years member i has spent working in the Fed system through meeting t. This includes years spent in the Fed before appointment

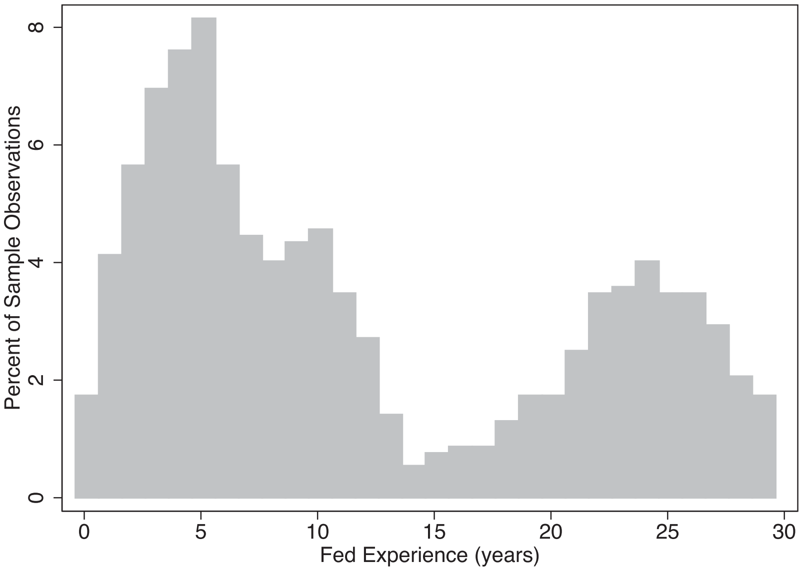


FIGURE VI

Histogram of Federal Reserve Experience (*$FedExp_{it}$*)

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

to the FOMC and years spent on the committee.[23] Clearly behavior while serving on the FOMC generates direct signals of expertise to all Fed observers. We also include years working in the Fed prior to FOMC appointment because rising through the ranks of the Fed system to its highest level can itself be taken as a strong performance signal that outsiders can use to infer high ability in monetary policy making. For example, Don Kohn was appointed to the FOMC in 2002 after spending more than 30 years in positions of increasing importance in the Fed. We maintain that the public should have less uncertainty about such members' expertise in determining and implementing appropriate monetary policy than that of a member who joins the FOMC having never worked in the Fed previously. As discussed in Section II, we expect career concerns to decline in $FedExp_{it}$. In Figure VI we plot the histogram of this variable across all members in our main sample period.

23. This information came from online sources and the *Who's Who* reference guides.

Our difference-in-differences specification for the baseline analysis is

$$(\boldsymbol{DinD})\ y_{it} = \alpha_i + \delta_t + \eta \boldsymbol{FedExp}_{it} + \phi \boldsymbol{D}(\mathit{Trans})_t \times \boldsymbol{FedExp}_{it} + \epsilon_{it},$$

where y_{it} is again one of our communication measures from section IV and $D(\mathit{Trans})$ is a transparency indicator.

Importantly, (DinD) includes both time and member fixed effects. Time fixed effects control for any time-specific factors affecting FOMC deliberation such as macroeconomic cycles or general trends in the deliberation. The inclusion of time fixed effects also renders the transparency dummy $D(\mathit{Trans})$ (1 after November 1993, 0 before) redundant in this regression.

Member fixed effects control for member-specific behaviors. This alleviates the objection to the experience proxy that there are at least a few notable exceptions of people who joined the committee as rookies (without prior Fed experience), but who had an exemplary reputation as macroeconomists and even as monetary economists. One prominent example, although not in our core sample of members present when the transparency regime changed, is Alan Blinder. Blinder joined the FOMC as a governor in 1994 with no prior years working in the Fed, although he had clearly become an expert on monetary economics through his academic work. However, the inclusion of member fixed effects controls for the initial reputation of person i: an Alan Blinder fixed effect in a regression would control for any communication pattern that his particular expertise generates on average.

(DinD) also includes FedExp_{it} as a control independent of transparency. This allows for experience itself to affect the nature of an individual's deliberation. As members serve for more time on the FOMC, their communication patterns may change for a variety of reasons beyond career concerns; for example, they may become more adept at technical analysis or more able to discuss multiple topics. Controlling for these effects is important since otherwise we might attribute observed changes after 1993 to the simple fact that all members in the core sample are becoming more experienced.

The main coefficient of interest to test the career concerns channel of transparency is ϕ. The inclusion of both member and time fixed effects in (DinD) means that the identification of ϕ relies on comparing the behavior of members based on their experience relative to their own average self, and the average in the meeting at time t. ϕ then measures the extent to which the

TABLE IX

DIFFERENCE-IN-DIFFERENCES RESULTS FOR ECONOMIC SITUATION DISCUSSION (FOMC1): COUNT MEASURES

Main regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Trans) × Fed experience	−0.18	0.015	0.0023	−0.21***
	[.912]	[.586]	[.863]	[.000]
Fed experience	1,492***	4.52*	2.29	29.2***
	[.000]	[.069]	[.344]	[.001]
Observations	920	920	920	920
Unique members	19	19	19	19
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting section	FOMC1	FOMC1	FOMC1	FOMC1
Rookie effect	0.5	−6.4	−3.3	48.1***

Notes. This table reports the results of estimating (DinD) on FOMC member statements from the economic situation discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p < .05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Trans)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

average marginal effect of an additional year of relative experience after transparency differs from the average marginal effect before. We attribute any significant difference as arising from career concerns. A positive (negative) ϕ indicates that members with greater career concerns do less (more) of whatever y_{it} is measuring.

We obtain a distribution of relative experience in each meeting since the committee composition changes over time. In a balanced panel, we could not identify ϕ because the time and member fixed effects would account for all variation in $FedExp_{it}$. Instead, we exploit the fact that members become more or less experienced relative to their colleagues as members enter and leave the FOMC. One criticism of using the restricted core sample of 19 members who were on the FOMC in late 1993 is that the ϕ estimates reflect their experience relative only to each other rather than to the FOMC as a whole for meetings in which not all 19 core sample members served. In Section VII we therefore present results from estimating (DinD) using all observations in the sample window and find our key results unaffected.

Tables IX and X present estimates for FOMC1, the section in which we expect discipline to affect behavior more than

TABLE X
DIFFERENCE-IN-DIFFERENCES RESULTS FOR ECONOMIC SITUATION DISCUSSION (FOMC1): TOPIC MEASURES

Main regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)
D(Trans) × Fed experience	0.00039	−0.00038***	0.00064*	0.00038***	0.0019**
	[.161]	[.005]	[.053]	[.005]	[.027]
Fed experience	0.10	−0.00042	0.075	0.079	0.24
	[.300]	[.984]	[.255]	[.126]	[.181]
# Stems	0.000068***	3.1e-06	1.7e-06	0.000033***	0.000059
	[.000]	[.557]	[.915]	[.000]	[.157]
Observations	920	920	920	920	920
Unique members	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Meeting section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Topics	P1	T4 & T23	P1	P1	P1
Similarity measure	—	—	Bhatt-acharyya	Dot product	Kullback-Leibler
Rookie effect	−4.7	24.3***	−1.4*	−7.0***	−5.9**

Notes. This table reports the results of estimating (DinD) on FOMC member statements from the economic situation discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (*** $p < .01$, ** $p < .05$, * $p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Trans)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

conformity. There are two main sets of results. First, less experienced members use significantly more numbers in their language, and significantly more references to data topics. (Recall that since career concerns decline with experience, the direction of the effect of career concerns is opposite in sign to the estimated coefficient.) This is consistent with discipline encouraging especially rookie members to gather additional data between meetings, which should appear in text data in the form of greater reference to numbers and quantitative indicators. Second, less experienced members discuss a relatively more diverse set of topics after transparency, which is again consistent with their collecting additional information between meetings. Instead of focusing on what their colleagues do, they tend to bring new dimensions of policy into their discussions.

