

The limits of forward guidance*

Jeffrey R. Campbell[†]
Jonas D. M. Fisher

Filippo Ferroni
Leonardo Melosi

February 15, 2019

Abstract

Forward guidance allows the Fed to influence private-sector expectations and thereby potentially improve macroeconomic outcomes. This tool's viability depends on the horizon over which the Fed is able to communicate its intentions and its influence on expectations over that horizon. We develop a tractable model of imperfect central bank communications and use it to measure how effectively the Fed has managed private-sector expectations about the future path of the federal funds rate and how its imperfect communications have influenced macroeconomic outcomes. Standard models assume the central bank has perfect control over the private sector's expectations about the policy rate up to an arbitrarily long horizon and this is the source of the so-called "forward guidance puzzle." Our estimated model suggests that the Fed's ability to affect expectations at horizons that are sufficiently long to give rise to the forward guidance puzzle is substantially limited. We also find that imperfect communication has influenced the propagation of forward guidance and is a source of macroeconomic volatility.

JEL Classification Numbers: E0

Keywords: monetary policy, forward guidance puzzle, central bank communication, business cycles

*The views expressed herein are the authors' and do not necessarily represent those of the Federal Reserve Bank of Chicago or the Federal Reserve System.

[†]All authors Federal Reserve Bank of Chicago. In addition Campbell is affiliated with CentER, Tilburg University, and Melosi is affiliated with the European University Institute. E-mail addresses are: jcampbell@frbchi.org; jfisher@frbchi.org; fferroni@frbchi.org; lmelosi@frbchi.org.

1. Introduction

Since the onset of the financial crisis informing the public about its future intentions has become central to Fed communications. Even before the crisis post-meeting statements and speeches by Fed officials included forward looking language intended to clarify its intentions about policy in the future. Such communications have become known as *forward guidance*. The Fed’s forward guidance during the period in which the policy rate was at its effective lower bound (ELB) is widely viewed to have been explicitly designed to lower expectations of future short term rates. By doing so far enough into the future such communications may have had a material effect on the interest rates that govern longer term investment projects and boosted economic activity more broadly.

In the standard New Keynesian (NK) framework forward guidance allows the central bank to influence private-sector expectations and thereby to potentially improve macroeconomic outcomes. However the viability of this policy tool ultimately depends on the horizon over which the Fed is able to influence expectations and the power of its communications to influence expectations within that horizon. In this paper we develop an empirically tractable NK model of central bank communications and use it to measure how effectively the Fed’s has influenced private sector expectations about the future path of the federal funds rate and the macroeconomic consequences of their communications over the period 1993-2016.

Standard NK models assume the central bank has perfect control over private sector expectations about the policy rate up to an arbitrarily long horizon. [Eggertsson and Woodford \(2003\)](#) and [Krugman \(1998\)](#) rely on this assumption to formulate policies to combat the severe consequences of monetary policy being constrained by the ELB. However such perfect control gives rise to implausible implications. In particular it is the source of the “forward guidance puzzle” highlighted in an empirical context by [Del Negro, Giannoni, and Patterson \(2015\)](#). The forward guidance puzzle arises from a theoretical implication common to all standard NK models: the near term effects of a commitment to maintain an interest rate peg for an additional period increases without bound with the horizon of the peg. That forward guidance would have such an effect is clearly implausible and this has sparked a large and growing literature exploring modifications to the NK model that dampen the ef-

fects of forward guidance as traditionally conceived. Key contributions include [Del Negro et al. \(2015\)](#) who emphasize finite lifetimes, [Gabaix \(2016\)](#) and [Angeletos and Lian \(2018\)](#) who consider bounded rationality, [McKay, Nakamura, and Steinsson \(2015\)](#) and [Hagedorn, Luo, Manovskii, and Mitman \(2019\)](#) who study incomplete markets, [Campbell, Fisher, Justiniano, and Melosi \(2017\)](#) who focus on preferences for government bonds, [Campbell and Weber \(2019\)](#) who study imperfect credibility, [Farhi and Werning \(2017\)](#) who explore the interaction of bounded rationality and incomplete markets, and [Kaplan, Moll, and Violante \(2018\)](#) who show study the implications of agent heterogeneity.

