U.S. Monetary Policy Since the 1950s and the Changing Content of FOMC Minutes

Pierre L. Siklos 5/12/2017

Not to be quoted or cited without permission

I am grateful for a CIGI-INET research grant for financial support, to the Hoover Institution where I was a National Fellow when research for this study was undertaken, and to Samantha St. Amand for excellent research assistance. Earlier drafts were presented at the Princeton University, Niehaus Center for Globalization and Governance workshop: "Monetary Policy and Central Banking: Historical analysis and contemporary approaches" and the Public Choice Society meetings. Comments on an earlier draft by Raymond Hicks and Ellen Meade are gratefully acknowledged.

ABSTRACT

Content analysis is used to analyze 60 years of FOMC minutes. Because of inherent problems with attempts to quantify language two different algorithms are used. Wordscores does not require the researcher to rely on a dictionary while DICTION is an alternative algorithm that is specifically designed to capture various elements that can be characterized as the sentiment or tone conveyed in a text. The resulting indicators are then examined empirically using vector autoregression techniques. The results broadly suggest substantive differences in the language of FOMC minutes across various Fed Chairs since the 1950s. A statistically significant link is found between the content of FOMC minutes and real GDP growth and changes in the fed funds rate. No significant links were found between the minutes and inflation. In addition, the relationship between content and macroeconomic conditions changes after 1993 when minutes are made public.

Pierre L. Siklos, Department of Economics, Wilfrid Laurier University, and Balsillie School of International Affairs, <u>psiklos@wlu.ca</u>

JEL Classification codes: E58, E52, E31, E37

Keywords: FOMC minutes, Wordscores, DICTION, monetary policy stance, vector autoregression

1. Introduction

Economists are generally inclined to evaluate monetary policy successes and failures via inferences drawn from data and model estimation. Generally, the data employed are observable macroeconomic and financial time series. Of course, there are risks in interpreting the motives and actions of policy makers based on this approach because these individuals make decisions in real time and it is not always possible for analysts to precisely recreate the data environment decision makers faced when they announce their monetary policy decisions. Moreover, monetary policy decisions are now communicated in the form of press releases, minutes of meetings, monetary policy reports, to give three examples. In principle, this allows policy makers to influence expectations without having to change the policy rate.

Economists have long known that what central bankers say and write can have implications for how the stance or strategy of monetary policy is interpreted (e.g., Bulíř et.al. 2014, Fracasso et.al. 2003, Schonhart-Bailey 2013, Holmes 2014). Central bank communication provides additional information about the views of policy makers. These publications offer a more detailed opportunity to understand how central bankers and their staff explain current and likely future economic developments. Recent events have served to heighten interest in what central bankers say and write. Whereas most of these developments are of a recent vintage the U.S. Federal Reserve has a long history of publishing information, often with a lag, that provides a window into the thinking of policy makers as well as insights about the conduct of monetary policy that is seen as complementing decisions about policy rates. There is a rich trove of documentary evidence concerning the internal discussions of the Fed's policy making committee, namely the federal Open Market Committee (FOMC).

This study analyzes the contents of FOMC deliberations since the early 1950s when, following the 1951 Fed-Treasury Accord, the Fed was freer to implement a monetary policy stance of its own choosing. The main objective of this study is to quantify the contents of FOMC minutes from 1952 to 2013 and to empirically examine its interaction with macroeconomic conditions. Generally, the existing literature that seeks to numerically code policy statements relies on far shorter samples.1

Needless to say, quantifying the contents of central bank minutes creates many challenges. Suffice it to say that since the personalities inside the FOMC have changed over time so has the content of the minutes. Controlling for these factors is far from straightforward. For example, as academics came to play a more important role in making monetary policy decisions in recent years so too have the expressions used to explain their actions. There have also been considerable improvements in our knowledge and ability to monitor economic developments over the decades. Both of these developments are likely to have an influence over the discourse of monetary policy.

The focus on FOMC minutes is motivated by several factors. First, given the long historical record of available FOMC minutes we can attempt to trace the evolution of what policy makers were thinking across a variety of policy regimes and momentous historical events over the past 60 years. This stands in contrast with the usual approach taken of, say, subjectively interpreting central bank communication by relying on a few chosen statements made by select FOMC members at specific moments in time. There is the risk that by 'cherry picking' statements we give undue emphasis to some passages over others than may have preceded them.

Second, the minutes offer a window into the thinking inside the FOMC as well as the extent to which its members' policy decisions are driven by data at their disposal, currently fashionable models, or other factors. In contrast, the language contained in other central bank publications (e.g., press releases or Inflation Reports) typically also incorporates staff or consensus views inside the FOMC. To be sure these publications are also interesting in their own right. However, we do not have a historical record that comes even close to matching the one contained in FOMC minutes.

Third, we are currently in an era when policy rates are hardly changing at all. As markets and the public wonder about when and whether policy ever resemble pre-crisis conditions this has

¹ I am not, of course, the first to mine the content of FOMC minutes (see below).

pushed central bankers to increasingly verbally explain their decisions. Even during the exit stage back to some new equilibrium fed funds rate level how future policy rate increases will be verbally communicated will be critical. In all of these circumstances, words rather than deeds play a critical role.

The present paper does not seek to determine the extent to which FOMC members understood, or applied, economic theories in vogue at the time the stance of monetary policy was determined or the level of sophistication of the debate. The paper also does not seek to determine whether the language used in FOMC statements over the years has become more or less readable (e.g., see Hernandez-Murillo and Shell 2014) although Fed clarity in its communication is a by-product of the exercise reported below. The algorithms used here are able to capture the content of the meetings along different dimensions and this should, in principle, be influenced by and, under certain circumstances, possibly influence the economic environment.

When choosing to study FOMC minutes over several decades the year 1993 is an important milestone since it became known that the Fed held transcripts of FOMC meetings previously not known to exist. The release of the transcripts, albeit with a five year lag, is thought to represent a 'structural' break of sorts in how FOMC members would henceforth conduct their deliberations (also, see Danker and Luecke 2005, and Hansen, McMahon and Prat 2014). Consequently, it has sometimes been alleged that since the transcripts would eventually be released this has served to stifle the discussion (e.g., Meade and Stasavage 2008). Not everyone, of course, shares this interpretation (e.g., Woolley and Gardner 2009). Therefore, a secondary hypothesis consists in asking whether the link between the content of FOMC minutes and macroeconomic outcomes changes after 1993.

I rely on two rather different methodologies quantify the content of the various texts considered. The texts that are analyzed are the minutes of the FOMC meetings, as they have been called since 1993, previously known as 'historical minutes' between 1952 and 1992. For convenience all such publications are henceforth referred to as minutes.

Next, a small standard econometric model is estimated with the aim of empirically investigating how central bank communication, in the form of the content of the minutes, may have influenced or responded to the stance of monetary policy. Both forward-looking and real time elements of monetary policy decisions are also accounted for where possible.

The rest of the paper is organized as follows. Section 2 provides a brief overview of the changing environment inside the FOMC since the 1950s as well as the challenges this poses in trying to gauge and quantify the evolution of the content of policy meetings. Section 3 outlines the two methodologies used in this paper to quantify the content of FOMC minutes. Some stylized facts are also provided as a prelude to the empirical evidence discussed in section 4. Section 5 concludes and summarizes the paper's findings and its potential implications.

Briefly, the paper finds that the impact of minutes does change after 1993. In addition, the most significant link found is between the proposed proxies for central bank communication and real GDP growth or changes in the stance of monetary policy, not inflation. The failure to find a statistically significant link between central bank communication and inflation, especially before 1993, may well reflect differences between policy recommendations made in real time versus ones obtained using ex post revised data that Orphanides (2001) highlighted. Alternatively, this result could also reflect changing views about the level of inflation consistent with price stability. A direct relationship between the content of the minutes and changes in the fed funds rate is also found. However, the size of the response of the fed funds rate to indictors of the content of the FOMC's deliberations is much smaller after 1993 than when the minutes were not made public.

2. The Evolving Role and Functions of FOMC Minutes

The FOMC, the principal policy making committee of the U.S. Federal Reserve, consists of 12 members. Seven are from the Board of Governors while the remaining five are Reserve Bank Presidents. The Reserve Bank Presidents serve one term on a rotating basis from the 12 reserve banks in the system sub-divided into four groups. Only the President of the New York Federal Reserve Bank has a permanent seat on the FOMC (see http://www.federalreserve.gov/monetarypolicy/fomc.htm). Members are appointed by the

President and confirmed by the Senate for 14 year terms. The Chair of the FOMC currently serves a four year renewable term. Members from the Reserve Banks also serve renewable terms (5 years) but they are appointed by the Board of Directors in their own districts. Full historical details concerning the FOMC's membership can be found at <u>http://www.federalreserve.gov/aboutthefed/bios/board/boardmembership.htm</u>.