To quantify the economic importance of the estimated coefficients on the interaction terms, we report for all communication

TABLE XI

DIFFERENCE-IN-DIFFERENCES RESULTS FOR POLICY STRATEGY DISCUSSION (FOMC2): COUNT MEASURES

Main regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Trans) × Fed experience	−2.53	0.082**	0.026**	−0.081**
	[.349]	[.010]	[.016]	[.017]
Fed experience	200	0.67	0.11	7.06**
	[.261]	[.776]	[.900]	[.038]
Observations	912	912	912	912
Unique members	19	19	19	19
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting section	FOMC2	FOMC2	FOMC2	FOMC2
Rookie effect	17.3	−33.7**	−52.1**	77.8**

Notes. This table reports the results of estimating (DinD) on FOMC member statements from the monetary policy strategy discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p <. 05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Trans)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

measures what we call the rookie effect. The first step in constructing this is to compute the estimated difference between how a less experienced member reacts to transparency compared to a colleague with 20 more years of Fed experience. (We choose 20 years as this is approximately equal to the difference in the modes of the distribution of experience presented in Figure VI.) For example, the estimated coefficient of −0.21 in column (4) of Table IX implies that the difference between the count of numbers in a rookie and a veteran text increases by $20 \times 0.21 = 4.2$ after transparency. The second step is to report this difference as a percent of the pretransparency average value of the communication measure. In the case of the count of numbers in FOMC1, this is 8.32. So the rookie effect is $100 \times \left(\frac{4.2}{8.32}\right) = 50.5$. We report rookie effects for all communication measures, and denote with stars the significance level of the coefficient used to calculate it. According to the rookie effect, the impact of transparency on behavior in FOMC1 is particularly strong on technical language.

Tables XI and XII present estimates for FOMC2, the section in which we expect conformity to operate in addition to discipline.

TABLE XII

DIFFERENCE-IN-DIFFERENCES RESULTS FOR POLICY STRATEGY DISCUSSION (FOMC2): TOPIC MEASURES

Main regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)	Pr (No Dissent) (6)
D(Trans) × Fed experience	−0.00077**	−0.00011	−0.00019	−0.00041***	−0.00040	−0.0015**
	[.014]	[.323]	[.222]	[.006]	[.377]	[.025]
Fed experience	−0.21***	−0.0035	−0.057	−0.11***	−0.22**	−0.41**
	[.000]	[.911]	[.140]	[.006]	[.045]	[.031]
# Stems	0.000023**	0.000018	0.000015**	0.000017***	0.000070***	−0.00011***
	[.048]	[.127]	[.030]	[.000]	[.001]	[.000]
Observations	910	910	910	910	910	910
Unique members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Meeting section	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2
Topics	P2	T4 & T23	P2	P2	P2	P2
Similarity measure	—	—	Bhatta-charyya	Dot product	Kullback-Leibler	—
Rookie effect	8.9**	5.6	0.4	5.5***	1.1	3.5**

Notes. This table reports the results of estimating (DinD) on FOMC member statements from the monetary policy strategy discussion. Dependent variable definitions are in Table IV. Coefficients are labeled according to significance (***$p < .01$, **$p < .05$, *$p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Trans)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

Table XI shows several statistically and economically significant results. Inexperienced members react to transparency by reducing the number of statements and questions more than experienced members. Overall there is no differential reaction in experience to transparency regarding total words. The implication is that after transparency rookie members disengage from the policy discussion, and instead tend to deliver their views in a limited number of long statements. At the same time, there is still a large rookie effect on the count of numbers, as in FOMC1.

Regarding the topic results of Table XII, we see an increase in topic concentration among less experienced members, and marginal evidence of less topic diversity (only the dot product measure is significant, but all estimated ϕ coefficients are negative). Unlike in FOMC1, in which more rookie members brought new topics into the dicussion, in FOMC2 they speak more like their colleagues and stick to a narrower agenda. Finally, our clearest measure of herding is the predicted probability of not dissenting in voice with Greenspan's proposal. Here we find that rookies are significantly more likely to not dissent after transparency relative to veterans, which strongly suggests the presence of a reputation-induced bias toward conformity during the policy discussion.[24]

Taken together, our results are consistent with the presence of both discipline and conformity effects on the FOMC in response to transparency. Regarding discipline, we find an increase in the use of numbers and data throughout the meeting and a greater diversity of topics discussed by more inexperienced members in the prepared statements on economic conditions in FOMC1. Regarding conformity, during the policy dicussion we find that more inexperienced members reduce their participation, discuss a more limited range of topics, and engage in more herding behavior.

V.C. Placebo Tests

As with all difference-in-differences strategies, an important identification assumption is that the heterogeneous responses in communication with respect to experience which we observe after transparency are in fact a response to transparency rather than a continuation of heterogeneous patterns that existed before transparency. To assess the appropriateness of this assumption, we conduct a placebo test on the pretransparency meetings from January 1988 through October 1993, again focusing on the set of members present on the FOMC during the natural experiment.[25] We define a placebo event in October 1990 and repeat the estimation of (DinD) with an indicator variable for this event in place of $D(Trans)_t$. If our identification assumption is valid, we should not find systematically significant estimates of ϕ.

Tables B.1 and B.2 in Appendix B present results. The only significant effect we find is on words in FOMC2, which

24. This result is distinct from that in Meade and Stasavage (2008) for two main reasons. First, we use a continuous measure of voiced dissent, albeit derived in part from the original multinomial voiced dissent variable. Second, we adopt a difference-in-differences approach, whereas Meade and Stasavage (2008) use a difference analysis. We have used the “no voiced dissent” categorical variable directly from Meade and Stasavage (2008) as a dependent variable in (DinD) but not obtained a significant result on the interaction term (although the sign is consistent with rookies dissenting less often).

25. We exclude the first few meetings of Greenspan’s tenure from the placebo to avoid any effects arising from the transition to a new chair. For example, in Greenspan’s first meetings, the separation between FOMC1 and FOMC2 was less clear than in those from 1988 onward.

is not an effect associated with the natural experiment—the estimated ϕ coefficient in column (1) of Table XI is insignificant. In terms of magnitudes, the rookie effect on numbers in FOMC1 in the placebo is in line with that in the natural experiment. Its insignificance in the placebo regression is consistent with pure noise. Certainly there seems to be no systematic sense in which less experienced members use more quantitative analysis in the pretransparency period: the placebo effect on quantitative topics in FOMC1 is negative (but positive in the natural experiment); the effect on numbers in FOMC2 is negative (ditto); and the effect on quantitative topics in FOMC2 is small in magnitude (as in the natural experiment).

Another possibility is that the reduced sample size in the placebo means that there is not enough power to estimate significant effects. To address this, we have repeated the placebo estimates using the full sample of votes in the period, which increases the sample size to 800. Again, we find hardly any significant effects. The rookie effect on numbers in FOMC1 is 38.4 but remains insignificant.

Overall, we are satisfied that our results are not driven by preexisting differential trends in communication depending on the experience level.

VI. Transparency and Influence

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.

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 setup, 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 IV.E. Let $\mathbf{W}_t$ be a within-meeting influence matrix. FOMC1 and FOMC2 share four policy topics: 1, 5, 26, and 39. Let $\boldsymbol{\chi}_{i,t,s}$ be the conditional distribution for speaker i in meeting t in section s over these four topics. $\mathbf{W}_t(i, j)$ is then the similarity between these distributions for member i in FOMC1 and j in FOMC2.

For the second step, use $\mathbf{W}_t$ to obtain a Markov matrix $\mathbf{W}'_t$ by way of the column normalization $\mathbf{W}'_t(i, j) = \frac{\mathbf{W}_t}{\sum_j \mathbf{W}_t(i,j)}$. From there, we measure the within-meeting influence of member i in meeting t as the ith element of the (normalized) eigenvector associated with the unit eigenvalue of $\mathbf{W}'_t$. Denote this value by W_{it}. Loosely speaking, W_{it} measures the relative contribution of member i's FOMC1 policy topics in shaping the policy topics of all members in FOMC2. Since 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.[26]

Some observers have argued that in fact influence across meetings is more important than influence within meetings.[27] We therefore define an across-meeting influence matrix $\mathbf{A}_t$ where $\mathbf{A}_t(i, j)$ is the similarity between member i's distribution over policy topics in FOMC2 in meeting t and member j's distribution over policy topics in FOMC2 in meeting $t + 1$. We then arrive

26. Meyer (2004) notes that "the Chairman exercised such disproportionate power that unless you could sway him over to your point of view, your view was not going to prevail."

27. 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."

