The Perceived Causes of Monetary Policy Surprises

Miguel Acosta*
Columbia University

December 8, 2021

[Click here for Latest Version]

Abstract

I estimate the macroeconomic effects of two critical aspects of Federal Reserve (Fed) communications: forward guidance regarding the path of interest rates and the provision of macroeconomic information. To estimate these effects, I identify two new series of shocks: monetary policy shocks and "information shocks." I recover the shocks by estimating a model of how Fed announcements determine interest-rate and GDP expectations in high frequency, using a measure of GDP forecast revisions I construct from the text of newspaper articles. To identify the model, I use a discrete change in the Fed's communication policy: the introduction of interest-rate forward guidance. I find that the identified monetary shock has macroeconomic effects that are consistent with New Keynesian models, and fall at the upper end of previous estimates. Additionally, information shocks resemble aggregate demand shocks and have effects of similar (absolute) magnitude as monetary shocks, which highlights the importance of the Fed's role in providing macroeconomic information.

Keywords: Monetary Policy, Communication, Natural Language Processing, Expectation Formation

*Columbia University, Department of Economics. m.acosta@columbia.edu I owe an enormous debt of gratitude to my advisors—Hassan Afrouzi, Jennifer La'O, Emi Nakamura, Jón Steinsson, and Michael Woodford—for numerous insightful discussions regarding this work. The paper has also benefited from discussions with Andrea Ajello, Paul Bouscasse, Lydia Cox, Emilien Gouin-Bonenfant, Juan Herreño, Isaiah Hull (discussant), Ellen Meade, José Luis Montiel-Olea, Christian Moser, Joseph Saia, Jesse Schreger, Yeji Sung, Pierre Yared, and seminar participants at Berkeley, Columbia, the 14th International Conference on Computational and Financial Econometrics, and the 2021 EEA Annual Congress. Part of this work was completed during a dissertation fellowship at the Federal Reserve Bank of San Francisco, although my discussions with the economists there were instrumental, the views expressed herein are mine and not those of the Fed. I owe a special thanks to Yasmin Saira at the Business and Economics Library at Columbia for help accessing data, especially when physical access to Columbia's libraries was limited due to the pandemic. Hannah Bingley and Blaine Helleloid provided outstanding research assistance. The work was supported by financial support from an Alfred P. Sloan Foundation Pre-Doctoral Fellowship in Behavioral Macroeconomics awarded through the NBER.

1 Introduction

The role of the Federal Reserve (Fed) has moved far beyond setting the level of the overnight interest rate. In the eyes of financial market participants, the Fed Chair is often seen as a fortune teller who communicates predictions regarding not only the path of future interest rates, but also forecasts of macroeconomic outcomes. An understanding of the Fed's role in the economy, then, crucially hinges on an understanding of the macroeconomic effects of both aspects of Fed communications. Two challenges arise when estimating these effects. First, communications regarding the path of interest rates and communications regarding the economic outlook are both highly endogenous with respect to economic fundamentals. Second, the Fed typically engages in both types of communications simultaneously, which complicates attempts to separately identify their effects.

In this paper I estimate the macroeconomic consequences of Fed communications—both the effects of monetary policy (explicitly communicating about future interest rates) and information provision policy (the effects of providing information about macroeconomic fundamentals). I contribute to a longstanding literature that studies the effects of monetary policy using market reactions to Fed policy announcements, and provide new estimates that overcome an important conceptual issue in the identification of these effects: Market-based measures of interest-rate expectations can respond to both types of policy, thereby identifying neither. I also take the view that information provision is an important component of the Fed's communication policy—rather than a statistical nuisance that challenges the identification of exogenous variation in monetary policy, as the literature has come to perceive it.

To estimate the effects of both aspects of Fed communications on macroeconomic outcomes, I identify two new series of shocks: perceived monetary policy shocks (shocks to the policy rule) and "information shocks" (shocks to beliefs about economic fundamentals).[1] I posit that an econometrician needs access to (at least) two measures of market reactions that respond differently to the two types of shocks in order to identify them separately. I show that interest-rate and GDP forecast revisions emerge as natural candidates to accomplish the task, based on the implications of standard New Keynesian theory. Intuitively, because interest-rate and GDP forecast revisions react differently in response to monetary policy and information shocks, observing their joint reactions to monetary policy announcements can provide useful information for identifying the prevalence of each shock.[2]

The foundation of my shocks consists of two high-frequency measures of macroeconomic forecast revisions: interest-rate surprises and a new, text-based measure of GDP forecast revisions. The former is standard in the empirical monetary literature. The latter I construct using newspaper articles written about each Fed policy meeting. Specifically, I compute a GDP "directionality index"

[1] I introduce the word "perceived" here because my shocks are based on data on expectations (as are estimates of monetary shocks in the literature that use interest rate futures data). The illustrative model in section 3.1 makes this observation precise.

[2] To give a familiar example, a contractionary monetary policy shock raises interest rates and lowers real GDP in standard New Keyensian models. Instead, a positive aggregate demand shock (which is what my information shocks resemble) causes both variables to increase.

for articles written in a one-day window around each announcement. The index is based on the difference between increasing and decreasing mentions of GDP. I anticipate that this high-frequency series of GDP expectations will be useful in other contexts. My proxy for GDP forecast revisions is the unpredictable component of post-meeting directionality vis-à-vis pre-meeting directionality. I assume that macroeconomic expectations measured shortly before each policy announcement reflect the Fed's communication expected by markets, given all macroeconomic events that have occurred up to the announcement. The difference between pre- and post-meeting expectations, then, should only arise from exogenous policy or information asymmetries. As such, the high-frequency construction of my variables addresses the general endogeneity of interest-rate and GDP forecasts with respect to observable economic fundamentals.

Because the Fed communicates simultaneously about interest rates and the macroeconomic outlook, I estimate a simultaneous-equations model of how markets update their interest-rate and GDP forecasts in response to Fed announcements, in order to recover the structural shocks from these forecast revisions. In the model, market participants are Bayesian forecasters whose model of the economy is a linear relationship between macroeconomic shocks (here, monetary policy and information shocks) and macroeconomic variables (here, interest rates and GDP). Despite this simple formulation—a system of two equations determined by two exogenous shocks—this is the forecasting model implied by the dynamic stochastic general equilibrium models that permeate macroeconomic analysis.

To identify the model, I use a discrete change in the Fed's communication policy: the introduction of interest-rate forward guidance in 2003. As highlighted by Lunsford (2020), prior to 2003 the Fed's post-meeting policy statements primarily described the economic outlook. In August 2003 the Fed began the practice of interest-rate forward guidance when it promised to keep interest rates low "for a considerable period." To see how this can help identify the model, consider the implications of this policy change. Before 2003, learning about the Fed's economic outlook was straightforward; in contrast, inference about the path of interest rates was possible only indirectly through the Fed's discussion of the economy. This observation allows me to make my formal identification assumption: Fed announcements induced market participants to update their expectations about the future path of interest rates, relative to their expectations about economic fundamentals, more completely after 2003. My data support this assumption: Before 2003, interest-rate and GDP forecast revisions were positively correlated. Intuitively, this regime primarily provides identification of the information shock. After 2003, the series are essentially uncorrelated. Therefore, this latter regime provides identification of the monetary shock by providing data whose variation is driven by a factor orthogonal to the information shock. This is "identification by heteroskedasticity," proposed by Rigobon (2003).[3] Formally, while the model is not identified within either regime

[3]Other papers (Wright, 2012; Arai, 2017; Nakamura and Steinsson, 2018) have used the heteroskedasticity-based identification assumptions in seeking to estimate the effects of monetary policy. The approach is typically seen as a method of purging "background noise" (or latent factors, in the case of Gurkaynak et al. (2020)) from OLS regressions. I use the approach completely differently in that I am interested in estimating two shocks, not simply purging one shock of interest of a nuisance component. Lewis (2019) identifies similarly-named shocks using a heteroskedasticity-based approach that provides only statistical identification.

separately, the added assumption that the shocks have the same effects on observables across the two regimes imposes enough parametric restrictions to jointly solve the model's implied moment conditions.[4] Importantly, while I estimate the model using a 3 year window around the 2003 policy change, I use the identified model to construct estimates of the two shocks over my entire 1999–2019 sample period.

I find that the monetary policy shock I identify has effects on macroeconomic outcomes and expectations that are consistent with New Keynesian macroeconomic models, a conclusion I reach without imposing such consistency a priori. A monetary shock that raises longer-term interest-rate expectations on impact leads to declines in industrial production and inflation. The responses of both variables are fairly delayed, with peak responses estimated between 2 and 3 years after the shock. Nominal and real interest rates increase on impact, while GDP and inflation expectations decrease. Notably, I use simple empirical specifications to identify these effects. When I substitute my measures for other estimates of contractionay monetary shocks in these specifications, the estimated responses of output are generally not in line with predictions from New Keynesian models. In terms of magnitudes, my results suggest that the effects of monetary policy are big, in the language of Coibion (2012), and similar to those estimated by Romer and Romer (2004): a 25 basis point increase in the policy rate leads to a roughly one percent decrease in industrial production.

Moving to the effects of information provision, I find that information shocks have effects on macroeconomic outcomes and expectations that are similar in (absolute) magnitude to those of monetary policy shocks. This remains true when controlling for recent macroeconomic news, which suggests first that the Fed plays an important role in its characterization of macroeconomic shocks.[5] In addition, my information shock creates a positive comovement of output and inflation (and expectations thereof, in low and high frequency), which suggests that the information primarily concerns demand-type factors.

The notion that monetary policy announcements can convey macroeconomic information, and thus contaminate estimates of exogenous monetary shocks, was put forth by Romer and Romer (2000). Campbell et al. (2012) and Nakamura and Steinsson (2018) highlight the fact that the presence of "information effects" can contaminate traditional high-frequency estimates of monetary policy shocks—my results confirm that the contamination is substantial: about 80% of a widely used high-frequency shock series is made up of information effects. Therefore, the empirical challenge became: How to disentangle information provision from exogenous monetary policy? My model shows that this challenge is the familiar problem of simultaneous determination encountered in supply and demand systems or structural vector autoregressions. Seen in this context, early work in this area imposed zero restrictions to identify monetary shocks, which are not warranted under the presence of information effects.[6] My work relaxes these restrictions. Nakamura and Steinsson

[4] This language comes from Stock and Watson (2017).

[5] This also addresses the critique of Bauer and Swanson (2020), whereby the positive correlation between interest rate surprises and macroeconomic expectations can result from both series' reaction to the same economic news.

[6] This is the case in Kuttner (2001) (who introduced surprises in the current-meeting interest rate) and Gürkaynak et al. (2005) (who introduced the notion of a shock to the path of interest rates). By not entertaining the possibility that changes in interest-rate futures could be driven by shocks other than exogenous monetary policy, the authors

(2018) estimate a structural model to overcome the identification challenge. In contrast to my work, their approach depends on all of the assumptions underlying the particular macroeconomic model.

Other papers cognizant of information effects have sought model-free approaches to identify high-frequency monetary policy shocks. Given that information effects are posited to stem from the Fed's private information, Miranda-Agrippino and Ricco (2021) and Handlan (2020) propose orthogonalizing high-frequency interest-rate surprises to the Fed's private information as captured by the Fed staff's presentation materials ("Greenbook forecasts").[7] This approach suffers from two conceptual shortcomings that I sidestep. First, it assumes that the staff's economic assessment spans that of the Fed's policymaking committee, which is ultimately tasked with policy communication. Romer and Romer (2008) show that these assessments generally do not align. Second, this approach requires choosing a set of variables that completely span the Fed's private information *when the announcement is made*. I avoid having to posit the variables over which the Fed has private information, which is difficult to know, given the vast number of indicators that inform Fed policy decisions.[8] In addition, by relying on staff-created reports, neither paper can control for events that occur shortly before policy announcements, and thus both papers potentially fail to control for endogenous macroeconomic events to which the Fed might respond. My high-frequency measures avoid this concern.

Cieslak and Schrimpf (2019) and Jarociński and Karadi (2020) assume theoretically motivated sign restrictions regarding the relationship between monetary shocks, information shocks, stock returns, and interest-rate surprises. These sign restrictions allow the authors to discuss the relative importance of—and identify, in the case of Jarociński and Karadi (2020)—monetary and information shocks. In high frequency, stock returns and interest-rate surprises are consistently negatively correlated, in contrast to my estimates of output expectations. This suggests a limited role for stock prices in differentiating between monetary and information shocks. In fact, it suggests a limited role for information effects.[9] Examining a variable whose response to each shock differs substantially from the responses of interest rates—e.g., output expectations—instead provides more

imposed the restriction that other shocks had zero effect on high-frequency interest-rate changes.

[7] Campbell et al. (2012) also attempt to control for information effects by controlling for professional macroeconomic forecasts. As Woodford's comment to that article notes, this approach requires that the control variables span the Fed's reaction function, and suffers from the possibility that not all relevant information may be captured by the lower-frequency forecasts. Hansen and McMahon (2016) also study the effects of both types of communication in a low-frequency setting.

[8] Doh et al. (2020) and Handlan (2020) are able to control more flexibly for the Fed staff's information than Miranda-Agrippino and Ricco (2021) by using machine-learning and text-based techniques applied to the Fed's alternative policy statements. Cai et al. (2021) and Lakdawala (2019) introduce alternative methods for controlling for the Fed's *private information* (the difference between the Fed's and the public's information) which, as discussed in Section 4.2.2, is necessary for properly "removing" information effects.

[9] This negative correlation does not serve as evidence "against" the presence of information effects—it only rejects the notion that information effects are the only shock operating when the Fed makes announcements. Put differently, without information effects, the negative correlation might be even stronger. Ultimately, the puzzle that suggested the presence of information effects was the effect of interest rate surprises on expectations of real macroeconomic variables (GDP, unemployment, etc.), not stock returns. D'Amico and King (2017) overcome this issue using a set-identification approach, However, they use lower-frequency data on expectations, which may not incorporate all relevant macroeconomic information to which the Fed responds.

power for identification.[10] By studying GDP expectations, I directly address one of the main puzzles in the high-frequency literature: that both output and output expectations increase in response to positive interest-rate surprises. Additionally, since my measure is not constructed using financial market data, I eliminate the potential confounding role of risk premia. Finally, my identification approach does not require the a priori imposition that the identified shocks have theoretically consistent effects. In contrast to my estimates, this approach only provides set identification.

The paper proceeds as follows. In Section 2, I present my text-based proxies of high-frequency macroeconomic forecast revisions. In Section 3, I present a simple theoretical framework that explains the identification challenge. In the context of that model, I lay out my identification assumptions and estimate the structural shocks. Section 4 contains evidence on the effects of these shocks on macroeconomic outcomes and expectations, along with a comparison with the effects estimated using existing measures. In Section 5, I dig deeper into both shocks and show that exogenous shocks to overnight interest rates are a thing of the past. Additionally, most of the "information effects" gleaned by markets reflect demand-type factors. Section 6 is the conclusion.

2 Data: Construction and Validation

A novel aspect of this research is the construction of a high-frequency (HF) proxy for output expectations. Why output? Output is a variable that features in nearly every macroeconomic model, and is the variable whose relationship to traditional HF estimates of monetary shocks is of first-order concern. Why high frequency? Measuring changes in high frequency allows me to isolate the source of the change: Here, the Fed. This ensures that the shocks I estimate are indeed Fed-based shocks and not confounded by information from other sources. Lacking access to existing high-frequency measures of output expectations, however, I use an alternative data source to create my proxy: newspaper articles.

At its core, the text analysis I employ in this paper is essentially a counting exercise: I count co-occurrences of output-related words with words that indicate whether an object is increasing or decreasing. Text-analysis methods for analyzing Fed announcements have used increasingly realistic models of natural language,[11] but with this has come a loss of interpretability and replicability. In contrast, my approach, by design, is conceptually simple, transparent, and easily replicable.

In Section 2.1, I describe the construction of my expectations proxy. The simplicity of my construction comes at a cost: It is subject to the criticism of being subjectively designed. I therefore also use a complementary construction that is a bit more complex but removes some subjectivity from the process. My results are nearly unchanged. I then turn, in Section 2.2, to an exercise that serves to validate my measure. I show that, in levels and differences, my proxies are positively

[10] This, combined with the fact that Jarociński and Karadi's approach only provides set identification of the shocks, allows the authors to (statistically) learn very little about the effects of Fed information—their identified set of impulse responses (that are robust to the chosen prior) includes zero for most horizons.

[11] Examples include the use latent Dirichlet allocation (Acosta, 2019; Hansen et al., 2019), whose topics are weighted averages of every word in a corpus; structurally rotated latent semantic analysis (Ter Ellen et al., forthcoming); and more general neural networks (Handlan, 2020).

correlated with existing lower-frequency measures of expectations. Finally, in Section 2.4 I briefly describe the traditional numerical data I use throughout the rest of the paper.

2.1 High-Frequency Text-Based Proxies of Macroeconomic Expectations

The construction of my index proceeds in three steps. I first construct a set of words related to economic output, a set of words that indicate something is increasing, and a set of words that indicate something is decreasing. I then collect a set of newspaper articles written in a one-day window around FOMC announcements, split them into articles written before and after the announcement, and compute a *directionality index* on pre- and post-meeting articles. I then construct my measure of output forecast revisions as the unpredictable component of post-meeting directionality vis-à-vis pre-meeting directionality. I discuss these steps in turn.

Words Lists My first set of words refer to economic output. This list is given by

$$\mathcal{Y} \equiv \{\texttt{economic\_growth}, \texttt{growth}, \texttt{economi}, \texttt{consumer\_spending}, \texttt{output}\}.$$

The origins of this list are the triplet `output`, `growth`, and `economy`. In order to show that the results of this paper are robust to expanding or modifying this word list, I trained the popular natural language model of Mikolov et al. (2013) on a large corpus of newspaper articles and extended the initial triplet by extracting—from the model—synonyms of the triplet. The list of synonyms, sorted by their proximity to the triplet, was sensible out to seven words. Thus, in my main analysis I retained the top five most similar words, and later show the robustness of my results to using the top three and top seven most similar words. The details of constructing the list of synonyms are provided in Appendix E.2.

The next two word lists come from the Harvard IV-4 dictionary. The first list contains words that indicate an object is increasing. This set, $\mathcal{I}$, consists of all words in the *increase* and *rise* lists from the Harvard dictionary. The second set, $\mathcal{D}$, consists of words indicative of an object decreasing: These are the words from the *decrease* and *fall* word lists. In both cases I retain the unique set of "stemmed" words (the lexical root of words). This ensures that I count all variants of these words in my newspaper articles. For example, `increase`, `increases`, and `increased` are all counted as mentions of `increase`.[12] These word lists are provided in Appendix E.1.

Newspaper Data I analyze all newspaper articles written the day before, day of, and day after each FOMC meeting in the *New York Times*, *Wall Street Journal*, and *Washington Post*. I collect these articles from Factiva, searching for articles with the keywords `Federal Reserve` and `FOMC`.[13] The 3-day window ensures that I capture the many articles written on the first day of two-day FOMC meetings (these typically start "The Fed begins a two-day policy meeting today"), and the print articles written on the day after the meeting. I found the timestamps in Factiva to

[12] I use the stemmer of Porter (1980).

[13] `Fed` is too general, and returns articles related to food and eating.

be fairly unreliable, so I manually sorted all articles on the day of the FOMC meeting based on whether they were written before or after the meeting took place. Similarly, I removed duplicate articles.[14]

By considering a 3-day window, I potentially include articles that are not primarily written about the policy meeting. To guard against this, I only retain articles that have `fed` or `rate` in the title or have one of {`fed`, `fomc`, `federal`} within five words of one of {`meet`, `meeting`, `policy`, `decision`, `rate`}. Thus, while the window is not as high frequency as tick-level data, the ability to narrow articles to those that specifically discuss the policy meeting reduces the possibility that other events influence my analysis.[15]

Finally, I concatenate all pre- and post-meeting articles into a single document for each meeting t, given by $\texttt{PRE}_t$ and $\texttt{POST}_t$, respectively. I stem all words to their lexical root using the algorithm of Porter (1980). Because my main measure counts co-occurrences of output words and modifiers, I follow Lucca and Trebbi (2009) and break each document into standalone "chunks" of text. This is performed by first tagging each word in a sentence with its part of speech, then using grammatical rules to separate the sentence into smaller constituent sentences (for long or run-on sentences) and various compound phrases. A chunk of text is either a complete sentence with no nested complete sentences or a noun phrase (e.g., `low inflation`) not contained in a complete sentence. The definition of a complete sentence follows rules of the English language and relies on the grammatical parsing of each sentence. I parse sentences using the algorithm of Manning et al. (2014). Though still an imperfect way to determine whether one word modifies another, this method (anecdotally) produces more sensible classifications than determining modification based on proximity (though my results are robust to a proximity-based approach).

Directionality Index Having gathered a list of words and documents to analyze, I then put these together to form a directionality index for economic output. For each aggregated document $d \in \{\texttt{PRE}_t, \texttt{POST}_t\}$ with sentence chunks indexed by c, the index is defined as

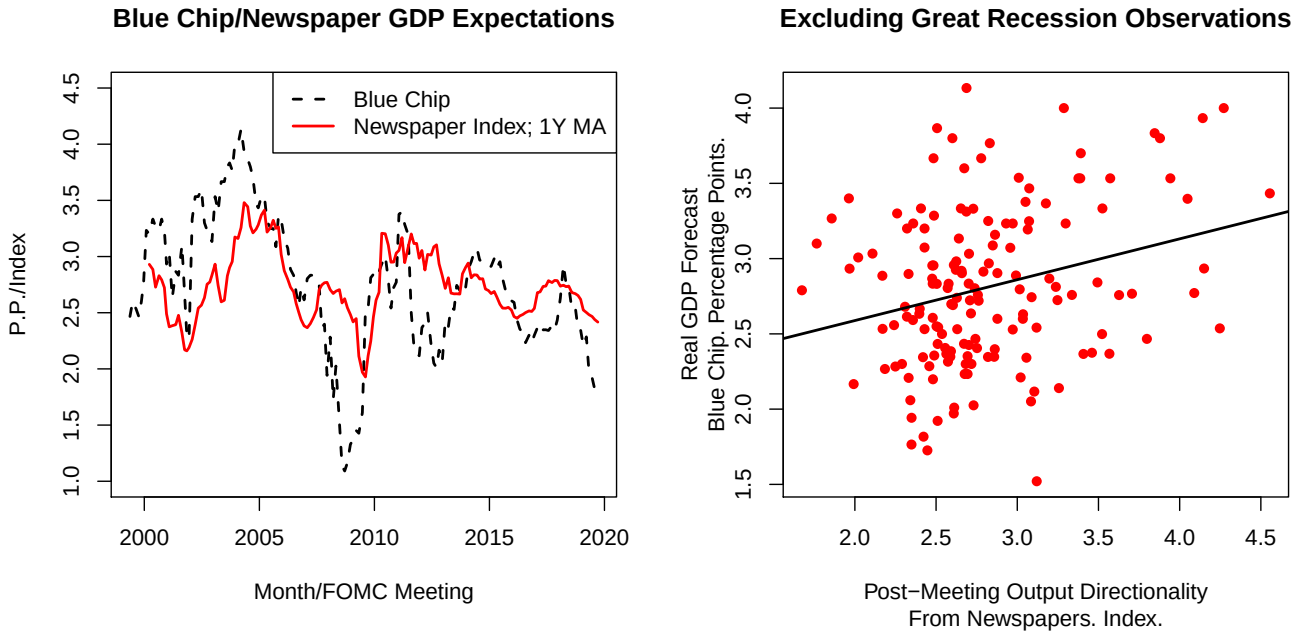
$$\omega_d \equiv \underbrace{\frac{\sum_{c\in d}\sum_{w\in\mathcal{T}_i}\mathbb{1}\{w\in c\}\left[\sum_{i\in\mathcal{I}}\mathbb{1}\{i\in c\}-\sum_{d\in\mathcal{D}}\mathbb{1}\{d\in c\}\right]}{\sum_{c\in d}\sum_{w\in\mathcal{T}_i}\mathbb{1}\{w\in c\}\left[\sum_{i\in\mathcal{I}}\mathbb{1}\{i\in c\}+\sum_{d\in\mathcal{D}}\mathbb{1}\{d\in c\}\right]}}_{\text{Sentiment Score}} \times \underbrace{\frac{1}{|d|}}_{\text{Normalization}} . \tag{1}$$

The "sentiment score" component of this measure is so named because it resembles sentiment scores found across the natural language processing literature, whose typical construction is the percent difference in counts of "positive" and "negative" words. Here, the score is the number of times a

[14] To do this, I created a vector representation of each article i, denoted by v_i, whose length was equal to the total number of unique words across all articles. The j-th element is the number of times term j appears in document i. I then calculate the pairwise distances between all articles using the cosine distance metric: For two documents v_i and v_k, this is $1 - v_i'v_k/[\sqrt{v_i'v_i} + \sqrt{v_k'v_k}]$. Plotting the distribution of cosine similarities, I found a second mode at 0.97, so I randomly chose one from each set of articles with mutual cosine distance above 0.97.

[15] Upon inspecting the articles, I found that many articles that fit the aforementioned criteria for being "relevant" instead discussed how another country's central bank might react to the Fed's announcement. Those articles almost always started with the name of a non-US city (i.e., reflecting where the article was written), so I exclude articles that have one of these non-US city names within the first 5% of the text. The list of these cities is in Appendix E.1.

Figure 1: Newspaper-based Output Directionality and Macroeconomic Expectations

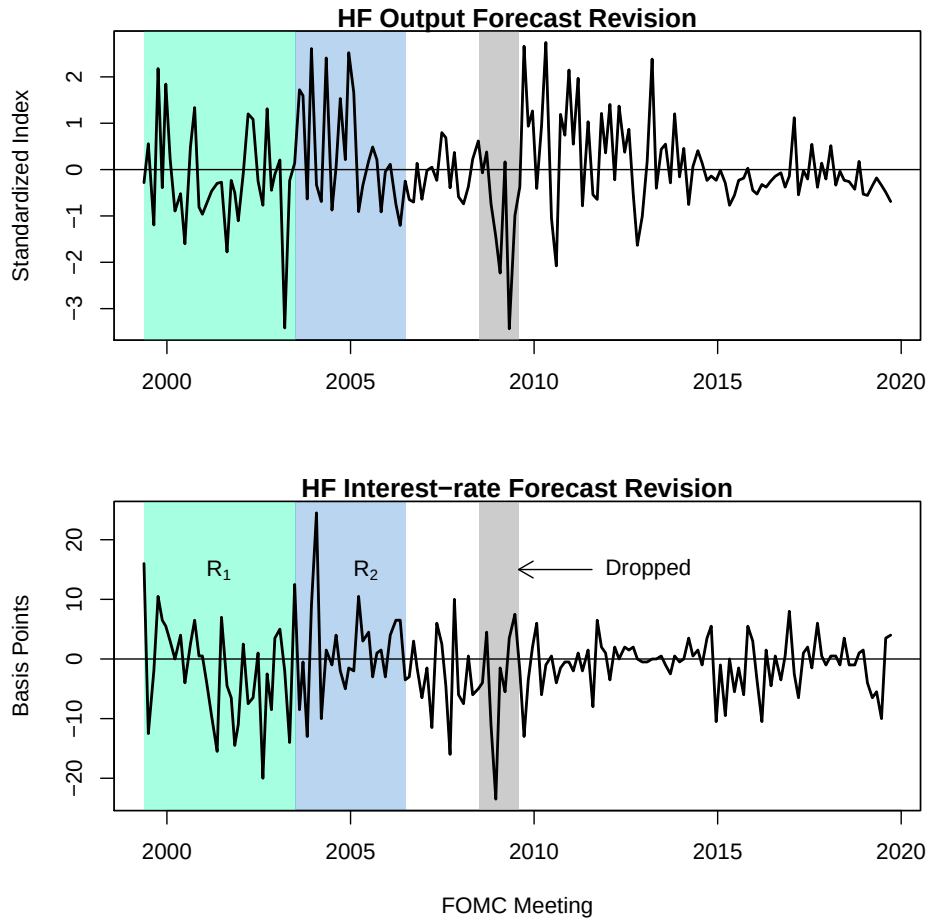


The left panel plots the average GDP forecast from the Blue Chip (black dashed line) described by equation (3) against the 1-year (eight-meeting) moving average of post-meeting output directionality ω_{POST_t} (solid red line). The right panel shows a scatter plot of the two series (in which each meeting month is merged with the corresponding Blue Chip month), using only observations that are used in the baseline empirical sample (excluding the July 2008–July 2009 period and meetings that occur in the first 7 days of each month). For ease of comparison, I normalize ω_{POST_t} to have the same mean and standard deviation as the corresponding $\mathbb{Y}_{t|t}$ variable.

word from the topic is modified by an increasing word less the number of times a word from the topic is modified by a decreasing word, divided by mentions in either direction. For chunks that contain a negation (`not` and `n't`), I flip the sign of the bracketed term in the numerator. Dividing by the total number of sentences (the "normalization") gives the index the average per-sentence (chunk) sentiment score.

In levels, I posit the economic growth directionality index as a proxy for macroeconomic expectations. The time series of ω_{POST_t} in the left panel of figure 1 (which is discussed in more detail in Section 2.2) supports this interpretation. In Appendix E.3, I create an alternative directionality index that makes this link explicit. Briefly, I count all pairwise co-occurrences of words in slightly expanded versions of $\mathcal{Y}, \mathcal{D}$, and $\mathcal{I}$ in a large corpus of newspaper articles. I estimate a LASSO regression to predict the level of real GDP expectations from the Blue Chip survey using these co-occurrence counts. This gives me a mapping from words to GDP expectations. I apply this mapping to pre- and post-FOMC articles to generate my alternative directionality indexes. The co-occurrences "selected" by LASSO are sensible. However, because my results are essentially unchanged between my baseline measure and this alternative, I opt to use the simpler index presented in equation (1).