The ultimate source of the effects of forward guidance in standard NK models as well as most of the extensions mentioned above is the assumption that the central bank is able to communicate to the public clear and credible commitments to particular interest rate paths. Is it plausible that the central bank has the ability to conduct such communications? If there are failures to communicate clearly and credibly what are their implications for the conduct of policy? Our framework allows us to address these questions. Consistent with the previous literature we conceive of forward guidance as communications about future deviations from the central bank’s policy rule. The central bank cannot perfectly communicate these deviations in advance but instead is limited to sending noisy signals about them each period. The public’s expectations are influenced as they learn from the signals about what the policy deviations actually will be.

We view the noise in the signals as reflecting two key challenges to the communication of forward guidance. The most obvious challenge is that the words used by central bankers may confuse the public. One example is the “taper tantrum” episode when bond markets seemingly over-reacted to remarks by Chairman Bernanke in May 2013 about winding down the Fed’s bond purchases. This reaction was puzzling to many observers since earlier Fed communications should had prepared markets for this eventuality. Another challenge is that the central bank simply does not know how it might want to deviate from its established rule in the future. For example, the policy rule assumed in standard models does not address risk management. Even when the policy rate is far from the ELB central banks might want to deviate from their implicit rules to guard against risks to the economic outlook. The history

of Fed communications is replete with examples of such behavior.¹

We reach our conclusions about central bank communications by embedding our proposed central bank communications technology within a medium scale NK model estimated using a rich array of macroeconomic data. We force our model to confront both data on aggregate activity and data on overnight interest rate futures. As a consequence expected future rates in our model and our estimates of forward guidance reconcile these data. One implication of this approach is that our estimated model does not exhibit the empirical forward guidance puzzle highlighted by [Del Negro et al. \(2015\)](#). Another implication of incorporating the futures rates in our estimation is that while we use linear methods to solve and estimate our model agents at no time expect future rates to violate the ELB.

Our estimated model suggests that the Fed’s ability to affect expectations at horizons that are sufficiently long to give rise to the forward guidance puzzle is substantially limited. The difficulties inherent in communicating complicated decisions about the future path of interest rates may be too great for forward guidance to be the powerful instrument predicted by standard NK models. We also find that imperfect communications heavily influence the propagation of monetary shocks. First, the response of the economy to a shock that changes the policy rate path is substantially different under imperfect communication. Second, unintended communication that contains no information about future policy deviations leads to sizable macroeconomic volatility. Finally, forward guidance raises macroeconomic volatility compared to no communication.

Our analysis exploits the observational equivalence established by [Chahrour and Jurado \(2018\)](#). In particular there is an observationally equivalent version of our model in which the central bank perfectly communicates *news* about future interest rates. The literature thus far has almost exclusively modeled forward guidance as news. News consists of the central bank perfectly communicating future deviations from the policy rule which are known to be subject to revisions at later dates. Why focus on signal extraction rather than a model with news? The news representation is easier to estimate for sure and we exploit that fact here. However, if one then wants to use the estimated model to evaluate the effectiveness of

¹See [Evans, Fisher, Gourio, and Krane \(2015\)](#) for examples of risk management considerations in statements and minutes of the Federal Open Market Committee.

forward guidance as a policy tool, the model of central bank communications we study seems more appropriate. The reason is that by focusing on news shocks one abstracts from how effectively central bank announcements affect private expectations. Our model addresses the communications challenges inherent in forward guidance directly and therefore allows us to shed light on the ability of the central bank to exploit the power of forward guidance. Of course observational equivalence means that the data cannot select one approach over the other. Nevertheless the signaling approach seems better suited to interpreting the evolution of expectations and therefore the nature of the communications challenges faced by central banks.