Figure 1 provides a stylized timeline of the evolution of key changes in the release of written and verbal documentation by the FOMC that are most germane to the present study. The minutes of the FOMC meetings are currently released 3 weeks after a policy decision. This has been the policy since December 2004. It was in 1993 when the Board of Governors decided to release minutes of the FOMC meetings shortly after it emerged that transcripts, the most detailed account of FOMC meetings, had been recorded. Until 1955 the minutes included the meetings of the Executive Committee which discussed the implementation of the FOMC's decisions. From 1955, when the Executive Committee ceased to exist (see <u>http://www.federalreserve.gov/monetarypolicy/fomc historical.htm</u>), until 1967 the text was sub-divided into a part that minuted the FOMC's discussion ("Memorandum of Discussion") while a separate portion provided information about attendance and actions taken by the committee.

Prior to 1967, the Record of Policy Actions and the Minutes of Actions were the only documents describing FOMC proceedings released to the public. The Memorandum of Discussion was discontinued beginning in May 1976. The degree to which the contents of the minutes represent a discussion of the background and the reasoning behind FOMC decisions has changed markedly over time. Indeed, until the late 1960s, the FOMC normally met every 3 weeks. Presently, the FOMC meets 8 times a year.² Until 1994 a press release did not even accompany FOMC decisions. Consequently, minutes provide a (filtered) view of the

² This does not include extraordinary meetings held on occasion, usually because of a crisis or other special factors.

deliberations inside the FOMC (e.g., see Acosta and Meade 2015, Meade, Burk, and Josselyn 2015).₃

Interpreting the words of policy makers inside the FOMC over a 60 year period raises several challenges. Minutes of meetings can reflect views that have already been altered or formed either because of pre-meetings or due to various forms of peer pressure (e.g., see Meyer 2004, Visser and Swank 2007, Meade and Stasavage 2008). Therefore, reading the minutes of meetings may not be as illuminating as observing the free flow of discussion inside the meeting. Indeed, this is what prompts a separate line of research (e.g., see Hansen, McMahon, and Prat 2014, Meade and Stasavage 2008) that focuses on the verbatim discussion inside the FOMC in order to discern how much conformity exists in the meetings. The upshot is that added transparency can be a mixed bag. It can stifle debate but perhaps not enough to prevent minutes from be kept out from informing us about the views of policy makers.

In 1993 former FOMC Chair Alan Greenspan stated: "People think reading the raw transcripts is a way of learning things; I would suggest that if they spend six or eight months reading through some of this stuff, they won't like it." 4 Ignored, however, is the more telling comment found earlier in the same paragraph of Greenspan's comment: "I think we've always argued that the Memorandum of Discussion--leaving aside the issues, which are not irrelevant, of its cost and the demand for such documents--is as good a record of what actually occurs in these meetings

³ In 1997, the Fed began to provide a numeric value for the 'intended' fed funds rate, the policy rate set by the Federal Open Market Committee FOMC. By 1999 the Fed began to announce, at the time of the press release announcing its policy decision, a 'bias' to provide some guidance for markets and the public about the direction in which the stance of monetary policy was leaning. In 2000, the 'bias' in the policy statement was replaced with an expression of the 'balance of risks' (e.g., see Thornton (2011), Wynne (2013)). In 2003, expressions indicating 'accommodation [in monetary policy] for a considerable period' began to appear in the FOMC's press releases. This would be followed later, at the height of the financial crisis, by statements indicating that the fed funds rate would remain 'exceptionally low for an extended period of time'. At the end of 2012, the Fed explicitly began to indicate that it would respond to unemployment rate levels, among other economic indicators, before deciding when to exit from the current extraordinary macroeconomic environment. Soon these forms of forward guidance would be abandoned. Unlike policy statements which represent the consensus view of the entire FOMC the minutes provide a far broader indication of the variety of views inside the committee and how economic conditions might influence them as well as their potential impact on expectations.

⁴ The quote has been used by several academics (e.g., Schonhardt-Bailey 2013, Chappell, McGregor, and Vermilyea 2005, p. 57) and is from а Conference Call meeting held in 1993. See http://www.federalreserve.gov/monetarypolicy/files/FOMC19931022confcall.pdf, page 3.

as you can get from the point of view of those who have a serious interest in monetary policy and the history of monetary policy."⁵ More recently, former Chair Ben Bernanke has been quoted as saying about the FOMC meetings: "They gets lots of attention...and most of them are deadly boring," (Bartash 2014) until, of course, with the hindsight of history they are not, as in the case of the October 2008 FOMC meeting that precipitated rapid decreases in the fed funds rate to the zero lower bound where it sat for over 6 years.6

Another challenge is the changing usage and meaning of language over time or the interpretation of the Fed's mandate that is defined in several pieces of legislation since at least the Employment Act of 1946. The Fed itself states that "[T]hese objectives include economic growth in line with the economy's potential to expand; a high level of employment; stable prices (that is, stability in the purchasing power of the dollar); and moderate long-term interest rates." (http://www.federalreserve.gov/pf/pdf/pf complete.pdf, pg. 2). Needless to say what potential output is, let alone agreement on what constitutes prices stability, has escaped a precise definition. One reason, of course, is that potential output is unobservable. In the case of inflation the norms of what are deemed acceptable, if not tolerable, has also eluded a precise definition let alone complete agreement about which price level definition one ought to focus on. It was not until 2012 that the FOMC judged 2% in the Personal Consumption Expenditures (PCE) index consistent with the stability as price mandate (http://www.federalreserve.gov/faqs/money 12848.htm). Indeed, for many years, during the Volcker era and before, Fed speak centered around the concepts of "stable" or "reasonable" inflation rates.7 Prior to the 2012 announcement defining of price stability in numerical terms FOMC Chairs expressed subtle differences in their definition of price stability. Greenspan's definition, for example, was expressed at a 1996 FOMC meeting, as follows: "is that state in

⁵ Meyer (2004), in a memoir of his time at the FOMC, notes that Greenspan was much clearer, and to the point, when he spoke at Committee meetings.

⁶ As Warsh (2016), a former FOMC member, points out the "tone at the top", that is, views expressed by the FOMC chair, as well as the knowledge that the minutes are published, are two of several potential factors that may have impacted the content of the minutes after 1993. Warsh also draws attention to the more frequent discussions of the Board of Governors, as distinct from the FOMC, and likely spillovers of information, in addition to any informal discussions among smaller groups of committee members outside FOMC meetings. See Swank and Visser (2007). 7 According to Timberlake (1993, Chapter 25) the consensus in the early to mid-1980s was that 3-4% inflation was

[&]quot;reasonable".

which expected changes in the general price level do not effectively alter business or household conditions." Previously, Paul Volcker, Greenspan's predecessor as FOMC Chair, publicly stated the following definition of price stability: "...a situation in which expectations of generally rising (or falling) over a considerable period are not a pervasive influence on economic and financial behavior." The subtle difference in the precision with which price stability is defined by the two Chairs is unmistakable. It is worth noting it was during the Volcker era that the notion of zero inflation as being desirable was put to a vote although the proposal was rejected (Silber (2012), Meltzer (2009)).

The evolution and sophistication of economic analysis may also have influenced the language used inside the FOMC for at least a couple of reasons. First, the background and education of FOMC members has changed substantially over time. Whereas academics are today regularly appointed and/or confirmed as FOMC or Board of Governor members, this was less common in the past though not unheard of.⁹ Moreover, our understanding and acceptance of the role of theories of monetary policy has no doubt also had an influence on FOMC deliberations. The skepticism about the uses (and misuses) of economic models by FOMC Chairs from McChesney Martin through at least Paul Volcker, is well-known (e.g., see Meltzer (2009), and Bremner (2004)). How these attitudes influenced the level of sophistication of deliberations inside the FOMC is unclear. For example, Romer and Romer (2004) offer evidence that monetary policy in the 1950s was "actually sophisticated in its thinking" while Meltzer (2009) strongly disputes such a claim.

It is also plausible that the personal experience, not to mention the education and background of policy makers, may also be reflected in influencing the stance of policy. For example, Malmendier and Nagel (2015), and Malmendier, Nagel, and Yan (2017) argue that FOMC and

⁸ As quoted in Silber (2013) from a speech Volcker gave at the annual meetings of the American Economic and American Finance Associations in December 1983.

⁹ Adolph (2013) is an interesting study of the career aspirations and strategies of senior central bankers. According to his data "[T]he most outstanding feature is the waning and waxing of financial sector experience." (op.cit., p. 78) In particular, he notes that economists and financiers have become more prominent in central banks since the 1950s.

Bank Presidents' experience with inflation influences their inflation expectations but may also dictate whether they are likely to be dovish or hawkish.

Nevertheless, some views or questions for debate have not changed, even if the language used to describe them have evolved over time. For example, there is plenty of evidence that FOMC Chairs since the 1950s have strongly supported the view that inflation rates are associated with business cycle movements. Another firmly held view over time is that an appropriate inflation rate must be balanced against adequate economic growth lest the FOMC be accused of failing to meet its statutory objective. This, of course, has been expressed in a number of ways over time. For example, the Fed's so-called "even keel" strategy, in place until the late 1960s, later referred to by some as the 'stop-go' strategy of monetary policy (e.g., Goodfriend (2005), Hetzel (2008)) represents the 1950s and 1960s version of what today would be called 'flexible' monetary policy (e.g., see Mishkin (2007), Svensson (2009)). The difference over time, of course, is that flexibility used to imply a delay in acting until inflationary pressures, in particular, ran away from the policy authorities. Rules, such as Taylor's rule, were supposed to constrain the discretion of the monetary authority. Instead, one can argue that, in the aftermath if the Great Recession of 2008-9, data dependence represents a new form of discretion.