Figure 2: High-frequency Forecast Revisions of Output and Interest Rates

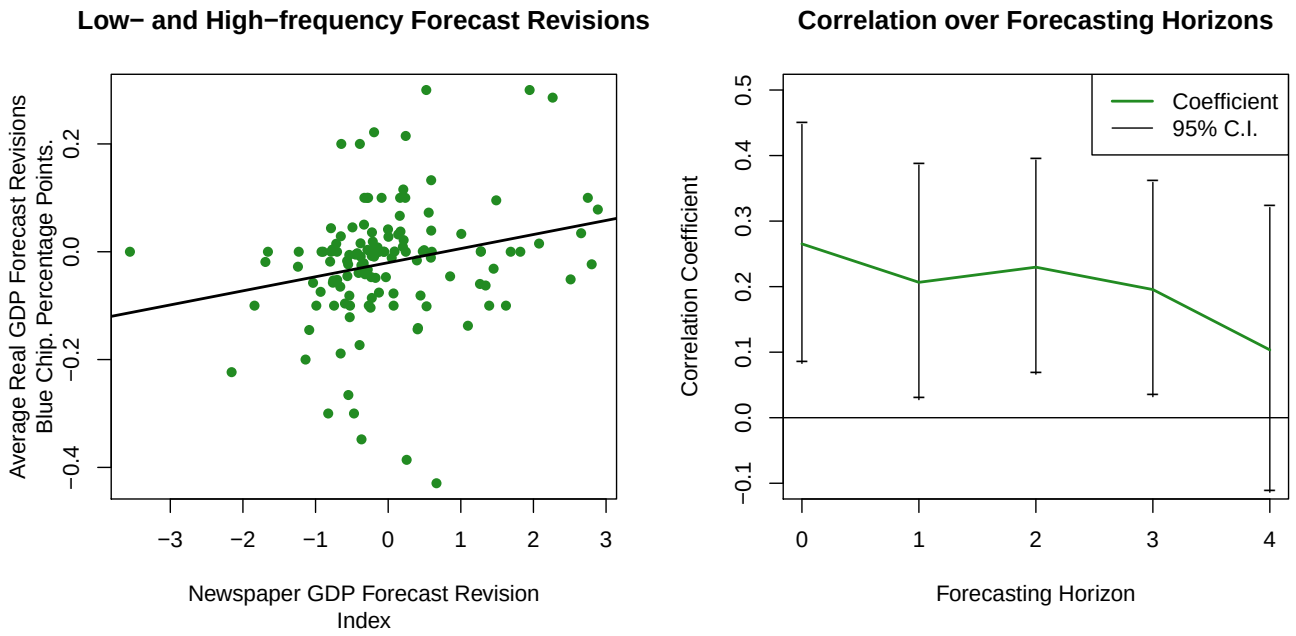


The top panel plots the high-frequency output forecast revision measure (from newspapers) described in Section 2.1 (the residual from equation 2). The bottom panel shows the 30-minute change in the price of the fourth-quarter Eurodollar futures, described in Section 2.4. The first two shaded areas present the regimes used for identification, described in Section 3.2. The third, gray shaded area highlights observations that are dropped for most analyses, discussed in Section 2.4.

Proxy for Changes in Macroeconomic Expectations The last step is to take the directionality indexes from levels to differences. One complication arises in this step: Namely, the indexes constructed above do not distinguish between descriptions of the Fed's announcement (i.e., the "level" of the announcement) and the surprise component of that announcement. I therefore construct my proxy for changes in output expectations as the unpredictable componenet of post-meeting directionality vis-à-vis pre-meeting directionality. Formally, I construct the proxy $\widehat{y}_t$ as the residual from the following regression:

$$\omega_{\text{POST}_t} = a + b\,\omega_{\text{PRE}_t} + c\,\omega_{\text{POST}_{t-1}} + \widehat{y}_t. \tag{2}$$

Figure 3: Blue Chip GDP Forecast Revisions and the High-frequency Proxy



The left panel plots the average Blue Chip forecast revision over the next three forecast horizons ($\mathbb{Y}_{t|t+1} - \mathbb{Y}_{t|t+1}$ from equation (3)) against my HF measure of output forecast revisions, $\widehat{y}_t$, standardized to have unit variance over the plotted sample. The slope of the line thus corresponds to the estimate of β from equation (4) in Table 1. The right panel shows this estimated slope using different Blue Chip forecast horizons—from zero to four quarters ahead. In that panel the right- and left-hand-side variables are standardized, so the coefficient is a correlation coefficient. The sample is as described in Section 2.4. Confidence intervals are generated using robust standard errors.

The constant and slope coefficients in the regression are 0.0002, 0.18, and 0.11.[16] This suggests that a fair amount of post-meeting articles discuss the surprise component of the policy announcement; to quantify this differently, the R^2 from the regression is only 0.13, which suggests that a fair amount of post-meeting coverage is unpredictable based on pre-meeting coverage. Without the inclusion of $\omega_{\text{POST}_{t-1}}$, the index exhibits a small degree of autocorrelation. Otherwise its inclusion is inconsequential for my estimates. The top panel of Figure 2 shows the time series of $\widehat{y}_t$.

2.2 Validation

This section contains several exercises meant to validate the indexes constructed in Section 2.1 as proxies for macroeconomic expectations. I show, first, that the directionality indexes (the ω_t variables) are positively correlated with the level of real GDP expectations from the Blue Chip survey of the same month. I then show that the high-frequency forecast revisions I construct, $\widehat{y}_t$, are positively correlated with lower-frequency forecast revisions. Finally, I perform a case study to show that the indexes are picking up features of the text that are noticeable to a human reader.

Figure 1 shows that the macroeconomic directionality indexes are positively correlated with

[16]The heteroskedasticity-robust t-statistics are 4.9, 3.6, and 1.5, respectively. I estimate this regression over my full sample of newspapers, which starts in 1995, but the regression coefficients are almost identical using the 1999+ sample (respectively, 0.0002, 0.18, and 0.19).

macroeconomic expectations taken from the Blue Chip survey. The dashed black line in the left panel is a summary statistic of GDP expectations I use often throughout the paper:

$$\mathbb{Y}_{t|\tau} \equiv \frac{1}{3}\sum_{h=1}^{3} \mathbb{E}_{\tau}^{\text{Blue Chip}}[\Delta\text{Real GDP}_{t+h}]. \tag{3}$$

This summary statistic—used by, e.g., Nakamura and Steinsson (2018)—is the average forecast of real GDP growth made in month τ over the year starting in month t (in Figure 1, I set $\tau = t$).[17] The solid blue line in the left panel is the 1-year (eight-meeting) moving average of post-meeting directionality ω_{POST_t}.[18] The lines are clearly positively correlated, and the right panel confirms that this correlation is not the result of one or two influential observations. The points in the right panel plot $\mathbb{Y}_{t|t}$ against ω_{POST_t} (not the moving average) for the months in my baseline sample.[19] The correlation of these series is 0.36, with a robust standard error of 0.08.[20] In summary, my output expectations are correlated with traditional measures of output expectations, in levels.

My index of output forecast *revisions* is also correlated with traditional measures of output expectation revisions. Figure 3 shows that the high-frequency proxy for macroeconomic forecast revisions, $\widehat{y}_t$, is also positively correlated with lower-frequency forecast revisions taken from the Blue Chip survey. The left panel plots the 1-month Blue Chip forecast revision surrounding each FOMC meeting ($\mathbb{Y}_{t|t+1} - \mathbb{Y}_{t|t}$) against the high-frequency proxy. This positive correlation is encouraging. I put this correlation to the test by estimating the regression

$$\mathbb{Y}_{t|t+1} - \mathbb{Y}_{t|t} = \alpha + \beta_t + e_t, \tag{4}$$

where X_t contains various explanatory variables of interest. The results are in Table 1. The first column shows that the correlation is indeed statistically significant—a finding that remains consistent across the columns. The correlation is not due to the construction of the Blue Chip variable in equation 4; the right panel breaks down the correlation by Blue Chip forecast horizon. This shows that $\widehat{y}_t$ is fairly evenly correlated with forecast revisions out to about 1 year.

Table 1 also highlights why $\widehat{y}_t$ is a useful addition to the study of monetary policy announcements. In the second column, I only include a traditional estimate of monetary shocks on the right-hand-side of equation (4)—HF changes in 1-year interest-rate expectations, denoted by $\widehat{i}_t$ (displayed in Figure 2 and described in Section 2.4). This column is a reproduction of a known, but puzzling, finding: Surprise interest-rate increases (i.e., a contractionary traditional HF shock) cause GDP expectations to increase. Interestingly, when the HF output revision is added to that

[17] Formally, the average forecast of GDP growth for quarters $q(t)$, $q(t)+1$, and $q(t)+2$, where $q(t)$ is the quarter of month t.

[18] The graph with pre-meeting directionality looks very similar.

[19] I exclude the observations from July 2008 through July 2009, which will be removed from my sample later in the paper because of the asset-pricing anomalies over this period discussed by Nakamura and Steinsson (2018). Also, to make the sample consistent with my later analysis, I exclude months with FOMC meetings that occur in the first 7 days of the month since, as Nakamura and Steinsson (2018) note, the exact timing of when Blue Chip respondents complete their surveys is not clear.

[20] The correlation with pre-meeting directionality (ω_{PRE_t}) is 0.37 (s.e. 0.08).

Table 1: Low- and High-frequency Forecast Revisions: Robustness

$\widehat{y}$ (HF newspaper)	0.0267		0.0258		0.0191
	(2.72)		(2.68)		(1.95)
Interest-rate surprise		0.00452	0.00439	0.00482	0.00207
		(3.17)	(3.20)	(2.85)	(1.71)
HF S&P 500 return				0.0574	
				(2.42)	
Jobs number					-0.00896
					(-7.30)
Jobs number, surprise					-0.0546
					(-0.50)
One-month S&P Return					0.00696
					(3.65)
Lagged BBK index					0.0410
					(3.82)
Observations	131	131	131	138	131
R^2	0.0566	0.0629	0.116	0.0684	0.373
$\overline{\lvert\text{LHS}\rvert}$	0.0775	0.0775	0.0775	0.0775	0.0775

t-statistics computed using robust standard errors are in parentheses, since the point estimates are so small.

This table shows estimates of equation (4). The left-hand-side is the 1-month forecast revision of GDP growth over the next three forecasting horizons, presented in equation (3), from the Blue Chip survey, in percentage points. The HF output forecast revision $\widehat{y}_t$ is from newspapers and standardized over the regression sample. The interest-rate measure ($\widehat{i}_t$) is the change in the four-quarter Eurodollar future contract, in basis points. The 1-month stock return is the 4-week percentage return in the S&P 500 (ending 1-day before the FOMC announcement). The lagged BBK index is the standardized value of the index of Brave et al. (2019) from one-month before the date t FOMC meeting. "Jobs number" and "Jobs number, surprise" are the level and surprise component of month t's release of the change in non-farm payrolls (in units of 100,000 jobs). Expectations of that release are from Bloomberg. The sample consists of all regularly scheduled FOMC meetings between May 1999 and October 2019, excluding July 2008–July 2009, that occur after the first week of the month. The row $\overline{\lvert\text{LHS}\rvert}$ is the average absolute GDP forecast revision over the regression sample (0.3 percentage point).

regression, the coefficient on the interest-rate surprise remains nearly unchanged. This is a hint that $\widehat{y}_t$ is not a useful variable for controlling for information effects—if it were, the coefficient on the interest-rate surprise should change. This also highlights (thinking about the omitted-variables-bias equation) that $\widehat{y}_t$ and $\widehat{i}_t$ are only weakly positively correlated (the correlation coefficient is 0.09, with a t-statistic of 1.3). The positive correlation suggests that information effects are still present in high frequency. The fact that the correlation is fairly weak also suggests a useful feature of bringing $\widehat{y}_t$ to this identification problem: It contains independent variation. This is less true with stock prices (Jarociński and Karadi, 2020); the high-frequency S&P 500 return around Fed announcements with $\widehat{i}_t$ is -0.36, with a t-statistic of -3.9.[21]

[21] One reaction to this negative correlation would be to ask "doesn't this refute the presence information effects?" It does not. This negative correlation is only useful for rejecting that there are *only* information effects (or that they

In the last column, I show that $\widehat{y}_t$ continues to be positively correlated with lower-frequency expectations when controlling for recent macroeconomic news. Bauer and Swanson (2020) estimate similar specifications using low-frequency measures of GDP forecast revisions. They find that recent news predicts both interest rate-surprises and GDP forecast revisions, and substantially mitigates the positive correlation between the two. They interpret these results as a suggestion that recent macroeconomic news induces a spurious positive correlation between interest-rate surprises and GDP forecast revisions, which argues against the presence of information effects. Following Bauer and Swanson, I include the most recent level of the non-farm payrolls release, the surprise component of that release, the 1-month stock return, and the lagged index of Brave et al. (2019). The coefficient drops somewhat, but remains marginally statistically significant. For transparency, this coefficient drops nearly to zero if the 13-week return of the S&P 500 return is included instead (the exact variable used by Bauer and Swanson). However, in Appendix H, I show that the 13-week return is the most potent return horizon over which stock returns "predict" news around monetary policy announcements, which calls into question the robustness of the conclusions of Bauer and Swanson

A benefit to performing this "automated narrative analysis" is that I can read newspaper articles to corroborate the largest shocks. For example, the largest negative Fed-induced change in expectations about GDP came in March 2004, when Grep Ip's post-meeting WSJ article state:

> The slightly less upbeat tone of the statement drove long-term bond yields down sharply. The Fed said risks to economic growth remain "roughly equal" while the risk of an "unwelcome fall in inflation" was "almost equal" to that of a rise in inflation.

In contrast, pre-meeting WSJ coverage states that yields were not "likely to fall much further, given countercurrents of strong economic growth" and that the Fed was not expected to "tweak significantly the language in the accompanying policy statement."

2.3 Predictability

When proposing an estimate of a macroeconomic shock, it is important to examine the extent to which that shock is predictable in some way. I test the predictability of my newly constructed estimate of output forecast revisions, $\widehat{y}_t$, and a traditional measure of monetary shocks, $\widehat{i}_t$ (described in Section (2.4)) in Table 2. Specifically, I regress each measure on a slew of controls. In the first two columns, I include two lags of the left-hand-side variables, the most recent surprise component of three macroeconomic news releases, and the two most recent changes in non-farm payrolls (using real-time data). All of these variables are observable to markets and Fed watchers before the meeting announcement. None of these right-hand-side variables have a statistically significant coefficient. Jointly the variables are statistically insignificant. The R^2 of both regressions are very small. The finding of minimal autocorrelation is consistent with the findings of Miranda-Agrippino and Ricco (2021).

are the strongest determinant of stock prices). Put differently, without information effects, this correlation would be much more negative.

Table 2: Predictability of High-frequency Output ($\widehat{y}_t$) and Interest-rate ($\widehat{i}_t$) Forecast Revisions

	$\widehat{y}_t$	$\widehat{i}_t$	$\widehat{y}_t$	$\widehat{i}_t$	$\widehat{y}_t$	$\widehat{i}_t$
Lagged LHS	0.035	-0.049	-0.008	-0.105	0.031	-0.055
	(0.082)	(0.093)	(0.080)	(0.087)	(0.083)	(0.092)
Twice-Lagged LHS	0.086	-0.093	0.071	-0.125	0.098	-0.099
	(0.095)	(0.096)	(0.097)	(0.099)	(0.098)	(0.098)
BB Surprise: CPI	0.052	-0.094	0.052	-0.077	0.052	-0.093
	(0.055)	(0.384)	(0.053)	(0.310)	(0.055)	(0.372)
BB Surprise: GDP	0.072	-0.160	0.058	-0.257	0.070	-0.179
	(0.041)	(0.285)	(0.039)	(0.266)	(0.042)	(0.282)
BB Surprise: Jobs	-0.473	-5.870	0.077	-4.321	-0.401	-5.577
	(1.574)	(11.600)	(1.703)	(10.676)	(1.576)	(11.603)
First ΔNFPR	1.905	9.573	0.358	2.702	1.794	8.943
	(1.372)	(7.180)	(1.570)	(6.528)	(1.399)	(6.975)
Second ΔNFPR	-1.181	-0.554	-1.323	-0.013	-1.102	0.056
	(1.136)	(5.452)	(1.197)	(5.321)	(1.142)	(5.365)
1-quarter stock ret.			0.016	0.188		
			(0.014)	(0.099)		
BBK Index$_{m(t)-1}$			0.377	1.259		
			(0.158)	(0.972)		
6-week stock ret.					0.018	0.111
					(0.018)	(0.136)
Constant	-0.039	-2.007	0.086	-1.806	-0.049	-2.106
	(0.114)	(0.737)	(0.136)	(0.729)	(0.113)	(0.744)
Observations	159	159	159	159	159	159
Adjusted R^2	0.032	0.012	0.108	0.078	0.032	0.012
F	1.277	1.056	2.453	2.008	1.273	1.146
p(F)	0.265	0.395	0.012	0.042	0.262	0.336

This table presents results from regressing the HF measure of output and interest-rate forecast revisions around FOMC announcements on several variables. The output measure ($\widehat{y}_t$) is from newspapers and standardized over the regression sample. The interest-rate measure ($\widehat{i}_t$) is the change in the 4-quarter Eurodollar future contract, in basis points. In each column two lags (at FOMC meeting frequency) of the left-hand-side variables are included. The three "BB surprise" variables are the difference between expected and announced annualized CPI inflation (basis points), annualized GDP growth (basis points), and monthly change in non-farm payrolls (100,000 jobs). I compute these forecast errors (from Bloomberg) from the most recent (pre-FOMC) announcement of each variable. The two NFPR variables are the most recent (pre-FOMC) and second most recent real-time releases of non-farm payrolls (100,000 jobs). The 1-quarter stock return is the 13-week percent return in the S&P 500 (ending one day before the FOMC announcement). The lagged BBK index is the standardized value of the index of Brave et al. (2019) from 1-month before the date t FOMC meeting. The sample consists of all regularly scheduled FOMC meetings between May 1999 and October 2019. Robust standard errors are in parentheses. The row labeled "F" is the F-statistic from a joint test that all coefficients (except the constant) are nonzero, and p(F) gives the associated p-value.

In the remaining columns I add additional variables considered by Bauer and Swanson.[22] Following the authors, in the third and fourth columns I add the 13-week return of the S&P 500 and the 1-month lagged index of Brave et al. (2019) ("the BBK index"). Here the evidence of predictability is stronger, with joint tests of significance smaller than 5% for both variables. In the fifth and sixth columns I examine these findings further. First, noting that the BBK index is not observable, I exclude it from the regression.[23] Second, recalling, from the discussion of table 1, the sensitivity of "predictability" regressions to the stock-market horizon, I change the return horizon to a six-week horizon. This approximates the inter-meeting stock return, since eight scheduled meetings are held per calendar year. The results in these columns again suggest a lack of predictability, again calling into question the robustness of the conclusion Bauer and Swanson, that both GDP and interest-rate forecast revisions are predictable from publicly available information.

2.4 Numerical Data

To measure the surprise component of monetary policy decisions, I use tick-level data on Federal Funds and Eurodollar futures. I use the 30-minute change in the current-month Federal Funds future rate (in the spirit of Kuttner (2001)) and the change in the price of the 4-quarter-ahead Eurodollar futures contract.[24] These series are well-known in empirical monetary economics and were extended by Acosta and Saia (In progress) through 2019 using tick-level data purchased from the CME group. The authors followed Nakamura and Steinsson (2018) exactly in the construction of these variables. The 4-quarter Eurodollar is highly (0.95) correlated with the shock used by Nakamura and Steinsson and the path factor of Gürkaynak et al. (2005), with a correlation coefficient of 0.91. Here, I measure high-frequency S&P 500 prices using the exchange traded fund SPY, and low-frequency prices from Yahoo finance.

I take data on macroeconomic expectations from Blue Chip Economic Indicators. In some regression specifications I control for the surprise components of macroeconomic news releases. I collect expectations and the announced values of these variables from Bloomberg—the data appendix of Acosta and Saia (In progress) describes these data in detail. When studying macroeconomic effects I use the shadow Federal Funds rate of Wu and Xia (2016) and other macroeconomic aggregates from FRED.

Unless noted otherwise, my sample consists of all regularly scheduled FOMC meetings between May 1999 and October 2019. May 1999 is the first meeting after which the Fed started to regularly release post-meeting statements. October 2019 is when my high-frequency data end. I exclude observations from July 2008 through 2009 because of the asset-pricing anomalies over this period discussed by Nakamura and Steinsson (2018).

[22] The authors also include recent changes in non-farm payrolls. Using real-time data, I find that those variables are not predictive of monetary shocks.

[23] The index did not exist for most of my sample. Even when it does exist, the index for month $m(t)$ (used in the regression) is not until month $m(t)+2$. Thus, when considering an FOMC meeting at time t, the value of the index for month $m(t)-1$ is not available to markets.

[24] For those familiar with the variable names in HF papers, this is ED4.

3 Shock Identification

In Section 3.1, I describe an illustrative version of the model I use to estimate my shocks, which models how expectations are revised in light of Fed announcements. This allows me to formalize the identification challenge: Forecast revisions about observable variables (e.g., interest rates) are linear combinations of forecast revisions about structural shocks (e.g., monetary and information shocks). This implies that no high-frequency change in expectations or prices will identify macroeconomic shocks *unless* additional assumptions are made. As a consequence, my new measure of high-frequency output expectations is neither an estimate of "information effects" nor can it be used as an instrument for information effects.

This is not to say that all hope is lost; only that more work has to be done. The second purpose of my illustrative model is to show that, together, my new measure, traditional estimates of monetary shocks, and historical information about the nature of the Fed's communication practices can be used to separately identify monetary and information shocks using the heteroskedasticity-based identification assumptions of Rigobon (2003). I pair the illustrative model with my data to show the intuition behind this procedure in Section 3.2. Finally, in Section 3.3 I describe the formal model and identification assumptions, then present the results of the estimation in Section 3.4.

3.1 The Identification Challenge in an Illustrative Model

My illustrative model describes how markets (and Fed watchers more generally) form macroeconomic expectations, which allows me to study how these expectations are revised in response to the Fed's policy announcements. Fed watchers use a simple macroeconomic model for this task. While this model, in its simplicity, serves its expositional purpose, when appropriate I discuss why the intuition gleaned from the simple model is consistent with expanded elaborations of the model.

Economy The macroeconomic model used by forecasters is described by output y_τ and the interest rate i_τ, which are related by the following two equations:

$$i_\tau = \phi y_\tau + \varepsilon_\tau \tag{5a}$$

$$y_\tau = -\gamma i_\tau + \eta_\tau. \tag{5b}$$

The exogenous shocks are a monetary shock, ε_τ, and a macroeconomic fundamental, η_τ. These serially uncorrelated shocks are normally distributed with zero mean and variances σ_ε^2 and σ_η^2. For the purposes of building intuition, I assume that the coefficients ϕ and γ are positive. I do not use these sign restrictions in my identification procedure. In appendix A, I show that this model nests the textbook three-equation New Keynesian model (Galí, 2015), in which I substitute out inflation and remove technology shocks (thus, the information shock here is a demand shock in the context of the New Keynesian model). As such, I refer to the equations as the forecaster's Taylor Rule (Taylor, 1993) and IS equation, respectively. The elimination of supply shocks is consistent with my findings below; whereby monetary announcements reveal very little about supply shocks. Appendix

C highlights the fact that the intuition from this static bivariate model extends to linear forward- and backward-looking general equilibrium models with noisy information as in, e.g., Blanchard et al. (2013).

Fed Announcements At discrete points in time, indexed by t, the Fed makes policy announcements. Empirically, I measure expectations about the observable variables, y_τ and i_τ, shortly before and shortly after each policy announcement. Denote the time of those pre- and post-announcement measurements by $\underline{t}$ and $\bar{t}$, respectively, and expectations taken at each time t about x at time τ as $x_{\tau,t} = \mathbb{E}_t[x_\tau]$, for $x \in \{y, i, \eta, \varepsilon\}$. Equations (5a) and (5b) show how these expectations of observable variables relate to perceptions of structural shocks before (left column) and after (right column) each announcement

$$i_{\tau,\underline{t}} = \phi y_{\tau,\underline{t}} + \varepsilon_{\tau,\underline{t}} \qquad i_{\tau,\bar{t}} = \phi y_{\tau,\bar{t}} + \varepsilon_{\tau,\bar{t}} \tag{6a}$$

$$y_{\tau,\underline{t}} = -\gamma i_{\tau,\underline{t}} + \eta_{\tau,\underline{t}} \qquad y_{\tau,\bar{t}} = -\gamma i_{\tau,\bar{t}} + \eta_{\tau,\bar{t}}. \tag{6b}$$

Denote high-frequency changes in expectations about $x_\tau \in \{y_\tau, i_\tau, \eta_\tau, \varepsilon_\tau\}$ by $\widehat{x}_{\tau,t} = \mathbb{E}_{\bar{t}}[x_\tau] - \mathbb{E}_{\underline{t}}[x_\tau]$. Then taking the difference between the left and right columns of equation (6a) reveals

$$\widehat{i}_{\tau,t} = \phi \widehat{y}_{\tau,t} + \widehat{\varepsilon}_{\tau,t}. \tag{7a}$$

Similarly, the IS equation in expectation-revision space is

$$\widehat{y}_{\tau,t} = -\gamma \widehat{i}_{\tau,t} + \widehat{\eta}_{\tau,t}. \tag{7b}$$

Combining equations (7a) and (7b) allows the observable variables to be expressed as linear combinations of the forecast revisions about perceived structural shocks:

$$\widehat{i}_{\tau,t} = \left(\frac{\phi}{1+\gamma\phi}\right)\widehat{\eta}_{\tau,t} + \left(\frac{1}{1+\gamma\phi}\right)\widehat{\varepsilon}_{\tau,t} \tag{8a}$$

$$\widehat{y}_{\tau,t} = \left(\frac{1}{1+\gamma\phi}\right)\widehat{\eta}_{\tau,t} - \left(\frac{\gamma\phi}{1+\gamma\phi}\right)\widehat{\varepsilon}_{\tau,t}. \tag{8b}$$

I refer to $\widehat{\varepsilon}_{\tau,t}$ as a perceived monetary shock, and $\widehat{\eta}_{\tau,t}$ as an information shock or information effects.

The Identification Problem Equations (8a) and (8b) are the crux of the identification challenge. First, equation (8a) shows that in general, traditional HF measures of monetary shocks ($\widehat{i}_{\tau,t}$) are contaminated by information effects ($\widehat{\eta}_{\tau,t}$). There is only a special case—when the Fed has no independent knowledge about the state of the economy (so that $\widehat{\eta}_{\tau,t} = 0,\ \forall t$)—in which traditional measures identify the (perceived) monetary policy shock $\widehat{\varepsilon}_{\tau,t}$. A testable implication of the model is that if traditional estimates are not contaminated by information effects, then the correlation between interest-rate and output expectations must be negative. Empirical evidence refutes this implication. Campbell et al. (2012) and Nakamura and Steinsson (2018) find a positive correlation

between output and interest-rate forecast revisions using low-frequency data, which I confirm using high-frequency data.

Second, equation (8b) shows that an estimate of output forecast revisions can neither be used to control for information effects nor to instrument for information effects. On the first point, in Appendix D I show that the residual from a regression of interest rate expectations on output expectations (i.e. a "cleaned" interest rate surprise) only identifies monetary shocks in the case that output expectations do not respond to monetary shocks which. This, in turn, is only the case with full monetary neutrality (i.e. $\phi = 0$) or there are no information effects in the first place.[25] (Interestingly, in the absence of information effects, any asset-price change in the window around a policy announcement identifies a monetary policy shock!) Next, output forecast revisions are not a valid instrument for information effects, because they are not exogenous with respect to monetary policy shocks.

These observations intuitively show the necessity for additional identification assumptions to be made in order to separately identify monetary and information shocks. Let $(\widehat{\sigma}_\nu, \widehat{\sigma}_\eta, \widehat{\rho}_{\eta,\nu})$ be the variance of $\widehat{\eta}_{\tau,t}$, the variance of $\widehat{\varepsilon}_{\tau,t}$, and their covariance, respectively. Formally, the identification problem is that the model contains five parameters $(\phi, \gamma, \widehat{\sigma}_\nu, \widehat{\sigma}_\eta, \widehat{\rho}_{\eta,\nu})$, but the data only provide three empirical moments: the variances of $\widehat{y}_t$ and $\widehat{i}_t$ and their covariance. I next discuss my identifying assumptions in the context of my data and illustrative model. These assumptions will allow me to estimate the linear mapping (a function of γ and ϕ) between the (unobserved) perceived structural shocks ($\widehat{\eta}_{\tau,t}$ and $\widehat{\varepsilon}_{\tau,t}$) and the (observed) forecast revisions ($\widehat{y}_{\tau,t}$ and $\widehat{i}_{\tau,t}$). I will then invert that linear mapping to recover the shocks from the revisions.

To this point I have remained agnostic as to the exact mathematical specification of the information communicated by Fed policy announcements. Such a formal structure is not necessary for my identification approach. However, I do make two assumptions about the information conveyed and how markets make use of that information. First, I assume that the information revealed allows forecast revisions about the structural shocks to be uncorrelated, i.e., $\widehat{\rho}_{\eta,\nu} = 0$. This need not be the case in general signal-extraction problems: If the Fed's communications are not sufficiently detailed in their discussion of each type of shock (economic fundamentals, η_t vs. monetary policy, ε_t), then markets will generally have to use their prior knowledge to parse the independent information revealed about each shock. Empirically, this assumption can be tested with overidentification tests. Theoretically, in Appendix B.1, I discuss a particular information structure under which this assumption would be valid. The second assumption is that markets make larger forecast revisions about a variable when signals about that variable are clearer—this is an implication, for example, of Bayes' rule. These assumptions, plus some historical knowledge described in the next section, allow me to identify the model's parameters.

[25] In the language of Angrist and Pischke (2008), $\widehat{y}_t$ is a "bad control" for information effects, since it is affected by monetary shocks.