The importance of communications in the transmission of monetary policy has long been acknowledged. For example, [Woodford \(2003\)](#) emphasizes that successful monetary policy is not so much a matter of effective control of overnight interest rates as it is the shaping of market expectations about how interest rates, inflation, and output evolve over the coming years. One of our objectives is to evaluate the Federal Reserve's ability to shape these expectations. There are several papers that study the central bank's ability to shape expectations. These papers are primarily theoretically oriented and generally eschew quantitative implications. Two key contributions are [Eusepi and Preston \(2010\)](#) and [Andrade, Gaballo, Mengus, and Mojon \(2018\)](#). The former develop a model with learning to study the link between central bank communication and the anchoring of inflation expectations. The latter study the normative implications of forward guidance in a NK model where agents have heterogeneous beliefs about the strength of the central bank's commitment.

There are a few papers that use structural models to study the quantitative implications of forward guidance. [Campbell et al. \(2017\)](#) use a medium scale NK framework along the lines of [Christiano, Eichenbaum, and Evans \(2005\)](#) and [Smets and Wouters \(2007\)](#) to quantify the macroeconomic effects of forward guidance after the onset of the Great Recession. Their analysis focuses on news about future interest rates and therefore has little to say about the ability of the central bank to communicate its future intentions. [Bianchi and Melosi \(2017\)](#) use a NK model to evaluate the welfare implications of forward looking communications. Their empirical framework is not as rich as ours and in particular does not exploit the information contained in market expectations of future interest rates which is central to our

analysis. Finally, [Nakamura and Steinsson \(2018\)](#) develop a stylized NK model in which they estimate the strength of Delphic forward guidance as defined by [Campbell, Evans, Fisher, and Justiniano \(2012\)](#). They do not pursue likelihood estimation of their model given its relatively small size. None of the papers in this literature empirically evaluate the central bank's ability to steer private sector expectations at different horizons as we do.

Another related line of research aims at drawing normative conclusions about how the central bank should communicate. [Morris and Shin \(2002\)](#) introduce a class of models with dispersed information to study the effects of central bank communications. They show that policy communication may hinder social welfare if private information is sufficiently accurate. [Angeletos and Pavan \(2007\)](#) show that this result emerges only under particular assumptions about the kind of externalities assumed in the payoff structure. This literature focuses on very stylized theoretical models. Furthermore, with the exception of [Melosi \(2016\)](#), it does not investigate these models empirically.

Our work is also related to the reduced form empirical literature launched by [Kuttner \(2001\)](#) which uses event studies to identify unexpected FOMC policy actions and their effects on macroeconomic outcomes and asset prices. [Gürkaynak, Sack, and Swanson \(2005\)](#) extended that methodology to identify forward guidance shocks which leave the current policy rate unchanged while impacting current expectations of its future values. They found substantial impacts of near-term (one to six quarters out) forward guidance shocks on two, five, and ten year Treasury bond yields, and [Campbell, Evans, Fisher, and Justiniano \(2012\)](#) verified that this pattern continued while the FOMC set its policy rate at the ELB. The large and growing event study literature has not settled on an interpretation of these long horizon effects. Our analysis strongly suggests that they are not due to the ability of the Fed to shape expectations over such long horizons.

The remainder of the paper proceed as follows. In the next section we describe our proposed communications technology. We then describe the fully specified model in which we embed this technology and how we estimate it. The remaining sections focus on two sets of findings. First we discuss the information flows implied by our estimates. Second we discuss the propagation of forward guidance shocks under various assumptions about communications.

2. Central bank communication

This section describes central bank technology for communicating to private agents that in the following section we embed in a fully specified dynamic general equilibrium model. The environment consists of identical utility maximizing agents who trade in a market for one-quarter government bonds that have a risk free rate of return set exogenously by the central bank. We assume the central bank sets the date t one-quarter interest rate on government bonds, R_t , according to

$$R_t = g_t(R_{t-1}, \pi_t^{gap}, y_t^{gap}) + \theta_t, \quad (1)$$

where g_t is the possibly time-varying *policy rule* which depends on the lagged policy rate and the central bank’s measures of the deviations of inflation from its objective and a variable called the output gap, π_t^{gap} and y_t^{gap} . Contemporaneous deviations of the policy rate from the rule, θ_t , are exogenous, stationary and zero mean Gaussian random variables which are serially correlated up to the H -th lag. We allow for serially correlated monetary policy shocks to be consistent with evidence of serially correlated deviations from estimated policy rules and evidence that changes in overnight interest rate futures before and after the release of FOMC statements to the public are correlated across their term structure. The private sector observes R_t and understands that policy is set according to (1). The central bank perfectly communicates to the private sector the policy rule g_t and its measures of the gaps.