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.[28]

Overall we obtain four measures of influence for each member and meeting: influence on the FOMC as a whole and influence on Greenspan, both within and across meetings. Moreover, each measure is constructed using the three different measures of similarity in Section IV.E. In the regression tables, Avg Infl $(\mathrm{X})_{i,\,t}$ denotes influence on the whole FOMC; within meetings this is W_{it} and across meetings this is A_{it}, each computed using similarity measure X (as before, B=Bhattacharyya, D=dot product and KL=Kullback-Leibler). Chair Infl $(\mathrm{X})_{i,\,t}$ is the influence on Greenspan defined above.

Table XIII displays the results for influence. For all similarity measures, average within-meeting influence for rookies rises significantly after transparency, and influence on Chairman Greenspan rises according to the Kullback-Leibler measure. The across-meeting influence results are even stronger, with every influence measure rising significantly more for rookies after transparency on the FOMC as a whole and on Greenspan. 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.

We have also conducted placebo tests on influence in the same way as described in Section V.C. The results are in Table B.3 in the Appendix. We again find no significant results on the placebo, and particularly small effects on across-meeting influence. The same is true in the larger sample that uses all members during the placebo period.

The influence results show that what inexperienced members speak about after transparency is more predictive of what others (and specifically the chair) speak about in the future. The presence of a net positive informational effect supports the conclusion

28. Table C.1 in the Appendix presents a ranking of members by their overall intermeeting influence and their intermeeting influence on Greenspan.

TABLE XIII

INFLUENCE RESULTS

Main regressors	Avg Infl (B) (1)	Avg Infl (D) (2)	Avg Infl (KL) (3)	Chair Infl (B) (4)	Chair Infl (D) (5)	Chair Infl (KL) (6)
Panel A: Influence within meeting						
D(Trans) × Fed experience	−0.000046**	−0.000086*	−0.00017***	−3.1e-06	−3.0e-06	−0.000012*
	[.042]	[.067]	[.010]	[.141]	[.614]	[.069]
Fed experience	−0.0062	−0.011	−0.017	−0.00033	2.6e-06	−0.00057
	[.235]	[.231]	[.249]	[.488]	[.998]	[.675]
# Stems	−4.1e-06***	8.4e-07	−0.000011***	−4.4e-07***	9.7e-08	−1.2e-06***
	[.000]	[.771]	[.000]	[.000]	[.768]	[.000]
Observations	910	910	910	910	910	910
Unique members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Topics	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2
Similarity measure	Bhattacharyya	Dot product	Kullback-Leibler	Bhattacharyya	Dot product	Kullback-Leibler
Rookie effect	1.8**	3.3*	6.7***	2.2	2.1	8.3*

TABLE XIII
CONTINUED

Main regressors	Avg Infl (B) (1)	Avg Infl (D) (2)	Avg Infl (KL) (3)	Chair Infl (B) (4)	Chair Infl (D) (5)	Chair Infl (KL) (6)
Panel B: Influence across meetings						
D(Trans) × Fed experience	−0.000022**	−0.00015***	−0.000081**	−2.4e-06*	−0.000019***	−9.9e-06**
	[.047]	[.000]	[.026]	[.051]	[.000]	[.019]
Fed experience	−0.00054	−0.019**	−0.000065	0.00011	−0.0013	0.00064
	[.861]	[.046]	[.995]	[.712]	[.177]	[.590]
# Stems	−1.6e-06***	−7.8e-07	−4.7e-06***	−1.4e-07*	7.0e-08	−3.6e-07
	[.001]	[.605]	[.007]	[.067]	[.803]	[.250]
Observations	892	892	892	892	892	892
Unique members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Topics	P2	P2	P2	P2	P2	P2
Similarity measure	Bhattacharyya	Dot product	Kullback-Leibler	Bhattacharyya	Dot product	Kullback-Leibler
Rookie effect	0.8**	5.5***	2.9**	1.6*	12.4***	6.3**

Notes. This table reports the results of estimating (DinD) for measures of influence defined in the main text. Panel A presents results for influence within an FOMC meeting, which are defined using the intersection of policy topics in FOMC1 and FOMC2 defined in Table III. Panel B presents results for influence across FOMC meetings, which are defined using the FOMC2 policy topics. Coefficients are labeled according to significance ($^{***}p < .01$, $^{**}p < .05$, $^{*}p < .1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Trans)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before November 1993. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

that the increase in information production due to higher effort is significantly larger than the reduction in information disclosure due to the desire to conform. Under this metric, the discipline effect of transparency appears to be stronger than the conformity effect.

VII. Robustness

In Tables D.1–D.3 in the Appendix, we explore the robustness of the main difference-in-differences results presented in the main text. In each table we report the estimated rookie effect labeled according to significance using *, **, or ***. The first line replicates the baseline results from the main text for comparison. As described below, the results for each robustness test are very similar to the baseline results: in nearly every case the sign of the rookie effect is the same. Although there is some variation in significance levels of the influence results depending on the similarity measure used—particularly for those robustness tests in which we reduce sample size—our main result that rookies become more influential after transparency remains.

VII.A. Committee Composition

First we consider issues related to the committee composition. The baseline analysis focuses on a core sample of 19 members who were present at the meetings immediately before and after the October 1993 revelation that the transcripts existed. This is to address the concern that the appointment of more prominent scholars from 1994, such as Alan Blinder and Janet Yellen, led to more academic and technical discourse, which may drive the results. The downside, as discussed already, is that we are estimating time fixed effects using a smaller-than-complete-FOMC subset of the membership in many of the meetings. As such, our first composition robustness check involves using all 19 members in each meeting in our baseline sample window. With more observations, but with changing composition of members, the results are not materially changed.

Another way to address any concern about the increasingly technical deliberation is to control, at the individual level, for the number of other FOMC members in each meeting who hold PhDs. We estimate this model on the full, noncore sample to capture the impact of the changing composition toward more academics. Although this control is often significant in its own right (not reported), it does not affect the estimates of the rookie effects following increased transparency.

Third, one may be concerned about the role of governors in driving the results. As described above, the Clinton administration appointed more technical governors during the 1990s. Another possibility is that during our sample Greenspan became increasingly dominant, especially among the Board of Governors, who were based in Washington, DC. Finally, concern may stem from the anecdotal evidence that the chairman engaged in informal, bilateral premeeting discussions with governors but not with presidents (see Section III.B). To address these concerns that the governors alone drive the results we find, we return to our core sample but drop all the nonchair governors. This reduces the sample size but our main results get stronger, which should alleviate such concerns.[29]

Fourth, we 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 Angell. According to the account in Lindsey (2003), they all found out earlier in 1993 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. The estimated effects tend to be larger than the baseline analysis.

Finally, we drop any member until they have served on at least four FOMC meetings. This is to address the concern, raised by Warsh (2014), that the only effect of relative inexperience would be noticeable in the first six months on the committee. Our results remain largely robust. It is perhaps unsurprising, and even reassuring, that removing those groups whom we think are most susceptible to the effects of changing transparency would weaken the results somewhat.

VII.B. Sample Selection

We switch attention to issues of the meeting sample selection. In this regard, the first robustness exercise that we carry

29. These results are consistent with presidents being more responsive to career concerns than governors. One explanation is that governors are accountable to Congress and the Executive Branch, but presidents are also accountable to their own boards, on top of likely being scrutinized by Congress.

out is to tighten the window by one year before and one year after the change. With such a sample, we consider a six-year window (November 1990 to September 1996) rather than the baseline eight-year window. This reduces the total number of meeting-speaker observations from 920 to 731; the statistical significance of some results, especially the influence measures, is reduced as a result. The estimates of the rookie effects are relatively unchanged.

Second, we exclude July 1993 to July 1994. Dropping meetings from July 1993 to October 1993 should confirm that it is not this period immediately before October 1993 that drives our results. Despite most members claiming (to each other in a conference call) that they did not know of the transcripts, we already mentioned that a few members certainly knew of them prior to October 1993.[30] We drop meetings from November 1993 to July 1994 because although October 1993 marked the decision to release the back catalog of transcripts, no decision to release post-1993 transcripts had been made. While it may have been relatively easy to predict that the FOMC would follow suit in releasing future transcripts, the internal FOMC committee set up to consider this issue only made the recommendation to do so in July 1994. Beyond the reduced significance of some of the influence measures, results are very similar to the baseline.