3.2 Regimes and Intuition for the Identification Assumptions

I bring historical knowledge about the nature of the Fed's communication practices to the identification problem. Intuitively, this knowledge provides the fourth moment needed to identify the model's four parameters. To be concrete, I rely on the episode discussed by Lunsford (2020): the August 2003 introduction of interest-rate forward guidance. To give context, in June 2003 the FOMC had lowered the Federal Funds rate to 1%. In August 2003, the post-meeting statement declared that this "policy accommodation can be maintained for a considerable period." This was the first instance of explicit forward guidance regarding the path of interest rates, and was used for reasons similar to what prompted its major re-emergence in 2008. In 2003, 1% was essentially seen as the effective lower bound on nominal interest rates. This episode thus gives me a natural place to split my sample. My first regime, $\mathcal{R}_1$, consists of all meetings from the start of my sample (May 1999) through June 2003. The second regime, $\mathcal{R}_2$, extends from August 2003 through the end of 2006.[26] These regimes are shown by the two left-most shaded regions of Figure 2.

To see how this episode can provide useful variation for identification, consider the change in the behavior of forecast revisions following the announcement, shown in Figure 4. Focusing on panel A, the behavior of forecast revisions clearly changed between the two regimes. In the first regime (left-most plot), GDP and interest-rate expectations tended to be revised in the same direction following Fed announcements. This statistically significant positive correlation disappears—and becomes even slightly negative—following the regime change.

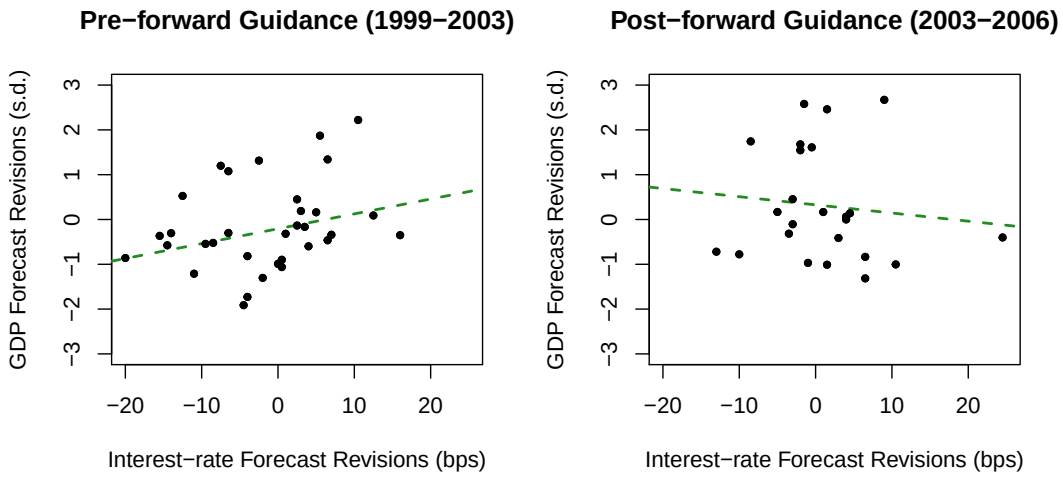
The change in the correlation of GDP and interest-rate forecast revisions is predicted by the illustrative model, given the nature of this episode. In the context of that model, perceived information shocks induce a positive correlation between GDP and interest-rate forecast revisions, while monetary shocks induce a negative correlation. This is evident from equations (8a) and (8b). Since both shocks are potentially present in both regimes, however, the correlation between the two series depends on which shock is larger, as determined by their variances. Lunsford (2020) highlights the fact that before 2003 the Fed's post-meeting statement primarily discussed the economic outlook and risks to that outlook. In the context of the illustrative model of Section 3.1, this can be formalized as clearer signals about η_t. Recalling the assumption that forecast revisions are larger for variables with clearer signals, this suggests a relatively high value of $\widehat{\sigma}_\eta$—i.e., information effects are relatively larger (than monetary shocks) in the first regime. Therefore, the model predicts that forecast revisions to both variables will primarily reflect information shocks, which is borne out by the positive correlation in the top-left panel of Figure 4. Similarly, a relatively clearer signal about future interest rates implies a larger role for perceived future monetary shocks. This role becomes larger after 2003, as shown by the slightly negative correlation.

The change in correlation thus suggests that the variance of the underlying perceived shocks changed in 2003. One assumption allows this observation to be used to identify the model's struc-

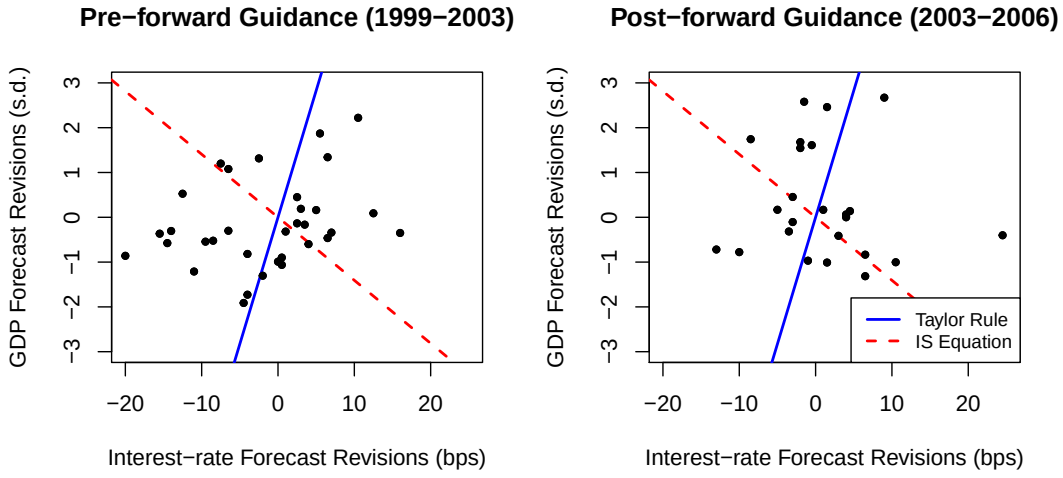
[26] These are slight elongations of the sample period studied by Lunsford (2020). I elongate the sample for statistical precision—my point estimates are nearly identical using Lunsford's exact regimes, but mildly less-precisely estimated. I will also note that I have estimated an over-identified system using the rest of my sample as a third regime, but again the point estimates are very similar.

Figure 4: The Identifcation Assumptions in Pictures

Panel A: Correlations in the two Regimes



Panel B: Estimated Data-Generating Process



Panels A and B contain scatter plots of high-frequency GDP and interest-rate forecast revisions around Fed announcements ($\widehat{y}_t$ and $\widehat{i}_t$ in the text). The left panel shows observations between May 1999 and June 2003; the right panel features observations between August 2003 and December 2006. The green line in panel A is the unconditional line of best fit, estimated separately in both regimes. The lines in panel B are those estimated to generate the data in both regimes (in a maximum-likelihood sense), estimated using data from both regimes using the identification procedure described in Section 3.4.

tural parameters. Namely, by assuming that the slopes in the forecasting model (ϕ and γ) remain unchanged across the policy regions, Rigobon (2003) shows that only a unique pair of slopes can generate the data in both regimes (in a maximum-likelihood sense). I display that unique set of slopes in panel B of Figure 4 (delaying a discussion of their estimation to the next section). To see that these slopes are those most likely to generate the data, consider tilting the solid blue line down in the bottom-left panel so that it better fits the points in the first regime. This would necessarily cause it to fit the points in the top of the bottom-right graph.

Intuitively, then, the first sample primarily contributes to the estimation of the information shock, which, from the perspective of macroeconomic theory, should induce a positive correlation between these variables (which can be seen by shifting the red dashed line, the IS equation, and observing that all points are assumed to lie at the intersection of the two lines). Having learned about the information shock, the low to negative correlation in the second regime allows the monetary shock to be identified, since that regime exhibits variation in the two variables that is orthogonal to the information shock. In the extreme case in which the first regime contains only information shocks and the latter only monetary shocks, this intuition would be exact and the green lines in panel A would provide structural slope coefficients. However, even with nonzero variances of each shock, the system can still be identified as long as the two shocks behave differently enough across the two regimes.

3.3 Formal Identification Assumptions

Model I estimate the following model:

$$\widehat{y}_{\tau,t} = m_{11}\xi_{1,t} + m_{12}\xi_{2,t} \qquad \xi_{1,t} \sim N(0, \sigma^2_{1,t}) \tag{9a}$$

$$\widehat{i}_{\tau,t} = m_{21}\xi_{1,t} + m_{22}\xi_{2,t} \qquad \xi_{2,t} \sim N(0, \sigma^2_{2,t}), \tag{9b}$$

where $i_{\tau,t}$ is the 30-minute change in four-quarter ahead Eurodollar futures (section 2.4) and $y_{\tau,t}$ is the high-frequency text-based proxy for output expectations (section 2.1). These observable variables are posited to be linear combinations of two independent Gaussian shocks, $\xi_{1,t}$ and $\xi_{2,t}$. I assume that the shocks exhibit heteroskedasticity of the form

$$\sigma_{1,t} = \begin{cases} 1 & t \in \mathcal{R}_1 \\ \sigma_1 & t \in \mathcal{R}_2 \end{cases} \qquad \sigma_{2,t} = \begin{cases} 1 & t \in \mathcal{R}_1 \\ \sigma_2 & t \in \mathcal{R}_2, \end{cases}$$

where, as described in section 3.2, the regimes, $\mathcal{R}_1$ and $\mathcal{R}_2$, consist of all regularly-scheduled FOMC meetings running from May 1999 through June 2003, and August 2003 through December 2006, respectively. The normalization of the variances of the shocks is without loss of generality—their levels are not identified, only their relative levels between the two regimes. The model can be

generalized to n variables and shocks, so for later reference I express the model in matrix form:

$$\underbrace{\widehat{\mathbf{x}}_{\tau,t}}_{n\times 1} = \underbrace{\mathbf{M}}_{n\times n}\underbrace{\boldsymbol{\xi}_t}_{n\times 1} \qquad\qquad \boldsymbol{\xi}_{\tau,t} \sim N(\mathbf{0}, \underbrace{\Sigma_t}_{n\times n}). \tag{10}$$

where

$$\mathbf{\Sigma}_t = \begin{cases} \Sigma_{(1)} & t \in \mathcal{R}_1 \\ \Sigma_{(2)} & t \in \mathcal{R}_2. \end{cases}$$

with the normalization $\Sigma_{(1)} = \mathbf{I}_n$.

The formal model just described is similar to the illustrative model used to generate intuition—this can be seen by comparing equations (9a) and (9b) to equations (8b) and (8a). Here, I posit that changes in expectations about observable variables are driven by two independent structural shocks. In the illustrative model, the shocks were given names and the coefficients were signed based on macroeconomic theory. Here, the coefficients are unrestricted and the shocks will based on the historical episode that I use to estimate them.

Assumption 1: Heteroskedasticity The assumption of heteroskedasticity is crucial. Without it, the model would not be identified. Only three empirical moments would be available from the two Gaussian observable variables—$\text{var}\left(\widehat{y}_{\tau,t}\right)$, $\text{var}\left(\widehat{i}_{\tau,t}\right)$, and $\text{cov}\left(\widehat{y}_{\tau,t}, \widehat{i}_{\tau,t}\right)$—but there would be four parameters to estimate, m_{11}, m_{12}, m_{21}, and m_{22}.[27] Rigobon (2003) developed a solution to this identification problem. The assumption of heteroskedasticity allows me to estimate those three empirical moments in both regimes, for a total of six moments. The model has six parameters to estimate: $m_{11}, m_{12}, m_{21}, m_{22}, \sigma_1$, and σ_2. Therefore, the system is just-identified, with moment conditions given by

$$\text{var}_1\left(\widehat{y}_{\tau,t}\right) = m_{11}^2 + m_{12}^2 \tag{11a}$$

$$\text{var}_1\left(\widehat{i}_{\tau,t}\right) = m_{21}^2 + m_{22}^2 \tag{11b}$$

$$\text{cov}_1\left(\widehat{y}_{\tau,t}, \widehat{i}_{\tau,t}\right) = m_{11}m_{21} + m_{21}m_{22} \tag{11c}$$

$$\text{var}_2\left(\widehat{y}_{\tau,t}\right) = m_{11}^2\sigma_1^2 + m_{12}^2\sigma_2^2 \tag{11d}$$

$$\text{var}_2\left(\widehat{i}_{\tau,t}\right) = m_{21}^2\sigma_1^2 + m_{22}^2\sigma_2^2 \tag{11e}$$

$$\text{cov}_2\left(\widehat{y}_{\tau,t}, \widehat{i}_{\tau,t}\right) = m_{11}m_{21}\sigma_1^2 + m_{21}m_{22}\sigma_2^2, \tag{11f}$$

where the notation $\text{var}_1\left(\widehat{y}_{\tau,t}\right)$ is the variance of $y_{\tau,t}$ in regime 1, $\text{cov}_2\left(\widehat{y}_{\tau,t}, \widehat{i}_{\tau,t}\right)$ is the covariance of the observable variables in regime 2, and so on. Those moments—on the left-hand-side of each equation—can be estimated directly in the data. These six equations therefore have six unknowns—the parameters on the right-hand-side of each equation. The solution can be expressed analytically (see Rigobon (2003)). As solutions to quadratic equations, however, the associated expressions are

[27] Without loss, the variances of the shocks could be normalized to unity.

not particularly enlightening.[28] An important condition for the system to be identified is that

$$\sigma_1 \neq \sigma_2. \tag{12}$$

In words, this rank condition states that the relative variances of each shock (between regimes) must be different. Suppose instead that the relative variances were equal, defining $c \equiv \sigma_1 = \sigma_2$. In this case, empirically, the six equations 11 would reduce to only three linearly independent equations, with (11a), (11b), and (11c) just c-scaled versions of (11d), (11e), and (11f). That parameter, c, could be (over-) identified, but not the m_{ij}.

Assumption 2: Stable Effects The second crucial assumption I make is that the m_{ij} coefficients remain unchanged across the two regimes. Without this assumption, four additional parameters (10 parameters total) would need to be estimated with only six equations. Empirically, this assumption can be tested by adding an additional regime and running over-identification (J) tests—I present results of such a test alongside my estimates. To assess its theoretical validity, I consider the assumption within the context of linear rational (RE) expectations models. In full-information RE models there is no role for information effects, so I focus my discussion on linear RE models with imperfectly-informed agents.

I consider first the class of models discussed by Blanchard et al. (2013). In these models, all agents are imperfectly informed about the economy's state variables. Models of this class take the form

$$\mathbf{A}\,\mathbf{x}_t + \mathbf{B}\,\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{C}\boldsymbol{\xi}_{t|t} = \mathbf{0} \tag{13}$$

where $\mathbf{x}_t$ is a vector of observable macroeconomic variables, $z_{t|\tau}$ denotes the mathematical expectation of z_t given information at time τ, and $\boldsymbol{\xi}_t$ is a vector of mutually independent structural shocks that evolves according to $\boldsymbol{\xi}_t = \mathbf{D}\boldsymbol{\xi}_{t-1} + \mathbf{F}\boldsymbol{\zeta}_t$, where $\boldsymbol{\zeta}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\zeta)$.[29] At the beginning of each period t, but before making decisions about $\mathbf{x}_t$, all agents receive the same noisy signal of the structural shock of the form $\mathbf{s}_t = \mathbf{G}\boldsymbol{\xi}_t + \mathbf{H}\boldsymbol{\nu}_t$, where $\boldsymbol{\nu}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\nu)$. Agents use this signal to form expectations about $\boldsymbol{\xi}_t$ using the Kalman filter (i.e., agents have RE). Agents are thus imperfectly informed in a symmetric way—they all share the same information set when making decisions.

In appendix C, I show that the model in equation (13) admits a solution of the form

$$\mathbf{x}_t = \mathbf{J}\boldsymbol{\xi}_{t|t} \tag{14}$$

along with a law of motion for perceived shocks, $\boldsymbol{\xi}_{t|t}$, where $\mathbf{J}$ depends on neither Σ_ζ not Σ_ν.

[28] In the (over-identified) case that one of the shocks is assumed to maintain the same variance between both regimes (Wright, 2012; Arai, 2017; Nakamura and Steinsson, 2018; Hébert and Schreger, 2017), an intuitive expression the "slope" coefficients m_{ij} emerges as essentially a change in OLS coefficients between the two regimes.

[29] The assumption that the shocks follow a VAR(1) is not restrictive—any finite VARMA can be re-written as a VAR(1). Adding the vector of lagged endogenous variables to equation (13) (i.e., $\mathbf{x}_{t-1}$) slightly changes the language, but not results, of this discussion. Specifically, rather than forecast revisions of endogenous variables being related linearly to revisions of structural shocks, it's forecast revisions of *surprises* (i.e., reduced form residuals) of endogenous variables that are related linearly to revisions of structural shocks.

Equations 14 and the law of motion for ξ_t reveal the result: the mapping between forecast revisions of observable variables and forecast revisions of structural shocks is fixed and linear:

$$\mathbb{E}_t[\mathbf{x}_{t+k}] - \mathbb{E}_{t-1}[\mathbf{x}_{t+k}] = \mathbf{J}\mathbf{D}^k[\boldsymbol{\xi}_{t|t} - \boldsymbol{\xi}_{t|t-1}].$$

This relationship resembles the relationship between reduced-form and structural shocks in the parlance of structural VARs, another macroeocnomic model in which the assumption of constant m_{ij} is valid.

The key ingredient for a linear RE model to feature a constant mapping between observable variables and structural shocks is that the model's structural equations can be solved independently of agents' Kalman filtering problem. The result would continue to hold, therefore, in models in which agents also observe the endogenous variables, $\mathbf{x}_t$. An important case in which this independence breaks is in linear RE models with dispersed information (as in the "islands" model of Lucas (1972)). Within the context of these models, it then becomes an empirical question: To what extent does the dependence of m_{ij} on the variance of the structural shocks affect the ability of the shocks to be recovered using the heteroskedasticity-based assumptions? In appendix I.1, I ask this question within the context of the model of Fed "signalling effects" posited and estimated by Melosi (2017), which features agents with dispersed information and realistic modeling of Fed signaling. Specifically, I simulate the "introduction of forward guidance" experiment within the model. I find that the shocks identified by heteroskedasticity uncover their structural counterparts remarkably well. In appendix I.2, I also discuss why the assumption of dispersed, rather than common, knowledge eliminates the aforementioned independence in the context of a simple asset pricing model that follows Townsend (1983).

Assumption 3: Naming In the illustrative model, the two structural shocks, $\widehat{\eta}_{\tau,t}$ and $\varepsilon_{\tau,t}$, were given names consistent with macroeconomic theory. The model of equations (9a) and (9b) instead is driven by two structural shocks, $\xi_{1,t}$ and $\xi_{2,t}$, with no natural names. Members of the high-frequency identification literature typically impose identification restrictions on the m_{ij} coefficients.[30] In this paper, I take the view that the m_{ij} are exactly the objects that require unrestricted estimation, since they ultimately determine the nature of the estimated shocks. Instead, I place assumptions on the size of the underlying shocks, motivated by the historical evidence. Here, the illustrative model and the historical evidence provide a natural solution: The ξ shock that I call the "information shock" is the shock whose variance becomes relatively smaller in the second regime. Formally, recalling that $\sigma_i = \text{var}_2(\xi_{i,t})/\text{var}_1(\xi_{i,t})$, I name the information shock as

$$\widehat{\xi}^I_t \equiv \text{information shock}_t \equiv \arg\min_{\xi_{i,t}}[\sigma_i]. \tag{15a}$$

[30] For example, Kuttner (2001) and Gürkaynak et al. (2005) impose zero restrictions on the m_{ij}, while (Jarociński and Karadi, 2020) impose sign restrictions.

Table 3: Heteroskedasticity-Based Estimates

Panel A: Structural Impact Matrix, $\mathbf{M}$

	Response of High-Frequency Forecast Revisions	
	Eurodollars (1Y)	GDP
Monetary Path Shock	-0.39	0.78
	[-1.17, 0.13]	[0.55, 1.28]
Information Shock	0.99	0.56
	[0.71, 1.43]	[0.02, 1.29]

Panel B: Relative Variance in Regime 2, $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$

Monetary Path Shock	Information Shock
1.37	0.63
[0.69, 2.00]	[0.38, 1.58]

This table shows the estimates of the system in equation (11). Panel A shows the structural impact matrix, which shows the effect that each structural shock (in the rows) has on forecast revisions made around FOMC announcements (in the columns). The shocks are normalized to have unit variance in the first regime (2000–2003). Panel B shows the variance of each shock in regime 2 relative to its variance in regime 1. 90% equal-tailed (studentized) confidence intervals are produced using 999 boostrap replications, where bootstrap samples are stratified by regime.

Conversely,

$$\widehat{\xi}_t^M \equiv \text{monetary path shock}_t \equiv \arg\max_{\xi_{i,t}}[\sigma_i]. \tag{15b}$$

I therefore name my shocks based on the size of their relative variance across the regimes. It then becomes an empirical question—not an assumption—whether these shocks have theoretically-consistent effects on high-frequency forecast revisions. This can be determined through the estimates of the m_{ij}. I carry out the estimation in the next section.

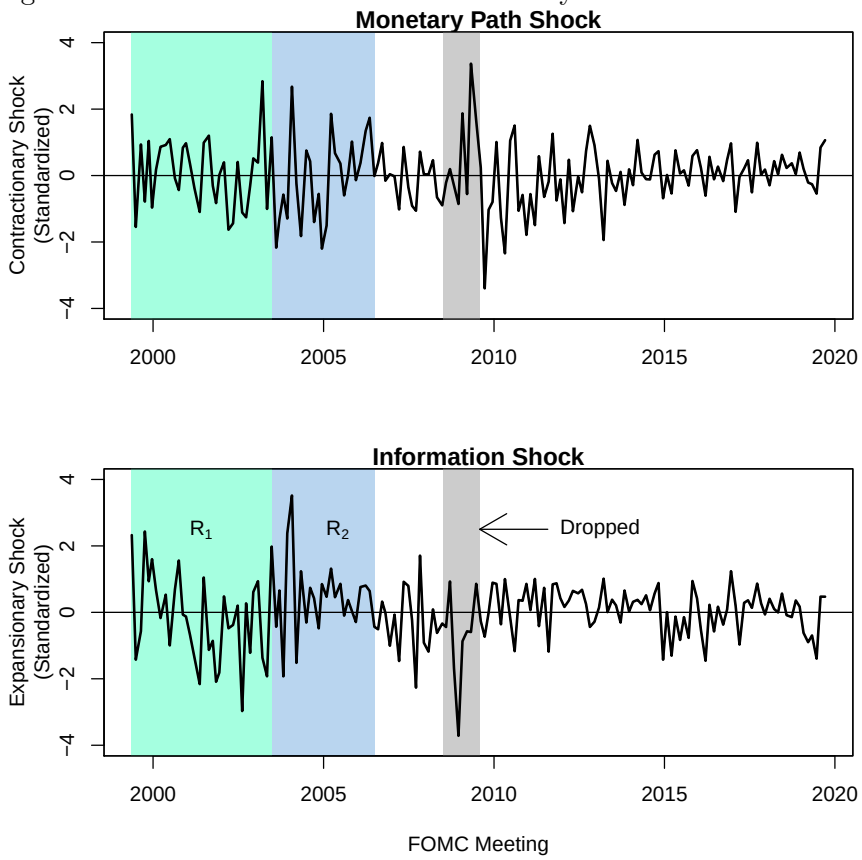
3.4 Estimation

I estimate the parameters using GMM and calculate bootstrapped standard errors and confidence intervals. In appendix F I describe the bootstrap procedure in detail. The procedure is mostly standard—nearly identical to that of Hébert and Schreger (2017)—though I design a method to handle the fact the columns of $\mathbf{M}$ are only identified up to order and sign.

I present the estimates in table 3. Panel A contains the estimated structural impact matrix, $\mathbf{M}$, and panel B shows the estimates of $\Omega_2\Omega_1^{-1}$—the variance of each of the structural shocks in $\mathcal{R}_2$ relative to $\mathcal{R}_1$.[31] Following the discussion of equation (15), I name the "information shock" based

[31] Recall that I am allowed the normalization of two parameters without loss—here I have normalized Ω_1 to be a 2×2 identity matrix.

Figure 5: Time Series of Estimated Monetary and Information Shocks



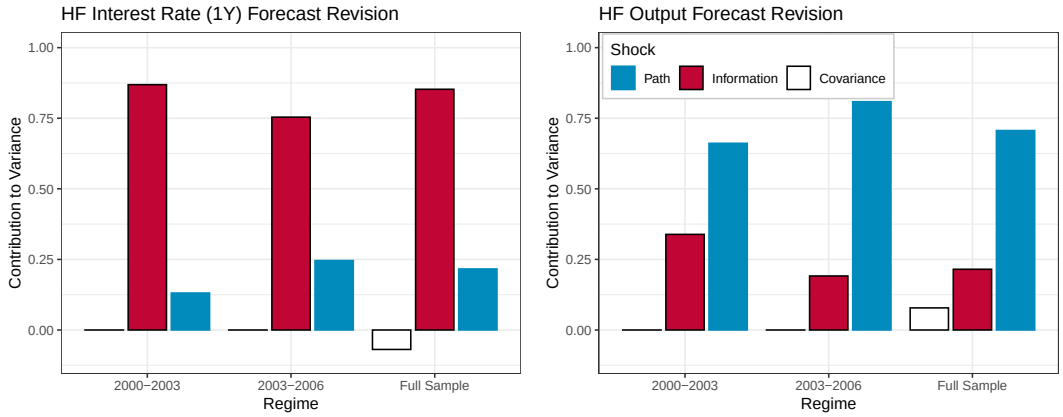
The top panel shows the identified monetary policy path shock described in section 3.4, and the bottom panel shows the identified information shock. Both shocks are standardized to have unit standard deviation over the plotted sample. The first two shaded areas present the regimes used for identification, described in section 3.2. The third, gray, shaded area highlights observations that are dropped for most analysis, discussed in section 2.4.

on the fact that it is relatively smaller in the second regime than in the second. This is consistent with the notion that the Fed's statement in the first regime focused exclusively on describing the state of the economy. Because I use interest rate expectations at a one-year horizon, I call the second shock a "monetary path shock"—an exogenous shock to the expected path of interest rates. Again, once the Fed starts to communicate explicitly about future interest rates, markets are able to make more-informed (larger) updates about expected future monetary shocks. Formally, I can reject the null hypotheses that the variance of the monetary shock is relatively smaller than the variance of the demand shock in the second regime with 90 percent confidence.[32]

Having named the shocks based on their relative variances, the estimates of $\mathbf{M}$ serve as the first check on the shocks. A monetary policy shock that lowers interest rates by 0.37 standard deviations *increases* GDP expectations by 0.76 standard deviations. Conversely, a monetary policy

[32]Let δ be the difference between the variance of the monetary and information shock. The 90% confidence interval for this test is (x, ∞), where x is $\widehat{\delta} - \text{s.e.}(\widehat{\delta})G_n^{*-1}(0.9)$, where $\widehat{\delta}$ is the point estimate of δ, $\text{s.e.}(\widehat{\delta})$ is the standard deviation across bootstrap replications, and G_n^* is the bootstrap distribution of t-statistics $(\widehat{\delta}^* - \widehat{\delta})/\text{s.e.}(\widehat{\theta}^*)$. Note that estimation $\widehat{\theta}^*$, requires an inner bootstrap (boostrapping the bootstrap) for which I also use 999 replications. Empirically, $x = 0.002$.

Figure 6: Variance Decomposition of HF Forecast Revisions



This figure shows which portion of the variance of HF interest rate forecast revisions, $\widehat{i}_t$ (left panel), and HF output forecast revisions, $\widehat{y}_t$ (right panel), can be explained by the monetary $\widehat{\varepsilon}_t$ and information $\widehat{\eta}_t$ shocks. Recalling from equations (9a) and (9b) that $\widehat{i}_t = \phi_{i\varepsilon}\widehat{\varepsilon}_t + \phi_{i\eta}\widehat{\eta}_t$ and $\widehat{y}_t = \phi_{y\varepsilon}\widehat{\varepsilon}_t + \phi_{y\eta}\widehat{\eta}_t$, the red outlined boxes for $j \in \{i, y\}$ show $\phi^2_{j\eta}\text{var}(\widehat{\eta}_t)/\text{var}(\widehat{j}_t)$; the blue non-outlined boxes show $\phi^2_{j\varepsilon}\text{var}(\widehat{\varepsilon}_t)/\text{var}(\widehat{j}_t)$; and the white-outlined boxes show the remaining covariance term (which is not restricted to be null in the full sample).

shock that lowers rates by 1 standard deviation *decreases* GDP expectations by 0.57 standard deviations. These impact responses are consistent standard New-Keynesian macroeconomic theory and the illustrative model. The time series of the estimated shocks, given by $\mathbf{M}^{-1}\widehat{\mathbf{x}}_t$, are shown in figure 5.[33]

To have a better sense of how the identified monetary and information shocks relate to interest rate and GDP forecast revisions, figure 6 shows the variance decomposition of the forecast revisions. The calculation during the period 1999–2006 follows immediately from the estimated system—the variance of forecast revisions is the weighted (by $\mathbf{M}$) sum of the variances of the underlying shocks. Both during the estimation period and in the full sample, it is apparent that the bulk of interest rate forecast revisions are driven by the information shock. This is not a surprising finding given the puzzling estimated effects of interest rate surprises, but these estimates allow me to quantify the extent of the contamination problem. Roughly about 70% of interest rate surprises (recall, these are traditional HF estimates of monetary policy shocks) are made up of information. This contamination was less severe during 2003–2006 than it was in 2000–2003, consistent with the illustrative model.[34]

[33] Here I make a less demanding form of an "invertibility" assumption, as discussed in the structural-VAR literature (see, e.g., Chahrour and Jurado (2021) or Fernández-Villaverde et al. (2007)). The illustrative model and discussion surrounding equation (14) provided the justification for its validity here. Note that my context—recovering *perceived shocks* from expectations data—is different from the VAR context, in which one seeks to identify structural shocks from reduced-form VAR residuals.