At each date t the central bank also communicates *forward guidance*. Forward guidance consists of providing information to the private sector about future deviations of the policy rate from the rule up to $H \geq 0$ periods ahead, θ_{t+h} , $h = 0, 1, 2, \dots, H$. These future deviations from the rule can be interpreted as the central bank implementing a modestly different policy path than what its policy rule would otherwise stipulate, along the lines discussed by [Leeper and Zha \(2003\)](#) and [Laséen and Svensson \(2011\)](#).² This notion of forward guidance is Odyssean in the sense described by [Campbell et al. \(2012\)](#). Since we have assumed the central bank perfectly communicates both its rule and measures of the gaps we exclude any role for [Campbell et al. \(2012\)](#)’s notion of Delphic forward guidance.

²We use the modifier “modestly” to indicate that the deviations from the rule are not sufficiently large to lead agents to change their view that the central bank is committed to the policy rule g_t .

The communications technology is as follows. Each period the central bank communicates to the private sector an $(H + 1) \times 1$ vector of noisy signals $s_t = [s_t^h]$ about the current ($h=0$) and up to H -period-ahead future deviations from the rule $\boldsymbol{\theta}_t = [\theta_{t+h}]$. The signals are specified as:

$$s_t = \boldsymbol{\theta}_t + v_t \tag{2}$$

where the vector $v_t = [v_t^h]$ denotes noise. The noise is Gaussian with mean zero and variance-covariance matrix Ξ_v . Since agents observe current and lagged policy rates as well as the central bank's gaps they observe θ_t as well so that $v_t^0 = 0$. The remaining noise should be interpreted as comprising two components. The first component represents the central bank's imperfect knowledge about what it will do in the future. For example, unforeseen events such as the near failure of Long-Term Capital Management in September 1998 or the onset of the First Gulf War in January 2001 led the Fed to reassess risks to the macroeconomic outlook and lower the federal funds rate seemingly below the level indicated by its contemporaneous measures of the output and inflation gaps. The second component represents miscommunication to the private sector about the central bank's true intentions.

The private sector observes the signals, knows the stochastic processes governing $\boldsymbol{\theta}_t$ and v_t , and updates its beliefs about $\boldsymbol{\theta}_t$ in a Bayesian fashion. The Gaussian structure of the shocks implies that it is optimal for agents to update expectations using the Kalman Filter with Kalman gain determined by the known and constant variance-covariance matrices of policy deviations and noise. Specifically, expectations following the release of date t signals are updated as follows

$$\begin{aligned} E_t \boldsymbol{\theta}_t &= E_{t-1} \boldsymbol{\theta}_t + \kappa \cdot (s_t - E_{t-1} \boldsymbol{\theta}_t) \\ &= E_{t-1} \boldsymbol{\theta}_t + \kappa \cdot (\boldsymbol{\theta}_t + v_t - E_{t-1} \boldsymbol{\theta}_t). \end{aligned} \tag{3}$$

where

$$\kappa = \Xi_\theta [\Xi_\theta + \Xi_v]^{-1} \tag{4}$$

and Ξ_θ denotes the variance-covariance matrix of current and future deviations from the rule $\boldsymbol{\theta}_t$ prior to receiving the signals s_t . At any date t conditional expectations are formed using

the full structure of the general equilibrium model specified in the next section.