A last issue on meeting sample selection is our decision not to include the FOMC conference calls. In the intermeeting period, the FOMC can meet via a phone call to discuss committee matters from economic news, issues relating to the Federal Reserve's engagement in international meetings and issues about the committee organization. We decided not to include these conference calls in the baseline sample for a number of reasons. First, these calls do not follow a fixed structure, and particularly they do not always have any discussion of monetary policy issues. Second, the transcript record for a number of these calls in the earlier years is missing as no minutes were necessarily produced as a result of the call; for these calls we know only what the planned agenda was rather than the precise contributions for each member. Finally, many of the conference calls involved mostly information being relayed from Greenspan and it seems, at least in terms of reading the transcripts, that back-and-forth discussion in such calls was

30. We have also followed Meade and Stasavage (2008) and excluded only 1993 from the estimation. The results remain robust.

especially lower than in a regular FOMC meeting; this could be driven by the conference call format or the specific agenda.

Nonetheless, we examined the conference call data. There are 35 conference calls that took place within our baseline window. Twenty-seven took place before November 1993 and eight after. October 1993 alone accounts for five conference calls to discuss the evolving situation with the House Banking Committee regarding transparency (we quoted from these calls already). Many are used to give an update of the economic situation in the intermeeting period, and seven calls relate to a decision to change monetary policy (either made in the meeting or Greenspan updating the FOMC of his decision to exercise a tilt directive given to him by the FOMC in the preceding meeting).

Regardless of whether a decision on monetary policy is involved in the call, substantive discussions of the economic environment in a conference call may have a bearing on the nature of the deliberation that takes place in the sections that we analyze (FOMC1 and FOMC2) of the following FOMC meeting. Although this should be picked up by the meeting fixed effect, it could affect the results if the conference call differentially affected those with more or less experience. Moreoever, there are more conference calls involving monetary policy discussion in our sample before the change in transparency (16) than after (5). As such, we run our analysis dropping any FOMC meeting that follows a conference call that discussed monetary policy issues (including an update of economic conditions). Dropping these meetings reduces the significance of the estimated rookie effects in some cases, but the results are unchanged qualitatively.

VII.C. LDA Model Selection

We now address two issues related to the LDA model selection. First, as discussed in the main text, we use a 40-topic LDA model in the baseline analysis for interpretability of the topics. We have also carried out the analysis using a 70-topic model. We selected this alternative as this is the size of topic model that provides the best out-of-sample fit (see Appendix A for a discussion of this).

Second, as discussed in the Online Technical Appendix, we choose for analysis the Markov chain that achieves the best average fit across sample draws, but this chain exhibits somewhat more volatility than the others. We therefore repeat all analysis for the chain with the lowest standard deviation in goodness-of-fit across draws. This also allows us to explore to what extent our results are driven by one particular Markov chain, which is a concern with LDA because the posterior has potentially many modes.

In both cases, the results are very similar. We lose significance of rookies using more quantitative topics in FOMC1, but gain significance of them using more quantitative topics in FOMC2. And with the 70-topic model, we lose significance of the increased probability of no dissent by rookies even though the magnitude of the rookie effect is similar.

The final issue we address related to LDA estimation is uncertainty arising from the sampling algorithm we use to estimate our communication measures. In our regressions, we use the average values of the measures across 80 draws from a Markov chain. We have also repeated all the baseline regressions for each draw, and thereby generated a distribution of rookie effects for each communication measure. The final row of Tables D.2 and D.3 in the Appendix reports the range of the 10th to 90th percentiles of these distributions.[31] In this exercise, we keep the set of policy topics fixed across draws. For this reason, we do not report sampling distributions for the probability-of-no-dissent regressions because this would require computing new policy topics on each draw.

VIII. 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 among less experienced members). 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 U.S. monetary policy making improved after 1993 as a result of transparency, but does suggest that transparency was responsible for changing policy makers' information sets in a meaningful way.

The main policy implication of our results is that central bank designers should seek to maximize the discipline effect and

31. Since the count measures are not based on a sampling algorithm, there is no distribution reported in Table D.1

minimize the conformity effect given that both are present in the data and have clear welfare implications. One example of how this insight might be implemented is the recently reformed disclosure policies at the Bank of England (Warsh 2014), whose Monetary Policy Committee (MPC) holds monthly two-day meetings. 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. Indeed, the Bank of England committed in August 2015 to publish transcripts from the second day of MPC meetings (with an eight-year delay) but not those from the first day.

Finally, our article highlights the value of machine learning in textual analysis. There are several approaches to automated text analysis (many of these are discussed in Gentzkow, Kelly, and Taddy 2017), but the economics literature to date has focussed primarily on keyword searches and counting words from prespecified lists. While these remain valuable tools, our article shows that machine-learning algorithms can uncover an interpretable latent space in large textual databases concerning the macroeconomy and faciliate the construction of rich communication measures. We believe this methodology has numerous potential applications beyond our work.

Appendix A: Model Selection

As discussed in the main text, we choose 40 topics primarily for interpretability, but an alternative is to choose the number of topics K based on a statistical criterion. We adopt perhaps the most popular approach—cross-validation—in which K is chosen based on the ability of the model to predict out-of-sample observations. We first randomly draw two-thirds of our sample of FOMC transcript interjections as training data, and fit an LDA model for various values of K beginning from $K = 10$. Then we take the estimated parameters and compute the goodness-of-fit for the test data (the held-out one-third of observations) using perplexity, a

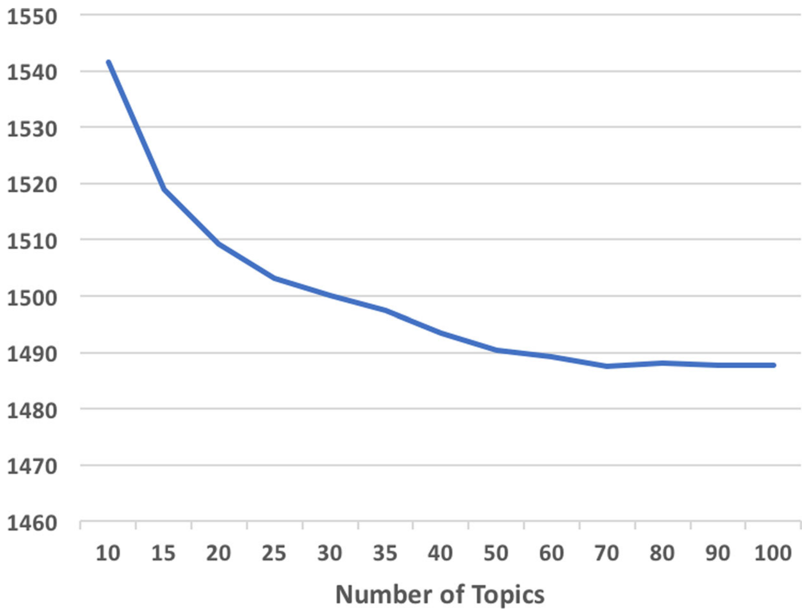


FIGURE A.1

Average Perplexity of Test Data for Different Topics

This figure shows the average perplexity, calculated according to the formula given by (1). These data show that as the number of topics increase, the goodness-of-fit of the model improves until around $K = 70$ after which it is flat.

standard measure in the machine learning literature given by

$$\exp\left[-\frac{\sum_d \sum_v x_{d,v} \log\left(\sum_k \beta_k^v \theta_d^k\right)}{\sum_d N_d}\right] \tag{A.1}$$

where $x_{d,\,v}$ is the count of term v in document d and N_d is the total number of terms in document d. Here the relevant documents are the test sample. We use the estimated value of β_k^v from the LDA estimation on the training data, and a uniform distribution for θ_d^k to compute perplexity as implied by the Dirichlet prior and as suggested by Grün and Hornik (2011). We repeat this procedure ten times, each time randomly drawing the training data. Figure A.1 reports the average perplexity computed on the test data across these ten draws. Lower values indicate better goodness-of-fit.

As we increase the number of pure behaviors, we can indeed better fit language patterns, as can be seen from the decreasing perplexity. Naturally, the most parsimonious model does not account for all the underlying correlations in the high-dimensional feature space. At the same time, the improvement in fit levels off fairly quickly, and the average perplexity stays flat after $K = 70$. For this reason, we choose $K = 70$ as the model that best fits the data as it does so with the fewest parameters.