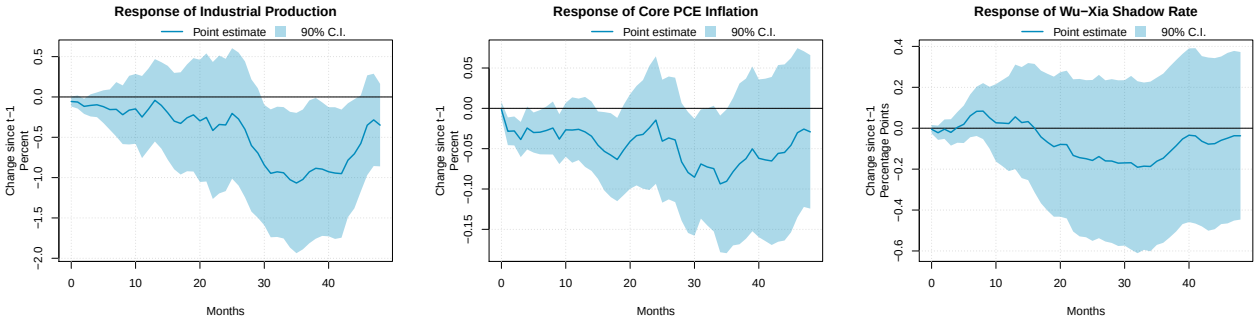
[34] To see this, notice that the variance of interest rate surprises is:

$$\text{var}\left(\widehat{i}_t\right) = \underbrace{\left(\frac{\phi}{1+\gamma\phi}\right)^2 \nu_\eta}_{\text{Var. due to info.}} + \underbrace{\left(\frac{1}{1+\gamma\phi}\right)^2 \nu_\varepsilon}_{\text{Var. due to MP}}.$$

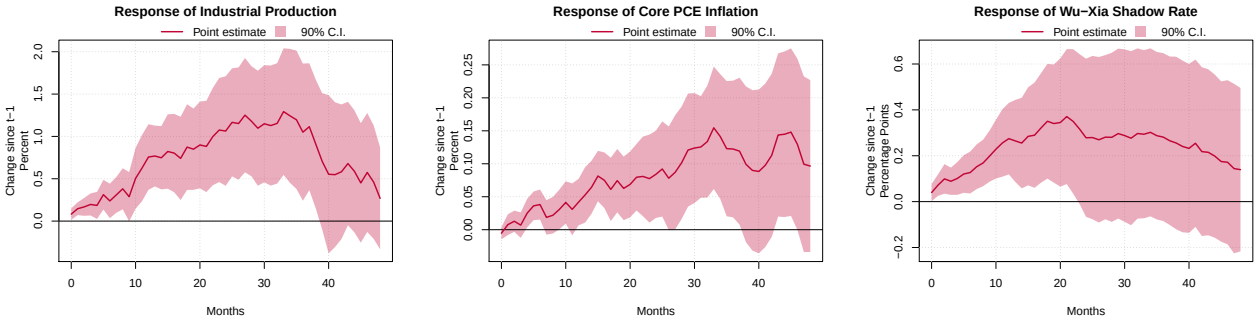
The relative contribution of information effects to interest rate forecast revisions is therefore $\phi^2\left(\frac{\nu_\eta}{\nu_\varepsilon}\right)$. In the model, a

Figure 7: Baseline Macroeconomic Effects

Panel A: Responses to a Contractionary Monetary Path Shock $\left(\beta_k^M\right)$



Panel B: Responses to an Expansionary Information Shock $\left(\beta_k^I\right)$



This figure plots the estimates of equation (16). Moving from left to right, the left-hand-side variables are k-period differences in 100 times the log of industrial production (FRED mnemonic INDPRO), 100 times the log of the core PCE prices index (PCEPILFE), and the shadow Federal Funds rate of Wu and Xia (2016) (in percent). The right-hand-side contains 12 lags of these variables, as well as the estimated monetary path shock and information shock, standardized to have unit standard deviation over the full sample period, which runs from May 1999 to October 2019, excluding July 2008–July 2009. Both shocks are also scaled to increase the 1-day change in the one-year treasury yield on impact—thus the monetary shock is contractionary, and the information shock is expansionary. I exclude months with no FOMC meeting. Confidence intervals are calculated using heteroskedasticity and autocorrelation-consistent asymptotic standard errors with the automatic lag selection method of Newey and West (1994), as implemented by Zeileis et al. (2020) and Zeileis (2004).

The "full sample" column in each panel serves as a check on the assumption that the shocks are uncorrelated. During the 1999–2006 period the shocks are forced to be orthogonal. During the remainder of the sample, this orthogonality condition is not imposed, so a covariance term enters the expression for the variance of forecast revisions. The covariance between the shocks contributes negligibly to the variance of observed forecast revisions, suggesting that the orthogonality restriction was warranted.

regime in which post-meeting statements focus on the Fed's outlook for the state of the economy is akin to increasing the precision of the information signal, i.e., lower $\sigma_{n,\eta}$, thus higher ν_η, which increases the relative contribution of information effects to interest rate surprises. Thus, a prediction of the model is that relative to 2003–2006, interest rate surprises during 2000–2003 should be more contaminated by information effects.

4 Effects of Monetary Policy and Information

I now turn to estimating the effects of monetary policy and Fed information. In Section 4.1, I present my baseline estimated effects on macroeconomic outcomes and expectations. My specifications are purposefully simple. I use local projections (Jordà, 2005) to estimate macroeconomic effects and OLS to estimate the response of macroeconomic expectations. Since part of my objective in this paper is to provide a credible and portable set of shocks, it is important that my shocks have sensible effects without depending on particular controls. In Section 4.2, I compare my estimated effects with those found using other measures in the literature. Along these lines, in section 4.3 I put my estimated effects in policy-relevant terms by estimating instrumental-variables versions of my baseline specification.

4.1 Baseline Results

Effects on Macroeconomic Outcomes In order to estimate the effects of my shocks on macroeconomic outcomes, I estimate local projections of the form

$$y_{t+k} - y_{t-1} = \alpha_k + \beta_k^M \xi_t^M + \beta_k^I \xi_t^I + \sum_{\ell=1}^{L} \Gamma'_{\ell,k} \Delta \mathbf{y}_{t-\ell} + \xi_{k,t}, \tag{16}$$

where ξ_t^M and ξ_t^I are the estimated monetary and information shocks.[35] The left-hand-side variable, y_t, is either the log of the industrial production index, the log of the core PCE price index, or the shadow Federal Funds rate of Wu and Xia (2016). The vector $\mathbf{y}_t$ contains the three y_t variables, of which I include $m = 12$ lags. I estimate the equation at monthly frequency, using the sample described in Section 2.4. The shocks are normalized to have unit standard deviation over the regression sample, with a positive effect on interest rates (the high-frequency change in the 4-quarter Eurodollar future). Thus, the coefficients of interest are β_k^M—the effects of a contractionary monetary policy shock—and β_k^I, the effects of an expansionary information shock.

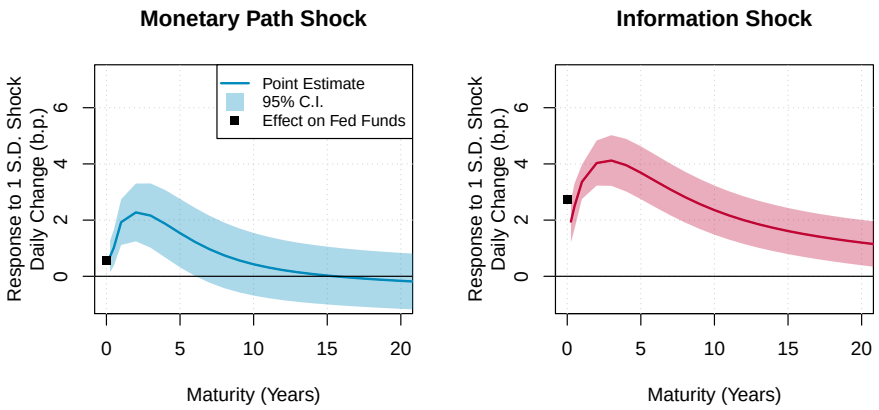
Figure 7 contains the estimated coefficients. The top panel shows the effect of a contractionary monetary policy shock. Encouragingly, both industrial production and inflation decrease. The delayed peak responses of these variables—at about 36 months—is consistent with the notion that monetary policy works with "long and variable lags."

The response of the shadow rate to a monetary policy shock is less clear-cut. In high frequency, a monetary shock leads to an upward revision of interest-rate expectations. As shown in Panel A Figure 8, nominal yields several years into the term structure also increase in a 1-day window around the policy announcement in response to a monetary shock. The effects at a monthly frequency are imprecisely estimated and fluctuate around zero. Figure 9 investigates whether this finding is robust across alternative measures of nominal rates. Specifically, I re-estimate equation (16) using longer-term interest rates in place of the shadow rate. While the estimates continue to be imprecise, the

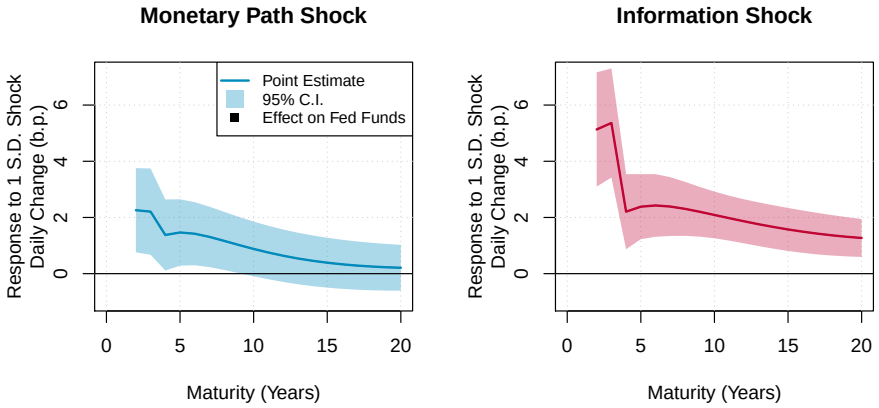
[35] Note, as described in Section 3.4, that the shocks are essentially uncorrelated over the sample, so their joint inclusion does not affect the point estimates, only the precision.

Figure 8: One-Day Changes in Yields

Panel A: Nominal Yields



Panel B: Real Yields



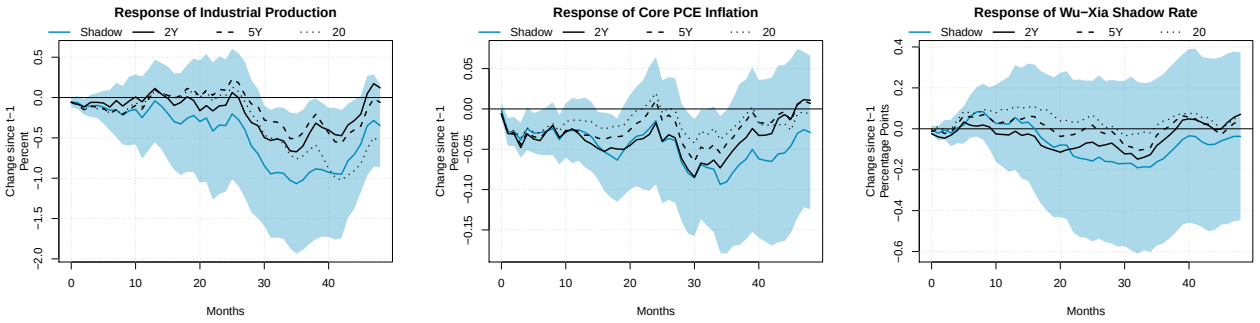
This graphs show the estimated slope coefficient from univariate regressions of 1-day changes in nominal (Panel A) and real (Panel B) interest rates of maturity m (Δi_t^m) on the monetary path shock (left panel) and the information shock (right panel), using my baseline sample. Regarding nominal rates, for maturities of at least 1 year, I use data from Gürkaynak et al. (2007) The Federal Funds, three- and six-month yields are from FRED, with mnemonics EFFR, DGS3MO, and DGS6MO, respectively. The zero lower bound leads to enormous standard errors in the Fed Funds rate regression, so I omit them. I take real rates from Gürkaynak et al. (2010). Confidence intervals are computed using robust standard errors.

negative effect on the interest rate diminishes as increasingly longer-term interest rates are used.

Panel B of Figure 7 shows the effects of an expansionary information shock. While the responses are more immediate than responses to a monetary shock, the peak effects on industrial production and prices are similar in magnitude (in absolute value) to the peak effects of a monetary shock. This suggests that the Fed plays an important role not only in setting interest rates, but also in the provision of macroeconomic information.[36]

[36]These estimates should prove useful for informing models of Fed information effects, e.g., that of Lepetyuk et al. (2021).

Figure 9: Effects of Monetary Shocks: Using Longer-term Interest Rates



These graphs show the response of industrial production, core PCE inflation, and interest rates to the identified monetary shock. These are estimates of equation (16), except that I replace the shadow rate of Wu and Xia (2016) with either the two-year treasury yield (solid line, FRED mnemonic DGS2), the five-year treasury yield (dashed line, DGS5), and the twenty-year treasury yield (dotted line, DGS20. See the note to table 7 for details regarding units, sample period, and confidence intervals.

The relative responses of industrial production and inflation also give a hint as to the nature of the shocks about which information is revealed. The positive comovement in response to an information shock suggests that the information primarily concerns demand-type factors. In Section 5.2 I provide further high-frequency evidence in support of this conjecture.

In Appendix G, I examine the robustness of the estimates of equation (16). I first add months with no FOMC meetings, setting the shocks in those months to zero. I next stop the estimation in 2015 in order to (i) obtain a consistent sample across all horizons and (ii) drop COVID observations. Similarly, I include a specification that retains Great Recession observations. I test the functional form of the controls by first controlling for no lags, then for 24 lags, and then add a linear trend. Finally, I use a 10-word word list $\mathcal{Y}$ for GDP to construct the shocks. Across these specifications, the responses are largely similar to the estimates shown in Figure 7.

Effects on Macroeconomic Expectations I close this section by studying the effects of monetary and information shocks on macroeconomic expectations. Specifically, I regress 1-month changes in macroeconomic expectations from the Blue Chip survey on the two shocks. The regression takes the form

$$\mathbb{X}_{t|t+1} - \mathbb{X}_{t|t} = \alpha + \beta^M \xi_t^M + \beta^I \xi_t^I + e_t, \tag{17}$$

where $\mathbb{X}_{t|t}$ is constructed analogous to equation (3), except now I study the responses of real GDP growth (as above), CPI inflation, and the unemployment rate. I present the results in Table 4. Again, the shocks are scaled to increase interest rates in high frequency—thus, the first row represents the effects of a contractionary monetary policy shock and the second represents the effects of an expansionary information shock.

The signs of the responses of all variables are consistent with both the macroeconomic effects estimated above, and standard macroeconomic theory. The information shock has larger (and more precisely estimated) effects on real variables, which highlights another way in which the Fed plays an important role in the provision of macroeconomic information. The information shock also

Table 4: Shocks and Macroeconomic Expectations

	GDP Growth	CPI Inflation	Unemp. Rate
Path Shock	-0.0126	-0.0152	0.0206
	(-1.36)	(-2.53)	(2.14)
Info. Shock	0.0389	0.0138	-0.0256
	(3.99)	(1.43)	(-3.69)
Observations	131	131	131
R^2	0.116	0.0614	0.105
$\overline{\lvert\text{LHS}\rvert}$	0.0775	0.0532	0.0625

This table shows estimates of equation (17). The left-hand-side is 1-month forecast revision of the variables listed atop the columns over the next three forecasting horizons, presented in equation (3), from the Blue Chip survey in percentage points. The estimated monetary path shock and information shock are standardized to have unit standard deviation over the full sample period. The sample consists of all regularly scheduled FOMC meetings between May 1999 and October 2019, excluding July 2008–July 2009, that occur after the first week of the month. The row $\overline{|\text{LHS}|}$ is the average absolute Blue Chip revision in the relevant column over the regression sample.

induces a positive correlation between these variables and inflation, which again suggests that this information concerns demand-type factors.

4.2 Comparison with Previous Estimates

While my paper is not the first that seeks to provide a credible measure of monetary policy shocks, my shocks differ in that they robustly exhibit theoretically consistent macroeconomic effects. In this section, I compare the estimated effects of alternative measures of monetary shocks found in the literature. I estimate the effects using the specification in equation (16), in which I replace $\left(\beta_k^M \xi_t^M + \beta_k^I \xi_t^I\right)$ with a single measure of monetary policy shocks.

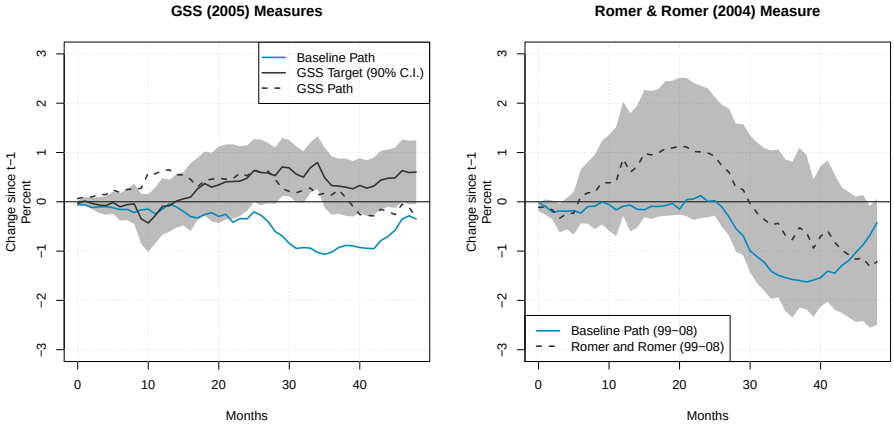
4.2.1 Traditional Measures

I first study the estimated effects of monetary policy using shocks that are pervasive in empirical macroeconomics. Specifically, I study the high-frequency shocks of Gürkaynak et al. (2005).[37] Those authors decompose high-frequency changes in interest rate expectations for maturities out to 1 year into a "target" and "path" factor. The shocks are identified and named with the assumption that the path factor has no effect on the current-meeting Federal Funds rate surprise. I updated these measures using the tick data described in Section 2.4. I also study the shocks of Romer and Romer (2004). Those shocks are constructed as the component of the change in the Federal Funds rate that cannot be predicted from the Fed staff's ("Greenbook") forecasts. I updated this series through 2012 (though I stop my estimation at the start of the 2008–2015 zero lower bound period).

I focus on the effects on industrial production and present the results in Figure 10. All shocks

[37] Other well-known high-frequency shocks are those of Kuttner (2001), Nakamura and Steinsson (2018), and Gertler and Karadi (2015). Those all differ by the maturity of the underlying interest rates. The shocks of Gürkaynak et al. span the maturity spectrum used by those other papers.

Figure 10: Response of Industrial Production to Traditional Contractionary MP Shocks



These graphs show the response of industrial production to various measures of monetary shocks. The blue line in the left panel is identical to that in Figure 7. All other lines are computed by estimating equation (16), replacing $\left(\beta_k^M \xi_t^M + \beta_k^I \xi_t^I\right)$ with βshock_t, where shock_t is either the target or path factor of Gürkaynak et al. (2005) (left panel), or the shock of Romer and Romer (2004) (right panel), described in the text. In the right panel I re-estimate the effects of my shocks with an abbreviated sample (1999–2008) for comparison with the effects of the shock of Romer and Romer. All shocks are standardized to have unit standard deviation over the sample period, and increase the 1-day change in the 1-year treasury on impact—they are all thus 1-standard deviation contractionary shocks. See the note to Table 7 for details regarding units, sample period, and standard errors.

are scaled to increase the 1-day change in the 1-year treasury yield, so the results can be interpreted as the responses to estimates of contractionary monetary policy shocks. The left panel shows the type of puzzling evidence suggested in the introduction: A contractionary high-frequency shock leads to increases in industrial production. This is the case whether regardless of the horizon of the interest-rate surprise (i.e., the target or path factor).

In the right panel of figure 10, I show the responses using the measure constructed by Romer and Romer (2004). I also re-estimate the effects of my monetary shock using the same (1999–2008) sample. Romer and Romer's responses are fairly noisy and fluctuate around zero, though all statistically significant estimates have the theoretically consistent sign. This last finding suggests that somehow adjusting for the Fed's private information is useful for estimating theoretically-consistent effects of monetary policy.

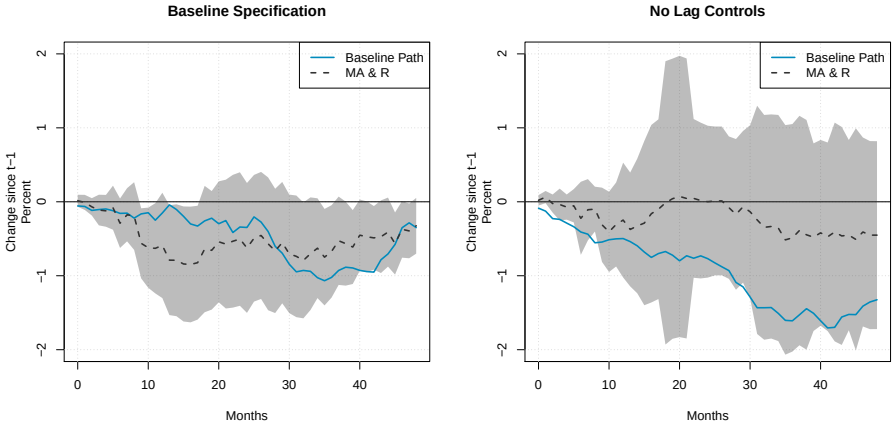
4.2.2 New Measures of Monetary Shocks

The findings in the previous study also inspired other authors to estimate "information free" measures of monetary policy shocks. The predominant method of point-identifying information free measures was developed by Miranda-Agrippino and Ricco (2021), who orthogonalize interest rate surprises to the Fed staff's forecasts—essentially a combination of the approach of Kuttner (2001) and Romer and Romer (2004).[38] I downloaded the shocks of Miranda-Agrippino and Ricco (2021) directly from the first author's website.

The left panel of figure 11 contains the result. The estimated effects of Miranda-Agrippino

[38] Hoesch et al. (2020) and Zhang (2020) take a similar approach.

Figure 11: Response of Industrial Production to "Information Free" Contractionary MP Shocks



These graphs show the response of industrial production to the shock of Miranda-Agrippino and Ricco (2021). The left panel includes 12 lags of observable variables as controls, while in the right panel no lags are included. See Figure 11 for other details.

and Ricco (2021) exhibit theoretically consistent effects, with much-more precise estimates than those of Romer and Romer (2004). In fact, the responses to the Miranda-Agrippino and Ricco shock are similar to my baseline estimates. At a first glance both seemingly-plausible identification approaches produce similar results.

Important differences arise when the specification is substantially simplified. In the right panel of Figure 11, I remove any controls from the regression in equation (16). The estimated effects using Miranda-Agrippino and Ricco's shock diminish substantially in magnitude and precision. The fact that my monetary shock, in contrast, provides similar estimates using this simpler specification speaks to its portability to other contexts. Put differently, my shocks can be used directly to estimate the effects of monetary policy, without requiring additional controls.[39]

4.3 Comparison with Previous Estimates: Magnitudes

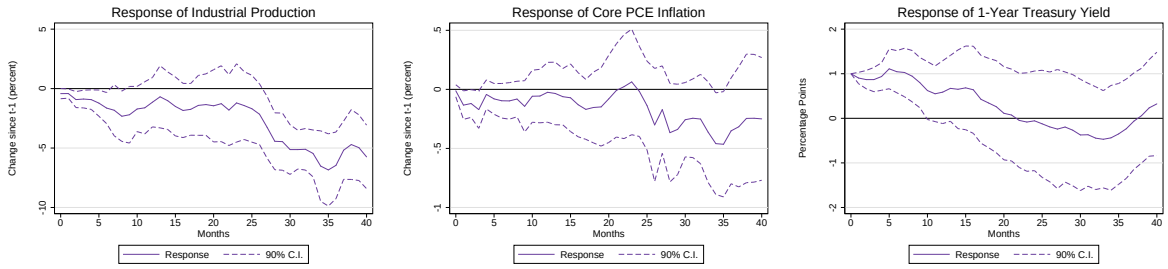
The estimated shocks have no interpretable magnitudes because they are linear combinations of two variables with different units.[40] The estimates presented thus far, therefore, speak to average observed influence of the Fed on macroeconomic aggregates, but do not put those effects in policy-relevant terms. In this section I follow the work of Gertler and Karadi (2015) and Ramey and Zubairy (2018) and estimate instrumental variables versions of equation (16). This specification,

[39]This result is unsurprising given the construction of Miranda-Agrippino and Ricco (2021)'s measure. It is important to note that information effects arise because the Fed and public have different information sets. Therefore, to "remove" information effects, it is imperative to control for the *difference* between the Fed's and public's information. By only controlling for the Fed's information, Miranda-Agrippino and Ricco's shock "leaves behind" the public's information. Thus, (at best) their measure is contaminated by (classical) measurement error that attenuates their estimated effects. The lags of observable variables in the expanded local projection likely span the public's information set, which thus allows for the unbiased estimates shown in the left panel.

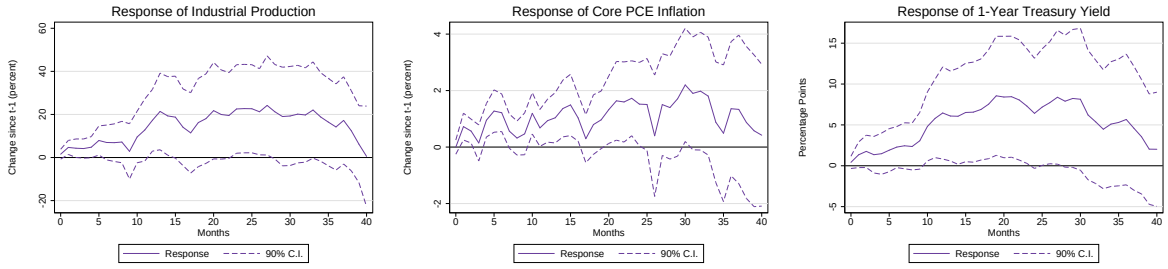
[40]One of those variables—the GDP forecast revision proxy—no interpretable units in the first place.

Figure 12: LP-IV Estimates

Panel A: Effects of Monetary Policy (1-year Treasury Yield)



Panel B: Effects of Information Provision (GDP Forecast Revision)



This figure plots the estimates of β_k from equation (18), where p_t and its instrument are described in the text. Moving from left to right, the left-hand-side variables are k-period differences in 100 times the log of industrial production (FRED mnemonic INDPRO), the 100 times the log of the core PCE prices index (PCEPILFE), and the 1-year Treasury yield in period $t+k$ in percent (DGS1, though the top-right panel shows the result of the level of 1-year Treasury yield, to make the effect on the instrument clear). Panel A has six lags of the first-differences in these three variables, panel B has twelve. The sample runs from May 1999 to October 2019, excluding July 2008–July 2009. I exclude months with no FOMC meeting. Confidence intervals are calculated using heteroskedasticity and autocorrelation-consistent asymptotic standard errors with the automatic lag selection method of Newey and West (1994), as implemented by Baum et al. (2010).

which I refer to as "LP-IV," takes the form

$$y_{t+k} - y_{t-1} = \alpha_k + \beta_k p_t + \sum_{\ell=1}^{L} \Gamma'_{\ell,k} \Delta \mathbf{y}_{t-\ell} + \xi_{k,t}, \tag{18}$$

where p_t is the policy-relevant variable of interest, for which I will instrument with the relevant policy shock. For estimating the effects of monetary policy, interest rates are a natural candidate, so I follow Gertler and Karadi (2015) and set p_t to the level of the 1-year Treasury yield. The first-stage F-statistic here is 14.[41] The policy-relevant variable or estimating the effects of information provision are less obvious, but following the framework of the rest of the paper, I posit GDP forecast revisions (from the Blue Chip) as a reasonable measure. Here the first-stage F-statistic is 24—an unsurprising result, given the results in Table 4, whose "GDP Growth" variable is the p_t used here.

[41] For the effects of monetary policy, I set $L = 6$. For the effects of information, I set $L = 12$. For reasons still under investigation, there are a handful of outlier observations that emerge when setting $L \approx 12$ in the monetary regression. These outliers cause a few of the point estimates, and their standard errors, to become orders of magnitude larger than those shown here. With outliers removed, the results are nearly identical to those shown here.

Figure 12 contains the results. To use the terminology of Coibion (2012)—who compares and reconciles the estimated effects of monetary policy from several prominent papers—the effects of monetary policy are big. The peak responses are similar in magnitude (about a 5% drop to a 1 p.p. increase in nominal rates) and timing (delayed by less than a year) as those estimated by Romer and Romer (2004). As Coibion points out, these estimates are on the upper end of estimates found in the literature. They are substantially larger than the effects found in the existing literature that uses high-frequency shocks in an LP-IV setting (Gertler and Karadi, 2015; Miranda-Agrippino and Ricco, 2021). Figure 8 helps to makes sense of these big effects. There, these increases in nominal interest rates are seen to translate nearly completely to increases in real interest rates.

While little existing empirical work serves as a reference point for understanding the effects of information provision, figure 12 suggests that the effects are big as well. A 1 p.p. increase in GDP expectations leads to a nearly 20% increase in industrial production. Care should be taken in interpreting these results, however. The average absolute GDP forecast revision is less than 0.1 percentage point, so these estimates are largely extrapolating outside of historical experience. The response of interest rates can help put the results in perspective—those increase by about 4p.p., about five-times less than industrial production.[42] In this sense, the results are also similar in magnitude to the effects of monetary policy.[43]

5 Expanding the Picture: Short-term Rates and Supply Factors

As discussed in the introduction, expanding the set of observable variables can give the econometrician greater insight into the nature of the shocks perceived by markets. In this section, I expand the set of variables along two dimensions. In Section 5.1, I study the role of short-term interest-rate surprises. The regimes I consider allow me to identify an additional shock—a monetary target shock—and study its behavior over time. In summary, following the intuition suggested by Ramey (2016), monetary policy shocks of this kind largely disappeared following the introduction of forward guidance. I also extend my model to include dynamic shocks in order to show how, simultaneously, path shocks can become larger while target shocks become smaller.