Today flexibility may demand patience but not necessarily at the expense of needing to act preemptively to prevent the un-anchoring of inflationary pressures (e.g., see Blinder (1998), White (2009)). Even so, few central bankers would object to McChesney Martin's definition of flexible monetary policy of the early 1960s as "leaning against the winds of deflation or inflation, whichever way they are blowing." (Bremner 2004).

Even the current challenge of central banks to incorporate a response to asset price developments in implementing monetary policy is not as novel as it might seem. Whereas the ZLB and 'unconventional monetary policy' are not expressions that can typically be found in the FOMC minutes prior to the latest global financial crisis the concern over asset price developments is, of course, not new (e.g., Bernanke and Gertler (1999), and Cecchetti, Genberg, Lipsky, and Wadhwani (2000)). For example, the FOMC minutes of a meeting held on July 7, 1954 refer to serious reservations expressed by then New York Fed President Alan Sproul about the state of the housing market which have an eerie resemblance to the debate inside the FOMC some fifty years later.

In a similar vein, while the word 'bubble' was not used, the state of the housing industry was described in terms that "...seem too good to be true" and would "..., spell some later difficulties if the funds were loosely lent..."¹⁰ More recently, the term 'bubble' became fashionable beginning in 2001 when the market for tech related equities collapsed subsequently leading to the twin worries about whether monetary policy should 'lean' or 'clean' (e.g., see White (2009)) and the possibility of a 'bad' deflation.¹¹ In both cases the potential connection between asset price movements and monetary policy was raised.

Finally, it is remarkable that for decades members of the FOMC have expressed the opinion that existing theories of inflation leave something to be desired. Even Chairman Bernanke pointed out, in 2008, during the celebration of the 50th year since the original Phillips curve paper was published, that "...much remains to be learned about both inflation forecasting and inflation control."¹² Indeed, one occasionally gets the impression, even if central bankers claim they have learnt the lessons of history,¹³ that policy makers can fall into the trap of believing that theirs heralds a new economic era, even though there were earlier echoes of the same phenomenon in the early 1960s. Then FOMC Chair McChesney Martin expressed the following view, in 1965, though tinged a hint of skepticism: "Then, as now, government officials, scholars, and businessmen are convinced that a new economic era has opened, an era in which business fluctuations have become a thing of the past." (McChesney Martin (1965)). The same sentiment

```
12 http://www.federalreserve.gov/newsevents/speech/bernanke20080609a.htm
```

¹⁰ Minutes of the executive committee of the FOMC, July 7, 1954, pages 3 and 4.

¹¹ Not surprisingly, this event prompted research into the advisability of central banks pre-empting asset price bubbles. See, for example, Bernanke and Gertler (1999), and Cecchetti, Gernberg, and Wadhwani (2002). On deflation, see Burdekin and Siklos (2004), and Borio and Filardo (2004).

¹³ Perhaps the most famous example of the lessons of history was Bernanke's apology to Friedman and SchwartzontheoccasionofMiltonFriedman's90thbirthday.Seehttp://www.federalreserve.gov/boarddocs/Speeches/2002/20021108/default.htm.

has been repeated about the period between 1984 and 2007, since known as the Great Moderation.

Other changes in the norms and usage of language likely stem from dramatic changes in our ability to diagnose economic problems, the strategies used to implement monetary policy, and changing views about transparency. Looking back it is unclear whether less judgment was used in setting the stance of monetary policy during, say, the 1950s than in the 2000s (Fischer 2017). Indeed, there has been a debate among current and former FOMC members about the extent to which the FOMC is data-driven versus reliant on judgment (e.g., Yellen 2006, p.3).

It is also the case that words that did not exist in some decades would become almost household words the next. It is inconceivable that policy makers would refer to the Great Moderation during the 1960s even if there is a resemblance with inflation and output growth rates achieved over two decades later. Similarly, expressions such as "irrational exuberance", "pre-emptive", "headwinds, tailwinds", or "adverse feedback loop" would not have been part of monetary policy vocabulary until the 2000s. In contrast, several expressions have fallen into disuse over time. For example, the principle that monetary policy ought to stimulate at times and accommodate at others (i.e., the stimulation and accommodation principles), language frequently used in McChesney Martin's era (e.g., see Timberlake (1993)), would not be encountered today.

Finally, the language of monetary policy has no doubt also been influenced by the strategy of monetary policy. Most obviously, the increased emphasis on inflation control, following the Great Inflation of the late 1960s through the mid-1980s (e.g., see Bordo and Orphanides (2013)), has pushed the Fed, if not most central banks around the world, to communicate the economic outlook more forcefully and transparently. This has forced central banks to tilt the debate to providing biases and forward guidance of various kinds that would not have been as central to the debate about the course monetary policy ought to take, say, during the 1950s or 1960s. As a result, monetary policy is conducted in a forward-looking manner, that is,

conditional on the central bank's economic outlook. Indeed, former Chair Ben Bernanke pointed out in 2006: "Future policy actions will be increasingly dependent on the evaluation of the economic outlook..." (Wessel (2009)). In earlier decades there is considerably less evidence that the FOMC developed a systematic way of thinking about the future (Meltzer (2009)). The foregoing is not meant to imply that monetary policy had no forward-looking elements at all prior to the availability, for example, of Greenbook forecasts.¹⁴ Nevertheless, the further one goes back in time, the more the economic outlook was based on less systematic analyses or a sophisticated understanding of what drove economic activity and inflation.

In spite of the significant changes in language or emphasis on the usage of some terms over others, the minutes of meetings convey a tone that ought to translate into FOMC members'¹⁵ views about current state of the economy, its outlook, the diversity of opinions inside the committee and, possibly, even the uncertainty about economic conditions over time. The challenges, if the maintained hypothesis of the paper is to be tested, is to quantify the content of FOMC meetings in a fashion that can be linked, statistically, to interest rate setting behavior and macroeconomic conditions more generally but is not beholden to a few words or expressions that do not stand the test of time.

One strand of research consists in quantifying in a number of ways not only what central banks say (e.g., in press conferences) and write (i.e., policy statements, reports of various kinds) but also how the media respond to central banks attempts to explain themselves. A few examples serve to illustrate this kind of research. For example, Bulíř et.al. (2014) are interested in the clarity of central bank communication which is assumed to be a function of the economic environment. The more uncertain the outlook the more difficult, it is claimed, it is to communicate clearly. If the hypothesis seems plausible it is far from obvious why overall economic uncertainty should make clarity more difficult. As the U.S. experience illustrates, the FOMC adjusts its communication when the underlying economic environment changes. In any

¹⁴ The role and accuracy of Greenbook forecasts has been studied extensively. See, for example, Romer and Romer (2000) and Messina, Sinclair and Steckler (2014).

^{15 &}quot;Members" in the minutes refer to the 12 voting members inside the FOMC. See Danker and Luecke (2005).

event, the authors conclude, based on data from several countries, that clarity does not suffer when the outlook becomes more uncertain. However, they also report that the level of understanding required to interpret central bank communications has increased since the onset of the global financial crisis.

Another series of papers attempt to interpret the textual material produced in several forms by central banks, code the results, usually in the form of a dummy variable that may or may not be ordered, to ask whether markets are moved by these attempts by central banks to communicate. Illustrations of this line of research include Ehrmann and Fratzscher (2007, 2009), and Berger, Ehrmann and Fratzscher (2011). Essentially, the aim is to ask how a perceived tightening or loosening of policies, as interpreted either by the central bank or certain media outlets, might affect, say, selected asset prices (e.g., stock returns, market interest rates). Rosa (2013), Bernanke, Reinhart, and sack (2004), and Romer and Romer (2004) are examples of studies that relyon a subjective interpretation of the content of written or verbal statements by central bankers.16

Tudor and Vega (2014) review a small sample of papers that create quantitative indicators based on qualitative data and report that, while extremely useful, there remain several problems that need addressing. First, the appropriateness of the dictionaries used to create content type indices is an important question since Loughran and McDonald (2011) demonstrate that some words that may have a negative connotation in one context may be neutral in others and that expressions in finance are not the same as in every day discussions. Moreover, whereas some words that might evoke positive or negative sentiment is a general context are included in standard dictionaries more recently used terms that have proved important might well be left out as they are uncommon. Arguably, however, a crisis can bring out the usage of new terminology to describe a particular stance or situation.

The foregoing only scratches the surface of techniques that have been deployed to interpret the content of documents published by central banks. Some resort to case studies as these can

¹⁶ Perhaps 'informed' might be a better term than subjective to describe how these authors interpret central bank texts although it is the latter term that has stuck in the literature.

illuminate how specific events put pressure on the form and content of communication (e.g., Meade et. al. 2015, Chappell, McGregor, and Vermilyea 2005, Havrilesky 1993). More recently, a growing number of algorithms have been applied that are able to quantify the content of press releases and other documents. These algorithms from simple word counts to more sophisticated programs that focus on changes in wording over time or other devices able to evaluate the semantic content of publications. Indeed, there exist well over a dozen algorithms of various kinds all of them claiming to most reliably quantify the content of textual material. Luca and Trebbi (2009) and Hansen and McMahon (2015) are two other examples of studies of FOMC policy statements, not the minutes of FOMC meetings, that rely on automated approaches to identify policy leanings of the Fed. In every case the samples investigated are short (e.g., since the late 1990s or later).