Appendix B: Placebo Tables

TABLE B.1

Placebo Results for Economic Situation Discussion (FOMC1)

(a) Count Measures

Main Regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Placebo) × Fed Experience	−2.45	0.068	0.023	−0.16
	[0.395]	[0.210]	[0.346]	[0.227]
Fed Experience	−870*	−1.14	−4.68	20.8
	[0.053]	[0.833]	[0.346]	[0.373]
Observations	598	598	598	598
Unique Members	19	19	19	19
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC1
Rookie effect	8.8	−21.9	−27	47.5

(b) Topic Measures

Main Regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)
D(Placebo) × Fed Experience	0.00029	0.00030	−0.00038	−0.00015	−0.00067
	[0.780]	[0.351]	[0.443]	[0.723]	[0.604]
Fed Experience	0.020	−0.0017	−0.17**	−0.11	−0.50**
	[0.867]	[0.968]	[0.028]	[0.133]	[0.016]
# Stems	0.000082***	5.2e-06	7.7e-07	0.000037***	0.000069
	[0.001]	[0.652]	[0.967]	[0.000]	[0.189]
Observations	598	598	598	598	598
Unique Members	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC1	FOMC1	FOMC1	FOMC1	FOMC1
Topics	P1	T4 & T23	P1	P1	P1
Similarity Measure	–	–	Bhattacharyya	Dot Product	Kullback-Leibler
Rookie effect	−3.6	−17.5	0.8	2.9	2

Notes. These tables report the results presented in Tables IX and X but under the placebo transparency change. See those tables for notes. The placebo transparency change is imposed as taking place October 1990. Coefficients are labeled according to significance (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Placebo)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before October 1990. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

TABLE B.2

PLACEBO RESULTS FOR POLICY STRATEGY DISCUSSION (FOMC2)

(a) Count Measures

Main Regressors	Words (1)	Statements (2)	Questions (3)	Numbers (4)
D(Placebo) × Fed Experienc	−8.47***	0.065	0.018	0.047
	[0.009]	[0.645]	[0.700]	[0.660]
Fed Experience	777	−0.91	−7.25	−0.96
	[0.105]	[0.918]	[0.276]	[0.938]
Observations	590	590	590	590
Unique Members	19	19	19	19
Member FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes
Meeting Section	FOMC2	FOMC2	FOMC2	FOMC2
Rookie effect	53.8***	−18.7	−24.3	−27.4

(b) Topic Measures

Main Regressors	Concentration (1)	Quant (2)	Avg Sim (B) (3)	Avg Sim (D) (4)	Avg Sim (KL) (5)	Pr(No Dissent) (6)
D(Placebo) × Fed Experience	0.00036	−7.2e-06	−0.00029	0.000090	−0.00095	0.000052
	[0.678]	[0.969]	[0.398]	[0.856]	[0.419]	[0.953]
Fed Experience	0.20**	−0.011	0.00051	0.066	0.066	0.049
	[0.043]	[0.540]	[0.994]	[0.274]	[0.756]	[0.775]
# Stems	0.000024	−4.8e-06	0.000010	0.000011	0.000066	−0.00016**
	[0.307]	[0.484]	[0.496]	[0.430]	[0.201]	[0.012]
Observations	590	590	590	590	590	590
Unique Members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Meeting Section	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2	FOMC2
Topics	P2	T4 & T23	P2	P2	P2	P2
Similarity Measure	–	–	Bhattacharyya	Dot Product	Kullback-Leibler	–
Rookie effect	−4	0.4	0.6	−1.2	2.5	−0.1

Notes. These tables report the results presented in Tables XI and XII but under the placebo transparency change. See those tables for notes. The placebo transparency change is imposed as taking place October 1990. Coefficients are labeled according to significance (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Placebo)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before October 1990. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

APPENDIX C: INFLUENCE RANKING

In Table C.1, we present a ranking of members by their overall FOMC 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

TABLE B.3
PLACEBO INFLUENCE RESULTS

(a) Influence within Meeting

Main Regressors	Avg Infl (B) (1)	Avg Infl (D) (2)	Avg Infl (KL) (3)	Chair Infl (B) (4)	Chair Infl (D) (5)	Chair Infl (KL) (6)
D(Placebo) × Fed Experience	−0.000066	−0.000091	−0.00021	−7.8e-06	−0.000010	−0.000025
	[0.277]	[0.422]	[0.217]	[0.263]	[0.555]	[0.233]
Fed Experience	0.00035	−0.014	−0.00012	−0.00081	−0.0046***	−0.0026
	[0.968]	[0.200]	[0.996]	[0.391]	[0.002]	[0.284]
# Stems	−3.6e-06***	4.3e-06	−8.8e-06***	−3.1e-07*	7.5e-07	−6.6e-07
	[0.001]	[0.104]	[0.006]	[0.055]	[0.223]	[0.203]
Observations	587	587	587	587	587	587
Unique Members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Topics	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2
Similarity Measure	Bhattacharyya	Dot Product	Kullback-Leibler	Bhattacharyya	Dot Product	Kullback-Leibler
Influence Horizon	Within	Within	Within	Within	Within	Within
Rookie effect	2.4	3.3	7.8	4.9	6.4	15.4

TABLE B.3
CONTINUED

(b) Influence across Meetings

Main Regressors	Avg Infl (B) (1)	Avg Infl (D) (2)	Avg Infl (KL) (3)	Chair Infl (B) (4)	Chair Infl (D) (5)	Chair Infl (KL) (6)
D(Placebo) × Fed Experience	−8.9e-06	0.000031	−7.0e-06	−3.1e-06	3.5e-06	−7.4e-06
	[0.756]	[0.828]	[0.950]	[0.434]	[0.828]	[0.607]
Fed Experience	−0.0091*	−0.0048	−0.032*	−0.00085	0.00049	−0.0037
	[0.058]	[0.815]	[0.056]	[0.233]	[0.837]	[0.147]
# Stems	−1.9e-06	−1.5e-06	−4.6e-06	−5.4e-08	−6.2e-08	4.0e-07
	[0.199]	[0.642]	[0.474]	[0.877]	[0.900]	[0.824]
Observations	576	576	576	576	576	576
Unique Members	19	19	19	19	19	19
Member FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Topics	P2	P2	P2	P2	P2	P2
Similarity Measure	Bhattacharyya	Dot Product	Kullback-Leibler	Bhattacharyya	Dot Product	Kullback-Leibler
Influence Horizon	Across	Across	Across	Across	Across	Across
Rookie effect	0.3	−1.1	0.2	1.7	−1.9	3.9

Notes. These tables report the results presented in Table XIII but under the placebo transparency change. See that table for notes. The placebo transparency change is imposed as taking place October 1990. Panel (a) presents results for influence within an FOMC meeting, which are defined using the intersection of policy topics in FOMC1 and FOMC2 defined in Table III. Panel (b) presents results for influence across FOMC meetings, which are defined using the FOMC2 policy topics. Coefficients are labeled according to significance (***$p < 0.01$, **$p < 0.05$, *$p < 0.1$) while brackets below coefficients report p-values calculated using Driscoll-Kraay standard errors. The rookie effect reports the estimated coefficient on $D(Placebo)_t \times FedExp_{it}$ multiplied by 20 (approximate difference in experience between the two modes in Figure VI) as a percentage of the average value of the dependent variable before October 1990. These effects carry the same star labels as the corresponding estimated coefficient on $D(Trans)_t \times FedExp_{it}$.