In Section 5.2, I present an estimate of high-frequency inflation forecast revisions. Around Fed announcements, inflation forecast revisions are unconditionally positively correlated with output forecast revisions. This is also true conditional on the estimated information and monetary path shocks. Both pieces of evidence suggest that markets do not learn about supply-type factors from the Fed. I also discuss why the regimes I use do not allow me to separately identify a "demand information" shock and a "supply information" shock. Intuitively, I can only identify shocks whose variance is induced to change across regimes; there is little evidence to support the notion that the introduction of forward guidance altered the relative precision of supply vs. demand signals.

[42] I do not use nominal interest rates as my policy-relevant instrument for information effects for primarily for practical reasons—I have no first stage.

[43] These large estimated effects of the information shocks intuitively suggest that there is some "real" information content of these shocks—the canonical interpretation of information effects—and not a measure of macroeconomic disagreement, as in Sastry (2021).

Table 5: Identifying a Target Shock

Panel A: Structural Impact Matrix, $\mathbf{M}$

	Response of High-frequency Forecast Revisions		
	Fed Funds	Eurodollars (1Y)	GDP
Monetary Path Shock	-0.03	0.39	-0.68
	[-0.11, 0.01]	[-0.02, 1.13]	[-1.14, -0.45]
Information Shock	0.04	0.94	0.64
	[-0.04, 0.15]	[0.67, 1.46]	[0.22, 1.25]
Monetary Target Shock	1.33	0.32	-0.17
	[1.01, 2.08]	[-0.06, 0.65]	[-0.44, 0.03]

Panel B: Relative Variance in Regime 2, $\mathbf{\Sigma}_2\mathbf{\Sigma}_1^{-1}$

Monetary Path Shock	Information Shock	Monetary Target Shock
1.65	0.65	0.05
[0.79, 2.42]	[0.36, 1.57]	[0.02, 0.27]

This table shows the estimates of the system in equation (11), where the vector of observable variables has been expanded to include the HF surprise component of the Federal Funds rate announcement, as in equation (19). Otherwise, everything is identical to Table 3, whose note can be referenced for more detail.

5.1 Short- vs. Long-term Interest Rates

In their original incarnation, high-frequency monetary shocks were the surprise component of the Fed's current-meeting interest-rate decision, in which the surprise was relative to expectations formed shortly before the decision was announced. This is the measure proposed by Kuttner (2001). Later work by Gürkaynak et al. (2005), cognizant of the 2003 introduction of forward guidance, sought to separately identify shocks to the interest-rate target from shocks to the interest-rate path. In this section I revisit this distinction, adding to my system the surprise in the current-meeting Fed Funds rate described in Section 2.4.

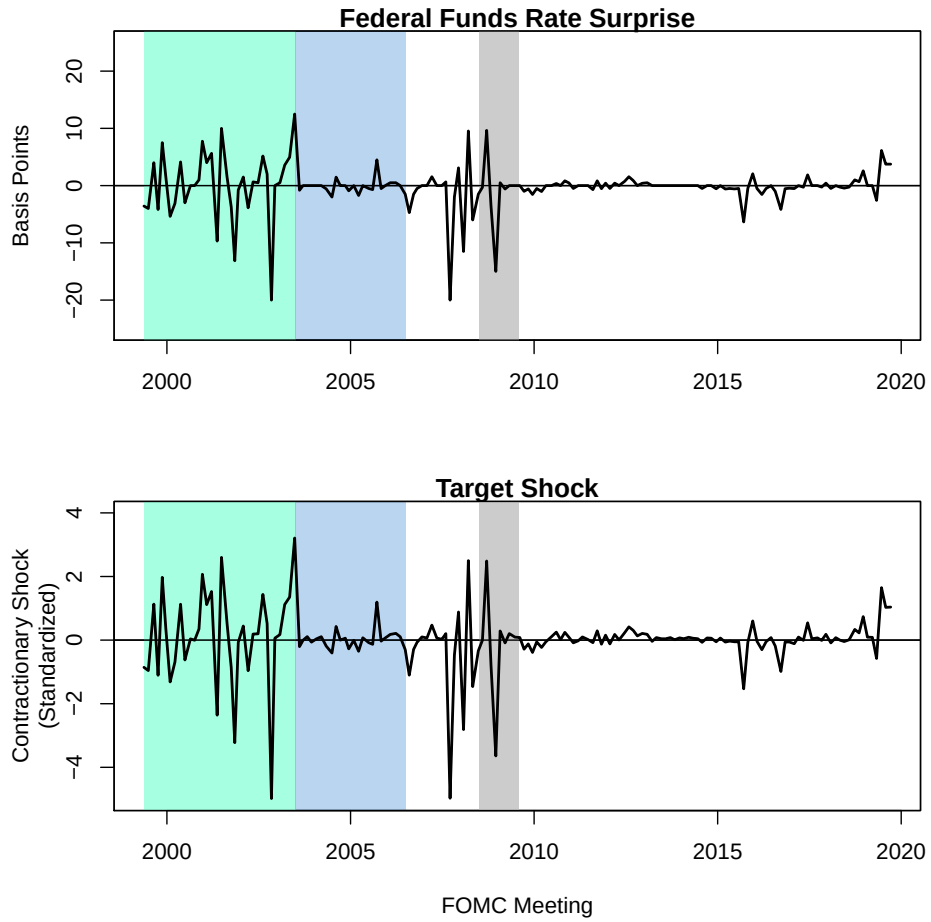
In Table 5, I show the estimates from estimating the moment conditions from equation (11), where now my observable variables are given by

$$\widehat{\mathbf{x}}_t \equiv \left[\widehat{\text{FF}}_t \quad \widehat{i}_t \quad \widehat{y}_t\right]', \tag{19}$$

where $\widehat{\text{FF}}_t$ is the 30-minute change in current-month Federal Funds rate expectations (the Kuttner measure), $\widehat{i}_t$ is the 30-minute change in the 4-quarter-ahead Eurodollar futures price, and $\widehat{y}_t$ is the high-frequency text-based proxy for output expectations (Section 2.1). The regimes are unchanged.

Before discussing the identified "monetary target shock," it is useful to examine the top

Figure 13: Federal Funds Rate Surprise and the Target Shock



The top panel shows the 30-minute change in the price of the current-month Federal Funds rate future, scaled by the percentage of days left in the month (see, e.g., Kuttner (2001)), in basis points. The bottom panel shows the identified monetary policy target shock described in Section 5.1 standardized to have unit standard deviation over the plotted sample. The first two shaded areas present the regimes used for identification, described in Section 3.2. The third, gray shaded area, highlights observations that are dropped for most analyses, discussed in Section 2.4.

two-by-two quadrant of panel A. There, notice that the relationship between the path shock and information shock with GDP and interest rate forecast revisions is nearly unchanged from Table 3.[44] This highlights the robustness of the identified information and path shocks. Neither shock has much of an effect on the surprise component of current-month interest rates.

The bottom row of panel A of Table 5 describes what I call a "monetary target shock." This shock creates a positive comovement in short- and longer-term interest-rate forecast revisions, though the response is much stronger for short-term rates. GDP expectations are revised in the opposite direction, consistent with a theoretical monetary policy shock. Panel B shows that after the introduction of forward guidance, monetary target shocks essentially disappeared. With a few exceptions in 2007–2008, this remained true for the rest of the sample.

Figure 13 plots the Federal Funds rate surprise (top panel) and the estimated Federal Funds

[44]Recall that the sign of the rows of Table 5 are not identified (only the relative signs between elements of a row).

target shock (bottom panel). The two series are fairly similar, consistent with the first column of panel A. This suggests that early estimates of monetary shocks were more closely aligned with exogenous monetary policy. The volatility of the two series also drops substantially in 2003, and never returns to its pre-2003 level. This disappearance causes a power problem: The estimated effects are far too noisy to draw any clear conclusions.

What drives this disappearance? Intuitively, once the Fed begins to communicate about future interest rates, it allows markets to forecast shorter-term interest rates better. Having received a signal about the Fed's time-t interest-rate decision at an earlier date, there is little room for markets to be surprised by the time-t interest-rate decision when it is announced. In Appendix B.2, I extend the illustrative model of Section 3.1 to a dynamic setting to show this point formally. Unlike the static illustrative model, in the dynamic model a perceived monetary policy target shock is the difference between the true shock revealed at time t and the pre-announced shock for time t made at meeting $t-1$. The variance of target shocks is therefore unambiguously decreasing in the clarity of forward guidance. In the extreme case that the time-t shock is revealed fully at time $t-1$, there are no perceived monetary shocks.

Ramey (2016) suggests that monetary shocks are now rare because the Fed conducts policy more systematically and concludes that this is "bad news for econometric identification." My findings corroborate the conclusion that (certain) monetary shocks are rare, but suggest an alternative mechanism and a different conclusion. The results in this section suggest that true monetary shocks—purely exogenous current-meeting interest rate surprises—*are* rare, as Ramey suggests. This conclusion may only be an artifact of the data that underlie my shocks: data on expectations.[45] With my data, these target-type monetary shocks largely disappear in response to a different mechanism: The Fed started announcing these shocks in advance in 2003, so their disappearance did not necessarily arise because the Fed has become more systematic. Few would argue, however, that over the last 50 years the Fed has become more systematic, so Ramey's mechanism is also likely at play.

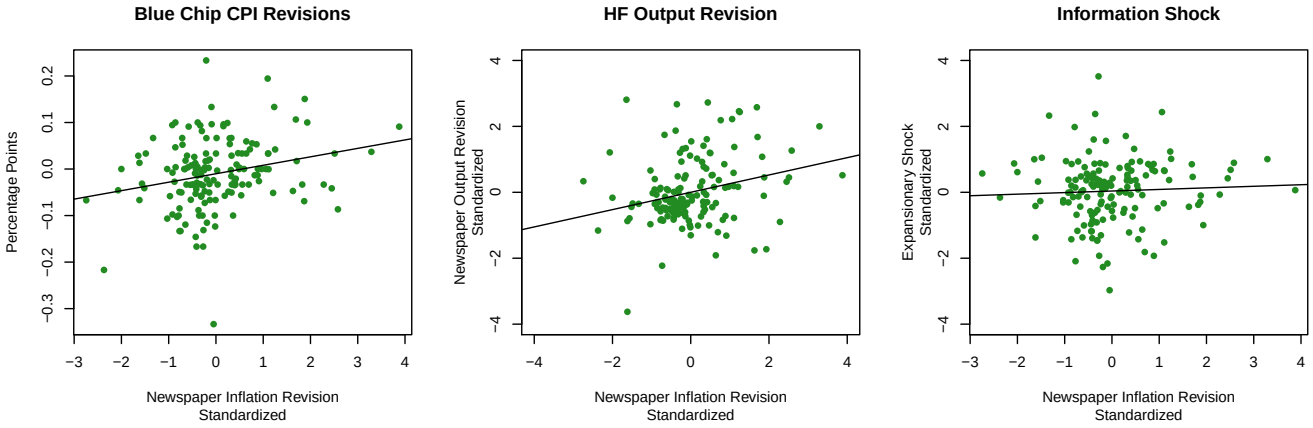
On the other hand, my results suggest that monetary path shocks—commitments to deviate from the policy rule's future prescriptions—are alive and well. These shocks and their effects only looked small because existing measures were swamped by information effects. Thus, rather than conclude that the Fed has put monetary econometricians out of business, I argue that finding new ways to measure the complexity of the Fed's communications policies should keep them in business for the foreseeable future.

5.2 Supply vs. Demand Factors

The last question I ask in this paper is: What type of shocks does the public learn about when the Fed makes a policy announcement. This is an important question from a theoretic perspective: Jia (2020) shows that optimal communications policy depends on the nature of the underlying shock

[45] As mentioned earlier in the paper, my shocks—alongside the shocks of any paper in the high-frequency literature—can only identify *perceived* shocks, since they are based on agents' perceptions (their expectations).

Figure 14: High-frequency Text-based Inflation Expectations



The left panel plots the average Blue Chip forecast revision of CPI inflation over the next three forecast horizons (analogous to $\mathbb{Y}_{t|t+1} - \mathbb{Y}_{t|t+1}$ from equation (3)) against my HF measure of inflation forecast revisions, $\widehat{\pi}_t$. The middle panel shows the relationship between my HF measure of inflation ($\widehat{\pi}_t$) and output forecast revisions ($\widehat{y}_t$). The final panel shows the relationship between $\widehat{\pi}_t$ and the information shock, ξ_t^I, identified in Section 3.4. Aside from Blue Chip forecast revisions, all other variables are standardized to have unit variance over the plotted sample.

(in particular, cost-push (supply) vs. natural-rate (demand) shocks). To answer this question, I construct a measure of high-frequency text-based inflation expectations following the construction in Section 2.1. The only difference between the inflation and output index is in the topic word list, $\mathcal{Y}$, which for inflation becomes

$$\mathcal{Y} = \{\texttt{inflat}, \texttt{price}, \texttt{oil\_prices}, \texttt{inflationari}, \texttt{deflat}\},$$

where I flip the sign of increasing and decreasing measures of `deflat`.[46] As with output, my measure of high-frequency inflation forecast revisions is the unpredictable component of post-meeting inflation directionality vis-à-vis pre-meeting directionality. I plot that index, $\widehat{\pi}_t$, in the left panel of Figure 14, against the summary statistic of CPI forecast revisions from the Blue Chip survey, analogous to the measure constructed in equation (3). The correlation coefficient of the two series is 0.26, with a robust standard error of 0.08.

In Section 4.1 I showed that the identified information shock induced a positive correlation between real GDP and inflation and in expectations thereof. The right panel of Figure 14 shows that this relationship holds in high frequency. Specifically, the figure shows a scatter plot of the expansionary information shock (which, recall, increases output expectations) and $\widehat{\pi}_t$. The two have a positive correlation of 0.13, with a robust standard error of 0.06.

The evidence thus leans against the notion that markets learn about supply-type factors from policy announcements. That does not imply that markets learn *nothing*, only that information about supply factors is less prevalent. My results suggest, additionally, that it may be difficult to

[46] The bigram `interest_rates` was also included as similar to my "seed" words of `inflation` and `prices`, but I removed it.

identify such a shock. The middle panel of Figure 14 shows that a fairly strong positive relationship exists between output and inflation forecast revisions ($\widehat{y}_t$ and $\widehat{\pi}_t$). The two series have a positive correlation of 0.27 (s.e. 0.10). That relationship is fairly consistent along my entire sample. Thus—at least given the heteroskedasticity-based approach to identification—there does not appear to be a natural regime that would help to separately identify a shock revealing supply-type factors from a shock revealing demand-type factors. When carrying out the estimation on a quad-variate system that includes high-frequency revisions in short-term rates, long-term rates, output, and inflation, the two non-monetary shocks have statistically equal variances across the two regimes. Identification by heteroskedasticity requires that the variance of different shocks differ across the two regimes, which means that the two information shocks are not identified. This is not surprising; nothing about the regime shift in 2003 suggests that markets could learn more about supply or demand shocks after 2003. That regime shift instead only allowed markets to learn relatively more about longer-term interest rates than about macroeconomic information in general.

Interestingly, the conclusion that markets learn little about supply-side factors from Fed announcements is consistent with the optimal policy prescriptions of Jia (2020). Jia highlights—in a much more realistic model than mine—the fact that optimal communications policy by the Fed reveals information about demand shocks, but obfuscates information about supply shocks (cost-push shocks, to be precise). Purposeful or not, my results suggest that the Fed's communication policies are in line with the optimal communications policy.

6 Conclusion

Estimating the macroeconomic effects of monetary policy is notoriously difficult, because interest rates are so highly endogenous with respect to macroeconomic variables. Despite several approaches in the literature aimed at identifying exogenous changes in interest rates (i.e., monetary policy shocks) using high-frequency data, the identification of a series that has theoretically consistent effects, without relying on a particular set of controls, has proven elusive.

In this paper I provide an estimate of monetary shocks that is free of the "information effects" that have been posited to plague previous high-frequency estimates.[47] To separately identify two shocks—an information shock and a monetary shock—I study two data series: high-frequency interest-rate and GDP forecast revisions. Not able to find the latter variable in the literature, I constructed one using newspaper articles written about Fed policy meetings.

The notion that two series do not identify two shocks, without additional identifying assumptions, is an old one in economics. I present a model of expectation revisions around Fed announcements, and show what assumptions have been made (largely implicitly), by those who have previously sought to identify monetary shocks in high frequency. Like zero restrictions in structural VARs, these are generally not supported by macroeconomic theory. My assumptions, identification by heteroskedasticity, are a bit more involved, but much less restrictive and better

[47] My results confirm that about 80% of a commonly used high frequency shock are made up of information effects.

suited to this context. The introduction of forward guidance by the Fed, a major regime change in Fed communications, drastically changed the shocks that markets could learn about from Fed policy announcements. Identification by heteroskedasticity provides a tool to turn this regime change into structural identification.

In terms of empirical findings, my monetary shock has macroeconomic effects that are consistent with standard New Keynesian models. I contrast my estimates with other leading alternatives, and show that my results do not depend on having particular controls in my regressions. This lends credibility to my estimated shocks and speaks to their portability. Given the widespread use of monetary shocks in the empirical macroeconomics literature, a credible and portable series that can be carried forward (i.e., not limited by zero lower bound constraints) is greatly needed.

My results reveal additional information on the nature of monetary policy shocks. It appears that unanticipated shocks to the current-month Federal Funds rate have become a thing of the past, largely ending in 2003, with the introduction of forward guidance. Instead, as a result of the Fed's clearer signals, markets were able to more completely update their longer-term interest-rate expectations, leading to an increase in the size of "path" shocks. My model highlights how the introduction of forward guidance can lead to both smaller target shocks and larger path shocks.

Finally, I find that the effects of Fed information shocks are essentially equal in absolute value to the effects of monetary policy shocks. My results suggest that Fed announcements primarily reveal information about aggregate demand. This is in keeping with the prior observation (Jia, 2020) that optimal central bank information provision policy primarily reveals information about demand shocks. Further research will serve to clarify why the Fed appears to play such an important role in providing macroeconomic information, and how such communications can be further optimized.[48]

[48]The large macroeconomic effects that I estimate require reconciliation with empirical work that shows that households know very little about what the Fed does and says (Binder, 2017; Coibion et al., Forthcoming).

References

Acosta, Miguel, "A New Measure of Central Bank Transparency and Implications for the Effectiveness of Monetary Policy," Working Paper 2019.

_ and Hassan Afrouzi, "Rationally Confused: On the Aggregate Implications of Information Provision Policies," Technical Report In Progress.

_ and Joseph M. Saia, "Estimating the Effects of Monetary Policy via High Frequency Factors," Technical Report In progress.

Angrist, Joshua D and Jörn-Steffen Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton university press, 2008.

Arai, Natsuki, "The effects of monetary policy announcements at the zero lower bound," *International Journal of Central Banking*, 2017, *13* (2), 159–196.

Bauer, Michael D. and Eric T. Swanson, "The Fed's Response to Economic News Explains the 'Fed Information Effect'," Working Paper 2020.

Baum, Christopher F, Mark E Schaffer, and Steven Stillman, "ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression.," 2010.

Binder, Carola, "Fed speak on main street: Central bank communication and household expectations," *Journal of Macroeconomics*, 2017, *52*, 238–251.

Blanchard, Olivier J., Jean-Paul L'Huillier, and Guido Lorenzoni, "News, Noise, and Fluctuations: An Empirical Exploration," *American Economic Review*, December 2013, *103* (7), 3045–70.

Brave, Scott A, R Andrew Butters, David Kelley et al., "A new "big data" index of US economic activity," *Economic Perspectives, Federal Reserve Bank of Chicago*, 2019, *1*.

Cai, Yong, Santiago Camara, and Nicholas Capel, "It's not always about the money, sometimes it's about sending a message: Evidence of Informational Content in Monetary Policy Announcements," Technical Report, Working Paper 2021.

Campbell, Jeffrey R, Charles L Evans, Jonas D.M. Fisher, and Alejandro Justiniano, "Macroeconomic effects of Federal Reserve forward guidance," *Brookings Papers on Economic Activity*, 2012, pp. 1–80.

Chahrour, Ryan and Kyle Jurado, "Recoverability and expectations-driven fluctuations," *Review of Economic Studies. Forthcoming*, 2021.

Cieslak, Anna and Andreas Schrimpf, "Non-monetary news in central bank communication," *Journal of International Economics*, 2019, *118*, 293 – 315.

Clarkson, Douglas B, "Estimating the standard errors of rotated factor loadings by jackknifing," *Psychometrika*, 1979, *44* (3), 297–314.

Coibion, Olivier, "Are the effects of monetary policy shocks big or small?," *American Economic Journal: Macroeconomics*, 2012, *4* (2), 1–32.

_ , **Yuriy Gorodnichenko, and Michael Weber**, "Monetary Policy Communications and their Effects on Household Inflation Expectations," *Journal of Political Economy*, Forthcoming.

D'Amico, Stefania and Thomas King, "What Does Anticipated Monetary Policy Do?," Technical Report, Working Paper 2017.

Doh, Taeyoung, Dongho Song, and Shu-Kuei Yang, "Deciphering Federal Reserve communication via text analysis of alternative FOMC statements," Technical Report, Working Paper 2020.

Ellen, Saskia Ter, Vegard H Larsen, and Leif Anders Thorsrud, "Narrative monetary policy surprises and the media," *Journal of Money, Credit, and Banking*, forthcoming.

Fernández-Villaverde, Jesús, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson, "ABCs (and Ds) of Understanding VARs," *American Economic Review*, June 2007, *97* (3), 1021–1026.

Galí, Jordi, *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*, Princeton University Press, 2015.

Gertler, Mark and Peter Karadi, "Monetary Policy Surprises, Credit Costs, and Economic Activity," *American Economic Journal: Macroeconomics*, January 2015, *7* (1), 44–76.

Gürkaynak, Refet S, Brian P Sack, and Eric T Swanson, "Market-based measures of monetary policy expectations," *Journal of Business & Economic Statistics*, 2007, *25* (2), 201–212.

_ , **Brian Sack, and Eric T Swanson**, "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*, 2005.

_ , _ , **and Jonathan H Wright**, "The TIPS yield curve and inflation compensation," *American Economic Journal: Macroeconomics*, 2010, *2* (1), 70–92.

Gurkaynak, Refet S, Burçin Kisacikoğlu, and Jonathan H Wright, "Missing Events in Event Studies: Identifying the Effects of Partially Measured News Surprises," *American Economic Review*, 2020, *110* (12), 3871–3912.

Handlan, Amy, "Text Shocks and Monetary Surprises: Text Analysis of FOMC Statements with Machine Learning," Working Paper 2020.

Hansen, Stephen and Michael McMahon, "Shocking language: Understanding the Macroeconomic Effects of Central Bank Communication," *Journal of International Economics*, 2016, *99.*

_ , _ , **and Matthew Tong**, "The long-run information effect of central bank communication," *Journal of Monetary Economics*, 2019, *108*, 185–202.

Hébert, Benjamin and Jesse Schreger, "The costs of sovereign default: Evidence from argentina," *American Economic Review*, 2017, *107* (10), 3119–45.

Hoesch, Lukas, Barbara Rossi, and Tatevik Sekhposyan, "Has the information channel of

monetary policy disappeared? Revisiting the empirical evidence," Working Paper 2020.

Jarociński, Marek and Peter Karadi, "Deconstructing Monetary Policy Surprises—The Role of Information Shocks," *American Economic Journal: Macroeconomics*, April 2020, *12* (2), 1–43.

Jia, Chengcheng, "The Informational Effect of Monetary Policy and the Case for Policy Commitment," Working Paper 2020.

Jordà, Òscar, "Estimation and inference of impulse responses by local projections," *American economic review*, 2005, *95* (1), 161–182.

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

Lakdawala, Aeimit, "Decomposing the effects of monetary policy using an external instruments SVAR," *Journal of applied econometrics*, 2019, *34* (6), 934–950.

Lepetyuk, Vadym, Lilia Maliar, Serguei Maliar, and John B Taylor, "Policy Anticipation Effects in New Keynesian Models," 2021.

Lewis, Daniel J, "Announcement-Specific Decompositions of Unconventional Monetary Policy Shocks and Their Macroeconomic Effects," *FRB of New York Staff Report*, 2019, (891).

Loughran, Tim and Bill McDonald, "When is a Liability not a Liability? Textual Analysis, Dictionaries, and 10-Ks," *The Journal of Finance*, 2011, *66* (1), 35–65.

Lucas, Robert E, "Expectations and the neutrality of money," *Journal of Economic Theory*, 1972, *4* (2), 103–124.

Lucca, David O. and Francesco Trebbi, ""Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements"," *NBER Working Paper Series*, 2009, *15367*.

Lunsford, Kurt G, "Policy language and information effects in the early days of Federal Reserve forward guidance," *American Economic Review*, 2020, *110* (9), 2899–2934.

Manning, Christopher D, Mihai Surdeanu, John Bauer, Jenny Rose Finkel, Steven Bethard, and David McClosky, "The Stanford CoreNLP natural language processing toolkit," in "Proceedings of 52nd annual meeting of the association for computational linguistics: system demonstrations" 2014, pp. 55–60.

Melosi, Leonardo, "Signalling effects of monetary policy," *The Review of Economic Studies*, 2017, *84* (2), 853–884.

Mikolov, Tomas, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean, "Distributed representations of words and phrases and their compositionality," in "Advances in neural information processing systems" 2013, pp. 3111–3119.

Miranda-Agrippino, Silvia and Giovanni Ricco, "The Transmission of Monetary Policy Shocks," *American Economic Journal: Macroeconomics*, July 2021, *13* (3), 74–107.

Morris, Stephen and Hyun Song Shin, "Inertia of forward-looking expectations," *American Economic Review*, 2006, *96* (2), 152–157.

Nakamura, Emi and Jón Steinsson, "High Frequency Identification of Monetary Non-Neutrality," *Quarterly Journal of Economics*, 2018, *133* (3), 1283–1330.

Newey, Whitney K and Kenneth D West, "Automatic lag selection in covariance matrix estimation," *The Review of Economic Studies*, 1994, *61* (4), 631–653.

Nimark, Kristoffer, "Dynamic higher order expectations," Technical Report, Working Paper 2017.

Petersen, K. B. and M. S. Pedersen, "The Matrix Cookbook," nov 2012. Version 20121115.

Porter, Martin F, "An algorithm for suffix stripping," *Program*, 1980.

Ramey, Valerie A, "Macroeconomic shocks and their propagation," in "Handbook of macroeconomics," Vol. 2, Elsevier, 2016, pp. 71–162.

_ **and Sarah Zubairy**, "Government spending multipliers in good times and in bad: evidence from US historical data," *Journal of Political Economy*, 2018, *126* (2), 850–901.

Rendahl, Pontus, "Linear time iteration," Technical Report, Working Paper 2017.

Rigobon, Roberto, "Identification through Heteroskedasticity," *Review of Economics and Statistics*, 2003, *85* (4), 777–792.

Romer, Christina D. and David H. Romer, "Federal Reserve Information and the Behavior of Interest Rates," *American Economic Review*, 2000, *90* (3), 429–457.

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

_ **and** _ , "The FOMC Versus the Staff: Where Can Monetary Policymakers Add Value?," *American Economic Review, Paper and Proceedings*, 2008, *98* (2), 230–35.

Sastry, Karthik, "Disagreement About Monetary Policy," Technical Report, Working Paper 2021.

Smets, Frank and Rafael Wouters, "Shocks and frictions in US business cycles: A Bayesian DSGE approach," *American economic review*, 2007, *97* (3), 586–606.

Stock, James H and Mark W Watson, "Twenty years of time series econometrics in ten pictures," *Journal of Economic Perspectives*, 2017, *31* (2), 59–86.

Taylor, John B., "Discretion versus Policy Rules in Practice," in "Carnegie-Rochester Conference Series on Public Policy," Vol. 39 1993, pp. 195–214.

Townsend, Robert M, "Forecasting the forecasts of others," *Journal of Political Economy*, 1983, *91* (4), 546–588.

Woodford, Michael, "Imperfect Common Knowledge and the Effects of Monetary Policy," *Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps*, 2003, p. 25.

_ , "Comment on 'Macroeconomic effects of Federal Reserve forward guidance'," *Brookings Papers on Economic Activity*, 2012, pp. 1–80.

Wright, Jonathan H, "What does monetary policy do to long-term interest rates at the zero lower bound?," *The Economic Journal*, 2012, *122* (564), F447–F466.

Wu, Jing Cynthia and Fan Dora Xia, "Measuring the macroeconomic impact of monetary policy at the zero lower bound," *Journal of Money, Credit and Banking*, 2016, *48* (2-3), 253–291.

Zeileis, Achim, "Econometric Computing with HC and HAC Covariance Matrix Estimators," *Journal of Statistical Software*, 2004, *11* (10), 1–17.

_ , Susanne Köll, and Nathaniel Graham, "Various Versatile Variances: An Object-Oriented Implementation of Clustered Covariances in R," *Journal of Statistical Software*, 2020, *95* (1), 1–36.

Zhang, Xu, "Disentangling the Information Effects in the Federal Reserve's Monetary Policy Announcements," Technical Report, Working Paper 2020.