This paper continues in the same vein. In common with all these techniques they have their advantages and disadvantages.¹⁷ However, unlike other comparable applications, I rely on two quite different algorithms as opposed to the single measurement approach common in other such studies. Hence, this paper offers a test of robustness in evaluating the content of FOMC minutes and over a much longer period than has heretofore been considered.

3. Quantifying the Content of FOMC Minutes: Two Strategies

FOMC minutes are available from the Board of Governors. Minutes for the 2009-13 period are available from http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm. Historical minutes are available from http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm. Historical minutes are available from http://www.federalreserve.gov/monetarypolicy/fomccalendars.htm. Historical minutes are available from http://www.federalreserve.gov/monetarypolicy/fomc_historical.htm. The complete text is used when constructing the indicators of tone or content of the minutes.

Wordscores is used for at least two reasons.¹⁸ First, it offers a means of estimating relative policy positions expressed in texts. Second, the approach is flexible enough to avoid the possibility that we are exclusively interpreting monetary policy of the 1950s through the words used, say, several decades later. Unlike traditional content methods, for example, which

¹⁷ Other candidates include Atlas, Alceste, General Inquirer, Leximancer, and even the narrative approach used in Romer and Romer (2004), and Holmes (2013). This list is far from exhaustive.

¹⁸ The software, one version of which is written for Stata, is available from <u>www.tcd.ie/Political</u> <u>Science/wordscores/</u>.

automate the hand coding of text or count the number of words, Wordscores compares texts where the policy position is known or is believed, a priori, to be known (e.g., tighter or looser monetary policy; hawkish or dovish attitudes about inflation).¹⁹ The latter is labelled the 'reference' text while the text to which it is being compared is called the 'virgin' text (Laver, Benoit, and Garry (2003)).

Every word in each text is recorded and the words are then grouped by frequency. The researcher can elect to omit certain words (e.g., the, or, that, and so on). Since words are used with different frequencies a wordscore is then a weighted average of the relative frequency with which different words are used.

Of course, it is not always obvious what the overall position of the FOMC is. That is, what may appear to some to be a tightening bias expressed by the committee may not be interpreted as such by others. Accordingly, I examine the sensitivity of the results by modifying assumptions made about overall monetary policy positions. Wordscores has the distinct advantage that one can use historically comparable texts to estimate policy positions.²⁰ Indeed, by changing the reference text we may also obtain a glimpse of how policy positions may have changed. To repeat, a policy position deemed hawkish in one era may not, in retrospect, appear so in a different one.²¹ By comparing word frequencies between reference and virgin texts Wordscores is effectively equivalent to a Bayesian reading of texts. In other words, we ask: given word frequencies in a reference text what is the likelihood that a virgin text expresses the same position (i.e., produces the same distribution of word frequencies)? Each wordscore is converted such that we can rank positions on a scale (e.g., hawkish to dovish).

As with all Bayesian forms of analysis there are at least a couple of drawbacks with the approach (e.g., see Clark, Evans and Scarborough (2007)). First, the analyst must choose a reference text. This problem is mitigated in the present context by selecting a large number of potential reference texts. This serves as a test of robustness. Second, one must take a stand

¹⁹ Of course, this is thanks in part to the benefit of hindsight. See, however, below.

²⁰ As pointed out in Danker and Luecke (2005, p. 178) "[T]he minutes follow a structure that is fairly consistent from one meeting to the next."

²¹ An excellent illustration is the case study of the 1994 and 2004 tightening of the monetary policy stance explored by Meade et. al. (2015).

concerning the bias or policy position of the reference text. The potential problem here can also be minimized by relying on (ex post) historical accounts to establish whether a consensus has developed around the positions taken by policy makers at the time.

Once the anchor or reference texts are selected wordscores calculations proceed as follows. Define a_{rt} as the position taken in reference text r at time t. Next, let f_{rt}^{w} represent the relative frequency of some word, w, in the reference text, again at time t. We care the f_{rt}^{w} estimate a probability that we are reading words w, from text r, which is written

Benoit and Garry (2003) define

$$WS_{rt} = \sum_{w} p_{rt}^{w} a_{rt}$$
(1)

where WS represents the weighted average of the reference text scores over all words *w*. Using a similar argument we can evaluate (1) based on the virgin text yielding WS_{vt}^{w} . The latter expression represents the *a priori* interpretation of the content of the text we are interested in. Laven, Benoit and Garry (2003) propose simple test statistics for whether or not the virgin and reference texts are statistically the same. However, to ensure comparability of the scores, it is necessary to transform the Wordscores of the virgin text so that both virgin and reference texts have the same standard deviations.²² In particular, they evaluate

$$WS_{vt}^{*} = (WS_{vt} - WS_{\overline{v}t})\frac{\sigma_{rt}}{\sigma_{vt}} + WS_{\overline{v}t}$$
(2)

where $WS_{\overline{vt}}$ is the mean wordscore from the set of virgin texts. Another advantage of the wordscores approach is that the analyst can estimate confidence intervals. The 'uncertainty' of $\sigma_{vt}^2 = \sum_w f_{vt}^w (WS_{wt} - WS_{vt})^2$ since the smaller the dispersion in

²² Since the universe of words is finite Laven, Benoit and Garry (2003) observe that virgin test scores are more clustered than reference test scores. Not everyone agrees that the resulting transformation is ideal. See, for example, Martin and Vanberg (2008).

wordscores over all possible words relative to the wordscores in the virgin text the less the 'uncertainty' around the estimated score.23

In quantifying the FOMC minutes I take the view that, other things equal, a rise in the fed funds rate (FFR) represents a tightening of, or more hawkish, monetary policy. It seems reasonable to assume that all FOMC meetings have in common the notion that a rise in the fed funds rate is seen as a signal of a more hawkish stance against inflation, an overheating economy, or both. The larger the change in the FFR the tighter is monetary policy (or looser, of course, in the event of a reduction in the FFR).24 Therefore, a score of +1 was assigned to a rise of 25bp, +2 in the event of a 50bp rise, and so on, with scores of -1, -2, and so forth assigned for the case of looser policy. In this manner we can identify the position (i.e., tighter or looser) of the reference texts using actual policy rate changes as the 'anchor'. I use the effective fed funds rate from FRED (Federal Reserve Bank of St. Louis). Prior to availability of this series the figure mentioned in the "Record of Policy Actions" was used. FOMC minutes were combined on an annual basis from 1955 to 2006, inclusive, and used as candidates for the reference text. There are no reference texts in 1993 because there were no changes in the FFR. The same is true of the minutes after 2008. Hence, the content of minutes of these meetings was not considered informative enough.25 Furthermore, while the minutes of meetings occur at irregular intervals, especially in the 1950s and 1960s, the 'data' are converted to the monthly frequency by simple arithmetic averaging so that the results are amenable to time series analysis which, as we shall see, relies on the quarterly sampling frequency. Once the data are compiled from the various text we can then construct the scores based on equations (1) and (2).

²³ This is equivalent to calculating the population variance of every word in a text.

²⁴ Whether an increase in the policy rate translates into a demonstrably tighter monetary policy ultimately depends, of course, on whether the real interest rate changes. Meltzer (2009, p. 587) has argued that the failure to understand the distinction between nominal and real interest rates was a "long-standing" problem at the Fed. Indeed, the failure to respond aggressively enough to inflation developments is viewed by some (e.g., Taylor 2010) as one of the failures of the Fed in the years that led up to the Great Recession of 2008-9 (see, however, Bernanke 2010).

²⁵ In deriving the values for WS_{rt} the reference text (r) changes annually from 1955 to 1993. Thereafter, given the relatively smaller number of fed funds rate changes averages of scores over a two year period (e.g., 1993-94, 1995-96, and so on) are used until 2006. The assumed position of the FOMC could also be based on an assessment, objective or subjective, of the Fed's position in the policy statement. Alternatively, after 2008, one could instead rely on estimates of the shadow fed funds rate as the 'anchor'. These options are left as extensions for future research.

The second algorithm used is DICTION (see Hart, Childers, and Lind (2013)).₂₆ The objective of this algorithm is to transform a collection of words that quantify the *tone* of the document. Tone is "..., a tool people use (sometimes unwittingly) to create distinct spcial impressions via word choice." (op. cit., p. 9) Meade, Burk, and Joesselyn (2015) argue that US policy makers view the minutes "...as providing insight about the breadth of views..." Furthermore, since meeting participants provide input before the final draft is published (Danker and Luecke 2005) the intention and, therefore, the tone of the document ought to reflect the views of FOMC members.