TABLE C.1

INTERMEETING INFLUENCE MEASURES BY MEMBER

Speaker	Core	Meetings under Greenspan	Average Influence Fixed Effect	Average Influence Average	Speaker	Core	Meetings under Greenspan	Greenspan Influence Fixed Effect	Greenspan Influence Average
Heller		14	0.0017	0.0696	Heller		14	0.00167	0.00610
Johnson		22	0.0000	0.0624	Johnson		22	0.00001	0.00451
Kohn		29	−0.0003	0.0599	Kohn		29	−0.00026	0.00408
Santomero		45	0.0008	0.0598	Santomero		45	0.00083	0.00395
Keehn	*	55	0.0013	0.0598	Keehn	*	55	0.00127	0.00391
Gramlich		62	0.0008	0.0586	Angell	*	51	−0.00014	0.00382
Angell	*	51	−0.0001	0.0585	Poole		64	−0.00022	0.00363
Guffey		32	0.0008	0.0576	Gramlich		62	0.00082	0.00362
Poole		64	−0.0002	0.0576	Greenspan	*	148	−0.00062	0.00354
Minehan		96	0.0006	0.0572	Guffey		32	0.00076	0.00345
Black		40	0.0008	0.0568	Minehan		96	0.00057	0.00345
Greenspan	*	148	−0.0006	0.0567	Parry	*	133	0.00093	0.00332
Parry	*	133	0.0009	0.0567	Black		40	0.00075	0.00332
Fisher		7	0.0013	0.0564	Corrigan		46	0.00054	0.00323
Corrigan		46	0.0005	0.0560	Syron	*	41	0.00070	0.00319
Syron	*	41	0.0007	0.0560	Forrestal	*	65	0.00069	0.00313
Forrestal	*	65	0.0007	0.0556	Kelley	*	113	0.00042	0.00310
Moskow		92	0.0008	0.0555	Melzer	*	82	−0.00005	0.00308
Kelley	*	113	0.0004	0.0554	Fisher		7	0.00130	0.00306
Mcdonough	*	79	0.0004	0.0554	Moskow		92	0.00077	0.00305
Ferguson		67	0.0004	0.0553	Meyer		45	0.00014	0.00305
Hoenig	*	116	0.0009	0.0553	Stern	*	147	0.00088	0.00305
Stern	*	147	0.0009	0.0552	Bies		34	0.00043	0.00304
Meyer		45	0.0001	0.0552	Hoenig	*	116	0.00093	0.00304
Guynn		81	0.0004	0.0551	Ferguson		67	0.00043	0.00303

TABLE C.1

CONTINUED

Speaker	Core	Meetings under Greenspan	Average Influence Fixed Effect	Average Influence Average	Speaker	Core	Meetings under Greenspan	Greenspan Influence Fixed Effect	Greenspan Influence Average
Boykin		27	0.0009	0.0550	Seger		28	0.00046	0.00303
Melzer	*	82	– 0.0001	0.0549	Boehne	*	99	0.00080	0.00302
Boehne	*	99	0.0008	0.0549	Geithner		18	0.00009	0.00302
Bies		34	0.0004	0.0548	Boykin		27	0.00086	0.00301
Broaddus	*	91	0.0003	0.0548	Mcdonough	*	79	0.00045	0.00301
Seger		28	0.0005	0.0547	Guynn		81	0.00041	0.00300
Yellen		34	0.0008	0.0547	Mcteer	*	110	0.00038	0.00300
Rivlin		24	0.0012	0.0546	Broaddus	*	91	0.00028	0.00299
Geithner		18	0.0001	0.0546	Lacker		14	– 0.00046	0.00298
Mcteer	*	110	0.0004	0.0545	Olson		34	0.00040	0.00297
Phillips	*	52	0.0013	0.0544	Phillips	*	52	0.00133	0.00296
Morris		9	0.0006	0.0543	Rivlin		24	0.00119	0.00295
Lacker		14	– 0.0005	0.0542	Laware	*	53	0.00082	0.00295
Laware	*	53	0.0008	0.0542	Morris		9	0.00057	0.00294
Mullins	*	29	0.0009	0.0541	Yellen		34	0.00083	0.00294
Olson		34	0.0004	0.0541	Mullins	*	29	0.00086	0.00294
Pianalto		25	0.0000	0.0539	Jordan	*	86	– 0.00064	0.00293
Hoskins		30	– 0.0003	0.0537	Pianalto		25	– 0.00003	0.00293
Blinder		13	– 0.0003	0.0537	Hoskins		30	– 0.00028	0.00288
Jordan	*	86	– 0.0006	0.0536	Blinder		13	– 0.00027	0.00286
Bernanke		22	0.0002	0.0533	Bernanke		22	0.00022	0.00285
Lindsey	*	41	0.0002	0.0530	Lindsey	*	41	0.00025	0.00282
Stewart		4	0.0012	0.0521	Stewart		4	0.00121	0.00267

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

influence measure in the table on time and member fixed effects (Avg Infl $(\mathrm{B})_{A,i,t}$/Chair Infl $(\mathrm{B})_{A,i,t} = \alpha_{it} + \delta_t + \epsilon_{it}$). 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. 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.

Appendix D: Robustness Tables

TABLE D.1

Comparison of Results for Different Robustness Checks I

Meeting Section	Words FOMC1 (1)	Statements FOMC1 (2)	Questions FOMC1 (3)	Numbers FOMC1 (4)	Words FOMC2 (5)	Statements FOMC2 (6)	Questions FOMC2 (7)	Numbers FOMC2 (8)
Baseline	0.5	−6.4	−3.3	48.1***	17.3	−33.7**	−52.1**	77.8**
Noncore sample	2	−6.5	−2.4	52.7***	17	−34.0***	−54.3**	79.9***
Control for Other PhDs	2	−6.5	−2.4	52.7***	17	−34.0***	−54.3**	79.9***
Only Presidents	−4.7	−5.3	−9.2	61.4***	32.7	−59.5***	−80.9***	189.0***
Drop Knowing	−2.4	−7	−1.2	38.2**	19.7	−35.4***	−54.8**	97.5**
Drop if <4 FOMC meetings	−0.5	−8.8	−6.5	41.6***	23.5	−33.6**	−37.2**	87.5**
Narrow	3.3	0.4	3.3	50.6***	4.3	−30.3**	−55.3**	67.8**
Drop Jul 93 - Jul 94	2.5	−18.6	−5.5	49.0***	−5.4	−43.9***	−56.4**	30.7
Drop if pre-CC on MP	0.2	−9.4	−5.2	43.2***	10.7	−30.1**	−42.8	59.6*
70 Topic Model	0.5	−6.4	−3.3	48.1***	17.3	−33.7**	−52.1**	77.8**
Min Variance Chain	0.5	−6.4	−3.3	48.1***	17.3	−33.7**	−52.1**	77.8**
Sampling Uncertainty	[.]	[.]	[.]	[.]	[.]	[.]	[.]	[.]

Notes. This table reports, for all the robustness tests as reported in the main text, the difference-in-differences rookie effect. The corresponding baseline estimates are fully described in the main text in Tables IX and X. As these count measures do not use the LDA measures, they are not subject to sampling uncertainty and hence those entries are listed as [.].

TABLE D.2

COMPARISON OF RESULTS FOR DIFFERENT ROBUSTNESS CHECKS II

Meeting Section	Concentration FOMC1 (1)	Quant FOMC1 (2)	Avg Sim (B) FOMC1 (3)	Avg Sim (D) FOMC1 (4)	Avg Sim (KL) FOMC1 (5)	Concentration FOMC2 (6)	Quant FOMC2 (7)	Avg Sim (B) FOMC2 (8)	Avg Sim (D) FOMC2 (9)	Avg Sim (KL) FOMC2 (10)	Pr(No Dissent) FOMC2 (11)
Baseline	−4.7	24.3***	−1.4*	−7.0***	−5.9**	8.9**	5.6	0.4	5.5***	1.1	3.5**
Noncore sample	−4.7	23.1***	−1.3*	−6.5***	−5.4**	7.0*	5.7	0.5	4.8**	1.3	3.3**
Control for Other PhDs	−4.7	23.1***	−1.3*	−6.5***	−5.4**	7.0*	5.7	0.5	4.8**	1.3	3.3**
Only Presidents	−2.4	17.6*	−0.5	−1.4	−2.1	9.3***	11.4*	0.8	7.9***	2	4.1***
Drop Knowing	−5.2	25.8***	−1.8**	−8.6***	−7.1***	6.9	10.5	0.6	5.9**	1.6	5.6***
Drop if <4 FOMC meetings	−4.8	19.7**	−1.2	−6.3**	−5.2*	10.5***	6.9	0.2	5.0*	0.6	2.6*
Narrow	−3.6	27.1***	−1.3*	−6.1**	−5.2**	9.0**	6.5	0.2	5.3**	0.1	3.5**
Drop Jul 93 - Jul 94	−2.6	29.7***	−2.5***	−11.1***	−9.4***	8.2*	2.4	0.6	5.7***	1.6	4.9***
Drop if pre-CC on MP	−3.4	22.8***	−1.5***	−7.1***	−6.1***	9.2***	9.1	0.6**	6.4***	1.9*	3.9*
70 Topic Model	−1	29.8***	−1.2*	−2.4	−4.7*	9.3***	25.6**	0.9	6.7**	4.7**	1.3
Min Variance Chain	−7.2	0.2	−2.8***	−16.2***	−9.8***	9.2**	6.5*	0.4	5.0*	1.3	3.1**
Sampling Uncertainty	[−8.3, −1.5]	[15.8, 32.8]	[−2.2, −0.8]	[−9, −5.3]	[−8.4, −3]	[6.1, 11.9]	[0.3, 11.8]	[−0.2, 1]	[3.3, 7.6]	[−1.3, 3.3]	[.]
Topics	P1	T4 & T23	P1	P1	P1	P2	T4 & T23	P2	P2	P2	P2
Similarity Measure	–	–	B	DP	KL	–	–	B	DP	KL	–

Notes. This table reports, for all the robustness tests as reported in the main text, the difference-in-differences rookie effect. The corresponding baseline estimates are fully described in the main text in Tables XI and XII. Similarity measures are Bhattacharyya (B), Dot Product (DP) and Kullback-Leibler (KL) as described in the text.