A Numerical Example: Textbook New Keynesian Model

In this appendix I show that the illustrative model in the text is a representation of the textbook New Keynesian model found in chapter 3 of (Galí, 2015, Ch. 3). I also show that the relevant structural impact matrix in the model (**M** in the text) has the same signs as what I estimate, and the type of solution that forms the notion of how "theoretically consistent" monetary shocks should behave. That model takes the form

$$\pi_t = c_\pi + \beta\,\mathbb{E}_t[\pi_{t+1}] + \kappa y_t - c_s s_{t|t} \tag{20a}$$

$$y_t = \mathbb{E}_t[y_{t+1}] - \frac{1}{\sigma}(i_t - \mathbb{E}_t[\pi_{t+1}] - \rho) + \frac{1}{\sigma}(1-\rho_d)d_{t|t} \tag{20b}$$

$$i_t = \rho + \phi_\pi \pi_t + \phi_y(y_t - y) + m_{t|t} \tag{20c}$$

where y_t is output, π_t is inflation, and i_t is the nominal interest rate. The model's three structural shocks are a supply shock, demand shock, and monetary policy shock which, respectively, follow the following AR(1) processes:

$$s_t = \rho_s s_t + \varepsilon_t^s \qquad \text{(Technology Shock)}$$

$$d_t = \rho_d d_t + \varepsilon_t^d \qquad \text{(Demand/Pref. Shock)}$$

$$m_t = \rho_m m_t + \varepsilon_t^m \qquad \text{(Monetary Shock)}$$

where the innovation terms ε_t^i are independent and normally distributed with means of zero and standard deviation σ_s, σ_d, and σ_m, respectively

Calibration I calibrate the model exactly as in chapter 3 of Galí (2015). The value of the structural parameters are

Parameter	Symbol	Value	Parameter	Symbol	Value
Discount factor	β	0.99	Risk aversion	σ	1
Inverse Frisch Elast.	φ	1	Cobb-Douglas	α	0.25
Consumption elast. of subs.	ϵ	9	Interest semielast. of mon. demand	η	4
Price stickiness	θ	0.75	Taylor rule inflation	ϕ_π	1.5
Taylor rule output	ϕ_y	0.125	Persistence of monetary shock	ρ_m	0.5
Persistence of demand shock	ρ_d	0.5	S.D. of demand shock	σ_d	1.0
S.D. of monetary shock	σ_m	0.5	S.D. of supply shock	σ_s	0

This implies the following values for the parameters used in the model of equation (20):

$$\begin{aligned}
\mu &= \log\left(\frac{\epsilon}{\epsilon-1}\right) \approx 0.117 \\
\Psi_y &= -\frac{(1-\alpha)(\mu-\log(1-\alpha))}{\sigma(1-\alpha)+\varphi+\alpha} \approx -0.152 \\
\Theta &= \frac{1-\alpha}{1-\alpha+\alpha\epsilon} = 0.25 \\
\lambda &= \frac{(1-\theta)(1-\beta\theta)}{\theta}\Theta \approx 0.021 \\
\kappa &= \lambda\left(\sigma+\frac{\varphi+\alpha}{1-\alpha}\right) \approx 0.057 \\
c_\pi &= -\kappa\Psi_y \approx 0.008 \\
c_s &= -\kappa\left(\frac{1+\varphi}{\sigma(1-\alpha)+\varphi+\alpha}\right) \approx -0.057 \\
\rho &= -\log(\beta) \approx 0.010
\end{aligned}$$

Solution Approach The model can be put into the form of equation (13)

$$\underbrace{\begin{bmatrix} -\kappa & 1 & 0 \\ 1 & 0 & \frac{1}{\sigma} \\ -\phi_y & -\phi_\pi & 1 \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} y_t \\ \pi_t \\ i_t \end{bmatrix} + \underbrace{\begin{bmatrix} 0 & -\beta & 0 \\ -1 & \frac{1}{\sigma} & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{B}} \mathbb{E}_t \begin{bmatrix} y_{t+1} \\ \pi_{t+1} \\ i_{t+1} \end{bmatrix} + \underbrace{\begin{bmatrix} -c_\pi \\ \frac{\rho}{\sigma} \\ -\rho+\phi_y y \end{bmatrix}}_{\overline{\mathbf{C}}} + \underbrace{\begin{bmatrix} 0 & c_s & 0 \\ -\frac{1}{\sigma}(1-\rho_d) & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} d_{t|t} \\ s_{t|t} \\ m_{t|t} \end{bmatrix} = \mathbf{0}$$

Given the simple format of the Taylor rule, this system can be reduced to a system in i_t and y_t by solving the Taylor rule for π_t. The new system takes the form

$$\begin{aligned}
&\underbrace{\begin{bmatrix} -\left(\frac{\phi_y}{\phi_\pi}+\kappa\right) & \frac{1}{\phi_\pi} \\ 1 & \frac{1}{\sigma} \end{bmatrix}}_{\mathbf{A}} \begin{bmatrix} y_t \\ i_t \end{bmatrix} + \underbrace{\begin{bmatrix} \beta\frac{\phi_y}{\phi_\pi} & -\frac{\beta}{\phi_\pi} \\ \frac{\phi_y}{\phi_\pi\sigma}-1 & -\frac{1}{\phi_\pi\sigma} \end{bmatrix}}_{\mathbf{B}} \mathbb{E}_t \begin{bmatrix} y_{t+1} \\ i_{t+1} \end{bmatrix} \\
&\qquad + \underbrace{\begin{bmatrix} -\frac{\rho}{\phi_\pi} - c_\pi - \beta\frac{\phi_y}{\phi_\pi}y + \frac{\rho\beta}{\phi_\pi} \\ -\frac{\phi_y}{\phi_\pi\sigma}y + \frac{\rho}{\phi_\pi\sigma} - \frac{\rho}{\sigma} \end{bmatrix}}_{\overline{\mathbf{C}}} + \underbrace{\begin{bmatrix} 0 & c_s & \frac{\beta\rho_m}{\phi_\pi} - \frac{1}{\phi_\pi} \\ -\frac{1}{\sigma}(1-\rho_d) & 0 & \frac{\rho}{\phi_\pi\sigma} \end{bmatrix}}_{\mathbf{D}} \begin{bmatrix} d_{t|t} \\ s_{t|t} \\ m_{t|t} \end{bmatrix} = \mathbf{0}
\end{aligned}$$

Full Information Numerical Solution Following the solution method described in section C, the solution of the model without noise consists of the exogenous processes for the demand and monetary shocks and the analogue to equation (27a):

$$\begin{bmatrix} y_t \\ i_t \end{bmatrix} = \begin{bmatrix} -0.15 \\ 0.01 \end{bmatrix} + \underbrace{\begin{bmatrix} 0.67 & -1.35 \\ 0.20 & 0.60 \end{bmatrix}}_{\mathbf{M}_{\mathrm{NK}}} \begin{bmatrix} d_{t|t} \\ m_{t|t} \end{bmatrix}, \tag{21}$$

where, given the full-information assumption, $d_{t|t} = d_t$ and $m_{t|t} = m_t$; in words, actual and perceived shocks coincide. Pre-multiplying both sides by $\mathbf{M}_{\text{NK}}^{-1}$ and then normalizing the coefficients on y_t and i_t to unity yields the reduced-form DIS curve and Taylor rules from this model.

$$\begin{aligned} y_t &= -0.17 - 2.26 i_t + 1.13 d_t \\ i_t &= 0.01 + 0.29 y_t + \varepsilon_t \,. \end{aligned}$$

B Formal Signaling Structures

In the text I remain agnostic as the nature of the information received by Fed watchers. In this appendix I provide simple formal examples that generate the intuitions described in the text.

B.1 Static Signaling Model

Assume that, for a Fed announcement occurring at time t, the Fed emits noisy signals about each shock given by

$$\begin{aligned} s_t^{\varepsilon} &= \varepsilon_\tau + n_t^{\varepsilon} \\ s_t^{\eta} &= \eta_\tau + n_t^{\eta}. \end{aligned}$$

The noise components n_t^{ε} and n_t^{η} are independently distributed normal variables with zero mean and respective variances $\sigma_{n,\varepsilon}^2$ and $\sigma_{n,\varepsilon}^2$. The formulation of these signals suggests that the Fed chooses "how noisy" to make its signals. This noise could arise, instead, from the noise with which the Fed perceives the underlying signals itself.

The signals s_t^{ε} and s_t^{η} are used to form expectations of the economy's structural shocks using Bayes' rule and knowledge of the model's parameters. Prior beliefs for each of the structural shocks are zero (the mean of the underlying distributions), so posterior expectations—equivalently, expectation revisions—are

$$\mathbb{E}_t[\varepsilon_\tau] = \underbrace{\left(\frac{\sigma_{\varepsilon}^2}{\sigma_{\varepsilon}^2 + \sigma_{n,\varepsilon}^2}\right)}_{\kappa_\varepsilon} s_t^{\varepsilon} \qquad\qquad \mathbb{E}_t[\eta_\tau] = \underbrace{\left(\frac{\sigma_{\eta}^2}{\sigma_{\eta}^2 + \sigma_{n,\eta}^2}\right)}_{\kappa_\eta} s_t^{\eta}. \tag{22}$$

The simple form of these expectations follows from the information structure above. That being said, I made two simplifications in formulating the information structure. The first concerns the lack of dynamics in the structural shocks (ε_τ and η_τ). With dynamic shocks, nearly identical formulas could replace equation (22) by positing that the signals were of the white-noise innovation to the structural shocks (or a lag-polynomial thereof, if $\tau > t$) instead of the structural shocks themselves. However, because the expectations revisions that I study are empirically uncorrelated, this simplification is warranted. The second simplification concerns the nature of the signals—specifically that the noise is independent across the structural shocks. If, instead, the signals were

of the observable variables (y_τ and i_τ), then the posterior expectations of the structural shocks would be correlated, as Acosta and Afrouzi (In Progress) demonstrate. Again, my empirical results do not suggest that forecast revisions regarding structural shocks are correlated, which alleviates this concern.[49]

B.2 Dynamic Extension

The discussion in the previous section was ambiguous with respect to the timing of the outcome variables. Empirically, the measures of output and interest rate expectations both captured the average value of those variables over the coming year. With measures of expectations now differentiated by time, this ambiguity must be resolved.

To do this, I assume that the monetary and information shocks follow exogenous first-order autoregressive processes, given by

$$\begin{aligned} \varepsilon_t &= \rho_\varepsilon \varepsilon_{t-1} + \mu_t^\varepsilon & \mu_t^\varepsilon &\sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right) \\ \eta_t &= \rho_\eta \eta_{t-1} + \mu_t^\eta & \mu_t^\eta &\sim \mathcal{N}\left(0, \sigma_\eta^2\right). \end{aligned}$$

I have simplified the timing conventions substantially here, assuming that t corresponds to calendar time and Fed meetings, and agents only receive information from the Fed. These simplifications are not necessary but help me to make my argument concisely. The next step is to allow the Fed to send a signal about future interest rates. For expositional purposes I restrict the model to allow the Fed to emit a signal only about next-period's interest rate. The information set of agents thus consists of all prior signals, along with

$$\begin{aligned} f_t^\varepsilon &= \mu_{t+1}^\varepsilon + n_t^f & n_t^f &\sim \mathcal{N}\left(0, \sigma_f^2\right) \\ s_t^\varepsilon &= \mu_t^\varepsilon + n_t^\varepsilon & n_t^\varepsilon &\sim \mathcal{N}\left(0, \sigma_{n,\varepsilon}^2\right) \\ s_t^\eta &= \mu_t^\eta + n_t^\eta & n_t^\eta &\sim \mathcal{N}\left(0, \sigma_{n,\eta}^2\right). \end{aligned}$$

With this information structure, the variance of forecast revisions of monetary policy innovations are given by

$$\text{var}\left(\widehat{\mu}_t\right) = \frac{\sigma_\varepsilon^4 \sigma_f^4}{(\sigma_\varepsilon^2 + \sigma_f^2)(\sigma_\varepsilon^2 \sigma_{n,\varepsilon}^2 + \sigma_\varepsilon^2 \sigma_f^2 + \sigma_{n,\varepsilon}^2 \sigma_f^2)} \qquad \text{var}\left(\widehat{\mu}_{t+1}\right) = \frac{\sigma_\varepsilon^4}{\sigma_\varepsilon^2 + \sigma_f^2},$$

where now the hat notation denotes changes between periods t and $t-1$, which is equivalent to high-frequency changes in this simple timing setup: $\widehat{x}_t = \mathbb{E}_t[x_t] - \mathbb{E}_{t-1}[x_t]$. These equations underlie the intuition for how the introduction of forward guidance—here a decrease in σ_f—can simultaneously

[49] This assumption seems less plausible if markets *only* observed the interest-rate decision itself. In reality, Fed announcements are composed of a post-meeting statement over my entire sample that explains the rationale behind the policy decision. Thus, the announcement is multidimensional. What's more, since the start of my sample several additional dimensions have been added to this signal: press conferences (with the opportunity to answer questions from the press) and economic forecasts.

make current-meeting interest-rate surprises smaller, but future-period interest-rate surprises larger. Signaling about tomorrow's monetary shock has two effects. Today, it allows markets to put more weight on that signal when forming forecasts about tomorrow's innovation, thus increasing the average size of forecast revisions (formally, $\text{var}\left(\widehat{\mu}_{t+1}\right)$ is decreasing in σ_f). When tomorrow, $t+1$, rolls around, agents will have already received a signal in period t, so they have less prior uncertainty about the current-period innovation. The signal is relatively less informative, dampening the size of tomorrow's revisions ($\text{var}\left(\widehat{\mu}_t\right)$ is increasing in σ_f). Note that, in the terminology of Gürkaynak et al. (2005), $\widehat{\mu}_t$ and $\widehat{\mu}_{t+1}$ are akin to target and path shocks, respecively—the former affects both $\widehat{i}_t$ and $\widehat{i}_{t+1}$, while the latter does not affect $\widehat{i}_t$.

This intuition continues to operate when considering the size of forecast revisions to the forward rates i_t and i_{t+1}. Combining equation (8a) with the fact that $\widehat{\varepsilon}_{t+1} = \widehat{\mu}_{t+1} + \rho\widehat{\mu}_t$ (and that $\mathbb{E}[\widehat{\mu}_{t+1}\widehat{\mu}_t] = 0$), it can be shown that when $\rho \leq 1$,

$$\frac{\partial \,\text{var}\left(\widehat{i}_t\right)}{\partial\, \sigma_f} > 0 \qquad\qquad \frac{\partial \,\text{var}\left(\widehat{i}_{t+1}\right)}{\partial\, \sigma_f} < 0. \tag{23}$$

In words, a clearer signal about future interest rates—i.e., a decrease in σ_f—leads to smaller current-period interest-rate revisions, but larger future interest-rate revisions. Note, however, that $\widehat{i}_{t+1}$ is a forward rate. Under the expectations theory of the term structure, the theoretical analog to my empirical variable is expected yield on a 2-period bond—under the expectations theory of the term structure, this is

$$i_{t,1} \equiv \frac{1}{2}\,\mathbb{E}_t[i_t + i_{t+1}].$$

Thus while the model offers clear predictions for how the size of forecast revisions for forward rates responds to the introduction of forward guidance, equation (23) shows that the response of the size of yield revisions is ambiguous. In particular, the response depends on the persistence of the monetary shock, ε_t. As the shock becomes increasingly transitory, yield forecast revisions $\widehat{i}_{t,1}$ become unambiguously larger in response to forward guidance, since current innovations carry through less to future shocks.[50] My finding that yield revisions increase following the introduction of forward guidance suggests, then, that monetary shocks are indeed fairly transitory.

The dynamic form of the baseline model also provides an explanation for why longer-term interest rates are more susceptible to information effects. To see this, note that longer-term interest-rate (forward) surprises are

$$\widehat{i}_{t+k} = \rho_\eta^k \eta_{t|t} + \rho_\varepsilon^{k-1}\varepsilon_{t+1|t}.$$

This can be used to express the contribution of information effects—relative to the contribution of

[50] This result is straightforward: If shocks aren't persistent, then today's interest-rate shock does not affect tomorrow's interest-rate forward.

monetary shocks—to the variance of interest rate surprises as

$$\lambda_k \equiv \frac{\rho_\eta^{2k}\text{var}(\widehat{\eta}_t)/\text{var}\left(\widehat{i}_{t+k}\right)}{\rho_\varepsilon^{2(k-1)}\text{var}(\widehat{\varepsilon}_t)/\text{var}\left(\widehat{i}_{t+k}\right)}.$$

Thus, as the interest-rate horizon increases, the relative contribution of information also increases *if* the information shock is more persistent, since

$$\frac{d\log(\lambda_k)}{dk} = 2(\log(\rho_\eta) - \log(\rho_\varepsilon)) > 0 \iff \rho_\eta > \rho_\varepsilon.$$

Empirically, Smets and Wouters (2007)—who estimate a DSGE of the US economy driven by a rich set of macroeconomic shocks—suggest that monetary shocks are fairly transitory.

C Solving the Noisy Information Model

This appendix contains the instructions to solve a noisy-information linear dynamic model following Blanchard et al. (2013), whose exposition I follow closely. The discussion is useful for considering models in which current decisions depend not only expected future decisions, but also on lagged decisions (an extension of the model in the text). The model takes the form

$$\mathbf{A}\,\mathbf{x}_t + \mathbf{B}\,\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{C}\,\mathbf{x}_{t-1} + \overline{\mathbf{C}} + \mathbf{D}\boldsymbol{\xi}_{t|t} = \mathbf{0} \tag{24}$$

where $\mathbf{x}_t$ is a vector of observable macroeconomic variables, $z_{t|\tau}$ denotes the mathematical expectation of z_t given information at time τ, and $\boldsymbol{\xi}_t$ is a vector of mutually independent structural shocks that evolves according to:

$$\boldsymbol{\xi}_t = \mathbf{H}\boldsymbol{\xi}_{t-1} + \mathbf{J}\boldsymbol{\varepsilon}_t \qquad\qquad \boldsymbol{\varepsilon}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon), \tag{25}$$

where $\boldsymbol{\Sigma}_\varepsilon$ is a diagonal matrix. At the beginning of each period t but before making decisions about $\mathbf{x}_t$, agents receive a noisy signal of the structural shock of the form

$$\mathbf{s}_t = \mathbf{F}\boldsymbol{\xi}_t + \mathbf{G}\boldsymbol{\nu}_t \qquad\qquad \boldsymbol{\nu}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_\nu), \tag{26}$$

which they use to form expectations about $\boldsymbol{\xi}_t$ using the Kalman filter. In the full-information case (i.e. $\boldsymbol{\Sigma}_\nu \to \mathbf{0}$), instead, structural shocks are observed perfectly and so $\boldsymbol{\xi}_{t|t} = \boldsymbol{\xi}_t$. I conjecture that the model satisfies all necessary stability conditions such that it admits a solution takes the following form

$$\mathbf{x}_t = \overline{\mathbf{x}} + \mathbf{L}\,\mathbf{x}_{t-1} + \mathbf{M}\boldsymbol{\xi}_{t|t} \tag{27a}$$

$$\boldsymbol{\xi}_{t|t} = (\mathbf{I} - \mathbf{K}\mathbf{F})\mathbf{H}\boldsymbol{\xi}_{t-1|t-1} + \mathbf{K}\mathbf{s}_t \tag{27b}$$

Plugging the conjectured guess of equation (27a) into (24) gives

$$\mathbf{A}(\overline{\mathbf{x}} + \mathbf{L}\,\mathbf{x}_{t-1} + \mathbf{M}\boldsymbol{\xi}_{t|t}) + \mathbf{B}\,\mathbb{E}_t[\overline{\mathbf{x}} + \mathbf{L}\,\mathbf{x}_t + \mathbf{M}\boldsymbol{\xi}_{t+1|t+1}] + \mathbf{C}\,\mathbf{x}_{t-1} + \overline{\mathbf{C}}\mathbf{D}\boldsymbol{\xi}_{t|t} = 0. \tag{28}$$

Plugging in the guess again for $\mathbf{x}_t$ and noticing that, using the law of iterated expectations and (25), $\mathbb{E}_t[\boldsymbol{\xi}_{t+1|t+1}] = \boldsymbol{\xi}_{t+1|t} = \mathbf{H}\boldsymbol{\xi}_{t|t}$, equation (28) becomes

$$\left[(\mathbf{A} + \mathbf{B}(\mathbf{I} + \mathbf{L}))\overline{\mathbf{x}} + \overline{\mathbf{C}}\right] + \left[\mathbf{AL} + \mathbf{BL}^2 + \mathbf{C}\right]\mathbf{x}_{t-1} + \left[\mathbf{AM} + \mathbf{BLM} + \mathbf{BMH} + \mathbf{D}\right]\boldsymbol{\xi}_{t|t} = \mathbf{0}.$$

To ensure that this equation holds with equality regardless of $\mathbf{x}_{t-1|t-1}$ and $\boldsymbol{\xi}_{t|t}$, it must be then that

$$\mathbf{AL} + \mathbf{BL}^2 + \mathbf{C} = \mathbf{0} \qquad \mathbf{AM} + \mathbf{BLM} + \mathbf{BMH} + \mathbf{D} = \mathbf{0} \qquad (\mathbf{A} + \mathbf{B}(\mathbf{I} + \mathbf{L}))\overline{\mathbf{x}} + \overline{\mathbf{C}} = \mathbf{0}$$

The first equation can be used to solve for $\mathbf{L}$ using the method of Rendahl (2017), the second is an "encapsulating sum" problem whose solution is given in Petersen and Pedersen (2012), and the last is linear in $\overline{\mathbf{x}}$. Solving for Kalman gain matrix $\mathbf{K}$ can also be done by iteration. Define the initial guess for the matrix $\mathbf{P}$ as $\mathbf{P}_0$. Then iterate on the following equations over i until convergence:

$$\mathbf{K}_i = (\mathbf{H}\mathbf{P}_{i-1}\mathbf{H}' + \mathbf{J}\boldsymbol{\Sigma}_\varepsilon\mathbf{J}')\mathbf{F}(\mathbf{F}(\mathbf{H}\mathbf{P}_{i-1}\mathbf{H}' + \mathbf{J}\boldsymbol{\Sigma}_\varepsilon\mathbf{J}')\mathbf{F}' + \mathbf{G}\boldsymbol{\Sigma}_\nu\mathbf{G}')^{-1} \tag{29a}$$

$$\mathbf{P}_i = (\mathbf{I} - \mathbf{K}_i\mathbf{F})(\mathbf{H}\mathbf{P}_{i-1}\mathbf{H}' + \mathbf{J}\boldsymbol{\Sigma}_\varepsilon\mathbf{J}') \tag{29b}$$

Having shown that equation (27a) can actually solve the model, I consider how we can interpret my formal empirical model in the context of this more elaborate model. The connection is as follow: let $\mathbf{z}_t \equiv \mathbf{x}_t - L\,\mathbf{x}_{t-1}$ be the reduced-form residual of $\mathbf{x}_t$: the unexpected change vis-à-vis the previous period's information. Then (27a) reveals that these reduced-form residuals, or data *surprises*, are related by a constant linear function to forecast revisions about structural shocks:

$$\mathbf{z}_{t|t} - \mathbf{z}_{t|t-1} = \mathbf{M}\left(\boldsymbol{\xi}_{t|t} - \boldsymbol{\xi}_{t|t-1}\right).$$

Thus, under this more-elaborate model I interpret my high-frequency forecast revisions as revisions about recent news about each variable, rather than revisions about the *levels* of the variables themselves.

D Proof of proposition 1

A first pass for identifying monetary policy shocks with these two measures would be to "purge" the monetary surprise $\widehat{i}_t$ of its information content (η_t) by orthogonalizing $\widehat{i}_t$ to $\widehat{y}_t$. Proposition 1 states the conditions under which this procedure identifies a monetary policy shock.

Proposition 1. *Denote by r_t the residual from a linear projection of $\widehat{i}_t$ on $\widehat{y}_t$. Unless $r_t = 0\ \forall\ t$, then r_t is independent of information effects if and only if output expectations do not respond to*

monetary policy shocks, i.e., $\omega_\varepsilon = 0$.

Let

$$\widehat{i}_t = \phi_\varepsilon \widehat{\varepsilon}_t + \phi_\eta \widehat{\eta}_t \tag{30}$$

$$\widehat{y}_t = \omega_\varepsilon \widehat{\varepsilon}_t + \omega_\eta \widehat{\eta}_t. \tag{31}$$

The residual r_t (i.e., the "clean" monetary shock) is

$$r_t = i_t - \widehat{\beta} y_t$$

where

$$\widehat{\beta} = \frac{\text{Cov}(\widehat{y}_t, \widehat{i}_t)}{\text{var}(y_t)} = \frac{\phi_\eta \omega_\eta \sigma_\eta^2 + \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2}{\omega_\eta^2 \sigma_y^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2}$$

This residual is then a linear combination of the monetary and information shocks:

$$\begin{aligned} r_t &= \widehat{i}_t - \widehat{\beta}\widehat{y}_t \\ &= \phi_\varepsilon\, \varepsilon_t + \phi_\eta \eta_t - \widehat{\beta}(\omega_\varepsilon\, \varepsilon_t + \omega_\eta \eta_t) \\ &= \underbrace{(\phi_\eta - \widehat{\beta}\omega_\eta)}_{c_\eta} \eta_t + \underbrace{(\phi_\varepsilon - \widehat{\beta}\omega_\varepsilon)}_{c_\varepsilon} \varepsilon_t\,. \end{aligned}$$

This strategy then only provides a "clean" shock if $c_\eta \equiv \phi_\eta - \widehat{\beta}\omega_\eta = 0$.

$$\begin{aligned} & 0 = \phi_\eta - \widehat{\beta}\omega_\eta \\ \iff & 0 = \phi_\eta - \omega_\eta \frac{\phi_\eta \omega_\eta \sigma_\eta^2 + \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2}{\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2} \\ \iff & 0 = \phi_\eta(\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2) - \omega_\eta(\phi_\eta \omega_\eta \sigma_\eta^2 + \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2) \\ \iff & 0 = \phi_\eta \omega_\eta^2 \sigma_\eta^2 + \phi_\eta \omega_\varepsilon^2 \sigma_\varepsilon^2 - \phi_\eta \omega_\eta^2 \sigma_\eta^2 - \omega_\eta \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2 \\ \iff & 0 = \phi_\eta \omega_\varepsilon^2 \sigma_\varepsilon^2 - \omega_\eta \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2 \\ \iff & 0 = \omega_\varepsilon \sigma_\varepsilon^2 (\phi_\eta \omega_\varepsilon - \omega_\eta \phi_\varepsilon) \end{aligned}$$

The strategy thus provides clean shock in three cases.

1. First, the case where $\sigma_\varepsilon^2 = 0$ means that that there are no monetary shocks, so $r_t = 0 \ \ \forall \ \ t$, violating our asusmptions.

2. The knife-edge case with $\phi_\eta \omega_\varepsilon = \omega_\eta \phi_\varepsilon$ also results in $r_t = 0 \ \ \forall \ \ t$. To see this, note that this

assumption also implies that $c_\varepsilon = 0$:

$$\begin{aligned} c_\varepsilon &= \phi_\varepsilon - \widehat{\beta}\omega_\varepsilon \\ &= \phi_\varepsilon - \omega_\varepsilon \frac{\phi_\eta \omega_\eta \sigma_\eta^2 + \phi_\varepsilon \omega_\varepsilon \sigma_\varepsilon^2}{\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2} \\ &= \frac{\phi_\varepsilon(\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2) - \omega_\varepsilon \phi_\eta \omega_\eta \sigma_\eta^2 - \phi_\varepsilon \omega_\varepsilon^2 \sigma_\varepsilon^2}{\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2} \\ &= \frac{\phi_\varepsilon \omega_\eta^2 \sigma_\eta^2 - \omega_\varepsilon \phi_\eta \omega_\eta \sigma_\eta^2}{\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2} \\ &= \frac{\omega_\eta \sigma_\eta^2 (\omega_\eta \phi_\varepsilon - \omega_\varepsilon \phi_\eta)}{\omega_\eta^2 \sigma_\eta^2 + \omega_\varepsilon^2 \sigma_\varepsilon^2} \\ &= 0 \end{aligned}$$

Thus, with $c_\eta = c_\varepsilon = 0$, we have $r_t = 0 \ \forall \ t$.

3. The final possibility is that output expectations do not respond to monetary policy shocks, i.e., $\omega_\varepsilon = 0$.

E Text Analysis Appendix

E.1 Word Lists

Increase words abound absorb absorpt acceler access accru accumul adjunct advanc ampli amplifi append arisen augment becam becom bloom blossom bolster boom boost boundless bounti branch broaden build capit collect comeback cultiv decor deepen develop doubl elabor embellish empow empower enhanc enrich exce excel expand expans extend flourish fortifi further garnish gener grow grown growth heap heighten hoard improv increas inflat intensifi inund lucr magnifi matur maxim momentum nourish overflow overwhelm peopl piec pile prolif promot prosper quicken radiat rais renaiss rise run shoot spread strengthen supplement surg sweeten thrive weight widen

Decrease Words abat allevi amput atrophi cheapen collaps contract corrod corros counteract cut decay declin decompos decreas deplet depreci detract dim diminish discount discourag dispel dispens drain dwindl eas empti engulf erad eras erod eros exasper exhaust extermin fade fall falter insuffici languish leakag lighten lower melt minim pass purifi ration reced recess reduc reduct refin retard revers rid rot scarciti shrank shred shrink shrivel shrunk slow subsid subtract sunder tatter vanish wane weaken wilt wither worsen

Rise Words aris aros ascend ascent blast climb come elev flew fli float flood jump leap outreach peak rais rise rose scale soar stretch surfac well

Fall Words bag buri burst cave collaps crash descend dip dive doubl drop fall fell knock lower parachut plung rain sank set sink slid slide slip slump spray sprinkl stagger stumbl submerg sunk sunken swoon toppl torrent trip tumbl

List of cities BASEL BEIJING, BRASILIA, FRANKFURT, HONG KONG, JAKARTA, LONDON, MEXICO CITY, MOSCOW, MUMBAI, MUMBAI, NEW DELHI, OTTAWA, PRETORIA, RIO DE JANIERO, SAO PAULO, SINGAPORE, SYDNEY, TOKYO, TORONTO, WELLINGTON, ZURICH, KUALA LUMPUR,

E.2 Word List Construction

I start with a set of three "seed" words: `output`, `growth`, and `economy`. I train the word2vec algorithm of Mikolov et al. (2013) on a subset of a large corpus of newspaper articles: the *The New York Times Annotated Corpus* from the University of Pennsylvania's Linguistic Data Consortium. The full corpus contains 1.8 million articles from the New York Times between 1987 and 2007, each manually tagged by library scientists. The word2vec algorithm consists of constructing vector representations of words that, via a neural network, can predict a word in a set of text given the surrounding words. The algorithm is thus well suited to finding synonyms, hence its employment here. I trained the algorithm on 94,601 articles that were tagged as related to either economic output, prices, or labor markets.[51] With a vector representation of every word in the corpus, I sort words based on their distance to the average vector of my seed words.[52] The resulting list, along with the distances from the seed vector, is listed in table table 6.