Content is interpreted on the basis of a dictionary of words that convey meaning along various dimensions. A total of 31 pre-defined dictionaries contains over 10,000 words. However, these were supplemented with words based on the set of financial dictionary used by Loughran and McDonald (2011). The actual words used in the minutes are then grouped into several master categories according to the message being conveyed. They are: *certainty*, that is, a collection of words indicating resoluteness and authority; *optimism*, namely language that endorses a position and is occasionally interpreted as indicting over-confidence or hubris; *activity*, are words suggest ideas or stances being implemented with inertia avoided; and, finally, *commonality*, that is, language that conveys the degree to which a common position is taken.²⁷

Unlike Wordscores, DICTION does not rely on a benchmark. Hence, the dictionary is based on language that is more commonly used by policy makers in recent years than, say, during the 1950s or the 1960s. I did experiment with alternative dictionaries that also includes words likely to be time-invariant (e.g., recession, expansion) but whose meaning can change depending on the context and the conclusions are unaffected. While it may seem desirable to augment existing dictionaries with words from a different era there is the danger of introducing

²⁶ Version 6.0 was used in the empirical work. See <u>www.dictionsoftware.com</u>.

²⁷ The measurement of tone is not limited to this set of words and their associated dictionaries. See, for example, Siklos (2014) for an illustration using DICTION 5.0 and Lombardi, Siklos, and St. Amand (2017) using DICTION 6.0. The period covered by these articles is the global financial crisis and a different set of words dealing with financial stability questions.

selectivity bias into the results. For this reason, I did not pursue expansion of the default dictionary along these lines.28

While statistics of the kind developed for Wordscores are not available in DICTION users have proposed some of their own relying on the fact that, for example, over time a change in the tone of a message can be identified by changes in the frequency or in the extent to which the balance or tilt of a text favours one of the content variables over another. Hence, since the content of statements have been coded over a long period of time one might view standardization relative to some mean level, perhaps under a particular FOMC Chair, as a relative indicator. As a result, departures one or more standard deviations from the mean give rise to a "normality" score. Similarly, standardizing individual scores for the components of the content of a statement and then summing these provides an indication of the general properties of a document across time.

4. Empirical Evidence

a. Stylized Facts and Quantifying FOMC Minutes

To investigate empirically the links between the content of minutes and macroeconomic outcomes I obtained as many time series as were available on a continuous basis, at least at the quarterly sampling frequency, since the mid-1950s from FRED (http://research.stlouisfed.org/fred2/). The series range from real and nominal GDP, CPI, PCE, commodity prices, money supply and banking reserves indicators, as well as interest rates and interest rate spreads. Next, in recognition of the important role played by real time data (e.g., Orphanides 2001), measures of the output gap are obtained from the Federal Reserve Bank of Philadelphia (http://www.philadelphiafed.org/research-and-data/real-time-center/real-time-

²⁸ The language used in the minutes is likely to contain a combination of financial and everyday language. Loughran and McDonald (2016) point out that the choice of words included in the dictionaries matters. Hence, it is possible that words with ambiguous meaning (e.g., risk, unemployment) could have an impact on the results. This is more likely to be the case, however, for short samples. Over longer samples such ambiguity tends to reflect the importance, in this case by members of the FOMC, they attach to these phenomena. Some experimentation along these lines (also, see Lombardi, Siklos, and St. Amand 2017) supports this interpretation.

<u>data/</u>). Finally, as the FOMC began to increasingly rely on forward looking indicators of economic activity I also consider a role for the Fed's Greenbook forecasts.²⁹

Next, I turn to some evidence based on content analysis. Figure 2 plots the fed funds rate and WS, the variable that Wordscores defines the position of the FOMC in the reference texts used to score rolling virgin texts over time, as previously explained. The taller are the bars the greater the bias in favor of a hawkish (positive value) or a dovish (negative value) interpretation of the stance of monetary policy. One might expect, the scores shown in Figure 2 should be in the same direction as the change in the fed funds rate. However, because the (Word)score relies on textual analysis the amplitude of the two series need not, of course, be the same. Hence, as in some of the narrative examples described earlier, the FOMC may be more or less aggressive in changing the policy rate than the discussion implied by the minutes. Moreover, actual policy rate changes represent a consensus of sorts, subject to dissent.₃₀

Figure 3 provides two sets of Wordscores again based on equation (2). The top figure considers the full 60 years of FOMC minutes analyzed. The bottom figure considers the Greenspan and Bernanke eras separately. The TS in the figures refer to the transformed score discussed earlier while the numbers indicate the year chosen for the reference text. For example, TS68 represents the Wordscore across time based on a reference text consisting of the minutes during the calendar year 1968. Strictly speaking, each series ought to be separately examined with a higher score for each case shown implying a more hawkish stance and vice-versa for a lower score.

²⁹ These are also available from the same source as the real time data. The quarterly data begin in 1965. A separate literature considers dissent inside the FOMC as an indicator of the changing positions of the committee over time (e.g., see Thornton and Wheelock (2014)). Dissent data were later dropped from estimation as the results described below were unaffected. The evidence on the impact of dissents is decidedly mixed. As Bernanke (2015, p. 239) points out:"...FOMC tradition called for consensus decision making, and in that context a "no" vote represents a strong statement of disagreement." Hence, even if dissents likely do matter it is unclear what the threshold is beyond which votes against the Chair are meaningful or threaten the credibility of the FOMC. A separate appendix provides some estimates of the simple correlations between dissenting votes and the various content indicators derived in this study.

³⁰ Simple cross-correlations (not shown) suggest that past Fed tightening leads to more dovish content up to 8 quarters in the future. Current tightening is correlated with more hawkish content 2 quarters later but this is reversed 5 quarters later.

The Wordscores indicate that policy deliberations were becoming more dovish during the Greenspan era based on the FOMC's position in 1987 which is the year of the stock market crash. Similarly, when the 1980 FOMC minutes are used as the reference text the content of subsequent minutes suggests a gradual reduction in hawkishness. This is the year when U.S. inflation reached its peak. 1980 marks the year when inflation in the CPI peaked as did the fed funds rate. This overall reduction is also apparent when 1968, or 1987 serve as the reference points for evaluating the stance of monetary policy. By this metric more hawkish content is clearly observed in the Martin era though it is far more gradual in nature if Volcker era minutes serve as the reference point than when 1968, namely toward the end of Martin's tenure as FOMC Chair and a politically turbulent year, is employed. When Greenspan minutes from 1987 are used the more hawkish minutes are only observed at the end of Martin's tenure. It is worth emphasizing that the content of the minutes express what FOMC members were thinking relative to some chosen benchmark.³¹ As noted previously, there need not be a one to one mapping between the views expressed by committee by committee members and decisions actually taken by the whole FOMC.

The bottom portion of the Figure considers the content of the minutes when the reference text changes from 2001 to 2006. Wordscores interprets monetary policy as becoming gradually more dovish during Bernanke's tenure if 2005 minutes are the benchmark while using 2001 as the reference point suggests views becoming slowly more hawkish. The year 2005 is when the fed funds rate peaked while 2001 is the year of the dot.com bubble. Statistically speaking, policy can be considered equally neutral in 2000, and equally loose in 2002 and 2003, and progressively tighter in the 2004 to 2006 period.³²

In a sense the results suggest that whether FOMC members, as a group, are more dovish or hawkish can depend on the benchmark one applies to gauge current views about monetary policy. Whether this reflects a 'this time is different' mentality is unclear. Nevertheless, the overall trends do suggest that the FOMC's thinking often changes gradually over time.

³¹ The rankings can change according to the chosen reference dates but are largely insensitive to such choices (results not shown).

³² Confidence intervals are not shown to avoid cluttering the figures.

I now turn to a discussion of the content of FOMC minutes as measured using the DICTION algorithm. These are shown in Figure 4. Given the manner in which the DICTION software estimates the relevant indicators the raw scores can be very volatile. While several measures can be applied to standardize the scores, as discussed previously, Figure 4 displays both the raw data as well as estimates obtained via exponential smoothing.³³ On this basis the degree to which the FOMC minutes convey certainty about members' views concerning the economic environment has varied considerably over time. Nevertheless, it is interesting to observe that certainty reaches its lowest ebb soon after Paul Volcker becomes Chairman. It also takes many years for certainty to reach levels previously seen in the 1960s. The certainty indicator is remarkably stable during the so-called Great Moderation period (approx. 1984-2007) but shows no signs of a large drop during the financial crisis. It is tempting to think that a long period of certainty in the language of the FOMC minutes is conducive to the complacency that produced conditions favorable to a financial crisis just as an earlier period of stable levels of certainty helped usher in the great inflation of the 1970s. Unfortunately, this interpretation can be questioned as a form of ex post rationalization.

The extent to which the FOMC minutes reveal the optimism of members suggests a sharp rise in the early 1960s, again until the great inflation of the second half of the 1960s when optimism declines and remains low until the early 1980s when, following a further decline during the two recessions in the early 1980s, optimism in the minutes experiences a steady rise though nowhere reaching the levels of the early 1960s. Thereafter optimism remains fairly steady until the tech bubble of the early 2000s when a sharp fall is recorded. Optimism never quite recovers and experiences a further negative shock in the fallout from the financial crisis of 2007-2009.