The conditional variance-covariance matrix of future deviations from the policy rule, Ξ_θ , does not vary over time because the agents' signal extraction problem is stationary. To see this notice that in every period t , agents have received one signal about the H -period-ahead deviation θ_{t+H} ; two signals about the $(H - 1)$ -period-ahead deviation θ_{t+H-1} (one today and the other yesterday), and so on. Given that we assume the noise is a stationary random variable, it follows the horizon of each central bank announcement is a sufficient statistic for tracking agents' beliefs about θ_{t+h} over time.³ Therefore the variance-covariance matrix of the vector of deviations θ_t conditional on publicly available information at date $t - 1$ is time-invariant.

3. General equilibrium model

We investigate how the communication of forward guidance influences its effects by embedding the communications technology described in the previous section within the medium-scale dynamic stochastic general equilibrium model described in [Campbell et al. \(2017\)](#). Since much of the model's specification is familiar and described in that paper we confine our discussion to an abbreviated description of the model and emphasize the key features that are integral to our measurement of forward guidance.

3.1. Households

The economy consists of a large number of identical, infinitely lived households with preferences described by the lifetime utility function

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[\varepsilon_t^b \frac{[(C_t - \varrho \bar{C}_{t-1}) (1 - \vartheta H_t^{1+\gamma_H})]^{1-\gamma_C} - 1}{1 - \gamma_C} + \varepsilon_t^s L \left(\frac{B_{t+1}}{P_t R_t} \right) \right].$$

Here C_t denotes the household's date t consumption purchased in the competitive final goods market at nominal price P_t , \bar{C}_t denotes *aggregate* per capita consumption (which is equal to

³The framework being linear and Gaussian is, of course, key for this stationary result to arise since uncertainty is not affected by the realization of signals.

C_t in equilibrium) and H_t denotes hours worked.⁴ The parameters γ_C and γ_H are both non-negative, $\beta \in (0, 1)$ and $\varrho \in [0, 1)$. The parameter ϑ is used to normalize hours worked to one in steady state. The argument of the increasing and concave function L , $B_{t+1}/(P_t R_t)$, is the consumption value of one-period government bonds purchased by the household at date t and carried into date $t + 1$ where B_{t+1} denotes the quantity of nominally denominated bonds.

The *discount factor shock* ε_t^b has been shown by [Justiniano, Primiceri, and Tambalotti \(2010a\)](#) and others to be a key driver of consumption fluctuations. It is often used, for example by [Eggertsson and Woodford \(2003\)](#), as a shock that drives the policy rate to the ELB and so it is particularly relevant for our analysis. We assume ε_t^b evolves according to the stationary process given by

$$\ln \varepsilon_t^b = \rho_b \ln \varepsilon_{t-1}^b + \eta_t^b, \eta_t^b \sim N(0, \sigma_b^2).$$

Including preferences for government bonds has important implications for our measurement of forward guidance.⁵ First, in standard NK specifications R_t equals the return to installed capital in steady state while including a preference for government bonds allows them to differ as they do empirically. As such they potentially lead to more plausible estimates of the model's structural parameters. Second, as discussed by [Campbell et al. \(2017\)](#) and [Fisher \(2015\)](#), a spread between private and government rates of return introduces discounting into the household's linearized inter-temporal Euler equation for consumption. This discounting mitigates the forward guidance puzzle. Finally, [Fisher \(2015\)](#) shows the exogenous shock to these preferences, ε_t^s , provides a simple micro-foundation for [Smets and Wouters \(2007\)](#)'s ad hoc shock to the consumption Euler equation. This shock is crucial to the identification of empirical NK models since it is one of the few sources of co-movement between consumption, investment and hours.

⁴[Campbell et al. \(2017\)](#) work with a more general specification of preferences. It turns out that this generality is not important for our empirical results and so we abstract from it here.

⁵These preferences were first introduced into an empirical NK model by [Campbell et al. \(2017\)](#). [Krishnamurthy and Vissing-Jorgensen \(2011\)](#) used them to study the market for government securities. They are beginning to get wider attention in the literature. See, for example, [Auclert, Rognlie, and Straub \(2018\)](#) and [Michaillat and Saez \(2018\)](#).