E.3 Machine Learning Approach

While transparent, the approach for constructing high-frequency GDP expectations in the paper required that several, potentially subjective, choices be made. In this appendix, I pursue a machine-learning based approach to overcome this potential criticism. Specifically, I take the following steps to measure high-frequency GDP expectations.

- First, I collected the subset of the *New York Times Annotated Corpus* described in the previous appendix. This corpus formed my training set.
- To construct features from the text, I collect my list of GDP words (which I expand to include the first 7 words from table 6, since after 7 the words are less-obviously related to GDP). Call this $\mathcal{Y}^+$. I also collect the rising, falling, increasing, decreasing, strong, and weak word lists

[51] Specifically, I retained articles labeled economic conditions and trends, united states economy, prices, wages and salaries, layoffs and job reductions, production, or labor. Later in the paper I also construct inflation and unemployment expectations in high frequency: thus the prices and labor tags.

[52] Mikolov et al. (2013) highlight that summing and distracting the vector representations of each word results in meaningful vectors. For example, the authors find that subtracting the vector for `man` from the vector for `king` results in a vector that is very similar to the vector for `queen`. To compute the distance of two words with vectors x and y, I compute the cosine similarity between them: $x'y/[\sqrt{(x'x)} + \sqrt{(y'y)}]$.

Table 6: Extended GDP Word List

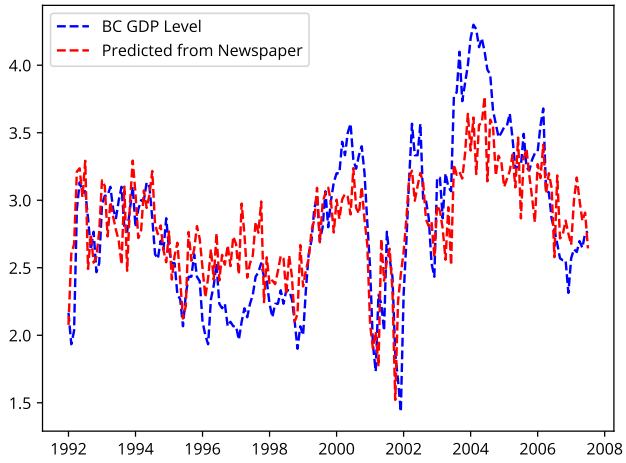
Term	Similarity to `output` + `growth` + `economi`
economic_growth	0.884380348
growth	0.844084993
economi	0.838640034
consumer_spending	0.768225532
output	0.73070604
recoveri	0.72648164
consumer_confidence	0.658283754
living_standards	0.642701291
anem	0.630283417
inflat	0.627934909

from the Harvard IV-4 dictionary. I collect the positive and negative word lists of Loughran and McDonald (2011). Call each of these lists ℓ_i, and their collection $\mathcal{L}$. Finally, I tag all words by their tense (past, present, future) using the grammatical sentence parser described earlier.

- I concatenate all articles written between the 8^{th} day of month $t-1$ and the 7^{th} day of month t. Features of these articles (described in the next bullet) are merged with Blue Chip expectations made in month t, which are typically made during the first week of month t.

- Within each month's concatenated list, I count the sum of all occurrences of words in each of the ℓ_i word list ($|\mathcal{L}|$ counts), all occurrences of words from the ℓ_i word lists that occur within 5 words of a word from $\mathcal{Y}^+$ ($|\mathcal{L}|-1$ counts here for each word from $\mathcal{Y}^+$), and all raw occurrences of words from $\mathcal{Y}^+$. When counting co-occurrences of $\mathcal{L}$ words and $\mathcal{Y}^+$ words, if `n't` or `not` occurs within the window, I flip the "sign" of the count (i.e., fall words become rise words). I count raw and co-occurrences (with words from $\mathcal{Y}^+$) of tense words. I also count three-way co-occurrences of tense $\times$ $\mathcal{L}$ words $\times$ $\mathcal{Y}^+$ words. All counts are normalized by the number of total sentences in a month. All in all, this leaves 263 features.

- I then use a LASSO regression to estimate the mapping from these features to the *level* of Blue Chip GDP expectations (the same summary statistic from the text: average forecasts over the next year). The penalty is chosen by 10-fold cross-validation. This figure shows the model fit:



The top 3 features that positively contribute to GDP expectations are

1. "Weak" mentions of "recovery," present tense
2. "Increasing" mentions of "output," present tense
3. "Strong" mentions of "consumer spending," any tense

and the top 3 features that contribute negatively are

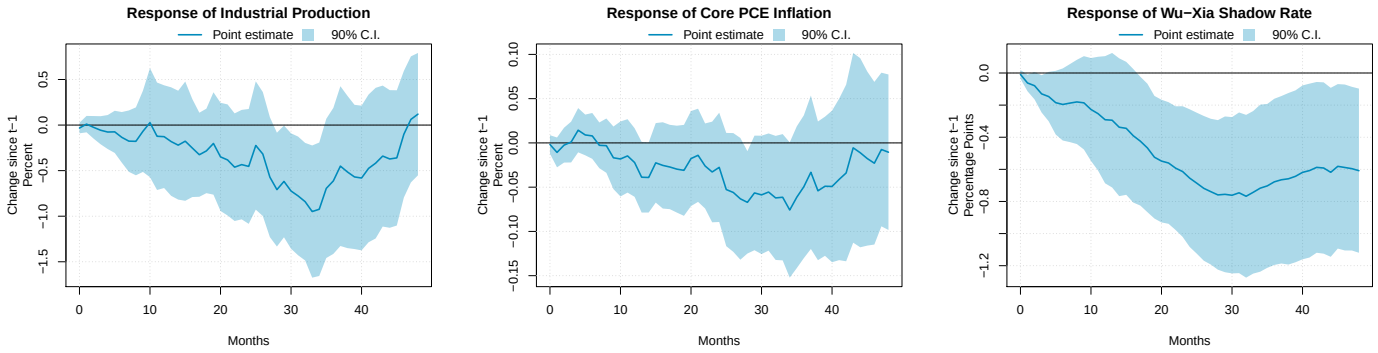
1. "Decreasing" mention of "economi," present tense
2. "Decreasing" word counts
3. "Weak" mention of "growth," future tense.

- I calculate the features that I calculated from the training set on all pre-meeting and post-meeting FOMC articles from my baseline Factiva dataset. I then apply the LASSO mapping to estimate implied GDP expectation for pre- and post-meeting articles—ω_{PRE_t} and ω_{POST_t}, respectively.
- Proceed the same way as the rest of the paper.

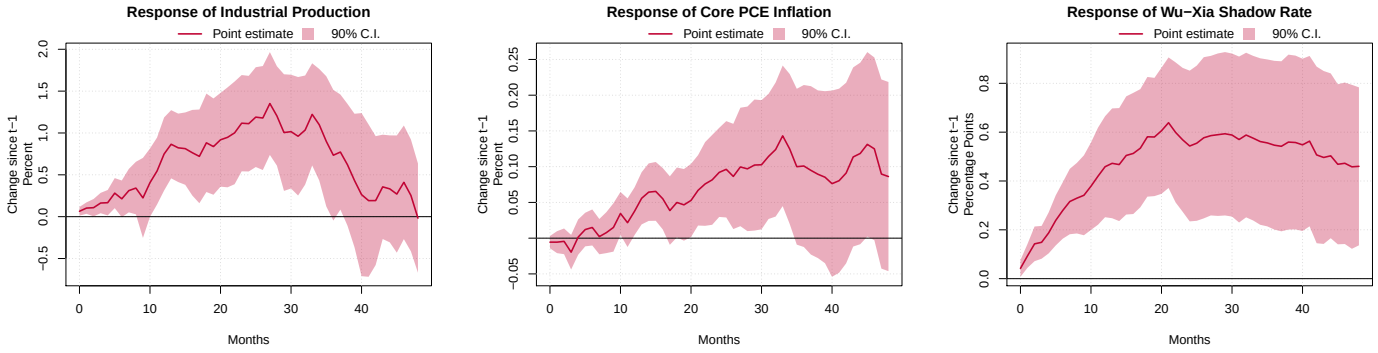
The structural impact matrix I estimate is

	Eurodollar	GDP
Monetary Path Shock	-0.38	0.86
Information Shock	1.18	0.68

where the variance in the second regime (relative to the first) of the monetary and path shocks are 1.03 and 0.69, respectively. Here are the impulse responses to the identified contractionary monetary path shock:



and the expansionary information shock



F Estimation and Bootstrap Procedure

Setup and Definitions My identification procedure begins by positing a model relating n observable variables in the vector x_t to n exogenous disturbances η_t given by

$$x_t = A\eta_t \qquad\qquad \eta_t \sim N(0, \Sigma_t),$$

where Σ_t is the (diagonal) covariance matrix of η_t. The time subscript on Σ_t indicates that the errors are heteroskedastic. Specifically, I posit that Σ_t can take two possible values depending on whether t falls in one of two disjoint regimes (formally, subsets $\mathcal{R}_1$ and $\mathcal{R}_2$) of my sample period. These values are given by

$$\mathbf{\Sigma}_1 \equiv \mathbb{E}[\eta_t \eta_t' \mid t \in \mathcal{R}_1]$$
$$\mathbf{\Sigma}_2 \equiv \mathbb{E}[\eta_t \eta_t' \mid t \in \mathcal{R}_2].$$

Identification This form of heteroskedasticity, together with the assumption that the $n \times n$ structural impact matrix A remains unchanged across regimes, implies the following moment conditions:

$$\Omega_1 = A\mathbf{\Sigma}_1 A' \tag{32a}$$

$$\Omega_2 = A\mathbf{\Sigma}_2 A' \tag{32b}$$

where $\Omega_i = \mathbb{E}[x_t x_t' \mid t \in \mathcal{R}_i]$ is the covariance matrix of the observable variables x_t in regime i. Thus, given estimates of Ω_1 and Ω_2 we can use the above moment conditions to identify $A, \mathbf{\Sigma}_1$, and $\mathbf{\Sigma}_2$.

The each of the symmetric matrices Ω_1 and Ω_2 each provide $n(n+1)/2$ empirical moments. With the normalization that $\mathbf{\Sigma}_1 = I_n$ (which is without loss of generality), there are n^2 parameters to identify in A and n in $\mathbf{\Sigma}_2$. Putting these together, there are $n(n+1)$ empirical moments and $n(n+1)$ parameters to estimate, so the system is exactly identified.

Estimation In practice, estimation of the moment conditions in equation (11) is complicated by the fact that ordering and sign of the columns (and elements of $\mathbf{\Sigma}_2$) are not identified. This does not present a challenge for point estimates, but it does present a challenge for inference. Thus, while a typical GMM estimation would rely on a numerical optimizer for estimation, the optimizer has no way of keeping "the same" shocks in the same order.

A numerical optimizer can, however, be circumvented (which is also nice for stability reasons). Notice that (again, maintaining the assumption that $\mathbf{\Sigma}_1 = I_n$)

$$\Omega_2\Omega_1^{-1} = A\mathbf{\Sigma}_2 A'(AA')^{-1} = A\mathbf{\Sigma}_2 A^{-1}.$$

The final expression resembles an eigendecomposition of $\Omega_2\Omega_1^{-1}$. In fact, any eigendecomposition of $\Omega_2\Omega_1^{-1}$ can be used as a starting point for estimation—this is convenient since fast and stable algorithms exist for performing these decompositions.

Specifically, let Q and Λ form an eigendecomposition of $\Omega_2\Omega_1^{-1}$:

$$\Omega_2\Omega_1^{-1} = Q\Lambda Q^{-1}. \tag{33}$$

Eigendecompositions are not in general unique; while the eigenvalues (the diagonal elements of Λ) are unique (up to ordering—an issue that is of no consequence for point estimates but will return when conducting inference), each eigenvector can be multiplied by a different scalar and the decomposition will be preserved.[53] However, information from the first regime can be used to pin down the unique scaling of the eigenvectors such that the moment conditions are satisfied. Set $A = QS$ where

$$S \equiv \sqrt{\operatorname{diag}\left[Q^{-1}\Omega_1 Q'^{-1}\right]},$$

[53]Let D be diagonal and $\widetilde{Q} \equiv QD$. Diagonal matrices commute, so $\widetilde{Q}$ and Λ form another eigendecomposition:

$$\widetilde{Q}\Lambda\widetilde{Q}^{-1} = QD\Lambda D^{-1}Q^{-1} = Q\Lambda DD^{-1}Q{-}1 = Q\Lambda Q^{-1} = \Omega_2\Omega_1^{-1}.$$

where the square root operates on each element of the embedded diagonal matrix. Then

$$
\begin{aligned}
AA' &= QSS'Q' && && \text{(34a)}\\
&= Q\text{diag}\left[Q^{-1}\Omega_1 Q'^{-1}\right]Q' && && \text{(34b)}\\
&= \Omega_1 && \text{see proposition 2.} && \text{(34c)}
\end{aligned}
$$

Letting $\mathbf{\Sigma}_2 = \Lambda$, we have

$$
\begin{aligned}
A\mathbf{\Sigma}_2 A' &= QS\Lambda S'Q' && \\
&= Q\Lambda SS'Q' && \text{diagonal matrices commute}\\
&= Q\Lambda\text{diag}\left[Q^{-1}\Omega_1 Q'^{-1}\right]Q' && \\
&= Q\Lambda\text{diag}\left[\Lambda^{-1}Q^{-1}\Omega_2 Q'^{-1}\right]Q' && \text{using (33)}\\
&= Q\Lambda\Lambda^{-1}\text{diag}\left[Q^{-1}\Omega_2 Q'^{-1}\right]Q' && \text{diagonal matrices commute}\\
&= \Omega_2 && \text{see proposition 2.}
\end{aligned}
$$

Thus, the moment conditions in equations (11) are satisfied by $A = QS$, $\mathbf{\Sigma}_1 = I$, and $\mathbf{\Sigma}_2 = \Lambda$.

As a brief aside, the fact that the heteroskedasticity-based identification (HBI) estimate of $\mathbf{M}$ consists of the eigenvectors of the ratio of the empirical covariance matrices $\Omega_2\Omega_1^{-1}$ reveals a connection to principal components analysis (PCA). To the best of my knowledge this relationship has not been described previously. The PCA estimate of $\mathbf{M}$ (factor loadings) consists of the eigenvectors of the covariance matrix of the observable variables—this covariance matrix is given by the full-sample Ω. Intuitively, this implies that PCA factors are designed to explain the largest amount of variation in the data with the fewest number of factors. By analogy, this means that the shocks estimated by HBI are those that can best explain the *relative* variance of the shocks in the two regimes, since they are based on the eigendecomposition of the *relative* covariance matrices.

Inference To conduct inference on the parameters estimates above (and functions of those estimates) I rely on a bootstrapping procedure. I begin by drawing observations from the two regimes with replacement, stratifying by the size of the regimes. Thus, I rely on a completely standard bootstrap, save for one computational difficulty. The estimation procedure above did not identify the ordering or sign of the columns of A.[54]

In order to increase the likelihood that the shocks from each bootstrap sample are in the same order and sign as the shocks that form my point estimates, I rely on an "aligning" procedure similar to that laid out in Clarkson (1979). Let $(A, \mathbf{\Sigma}_2)$ be the point estimates of the system and $(A^b, \mathbf{\Sigma}_2^b)$ be the estimates from a particular bootstrap sample b. I search over the set of all column permutation matrices $\mathcal{P}$ and the set $\mathcal{S}$ of n-dimensional diagonal matrices with elements in $\{-1, 1\}$

[54]The eigendecomposition did not identify the order, and any version of A with flipped column signs would have satisfied the moment conditions since these cancel out when taking AA'.

and define

$$\begin{aligned}\left(A^{b*}, \boldsymbol{\Sigma}_2^{b*}\right) = \underset{P\in\mathcal{P}, S\in\mathcal{S}}{\operatorname{argmin}} (1-\lambda)\left\|\widetilde{A}^b - A\right\|_F + \lambda\left\|\widetilde{\boldsymbol{\Sigma}}_2^b - \boldsymbol{\Sigma}_2\right\|_F \\ \text{s.t. } \widetilde{A}^b = A^b PS \\ \widetilde{\boldsymbol{\Sigma}}_2^b = \boldsymbol{\Sigma}_2^b P\end{aligned}$$

for some $\lambda \in [0,1]$. The elements $\left\{A^{b*}, \boldsymbol{\Sigma}_2^{b*}\right\}_b$ form my bootstrap distribution. In words, I rearrange the order and sign of the columns of A^b (and the corresponding order of the elements of $\boldsymbol{\Sigma}_2^b$) such that the distance between the rearranged matrices are closest (under the Frobenius norm) to the point estimates. In practice the λ matters little, as long as it is interior. I calculate confidence intervals using the percentiles of the bootstrap distribution.

I conclude the section by proving a matrix equality used earlier in this appendix.

Proposition 2. *In the context of equation* (34)*, the following equality holds:* $Q\,diag\left[Q^{-1}\Omega_1 Q'^{-1}\right]Q' = \Omega_1$ *as long as the elements of* Λ *are unique (i.e.* $\Lambda_{ii} \neq \Lambda_{jj} \forall i \neq j$*).*[55]

Proof. We first need access to the following relation for a diagonal matrix D with distinct diagonal d_i and matrix with unknown properties X:

$$DXD^{-1} = X \Longrightarrow \ X \text{ is a diagonal matrix.} \tag{35}$$

To see this, suppose that for $i \neq j$, $x_{ij} \neq 0$. Note that the ij^{th} element of DXD^{-1} is given by $d_i d_j^{-1} x_{ij}$. By assumption, $d_i d_j^{-1} x_{ij} = x_{ij}$. Since $x_{ij} \neq 0$, we can divide both sides by x_{ij} to get $d_i = d_j$ which contradicts the assumption that the diagonal elements of D are unique.

This result allows us to see that $Q^{-1}\Omega_1 Q'^{-1}$ is diagonal. Performing two rearrangements of equation (33), we have

$$\begin{aligned}Q'^{-1} &= (Q^{-1})' = (\Lambda^{-1}Q^{-1}\Omega_2\Omega_1^{-1})' = \Omega_1^{-1}\Omega_2 Q'^{-1}\Lambda^{-1} \\ Q^{-1} &= (\Omega_2\Omega_1^{-1}Q\Lambda^{-1})^{-1} = \Lambda Q^{-1}\Omega_1\Omega_2^{-1}.\end{aligned}$$

With these we can see that

$$\begin{aligned}Q^{-1}\Omega_1 Q'^{-1} &= \Lambda Q^{-1}\Omega_1\Omega_2^{-1}\Omega_1\Omega_1^{-1}\Omega_2 Q'^{-1}\Lambda^{-1} \\ &= \Lambda Q^{-1}\Omega_1 Q'^{-1}\Lambda^{-1}\end{aligned}$$

which, by equation (35), implies that $Q^{-1}\Omega_1 Q'^{-1}$ is diagonal.

[55] Note that this condition on Λ is implied by the heteroskeasticity-based assumptions that no two shocks change in the same proportion.

Next we establish notation. Let E_k be an $n \times n$ matrix whose elements e_{ij} are given by

$$e_{ij} = \begin{cases} 1 & i = j = k \\ 0 & \text{otherwise.} \end{cases}$$

For intuition, $E_k X$ zeros out all but the i^{th} row of X, and XE_k zeros out all by the i^{th} column of X. Note that for a diagonal matrix D, when $i \neq j$,

$$E_i D E_j = E_j D E_i = 0_n \tag{36}$$

an $n \times n$ matrix of zeros. Note further that

$$\text{diag}(X) = \sum_{i=1}^{n} E_i X E_i.$$

With this notation and the above-established result, we can proceed to the following derivation.

$$\begin{aligned} Q\text{diag}\left[Q^{-1}\Omega_1 Q'^{-1}\right] Q' &= Q\left[\sum_i E_i Q^{-1}\Omega_1 Q'^{-1} E_i\right] Q' \\ &= \sum_i Q E_i Q^{-1}\Omega_1 Q'^{-1} E_i Q' \\ &= \left[\sum_i Q E_i\right] Q^{-1}\Omega_1 Q'^{-1} \left[\sum_i E_i Q'\right] \quad \text{by (36), since } Q^{-1}\Omega_1 Q'^{-1} \text{ is diagonal} \\ &= Q Q^{-1}\Omega_1 Q'^{-1} Q' \qquad\qquad \text{definition of } E_i \\ &= \Omega_1, \end{aligned}$$

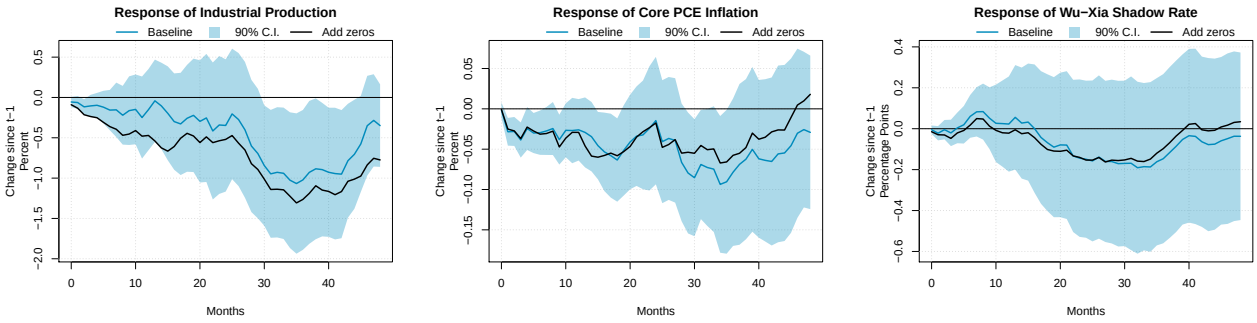
which is what we wanted to show. □

G Local Projections: Robustness

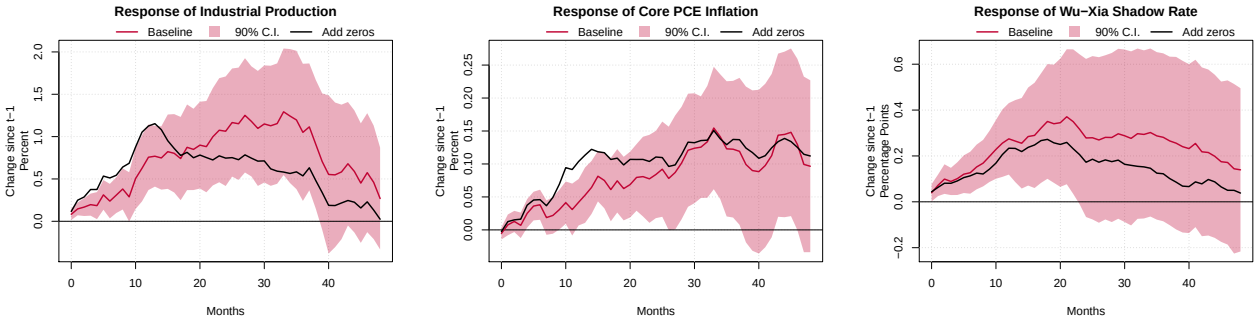
This section contains several variations on the choices made to estimate equation (16).

1. Adding months with no FOMC meetings (setting shocks to 0)

Responses to Contractionary Monetary Path Shock $\left(\beta_k^M\right)$

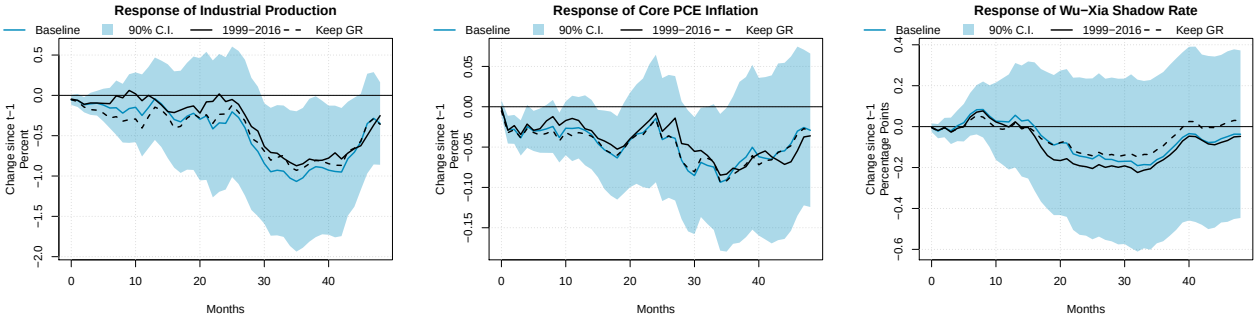


Responses to Expansionary Information Shock $\left(\beta_k^I\right)$

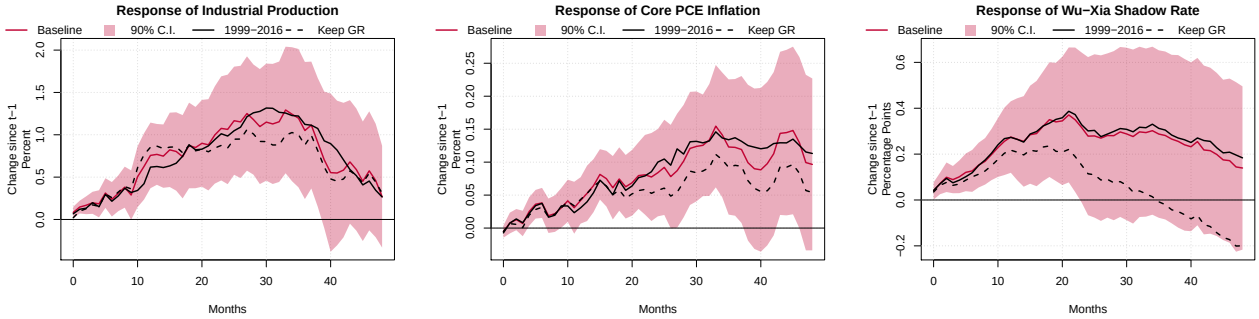


2. Sample period Stopping the estimation in 2015 ensures the same number of observations for each regression horizon, and drops COVID observations). The "Keep GR" response does not drop the 07/08–07/09 dates.