If activism in the language of FOMC minutes expresses a desire for discretion in policy making among FOMC members then the mid-1950s reveal a sharp rise, perhaps as a partial response to the autonomy granted under the Fed-Treasury Accord of 1951. This state of affairs continues throughout the 1960s, curiously until the great inflation period. As the collection of papers in Bordo and Orphanides (2013) suggests, the FOMC took the view that inflation was, to some

³³ Other smoothing functions were considered but produced comparable results.

extent, inevitable, and this may be reflected in less activist language until the early 1970s when, in the aftermath of the two oil price shocks, there is once again a steady decline in activist language. Activism, at least as measured from the content of the FOMC minutes, remains steady throughout the 1990s and suffers another blow, again at the same time as the tech bubble and the first recession of the 2000s. Activism then steadily declines in parallel with the decline of the fed funds rate to the ZLB. This raises the question why the minutes of FOMC minutes during and after the crisis do not appear to capture the activism inherent in the shift to unconventional monetary policy.

Finally, a plausible way to interpret the commonality variable is an alternative indicator of dissent inside the FOMC. There is little dissent in the 1950s and the first half of the 1960s, that is, during the Martin era, a sharp rise in Burns' FOMC. A steady decline in dissent takes place during Volcker's Chairmanship although it is considerably higher than in Martin's era. Only shortly after Greenspan takes over is the level of differences inside the committee, as reflected in the language of the minutes, lower than levels seen since the 1950s. Notable also is the sharp drop at the end of the sample in 2012 and 2013 when QE policies are in full swing.³⁴

A priori it is unclear whether the individual indicators of the hawkish or dovish tone of FOMC minutes ought to be considered separately or, since statistical evidence reveals the characteristics to be related (results not shown), these should somehow be combined. As a consequence, a factor model is estimated based on the 4 chosen characteristics from DICTION. A rise in the factor model generated indicator signals a more hawkish tone and vice-versa.

³⁴Some correlations and regression results were obtained to determine whether potential determinants of changes in the language of FOMC minutes include recessions, the tenure of individual FOMC Chairs, and Greenbook forecast errors (not shown). Moreover, I was also interested in whether the various characteristics of content are correlated with each other and with observed levels of agreement or dissent (Thornton and Wheelock (2014)). For example, dissenting votes by Governors are negatively and significantly correlated with 'certainty' (-0.21) and mildly, but significantly, positively correlated with 'optimism' (0.09). The results are comparable if we include dissenting votes by Fed bank presidents. There is, however, no significant correlation between dissents by Governors and 'commonality' (0.04) while the same variable is positively, and significantly, correlated with dissents by Fed bank presidents (0.17). Broadly speaking, the correlation matrices confirm that there are substantive differences in the language of the FOMC minutes across various fed Chairs. Also evident is the relatively higher level of conflict and dissent during the Burns era. Other results suggest, on average, that FOMC language is less optimistic during recessions and there is less commonality as well. There is also, again on average, consistently more optimistic language in the Greenspan and Bernanke led FOMC. Forecast errors in real GDP growth and inflation produce more activist language while larger inflation forecast errors results in less commonality.

Figure 5 repeats the earlier exercise conducted using Wordscores (see Figure 6). There is no apparent correlation between the content of the minutes and fed funds rate changes.³⁵ However, there is a much more statistically significant cross-correlation relationship between these same variables. Past tightening of monetary policy is highly correlated with a more dovish tone in current deliberations in the minutes which suggests that past rises in the policy rate lead to subsequent reversal in the hawkish tone of the minutes. In contrast a current monetary policy tightening leads to persistently more hawkish tone in the minutes (results not shown). This is consistent with the notion that Fed tightening is gradual.

b. Econometric Evidence

A series of small scale VARs are estimated for a variety of samples motivated by the timeline given in Figure 1. The object of the exercise is to ascertain whether traditionally estimated links between changes in monetary policy, via the fed funds rate, and macroeconomic conditions are influenced by the addition of the content of FOMC minutes. Moreover, while the content of FOMC minutes post 1994 can conceivably influence economic conditions via expectations the same cannot be true for the period when only 'historical' minutes are available. In that case, it is still of interest to ask whether there is a statistical relationship between the content of minutes and the stance of monetary policy.

For the pre-1994 sample I assume that while economic conditions may influence the content of the minutes the latter cannot influence macroeconomic conditions since there is no mechanism for the minutes to influence expectations due to their release with a long lag if available at all. Consequently, although the content of minutes is treated as an endogenous variable it is ordered below the usual macroeconomic time series in a VAR subject to a Choleski decomposition. Estimates that rely on Wordscores are conventional VARs while the results that rely on DICTION are in effect factor-augmented or FAVARs because the various characteristics

³⁵ The factor loadings were as follows: -0.26 (activity), -0.53 (certainty), 0.73 (optimism), 0.88 (commonality). Therefore, more commonality and optimism loads positively indicting a more hawkish stand while more activity and certainty load into a more dovish tone.

that make-up the tone of FOMC minutes are combined to create a single factor that describes the content of the minutes (see Figure 5). 36

The pre-1994 VARs and FAVARs are based on the following vector of endogenous variables

PRE 1994:
$$[\dot{y}_t, \Delta \pi_t^{PCE}, \Delta i_t^{FFR}, CM_t]$$
 (3)

 \dot{y} , $\Delta \pi^{PCE}$, Δi^{FFR} , CM are, respectively, real GDP growth, the change in inflation in the Personal Consumption Expenditures index (both annualized)³⁷, the change in the fed funds rate, and the proxy for the content of central bank minutes.³⁸ As discussed above, the factor model proxy from DICTION and WS from Wordscores are used.

The basic VAR and FAVARs are also augmented by a vector of exogenous variables including oil price inflation lagged one period, accumulated revisions to the output gap so lagged one period³⁹, as well as dummy variables for the McChesney Martin and Burns chairmanships.

The VARs and FAVARs for the post 1994 sample are similar to equation (3). However, since FOMC minutes are now made public these are now not only influenced by the existing state of the world but, in turn, may impact macroeconomic variables directly or indirectly via expectations.

POST 1994:
$$[CM_t, \dot{y}_t, \pi_t^{PCE}, \Delta i_t^{FFR}]$$
 (4)

Economic theory is insufficiently clear about where the CM variable stands in the Choleski decomposition. However, the earlier discussion would seem to suggest that all other shocks in

³⁶ The VARs were also estimated by separately including each one of the characteristics that define the tone of a document. The results of these VARs are relegated to an appendix.

³⁷ One potential concern is the stationarity of some of the series. This is what prompted me to use the change in the fed funds rate as opposed to its level and the change in inflation in the pre-1994 period. The impulse responses shown below are only marginally impacted when $\Delta \pi^{PCE}$ is used instead of its level (not shown).

³⁸As discussed in Danker and Luecke (2005) since 2004 the minutes are released after 3 weeks. The lags were longer previously but within the quarterly sampling frequency.

³⁹Using the accumulated revisions to real GDP growth did not change the conclusions and the bank reserves to GDP is also excluded as he results were not improved (not shown). It made no difference to the conclusions when real oil prices are used. Hence, estimates rely on the rate of change in nominal oil prices. The same is true when changes in unemployment or an estimate of the unemployment gap is used instead of real GDP growth (or changes in the output gap).

the model are likely to impact the deliberations and, hence, either the DICTION or the Wordscores indicators of the content of the minutes. In other words, the content of the of the minutes are assumed not to respond contemporaneously to the real and monetary policy variables in the model.⁴⁰

The same exogenous variables as in the pre-1994 sample are added including a Greenspan dummy serves as the control for the period covered by his tenure as FOMC chair, oil price inflation, the Greenbook forecasts for inflation and real GDP growth, and the accumulated revisions in the output gap. Finally, lag lengths were chosen on the basis of AIC (Akaike Information Criterion) or FPE (Final Prediction Error) metrics. When these tests produced conflicting results VARs and FAVARs with different lag lengths were estimated. The results presented below are robust to changes in lag lengths.

Figures 6 and 7 display selected impulse response functions. Complete details of other estimates are relegated to the appendix. Figure 6A displays the results using Wordscores for a sample that ends just as the Great Recession begins. A positive real GDP shock raises the score which, under the circumstances, implies a shift towards tightening relative to the reference FOMC minutes. The impact is more modest than the one found relying on the DICTION indicator (see below) despite the fact that the two measures are quite different from each other. The middle impulse responses also indicate that a tightening via a shock coming from the fed funds rate also changes the content of FOMC minutes in the same direction that is, in a more hawkish direction. Finally, in the pre-1994 sample, a relative shift towards tightening in virgin documents mirrors positive shocks from the fed funds rate but the effect is noticeably larger and persists over 5 quarters unlike the DICTION counterpart (Figure 7) which vanishes after 1 quarter as we shall see. Figure 6B displays what happens when the post-1994 sample is extended to the end of 2012. The response of WS to changes in the fed funds rate now

⁴⁰ This could be interpreted as the assumption of data dependence or caution in responding to the most recently available data. As Bernanke (2015, p. 503) has also pointed out: "...economists are criticized for not being able to predict the future but, because the data are incomplete and subject to revision, we cannot even be sure what happened in the recent past." In (3) and (4) I also experimented with placing CM before Δi^{FFR} which effectively means that changes in the fed funds rate respond to changes in the content of the minutes. None of these variations materially affected the results discussed below.

resemble that of the pre-1994 sample although the size of the responses is roughly half of the ones obtained for the more recent sample. There is now a very modest but significant positive response of WS to a positive real GDP shock.