TABLE D.3

COMPARISON OF RESULTS FOR DIFFERENT ROBUSTNESS CHECKS III

Influence Horizon	Avg Infl (B) Within (1)	Avg Infl (D) Within (2)	Avg Infl (KL) Within (3)	Chair Infl (B) Within (4)	Chair Infl (D) Within (5)	Chair Infl (KL) Within (6)	Avg Infl (B) Across (7)	Avg Infl (D) Across (8)	Avg Infl (KL) Across (9)	Chair Infl (B) Across (10)	Chair Infl (D) Across (11)	Chair Infl (KL) Across (12)
Baseline	1.7**	3.2*	6.5***	2.1	2.1	8.2*	0.8**	5.5***	2.9**	1.6*	12.4***	6.3**
Noncore sample	1.5*	2.4	5.8***	1.8	1.1	7.2*	0.9**	5.4***	3.1**	1.7**	12.1***	6.5**
Control for Other PhDs	1.5*	2.4	5.8***	1.8	1.1	7.2*	0.9**	5.4***	3.1**	1.7**	12.1***	6.5**
Only Presidents	3.4***	7.0*	12.0***	5.9**	11.2	20.6**	1.5**	9.0***	5.6***	2.6**	18.1***	10.3**
Drop Knowing	1.3	2.6**	5.3**	1.1	0.2	5.4	0.8*	5.0***	2.8**	1.8**	12.4***	7.0**
Drop if <4 FOMC meets	1.8**	3.6*	6.7**	2.1	2.3	8.1	0.6	5.5***	2.4*	1.4	12.9***	5.5**
Narrow	1.3	1.8	4.9*	1.2	−1	4	0.6	5.7***	2.4	1	12.1***	4.4
Drop Jul 93 - Jul 94	1.8**	3.1*	7.3***	1.1	−2.7	5.2	0.9***	5.1***	3.0***	1.3	10.7***	5.1
Drop if pre-CC on MP	2.0**	3.5**	7.4***	2.6	3.4	9.6*	0.9**	6.1***	3.4**	1.7*	13.1***	6.4**
70 Topic Model	1.9***	1.7	5.1**	3.0**	1.7	6.9*	0.1	6.0***	2	0.7	14.4***	5.9
Min Variance Chain	0.7	2.3	2.8	0.5	1.2	3.1	0.8***	4.9***	2.8**	2.3***	13.1***	8.6***
Sampling Uncertainty	[0.5, 2.8]	[1.7, 5.1]	[2.7, 10]	[−0.5, 4.7]	[−3.9, 7.9]	[−1, 16.5]	[0.2, 1.4]	[4, 7.4]	[1.1, 4.9]	[0.4, 2.8]	[8.8, 16.4]	[2, 11]
Topics	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P1 ∩ P2	P2	P2	P2	P2	P2	P2
Similarity Measure	B	DP	KL	B	DP	KL	B	DP	KL	B	DP	KL

Notes. This table reports, for all the robustness tests as reported in the main text, the difference-in-differences rookie effect. The corresponding baseline estimates are fully described in the main text in Table XIII. Similarity measures are Bhattacharyya (B), Dot Product (DP) and Kullback-Leibler (KL) as described in the text.

Supplementary Material

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online. Data replicating tables and figures in this article can be found in Hansen, McMahon, and Prat (2017), in the Harvard Dataverse, doi:10.7910/DVN/XAR1WZ.

University of Oxford, Alan Turing Institute, and Centre for Economic Policy Research

University of Oxford, Centre for Economic Policy Research, Centre for Competitive Advantage in the Global Economy, Centre for Macroeconomics, and Centre for Applied Macroeconomic Analysis

Columbia University, Centre for Economic Policy Research

References

Acosta, Miguel, "FOMC Responses to Calls for Transparency," Finance and Economics Discussion Series 2015-60, Board of Governors of the Federal Reserve System (U.S.), 2015.

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

Avery, Christopher N., and Judith A. Chevalier, "Herding over the Career," *Economics Letters*, 63 (1999), 327–333.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, "Measuring Economic Policy Uncertainty," *Quarterly Journal of Economics*, 131 (2016), 1593–1636.

Bandiera, Oriana, Stephen Hansen, Andrea Prat, and Raffaella Sadun, "CEO Behavior and Firm Performance," NBER Working Paper No. 23248, 2017.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 28 (2014), 29–50.

Bird, Steven, Ewan Klein, and Edward Loper. *Natural Language Processing with Python* (CA, USA: O'Reilly Media, 2009).

Blei, David, "Probabilitic Topic Models," *Communications of the ACM*, 55 (2012), 77–84.

Blei, David, and John Lafferty, "Dynamic Topic Models," in *Proceedings of the 23rd Internaional Conference on Machine Learning,* 2006, 113–120.

———, "A Correlated Topic Model of Science," *Annals of Applied Statistics*, 1 (2007), 17–35.

———, "Topic Models," in *Mining: Classification, Clustering, and Applications*, A. Srivastava and M. Sahami, eds. (London: Taylor & Francis, 2009).

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

Bligh, Michelle C., and Gregory Hess, "Deconstructing Alan: A Quantitative Assessment of the Qualitative Aspects of Chairman Greenspan's Communication," in *Central Bank Communication, Decision Making, and Governance: Issues, Challenges, and Case Studies*, P. L. Siklos and J.-E. Sturm, eds. (Cambridge, MA: MIT Press, 2013), 123–147.