Responses to Contractionary Monetary Path Shock $\left(\beta_k^M\right)$

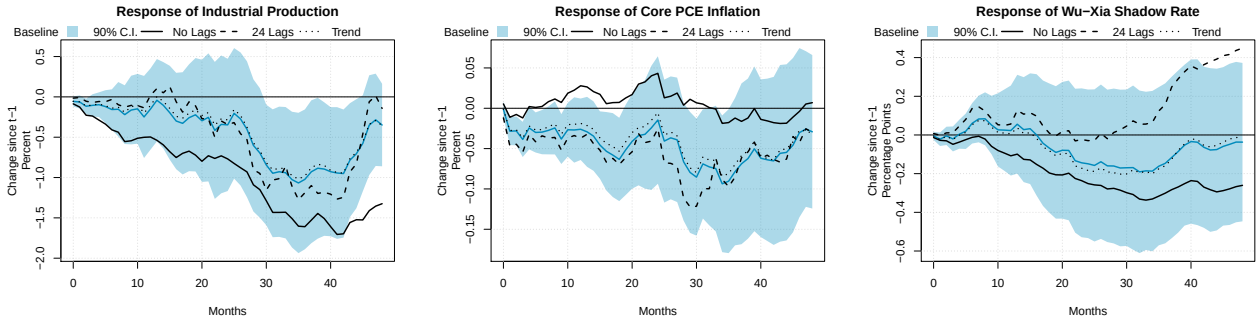


Responses to Expansionary Information Shock $\left(\beta_k^I\right)$

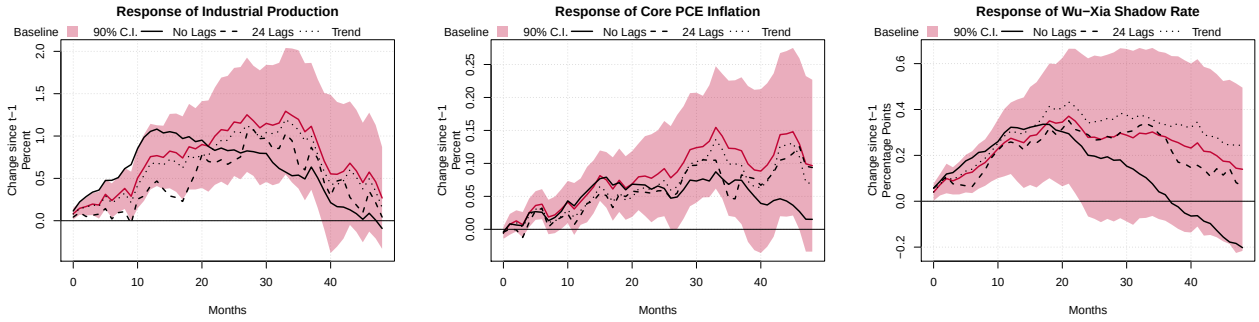


3. More (24) and fewer (0) lags as controls; adding a trend

Responses to Contractionary Monetary Path Shock $\left(\beta_k^M\right)$

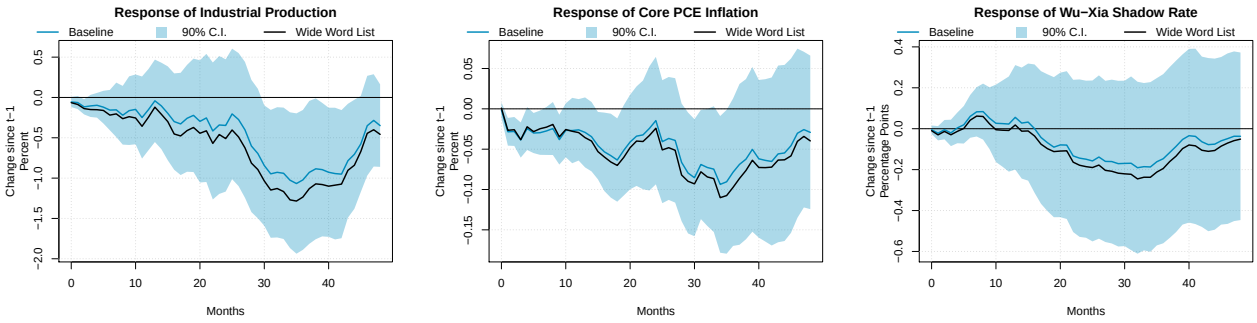


Responses to Expansionary Information Shock $\left(\beta_k^I\right)$

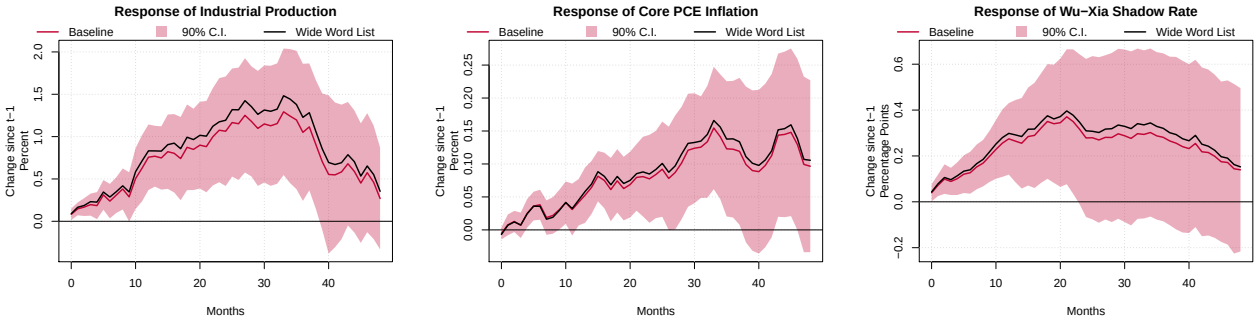


4. Using the ten-word word list

Responses to Contractionary Monetary Path Shock $\left(\beta_k^M\right)$

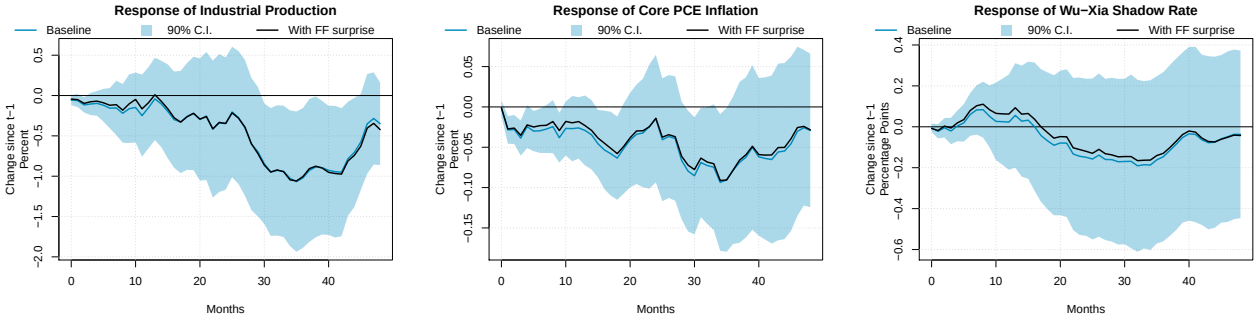


Responses to Expansionary Information Shock $\left(\beta_k^I\right)$

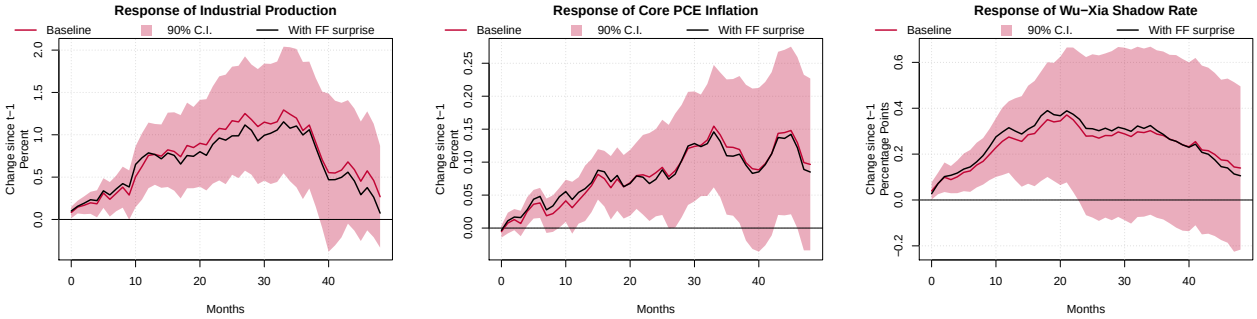


5. Tri-variate high-frequency system in Eurodollar futures, output forecast revisions, and the surprise in the Federal Funds rate, from section 5.1

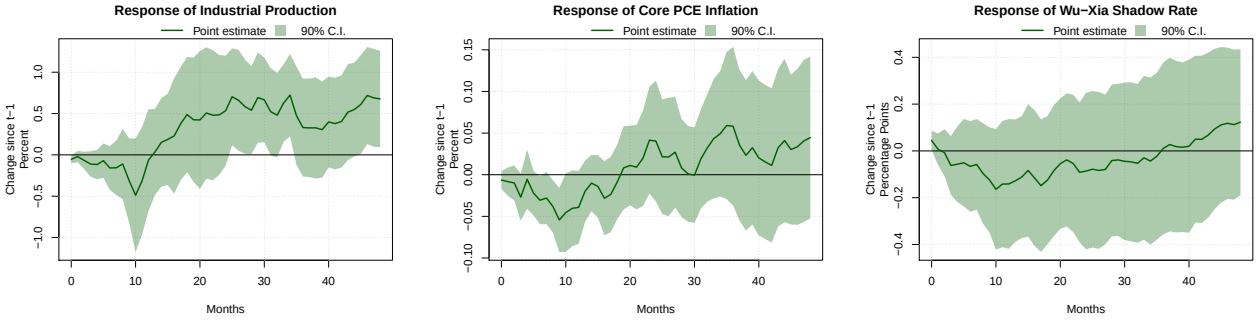
Responses to Contractionary Monetary Path Shock $\left(\beta_k^M\right)$



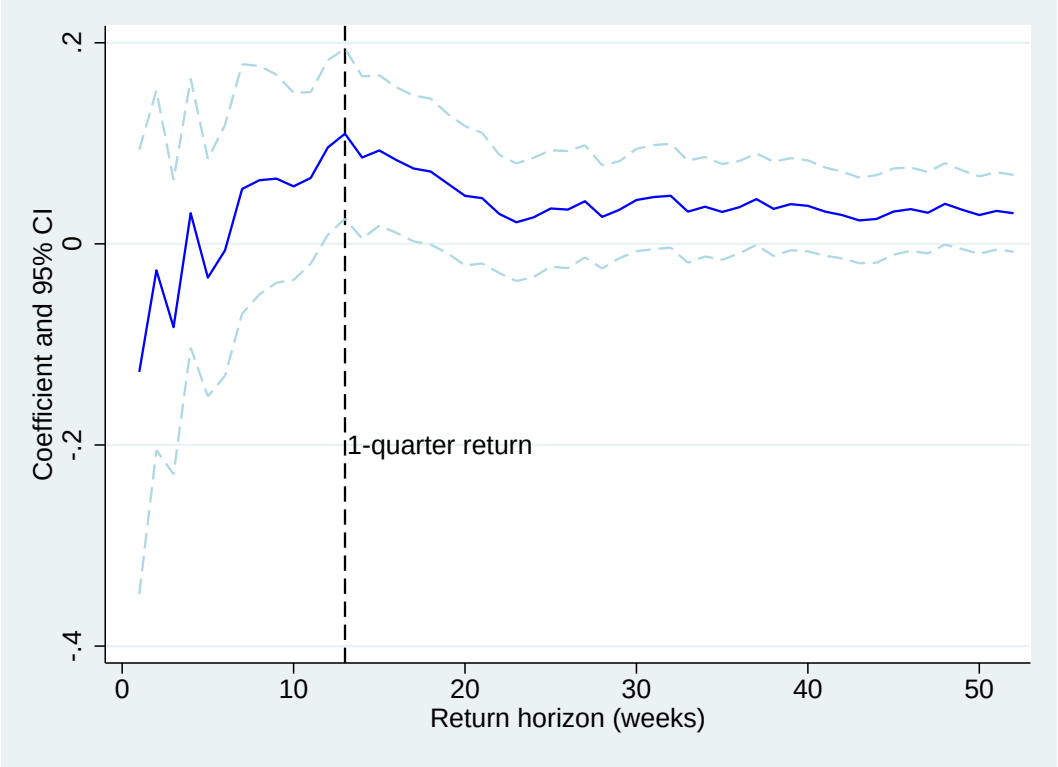
Responses to Expansionary Information Shock $\left(\beta_k^I\right)$



Responses to Contractionary Target Shock



H Predictability of Interest-Rate Surprises: A Comment on Bauer and Swanson (2020)



This figure presents estimates of the following regression

$$\widehat{i}_t^{\text{NS}} = \alpha_h + \beta_h(p_{t-1} - p_{t-7h}) + e_t$$

where $\widehat{i}_t^{\text{NS}}$ is the high-frequency interest-rate surprise of Nakamura and Steinsson (2018) (from their replication materials), and p_t is the log of the S&P 500 index on day t's market close. The sample consists of all days (t) with regularly-scheduled FOMC meetings between 1995 and 2015 (excluding the July 2008–July 2009 period) and after the first week of the month (i.e., the observations used when testing for the presence of information effects with Blue Chip data). The right-hand-side variable, $p_{t-1} - p_{t-7h}$, is thus the h-week return in the S&P 500 ending the day before each FOMC meeting. I estimate the regression for each h, which is shown on the x-axis. Bauer and Swanson (2020) present results using the 13-week return, which I highlight with a dashed line. This return horizon has the largest and most statistically significant coefficient.

I Constant Mapping between Shocks and Observables

I.1 Empirical Validity: Melosi (2017)

In this appendix, I explore whether the assumptions I made to identify structural shocks in my empirical application allow me to identify the relevant structural shocks in a theoretical context in which those assumptions are known not to hold. To that end, I study an extension of the model of Melosi (2017).

Economy The log-linear model equations are

$$y_t = \mathbb{E}_t[y_{t+1}] - (i_t - \mathbb{E}_t[\pi_{t+1}]) + (g_t - \mathbb{E}_t[g_{t+1}]) \tag{IS}$$

$$i_t = \phi_\pi (\pi_t + \xi_{\pi,t}) + \phi_x (x_t - a_t + \xi_{x,t}) + \xi_{m,t} \tag{TR}$$

$$\pi_t = (1-\theta)(1-\beta\theta)\sum_{k=1}^{\infty}(1-\theta)^{k-1}\mathrm{mc}_{t|t}^{(k)} + \beta\theta\sum_{k=1}^{\infty}(1-\theta)^{k-1}\pi_{t+1|t}^{(k)}. \tag{ICKPC}$$

where $\mathrm{mc}_{t|t}^{(k)} \equiv y_{t|t}^{(k)} - a_{t|t}^{(k-1)}$. The model is evidently similar to the standard three-equation New-Keynesian model, describing the evolution of GDP (y_t), interest rates (i_t), and inflation (π_t) according to, in order, an IS equation (IS), Taylor Rule (TR), and Phillips curve (ICKPC). The model's fundamental shocks are the household's discount factor shock (the "demand" shock, g_t), an aggregate productivity shock (a_t), an exogenous monetary shock (ξ_m), and two shocks reflecting the monetary authority's imperfect measurement of inflation and the output gap ($\xi_{\pi,t}$ and $\xi_{x,t}$, respectively). All shocks evolve according to mutually uncorrelated first-order autoregresive processes. The IS equation is standard, determined by perfectly-informed households' optimal consumption-saving decision and the equilibrium condition that firms must meet household demand in each period, having set prices in the beginning of each period. The Taylor rule is mostly standard, save for the Fed's imperfect reading of the state of the economy. The imperfect common knowledge Phillips curve is determined by firms' (indexed by j) optimal pricing behavior, based on their (imperfect) information sets ($\mathcal{I}_{jt}$, to be described shortly). Because firms are monopolistically competitive, each firm's optimal price is a function of aggregate demand *and* the average price of their competitors. As introduced and described by Woodford (2003), not having access to the their competitors' information sets, firms are left not only to form expectations of the current state of the economy, but also about their competitors' expectations which in turn depend on their competitors' expectations of their competitors' expectations, and so on. Thus the notation for these "higher-order expectations," used above, for a generic variable z_t:

$$z_{t|t}^{(0)} = z_t \qquad z_{t|t}^{(1)} = \int \mathbb{E}_{j,t}\left[z_t^{(0)}\right] d_j = \int \mathbb{E}_{j,t}[z_t] d_j \qquad \dots \qquad z_{t|t}^{(k)} = \int \mathbb{E}_{j,t}\left[z_t^{(k-1)}\right] d_j.$$

Of particular relevance for my empirical application is $z_{t|t}^{(1)}$—average expectations about z_t.

Information The information structure for a firm j is thus central for determining the behavior of aggregate variables. If, as traditionally assumed, firms were perfectly informed, the economy's Phillips curve would collapse to the familiar $\pi_t = \kappa \text{mc}_t + \beta \mathbb{E}_t[\pi_{t+1}]$. Firms are not perfectly informed, however, and use the Kalman filter to update expectations based on the information available to them.[56] I assume that each discrete period, indexed by integers t, is split into the following segments:

- at time t, all shocks are realized, and the Fed sets interest rates;
- at time t_F, the Fed makes its policy announcement, endowing firms with independent signals of the monetary (ξ_m) and demand (g_t) shocks, each buffeted by Gaussian noise, and firms update their expectations;
- at time t_P, firms observe their private productivity (a Gaussian deviate of aggregate productivity) and a private signal about demand, update their expectations, and set prices; and
- at time t_E, the representative household becomes perfectly informed and makes its consumption/savings decision, and firms produce to meet household demand

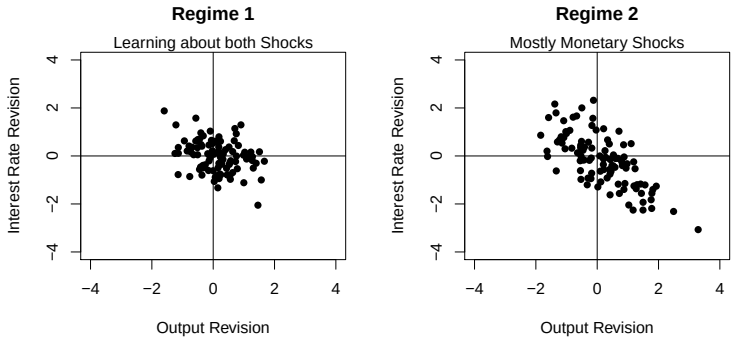
where $t < t_F < t_P < t_E < t + 1$. This setup follows that of Melosi, with two notable modifications. First, Melosi assumes that firms observe the Fed's announcement and their private signals simultaneously. Splitting these phases up allows me to measure how expectations are revised following the Fed's policy announcements. Second, more fundamentally, I make the Fed's announcements more-detailed than Melosi, who construes of those announcements as simply the revelation of i_t. At least within the context of my 1999–2019 empirical sample, that is an outdated representation. The specific shocks that I assume are signaled at t_F are consistent with my empirical findings—markets appear to learn primarily about demand and monetary shocks, and are able to distinguish the information provided about both shocks.[57]

Calibration My baseline calibration closely follows Melosi's estimates. While those estimated values speak to the *totality* of the information firms have, they cannot speak to the *source* of the information. By splitting up firm's information streams into the Fed announcement and private revelation, I thus have to take a stance on what information is learned when. I thus assume, as estimated by Melosi, that firms learn very little about demand from their private signal (signal-to-noise ratio of ≈ 10) whereas Fed announcements reveal demand with twice the precision of those

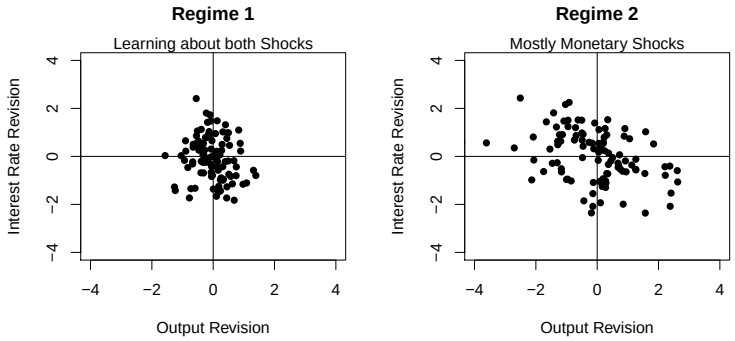
[56] This information structure is not necessarily the optimal one that firms would chose if they had a constraint on the extent to which they can process information. If we assume, however, that firms "in the real world" collect information in this way, then it may be reasonable to assume that Melosi's estimates capture the optimal values of parameters chosen by such rationally inattentive firms.

[57] If firms only observed i_t, they would have to infer which fundamental shock (demand, technology, or monetary policy) drove the interest rate decision, based on other private information they observe and their priors. Forecast revisions of these (mutually uncorrelated) structural shocks would therefore be correlated. My empirical results suggest that these forecast revisions are essentially uncorrelated (see figure 6).

Figure 15: Heteroskedasticity in the Model
Panel A: Baseline Model



Panel B: Common Knowledge Model



private signals. My baseline calibration for the signal-to-noise ratio of the monetary shock signal is about ten. I chose these values to broadly match my data on high-frequency forecast revisions.[58]

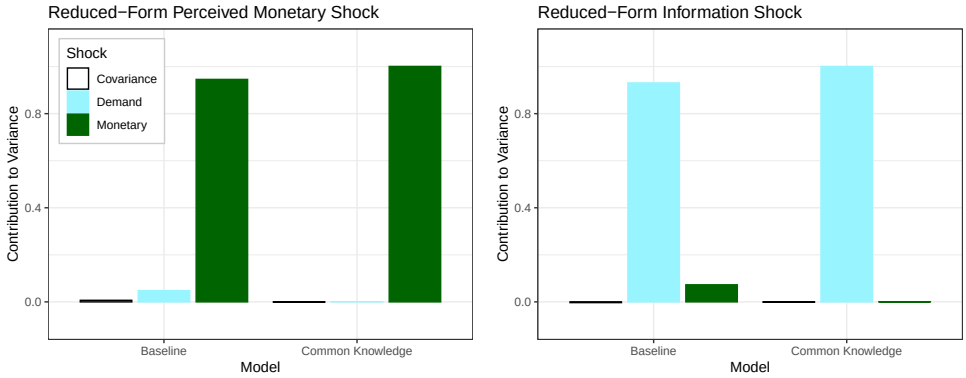
Experiment Ultimately I am interested in measuring the extent to which the shocks identified using the heteroskedasticity-based identification (HBI) assumptions resemble their theoretical counterparts. To assess this, I simulate the model under two regimes. For 5000 periods, I simulate the model using the baseline calibration described above. I then reduce the variance of the noise with which the Fed communicates about the monetary shock by a third, and simulate the model for 5000 more periods. This is meant to capture the "introduction of forward guidance" that I study empirically.[59] In each simulated series, I calculate forecast revisions made about the model's endogenous variables and exogenous shocks around the Fed's announcement.[60] Denote the changes in expectations about output and interest rates by $\widehat{y}_t$ and $\widehat{i}_t$. Note that firms only update expec-

[58] A more-satisfactory approach would be to measure, in Melosi's model, the mutual information that firms' information sets provide about each structural shock. I could then keep this level of mutual information fixed and chose how to partition it between t_F and t_P.

[59] Since the monetary shock is highly persistent in this model, this is a reasonable proxy for explicit forward guidance, since it signals future deviations from the systematic component of the Taylor rule.

[60] Formally, for a variable η_t, I calculate $\eta^{(1)}_{t|t_F} - \eta^{(1)}_{t|(t-1)_P}$—the change in the average forecast—consistent with my data.

Figure 16: Composition of the Heteroskedasticity Estimator



tations about the demand and monetary shocks following Fed announcements, since these are the only shocks that the Fed discusses.

Next, I append the two sets of simulated data and estimate two shocks using the HBI assumptions, applied to the vector of observable variables $\left[\widehat{y}_t, \widehat{i}_t\right]$. These are plotted in panel A of figure 15. Recall that the model's endogenous variables—including output and interest rates—will be a function of not only exogenous shocks, but increasingly higher-order expectations of those exogenous shocks. Thus, forecast revisions of y_t and i_t—and the resulting HBI shocks—will be linear combinations of forecast revisions of not only demand and monetary shocks, but also higher-order expectations thereof. The question is: *how much* of the variation in each HBI shock comes from forecast revisions about each structural shock? Because I can directly measure forecast revisions about $\{\xi_{m,t}, g_t, \xi_{m,t}^{(1)}, g_t^{(1)}, \xi_{m,t}^{(2)}, g_t^{(2)}, \ldots\}$, I project each HBI shock onto all of these forecast revisions, and can thus decompose the variance of each shock into (1) the amount explained by forecast revisions about the monetary shock (and higher-order expectations thereof), (2) the amount explained by forecast revisions about the demand shock (and higher-order expectations thereof), and (3) any covariance between (1) and (2).[61]

Figure 16 shows the variance decomposition of the "reduced-form perceived monetary shock" identified by heteroskedasticity (the shock with the larger variance in the second regime) and the "reduced-form information shock." In the baseline model, it is apparent that these HBI shocks uncover their structural counterparts remarkably well. The information shock is primarily made up of forecast revisions about the structural demand shock, while the perceived monetary shock primarily measures the structural monetary shock. That the covariance term is negligible is expected given the fact that firms observe independent signals about each structural shock.

[61] Formally, let $\widehat{s}_t$ be an HBI shock. I estimate the OLS regression $\widehat{s}_t = \underbrace{\sum_{k=0}^{10} \phi_{m,k} \widehat{\xi}_{m,t}^{(k)}}_{\text{monetary factors}} + \underbrace{\sum_{k=0}^{10} \phi_{g,k} \widehat{g}_t^{(k)}}_{\text{demand factors}}$. (Note that this projection has no error term.) I can then use the estimated ϕ coefficients to decompose the variance of $\widehat{s}_t$ into the variance arising from monetary factors, the variance arising from demand factors, and the covariance between monetary factors and demand factors.

Discussion: The Common Knowledge Benchmark To understand why the shocks identified with the HBI assumptions do not perfectly measure the model's structural shocks, it is useful to consider a variant of the baseline in which firms have common-knowledge, only forming expectations based on public signals.[62] There, the HBI and structural shocks perfectly coincide. This coincidence is verified in figure 16, where the HBI shocks were estimated using the data in panel B of figure 15.[63] To see why the shocks are correctly identified in this model, it is useful to consider how the model is solved. The solution of the model is assumed to take the form

$$x_t = M\eta_{t|t} \tag{37}$$

where x_t are the model's endogenous variables (y_t, i_t, and π_t), and $\eta_{t|t}$ are agents' perceptions of the model's structural shocks, and high-order expectations of those shocks.[64] Those expectations are assumed to follow an VAR(1) process

$$\eta_{t|t} = L\eta_{t-1|t-1} + N\xi_t \tag{38}$$

where ξ_t are innovations to the economy's structural shocks. The matrices M, L, and N form the model's solution, and are found by iterating back and forth between equations (37) and (38) until convergence. If firms do not observe endogenous variables, the solution method is even simpler: solve for L and N by solving firms' Kalman filtering problem, then solve for M a solver for linear rational expectations models (see Appendix C).

Writing equation (37) in expectations-revisions space allows us to see that forecast revisions about observable variables are linear combinations about structural shocks:

$$\widehat{x}_t = M\widehat{\eta}_{t|t}.$$

The question is therefore: does changing the noise with which signals are communicated affect M? In the common-knowledge model, the answer is no. The change in signal clarity affects firms' abilities to infer signals, but this does not change anything fundamental about their response to a *perceived shock of the same magnitude*. Instead, in the baseline (imperfect common knowledge) model, expectation formation plays a fundamental role, altering the relationship between firm j's price and (perceived) aggregate demand. In that model, the presence of idiosynratic signals causes

[62] The common-knowledge model therefore resembles that of Jia (2020). Currently I have implemented the common-knowledge model by replacing (ICKPC) with $\pi_t = \kappa \text{mc}_{t|t} + \beta\pi_{t+1|t}$ as in Melosi's perfect information model. Alternatively, I could follow Jia exactly and replace (ICKPC) with $\pi_t = \beta\theta\pi_{t+1|t} + (1-\theta)\pi_{t|t} + \kappa\theta\widetilde{y}_t$. For reasons I discuss below, this will not change my conclusions—HBI will still identify forecast revisions of the corresponding structural shocks.

[63] I make a few modifications to the baseline parameterization so that the simulated data continue to resemble the data from the baseline model. I shut off $\{a_t, \xi_{\pi,t}, \xi_{x,t}\}$, though this is likely not crucial and something to revisit. More pertinently, I reduce the autocorrelation of the monetary shock from 0.94 to 0.3—similar to Melosi's estimate of this parameter in his perfect information model. Without this, positive monetary shocks lead to decreases in i_t (an expected result—see Galí (2015)), so monetary and demand shocks "look" the same in terms of the sign of their impact responses on i_t and y_t. Finally, I reduce the signal-to-noise ratio of the monetary shock from 1.25 to 0.25.

[64] Thus, in principle $\eta_{t|t}$ is infinite-dimensional. Following Melosi, I cap higher-order expectations are at 10.

firms to remain confused as to the source of those signals, unclear as to whether they reflect aggregate or idiosynratic shocks. This results in firms placing different weights on expectations of different higher-orders when the precision of their signals changes, altering the "slope" of the Phillips curve. It is this sense in which the change is "fundamental" in the baseline model. The next section explores these issues in more detail.

I.2 Dispersed vs. Common Knowledge

The discussion in the text, and in the previous appendix, noted that models with dispersed information do not feature a constant mapping between forecast revisions of observable variables, and forecast revisions of structural shocks. In this appendix I analyze a simple model to explain this point. Note that the model in the previous appendix featured dispersed information. However, I showed that the shocks identified by heteroskedasticity uncovered their structural counterparts remarkably well.

I consider a model of the evolution of an asset price p_t, whose determination depends on the average expected future price of the asset and an exogenous fundamental:

$$p_t = \beta \int \mathbb{E}[p_{t+1} \mid \Omega_{j,t}]dj + \theta_t \qquad \theta_t = \rho\, \theta_{t-1} + \zeta_t \qquad \zeta_t \sim N(0, \sigma_\zeta^2) \tag{39}$$

where $\Omega_{j,t}$ is the information set of agent j for $j \in (0,1)$, and $\beta, |\rho| \in [0,1)$.[65] All agents receive mutually independent signals $s_{t,j}$ regarding the fundamental, which they never forget:

$$s_{t,j} = \theta_t + \nu_{t,j} \qquad \nu_{t,j} \sim \mathcal{N}(0, \sigma_\nu^2) \qquad \Omega_t = \Omega_{t-1} \cup \{\nu_{t,j}\}.$$

Introducing the notation $x^{(k)}_{t+j|t} = \int x^{(k-1)}_{t+j|t} dj$ with $x^{(0)}_{t+j|t} = x_t$, the model can be expressed by recursive substitution as

$$p_t = \sum_{k=0}^{\infty} \beta^k \theta^{(k)}_{t+k|t}. \tag{40}$$

The question of relevance for my empirical results is whether a change in $\nu_{t,j}$ (or σ_ζ^2) changes the mapping between forecast revisions regarding p_t, $p^{(1)}_{t|t} - p^{(1)}_{t|t-1}$, forecast revisions regarding ζ_t, $\theta^{(1)}_{t|t} - \theta^{(1)}_{t|t-1}$. The fairly complex form of (40) (specifically, the θ terms therein) suggests that this need not be the case. However, in the common knowledge case, with $\nu_{t,j} = \nu_{t,i},\ \forall (i,j)$, the mapping is constant. To see this, notice that in this case the model reduces to

$$p_t = \sum_{k=0}^{\infty} \beta^k \theta^{(1)}_{t+k|t} = \sum_{k=0}^{\infty} \beta^k \rho^k \theta^{(1)}_{t|t} = \frac{1}{1-\beta\rho} \theta^{(1)}_{t|t}$$

[65] This dynamic asset pricing problem shares features with to models studied by Townsend (1983), Morris and Shin (2006), and Woodford (2003). Nimark (2017) is a fantastic reference for understanding models of this form. The author proposes a solution method for generalized models of the form I present in equation (39), a special case of which is the model of Melosi (2017).

so that forecast revisions are

$$p_{t|t}^{(1)} - p_{t|t-1}^{(1)} = \frac{1}{1-\beta\rho}\left[\theta_{t|t}^{(1)} - \theta_{t|t-1}^{(1)}\right],$$

where the mapping $\frac{1}{1-\beta\rho}$ evidently does not depend on the variance of the fundamental or noise.[66]

[66]This is not to say that the relationship between the fundamental itself and the price level remains unchanged. Setting $\rho = 0$, the model's solution is $p_t = \left(\frac{1}{1-\beta\rho}\right)\left(\frac{\sigma_\zeta^2}{\sigma_\zeta^2+\sigma_\nu^2}\right)(\theta_t + \nu_t)$. The conclusion without $\rho = 0$ reveals a solution whereby p_t is a function of all realized fundamental and noise shocks, with a mapping onto prices that also depends on the variances of those shocks.