Figure 7A reveals that a shock to real GDP growth raises the DICTION indicator through 5 quarters indicating that a more hawkish tone prevails inside the FOMC suggests in response to a growing economy. This result is not too dissimilar to the one reported using Wordscores. None of the other shocks were found to have a significant impact on FOMC minutes.⁴¹ Before the release of the minutes (bottom plot) the only modest response inside the FOMC is to a rise in the fed funds rate but only in the first quarter. This suggests that when the Fed tightened the policy stance (i.e., a positive fed funds rate shock) this was reinforced by a comparable change in the tone of FOMC minutes. Figure 7B then extends the results to the end of 2012. The content of FOMC minutes no longer responds to real GDP growth, that is, one cannot detect either a more hawkish or dovish tone. The link between the content of FOMC minutes and changes in the fed funds rate is only very modestly stronger than in the pre-1994 sample but the difference is hardly convincing.⁴²

Two broad conclusions emerge. First, in spite of substantive differences in how the content of FOMC minutes are quantified, the results are broadly comparable for each sample. Second, ignoring a role for the content of FOMC minutes is tantamount to omitting a significant variable. This result holds both when FOMC minutes are public as well as when they were not available to outside observers. Perhaps the biggest surprise is the lack of responsiveness of the central bank communication variables to inflation. Changes in both the definitions of inflation that the Fed was concerned about and, until 2012, the absence of an explicit inflation objective likely contribute to this finding.

⁴¹ A fairly general result is that there is persistence in the DICTION shocks indicating that a tightening tone this quarter is likely to tilt the tone toward more tightening in future quarters. However, the persistence dies down very quickly, sometimes after only 2 quarters. Moreover, persistence is lower post-1994 than when the minutes were not publicly released (results not shown).

⁴² There were no noticeable changes in the impulse responses when Wu and Xia's (2016) shdow fed funds rate data replace the observed policy rate from 2009 to 2012. Given that this represents only 4 years in a sample of almost 20 years of data may explain this result.

5. Conclusions

What central bankers say and think, as reflected in the minutes of their meetings, provides an additional piece of information that can supplement the usual approach of relying solely on movements in policy rates to assess changes in the stance of monetary policy. However, in estimating macro models, the challenge is to quantify words and their meaning in a manner that faithfully captures the content of policy making committee documents. The usual approach is to adopt a single methodology very often based on an interpretation or a particular coding of a document. In this paper two very different methodologies are used not only as a test of robustness but also to provide different perspectives about the influence of the content of central bank communication.

The minutes of FOMC meetings since the early 1950s are quantified and used as an additional determinant in small scale macro models. Because FOMC minutes were not released to the public before 1993 separate estimates are generated. Both VARs and FAVARs are estimated and these reveal that proxies for the content of FOMC minutes are significantly influenced by the overall state of the economy. The strongest links are found between the proxies for the tone and content of FOMC meetings and either real GDP growth or changes in the fed funds rate. No statistically significant link was found between the content of central bank minutes and inflation.

At least three extensions of the present study ought to be considered. First, I have not considered the possibility that the proxies for the minutes are asymmetrically related to the state of the macro economy. What central bankers think is the appropriate stance of policy in upturns versus recessions may well be different. Second, it is possible that additional endogenous variables could be included in the VARs and FAVARs to better condition the estimates on structural breaks or other regime shift like events. Finally, it might be possible to investigate the potential links between the content of FOMC minutes and policy uncertainty or uncertainty more generally. I leave these extensions for future research.

References

Acosta, M., and E. Meade (2015), "Hanging on Every Word: Semantic Analysis of the FOMC's Postmeeting Statement", FEDS Notes, 30 September.

Adolph, C. (2013), *Bankers, Bureaucrats, and Central Bank Politics* (Cambridge, Mass.: Cambridge University Press).

Axilrod, S. H. (2009), Inside the Fed (Cambridge, Mass.: The MIT Press).

Bartash,, J. (2014), "Bernanke Reveals What Really Happens at Fed Meetings", *MarketWatch* 30 December.

Berger, H., M. Ehrmann, and M. Fratzscher (2011), "Monetary Policy in the Media", *Journal of Money, Credit and Banking* 43 (4): 689-709.

Bernanke, B.S. (2015), *The Courage to Act: A Memoir of a Crisis and Its Aftermath* (New York: W.W. Norton).

Bernanke, B. (2013), *The Federal Reserve and the Financial Crisis* (Princeton, N.J.: Princeton University Press).

Bernanke, B. (2010), "Monetary Policy and the Housing Bubble", speech at the Annual meetng of the American Economic Association, Atlanta, January.

Bernanke, B., and M. Gertler (1999), "Monetary Policy and Asset Price Volatility", Federal Reserve Bank of Kansas City *Economic Review* 84 (4): 17-51.

Bernanke, B. S., V. R. Reinhart, and B. P. Sack (2004): "Monetary Policy Alternatives at the Zero Bound: An Empirical Assessment," *Brookings Papers on Economic Activity*, 70 (2004-2), pp. 1–100.

Bligh, M., and G. Hess (2013), "Deconstructing Alan: A Quantitative Assessment of the Qualitative Aspects of Chairman Greenspan's Communication", in P.L. Siklos and J-E. Sturm (Eds.), *Central Bank Communication, Decision Making, and Governance* (Cambridge, Mass: The MIT Press), pp. 123-148.

Blinder, A. (1998), Central Banking In Theory and Practice (Cambridge, Mass.: The MIT Press).

Bordo, M.D., and A. Orphanides, Eds. (2013), *The Great Inflation: The Rebirth of Modern Central Banking* (Chicago and London: University of Chicago Press for NBER).

Borio, C., and A. Filardo (2004), "Looking Back at the International Deflation Record", North American Journal of Economics and Finance 15 (December): 287-311.

Bremner, R. (2004), Chairman of the Fed (New Haven, Conn.: Yale University Press).

Bulíř, A., Hurnik, J., and K. Šmídková (2014), "Inflation Reports: How Well Do Cetral Bankers Really Write?, IMF working paper 14/91.

Burdekin, R., and P. Siklos, Eds. (2004), *Deflation: Current and Historical Perspectives* (Cambridge, Mass.: Cambridge University Press).

Cecchetti, S., H. Genberg, J. Lipsky, and S. Wadhwani (2000), *Asset Prices and Central Bank Policy*, Geneva Report on the World Economy No. 2, May.

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

Dale, S., A. Orphanides, P. Österholm (2011), "Imperfect Central Bank Communication: Information Versus Distraction", *International Journal of Central Banking* 7 (June): 3-39.

Danker, D.J. and M.M Luecke (2005), "Background on FOMC Meeting Minutes", *Federal Reserve Bulletin* (Spring): 175-179.

Ehrmann, M., and M. Fratzscher (2009),"Explaining Monetary Policy in Press Conferences", *International Journal of Central Banking* 5 (June): 41-84.

Ehrmann, M., and M. Fratzscher (2007), "The Timing of Central Bank Communication", *European Journal of Political Economy* 23 (March): 124-145.

Eusepi, S., and B. Preston (2010), "Central Bank Communication and Expectations Stabilization", *American Economic Journal: Macroeconomics* 2 (July): 235-271.

Fischer, S. (2017), "Committee Decisions and Monetary Policy Rules", at 'The Structural Foundations of Monetary Policy', Hoover Institution Monetary Policy Conference, 5 May.

Fracasso, A., H. Genberg, and C. Wyplosz (2003), *How Do central Bankers Write: An Evaluation of Inflation Targeting Central Banks*, Geneva Reports on the World Economy Special Report 2.

Goodfriend, M. (2005), "The Monetary Policy Debate Since October 1979: Lessons for Theory and Practice", *Review* of the Federal Reserve Bank of St. Louis 87 (March): 243-262.

Hansen, S., M. McMahon, and A. Prat (2014), "Transparency and Deliberation within the FOMC: A Computational Linguistics Approach", CEPR discussion paper 9994, May.

Hart, R.P., J.P. Childers, and C.J. Lind (2013), *Political Tone: How Leaders Talk and Why* (Chicago: University of Chicago Press).

Havrilesky, T. (1993), *The Pressures on American Monetary Policy* (London: Kluwer Academic Publishers).

Hernández-Murillo, R., and H.G. Shell (2014), "The Rising Complexity of the FOMC Statement", *Economic Synopses*, Federal Reserve Bank of St. Louis, No.23.

Hetzel, R.L. (2008), *The Monetary Policy of the Federal Reserve: A History* (Cambridge, Mass.: Cambridge University Press).

Holmes, D.R. (2014), *Economy of Words* (Chicago: University of Chicago Press).

Jansen, D-J., and J. de Haan (2013), "An Assessment of the Consistency of ECB Communication", ", in P.L. Siklos and J-E. Sturm (Eds.), *Central Bank Communication, Decision Making, and Governance* (Cambridge, Mass: The MIT Press), pp. 183-202.

Klemmensen, R., S. B. Hobolt, and M. E. Hansen (2007), "Estimating Policy Positions Using Political Texts: An Evaluation of the Wordscores Approach", *Electoral Studies* 26: 746-755.

Laver, M., K. Bnoit, and J. Garry (2003), "Extracting Policy Positions from Political Texts Using Words as Data", *American Political Science Review* 97 (May): 311-331.