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

Boukus, Ellyn, and Joshua V. Rosenberg, "The Information Content of FOMC Minutes," Mimeo, Federal Reserve Bank of New York, 2006.
Budak, Ceren, Sharad Goel, Justin M. Rao, and Georgios Zervas, "Do-not-Track and the Economics of Third-Party Advertising," Boston University, School of Management Research Paper, 2505643, 2014.
Chang, Jonathan, Sean Gerrish, Chong Wang, Jordan L. Boyd-graber, and David M. Blei, "Reading Tea Leaves: How Humans Interpret Topic Models," in *Advances in Neural Information Processing Systems,* Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams and A. Culotta, eds. (Curran Associates, 2009), 288–296.
Chappell, Henry W., Thomas M. Havrilesky, and Rob Roy McGregor, "Monetary Policy Preferences of Individual FOMC Members: A Content Analysis of the Memoranda of Discussion," *Review of Economics and Statistics*, 79 (2000), 454–460.
Chappell, Henry W., Rob Roy McGregor, and Todd A. Vermilyea, *Committee Decisions on Monetary Policy: Evidence from Historical Records of the Federal Open Market Committee* (Cambridge, MA: MIT Press, 2005).
Cieslak, Anna, and Annette Vissing-Jørgensen, "The Economics of the Fed Put," SSRN Research Paper, 2017, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2951402.
Driscoll, John C., and Aart C. Kraay, "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *Review of Economics and Statistics*, 80 (1998), 549–560.
Egesdal, Michael, Michael Gill, and Martin Rotemberg, "How Federal Reserve Discussions Respond to Increased Transparency," SSRN Research Paper, 2015, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2676429.
Federal Reserve Bank of Philadelphia, "A Day in the Life of the FOMC: An Insider's Look at the Federal Reserve's Monetary Policy Body," 2008, https://www.philadelphiafed.org/education/teachers/resources/day-in-life-of-fomc.
Fehrler, Sebastian, and Niall Hughes, "How Transparency Kills Information Aggregation: Theory and Experiment," *American Economic Journal: Microeconomics*, forthcoming.
Financial Times, "ECB Heads Towards More Transparency over Minutes," August 1, 2013, http://www.ft.com/cms/s/0/a678451e-fa99-11e2-87b9-00144feabdc0.html#axzz2fou0Xcm5.
Fligstein, Neil, Jonah Stuart Brundage, and Michael Schultz, "Why the Federal Reserve Failed to See the Financial Crisis of 2008: The Role of 'Macroeconomics' as a Sensemaking and Cultural Frame," Mimeo, University of California Berkeley, 2014.
Friedman, Jerome, Trevor Hastie, and Robert Tibshirani, "Regularization Paths for Generalized Linear Models via Coordinate Descent," *Journal of Statistical Software*, 33 (2010), 1–22.
Gentzkow, Matthew, Bryan T. Kelly, and Matt Taddy, "Text as Data," NBER Working Paper No. 23276, 2017.
Gentzkow, Matthew, and Jesse M. Shapiro, "What Drives Media Slant? Evidence from U.S. Daily Newspapers," *Econometrica*, 78 (2010), 35–71.
Geraats, Petra M., "Central Bank Transparency," *Economic Journal*, 112 (2002), 532–565.
———, "Trends in Monetary Policy Transparency," *International Finance*, 12 (2009), 235–268.
Gersbach, Hans, and Volker Hahn, "Information acquisition and transparency in committees," *International Journal of Game Theory*, 41 (2012), 427–453.
Greenspan, Alan, "Comments on FOMC Communication," in U.S. House of Representatives, Hearing before the Committee on Banking, Finance and Urban Affairs, 103rd Congress, October 19, 1993.
Greider, William, *Secrets of the Temple* (New York: Simon Schuster, 1987).
Griffiths, T. L., and M. Steyvers, "Finding Scientific Topics," *Proceedings of the National Academy of Sciences*, 101 (2004), 5228–5235.

Grün, Bettina, and Kurt Hornik, "topicmodels: An R Package for Fitting Topic Models," *Journal of Statistical Software*, 40 (2011).
Hansen, Stephen, and Michael McMahon, "First Impressions Matter: Signalling as a Source of Policy Dynamics," *Review of Economic Studies*, 83 (2016), 1645.
Hansen, Stephen, Michael McMahon, and Andrea Prat, "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach," CEPR Discussion Paper 9994, 2014.
———, "Replication Data for: 'Transparency and Deliberation within the FOMC: A Computational Linguistics Approach'," Harvard Dataverse, doi:10.7910/DVN/XAR1WZ, 2017.
Hansen, Stephen, Michael McMahon, and Carlos Velasco, "How Experts Decide: Preferences or Private Assessments on a Monetary Policy Committee?," *Journal of Monetary Economics*, 67 (2014), 16–32.
Hazen, Timothy, "Direct and Latent Modeling Techniques for Computing Spoken Document Similarity," 2010 IEEE Spoken Language Technology Workshop, Berkeley, CA, 2010, 366–371.
Hendry, Scott, "Central Bank Communication or the Media's Interpretation: What Moves Markets?" Working Paper 12-9, Bank of Canada, 2012.
Hendry, Scott, and Alison Madeley, "Text Mining and the Information Content of Bank of Canada Communications," Working Paper 10–31, Bank of Canada, 2010.
Hofmann, Thomas, "Probabilistic Latent Semantic Indexing," in *Proceedings of the 22nd International Conference on Research and Development in Information Retreival*, (New York: ACM, 1999), 50–57.
Holmström, Bengt, "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66 (1999), 169–182.
Hong, Harrison, Jeffrey D. Kubik, and Amit Solomon, "Security Analysts' Career Concerns and Herding of Earnings Forecasts," *RAND Journal of Economics*, 31 (2000), 121–144.
Justeson, John S., and Slava M. Katz, "Technical Terminology: Some Linguistic Properties and an Alogrithm for Identification in Text," *Natural Language Engineering*, 1 (1995), 9–27.
Levy, Gilat, "Anti-Herding and Strategic Consultation," *European Economic Review*, 48 (2004), 503–525.
———, "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, 97 (2007), 150–168.
Lindsey, David E., *A Modern History of FOMC Communication: 1975–2002* (Washington, DC: Federal Reserve Board of Governors, 2003).
Lombardelli, Clare, James Proudman, and James Talbot, "Committees versus Individuals: An Experimental Analysis of Monetary Policy Decision-Making," *International Journal of Central Banking*, 1 (2005), 181–205.
Lucca, David O., and Francesco Trebbi, "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements," NBER Working Paper No. 15367, 2009.
Mankiw, N. Gregory, "U.S. Monetary Policy during the 1990s," NBER Working Paper No. 8471, 2001.
Meade, Ellen, "The FOMC: Preferences, Voting, and Consensus," *Federal Reserve Bank of St. Louis Review*, 87 (2005), 93–101.
Meade, Ellen E., and David Stasavage, "Publicity of Debate and the Incentive to Dissent: Evidence from the US Federal Reserve," *Economic Journal*, 118 (2008), 695–717.
Meade, Ellen E., and Daniel L. Thornton, "The Phillips Curve and US Monetary Policy: What the FOMC Transcripts Tell Us," *Oxford Economic Papers*, 64 (2012), 197–216.
Meyer, Laurence. *A Term at the Fed: An Insider's View* (New York: Collins, 2004).
Mueller, Hannes, and Christopher Rauh, "Reading between the Lines: Prediction of Political Violence Using Newspaper Text," *American Political Science Review*, (2017), forthcoming.

Nimark, Kristoffer P., and Stefan Pitschner, “Delegated Information Choice,” CEPR Discussion Paper, 11323, 2016.
Ottaviani, Marco, and Peter Norman Sørensen, “Reputational Cheap Talk,” *RAND Journal of Economics*, 37 (2006), 155–175.
Palacios-Huerta, Ignacio, and Oscar Volij, “The Measurement of Intellectual Influence,” *Econometrica*, 72 (2004), 963–977.
Prat, Andrea, “The Wrong Kind of Transparency,” *American Economic Review*, 95 (2005), 862–877.
Prendergast, Canice, and Lars Stole, “Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning,” *Journal of Political Economy*, 104 (1996), 1105–1134.
Romer, Christina D., and David H. Romer, “A New Measure of Monetary Shocks: Derivation and Implications,” *American Economic Review*, 94 (2004), 1055–1084.
Scharfstein, David S., and Jeremy C. Stein, “Herd Behavior and Investment,” *American Economic Review*, 80 (1990), 465–479.
Schonhardt-Bailey, Cheryl, *Deliberating American Monetary Policy: A Textual Analysis* (Cambridge, MA: MIT Press, 2013).
Taylor, J. B., “Diecretion versus Policy Rules in Practice,” *Carnegie-Rochester Conference Series on Public Policy*, 39 (1993), 195–214.
Thorsrud, Leif Anders, “Words Are the New Numbers: A Newsy Coincident Index of Business Cycles,” Technical Report 21/2016, Norges Bank Working Paper, 2016.
Tibshirani, Robert, “Regression Shrinkage and Selection via the Lasso,” *Journal of the Royal Statistical Society Series B*, 58 (1996), 267–288.
Toutanova, Kristina, Dan Klein, Christopher D. Manning, and Yoram Singer, “Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network,” in *Proceedings of HLT-NAACL 2003* (PA, USA: Association for Computational Linguistics 2003), 173–180.
Transparency International, “Improving the Accountability and Transparency of the European Central Bank,” Position paper, 2012.
Visser, Bauke, and Otto H. Swank, “On Committees of Experts,” *Quarterly Journal of Economics*, 122 (2007), 337–372.
Warsh, Kevin, “Transparency and the Bank of England’s Monetary Policy Committee,” Bank of England technical report, 2014.
Woolley, John T., and Joseph Gardner, “The Effect of “Sunshine” on Policy Deliberation: The Case of the Federal Open Market Committee,” *The Social Science Journal*, 54 (2017), 13–29.