Lombardi, D., P. Siklos, and S. St. Amand (2017), "Government Bond Yields At the Effective Lower Bound: International Evidence", *Contemporary Economic Policy* (forthcoming).

Loughran, T., and B. McDonald (2016), "Textual Analysis on Accounting and Finance: A Survey", *Journal of Accounting Research* 54 (September): 1187-1230.

Loughran, T., and B. McDonald (2011), "When Is A Liability Not a Liability? Textual nalaysis, Dictionaries, and 10-Ks", *Journal of Finance* 66 (1): 35-65.

Malmendier, U., S. Nagel, and Z. Yan (2017), "The Making of Hawks and Doves: Inflation Experiences on the FOMC", CEPR discussion paper 11902, March.

Malmendier, U., and S. Nagel (2016), "Learning from Inflation Experiences", *Quarterly Journal of Economics* 131 (1): 53-87

Martin, L., and G. Vanberg (2008), "A Robust Transformation Procedure for Interpreting Political Texts", *Political Analysis*, 16 (1): 93-100.

McChesney Martin, W. (1965), "Does Monetary History Repeat Itself?", address before the Commencement Day Luncheon of the Alumni Federation of Columbia University, 1 June, available from FRASER (Federal Reserve Archive).

Meade, E.E., N. A. Burk, and M. Josselyn (2015), "The FOMC Meeting Minutes: As Assessment of Counting Words and the Diversity of Views", FEDS Notes, 26 May.

Meade, E. E., Y. Nozawa, L. Petrasek, and J.K. Zickler (2015), "The Effects of FOMC Communications Before Policy Tightening in 1994 and 2004", FEDS Notes, 24 September.

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

Meltzer, A. (2009), *A History of the Federal Reserve, vol. 2* (Chicago: University of Chicago Press).

Messina, J., T.M. Sinclair, and H.O. Steckler (2014), "What Can We Learn From Revisions to the Greenbook Forecasts?", working paper, George Washington University.

Meyer, L. H. (2004), A Term at the Fed (New York: Harper Business).

Mishkin, R. (2007), "Inflation Dynamics", International Finance 10 (January): 317-334.

Morris, S., amd H. S. Shin (2002), "Social Value of Public Information", *American Economic Review* 92 (December): 1521-1534.

Nelson, E. (2013), "Miton Friedman and the Federal Reserve Chairs, 1951-1979", working paper, October.

Orphanides, A. (2001), "Monetary Policy Rules Based on Real-Time Data", *American Economic Review* 91 (September): 964-985.

Romer, C.D., and D.H. Romer (2004), "A New Measure of Monetary Shocks: Derivation and Implications", *American Economic Review* 94(5): 1055-1084.

Romer, C., and D. Romer (2000), "Federal Reserve Information and the Behavior of Interest Rates", *American Economic Review* 90 (June): 429-457.

Rosa, C. (2013), "The Financial Market Effects of FOMC Minutes", *FRBNY Economic Policy Review* (December): 67-81.

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

Siklos, P.L. (2014), "The Global Financial Crisis and Central Bank Speak", in R.P. Hart (Ed.), *Communication and Language Analysis in the Corporate World* (Hershey, PA: IGI Global), pp. 293-314.

Silber, W.L. (2012), Volcker: The Triumph of Persistence (New York: Bloomsbury Press).

Spicer, J., and L. Kihara (2014), "Central Banks Ending Era of Clear Promises, Return to 'Artful' Policy", Reuters, 7 July.

Staiger, D., J.H. Stock, and M.W. Watson (1997), "The NAIRU, Unemployment and Monetary Policy", *Journal of Economic Perspectives* 11 (Winter): 33-49.

Svensson, L. (2009), "Flexible Inflation Targeting – Lessons from the Financial Crisis", at the workshop 'Towards a New Framework for Monetary Policy? Lessons from the Crisis', Netherlands Bank, 21 September.

Taylor, J. (2010), "Getting Back on Track: Macroeconomic Policy Lessons from the Financial Crisis", *Review* of the Federal Reserve Bank of St. Louis (may/June): 165-176.

Thornton, D.L., and D. C. Wheelock (2014), "Making Sense of Dissents: A History of FOMC Dissents", *Review* of the Federal Reserve Bank of S. Louis (Third Quarter): 213-228.

Thornton, D.L. (2011), "What Does the Change in the FOMC's Statement of Objectives Mean?", *Economic Synopses*, Federal Reserve Bank of St. Louis, No.1.

Thornton, D., and D.C. Wheelock (2000), "A History of the Asymmetric Policy Directive", *Review* of the Federal Reserve Bank of St. Louis 82(5): 1-16.

Timberlake, R. H. (1993), *Monetary Policy in the United States* (Chicago: University of Chicago Press).

Tudor, C., and C. Vega (2014), "A Review of Textual Analysis in Economics and Finance", in R.P. Hart (Ed.), *Communication and Language Analysis in the Corporate World* (Hershey, PA: IGI Global), pp. 122-139.

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

Warsh, K. (2016), "Institutional Design: Deliberations, Decisions, and Committee Dynamics", in *Central bank Governance and Oversight Reform*, edited by J. Cochrane and J. Taylor (Stanford, CA: Hoover Institution Press), pp. 173-193.

Wessel, D. (2009), In Fed We Trust (New York: Crown Business).

Woolley, J., and J. Gardner (2009), "Does Sunshine Reduce the Quality of Deliberation? The Case of the Federal Open Market Committee", working paper, UC Santa Barbara.

White, W. R. (2009), "Should Monetary Policy "Lean or Clean"?" Globalization and Monetary Policy Institute working paper 34, August.

Woodford, M. (2012), "Methids of Policy Accommodation at the Interest-Rate Lower Bound", in *The Changing Policy Landscape*, Economic Policy Symposium Proceedings (Kansas City: Federal Reserve Bank of Kansas City), pp. 185-288.

Woodford, M. (2005), "Central Bank Communication and Policy Effectiveness", in *The Greesnpan Era: Lessons for the Future*, Economic Symposium, Federal Reserve Bank of Kansas City (Kansas City: Kansas City Federal Reserve), pp.399-474.

Wu., C. and F. Xia (2016), "Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound", *Journal of Money, Credit, and Banking* 48(2-3): 253-291.

Wynne, M.A. (2013), "A Short History of FOMC Communication", *Economic Letter*, Dallas Fed, 8 (September): 1-3.

Yellen, J. (2012), "Evolution and Revolution in Central Bank Communication", at the Haas School of Business, UC Berkeley, 13 November,

http://www.federalreserve.gov/newsevents/speech/yellen20121113a.htm.

Yellen, J. (2006), "Minutes of the Federal Open Market Committee, June 28-29", <u>http://www.federalreserve.gov/fomc/minutes/20060629.htm</u>.

Figure 1 Stylized Timeline of FOMC Releases



Note: From <u>https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm</u>. Prior to 1994 the transcripts made from audio recordings were lightly edited but not reviewed by committee members. After 1993 the transcripts are lightly edited with review from committee members. Between 1993 and 2004 the minutes were released 3 days after the subsequent FOMC meeting; beginning in 2005, 3 weeks after the decision. The Bluebook was the document where monetary policy alternatives are discussed; the Greenbook provides a discussion of current economic and financial conditions (and forecasts). These were merged into the Tealbooks in 2010.



Sources: Author's calculations and FRED. WS represents a variable constructed as multiples of 25bp changes based on Wordscores based on rolling estimates, year by year between 1955 and 1993 and then in two year intervals therafter, of equation (2). See the text for the details. Data are quarterly.



Figure 3 Selected Wordscores, 1952-2013

NOTE: TS refers to the transformed Wordscores described in the text (see equation (2)). The figures refer to the calendar year used to construct the reference score from FOMC minutes. Data are monthly in the top figure, quarterly in the bottom figure.



Note: See text for the description of the quantification and interpretation of the tone of the FOMC minutes. Data are monthly. The vertical shaded areas are the NBER recession dates.



Figure 5 DICTION and the Federal Funds Rate

Note: The DICTION indicator is based on a factor model that includes activity, certainty, optimism, and commonality. A single factor is estimated via the Kaiser-Guttman approach using maximum likelihood. The vertical dashed line indicates when FOMC minutes were made public. The left hand scale is the fed funds rate (in %).

Figure 6A Select Impulse Responses: Communication (Wordscores) and Economic Activity:

Post-1994 Sample Ends in 2008Q4



Note: Based on a VAR for the 1955Q1-1993Q4 (pre) and 1994Q1-2008Q4 (post) samples. The text provides the details of the VARs. 1 lag is used for the post sample and 5 lags in the pre 1994 sample.

Figure 6B Select Impulse Responses: Communication (Wordscores) and Economic Activity:



Post-1994 Sample Ends in 2012Q4

Note: same VAR as in Figure 6A except 4 lags are used.

Figure 7A Select Impulse Responses: Communication (DICTION) and Economic Activity:



Post-1994 Ends 2008Q4

Note: Based on a VAR for the 1955Q1-1993Q4 (pre) and 1994Q1-2008Q4 (post) samples. The text provides the details of the FAVARs. 1 lag is used for the post sample and 5 lags in the pre 1994 sample.





Response of DICTION to real GDP growth: post-1994

Note: same VAR as in Figure 7A except 2 lags are used.