We assume the *liquidity preference shock* to the preference for “safe and liquid” government bonds evolves according to the stationary process given by

$$\ln \varepsilon_t^s = (1 - \rho_s) \varepsilon_*^s + \rho_s \ln \varepsilon_{t-1}^s + \eta_t^s, \eta_t^s \sim N(0, \sigma_s^2).$$

The parameter ε_*^s determines the steady state spread between the rates of return on government and private bonds.

Households own the installed capital stock K_t which they accumulate using the technology

$$K_{t+1} = [1 - \delta(U_t)] K_t + \varepsilon_t^i \left[1 - S \left(\frac{I_t}{q_t I_{t-1}} \right) \right] I_t,$$

where I_t denotes gross investment purchased from investment good producers described below and S is an adjustment cost function that has the usual properties. The term q_t , defined below, corresponds to the growth rate of investment’s stochastic trend. The owners of installed capital control the intensity with which it is utilized, U_t , so that the effective supply of capital services in period t is $K_t^e = U_t K_t$. Increasing capacity utilization entails a cost in the form of faster depreciation via the function δ . This function is specified as in [Campbell et al. \(2017\)](#).

The technology for transforming investment goods into installed capital is subject to the shock ε_t^i . We assume this *investment-demand shock* evolves according to the stationary process given by

$$\ln \varepsilon_t^i = \rho_i \ln \varepsilon_{t-1}^i + \eta_t^i, \eta_t^i \sim N(0, \sigma_i^2).$$

[Justiniano, Primiceri, and Tambalotti \(2010b\)](#) find that this shock explains a substantial fraction of business cycle fluctuations in investment.

3.2. Goods Markets

Households own all goods producers. Final goods are produced using differentiated intermediate inputs with the usual Dixit-Stiglitz technology that is subject to shocks to the elasticity of substitution. Intermediate goods producers are monopolistic competitors who maximize profits subject to a standard Calvo pricing scheme with indexing. This and the shocks to

the elasticity of substitution translate to a *price mark-up shock*, λ_t^p , which evolves according to the stationary process given by

$$\ln \lambda_t^p = (1 - \rho_p) \ln \lambda_*^p + \rho_p \ln \lambda_{t-1}^p - \phi_p \eta_{t-1}^p + \eta_t^p, \eta_t^p \sim N(0, \sigma_p).$$

The parameter λ_*^p denotes the steady state mark-up. Notice that we allow for innovations to price markups to be a first-order moving average process.

Intermediate goods producer i produces its output Y_{it} using the technology:

$$Y_{it} = (K_{it}^e)^\alpha [A_t^Y H_{it}]^{1-\alpha} - A_t \Phi, \quad (5)$$

where H_{it} is composite labor rented from labor compositors (described below) in a competitive market, $\alpha \in (0, 1)$, and $\Phi > 0$ is the fixed costs of production, paid in final goods.⁶ The term A_t^Y is the level of the *neutral technology*. This is a non-stationary process with growth rate $\nu_t \equiv \ln(A_t^Y/A_{t-1}^Y)$ that evolves according to the stationary process given by

$$\nu_t = (1 - \rho_\nu) \nu_* + \rho_\nu \nu_{t-1} + \eta_t^\nu, \eta_t^\nu \sim N(0, \sigma_\nu^2),$$

where ν_* is the steady state growth rate of the neutral technology. We refer to ν_t as the *neutral technology shock*. The term A_t in (5) is the stochastic trend of equilibrium consumption and output measured in consumption units which equals $A_t^Y (A_t^I)^{\alpha/(1-\alpha)}$, where A_t^I is the level of the investment-specific technology described below with log growth rate denoted ω_t . The log growth rate of A_t is $z_t = \nu_t + \alpha\omega_t/(1 - \alpha)$.

Perfectly competitive firms supply investment goods to households using a linear technology that transforms final goods into investment goods at rate A_t^I . The growth rate of A_t^I evolves according to the stationary process given by

$$\omega_t = (1 - \rho_\omega) \omega_* + \rho_\omega \omega_{t-1} + \eta_t^\omega, \eta_t^\omega \sim N(0, \sigma_\omega^2).$$

The parameter ω_* is the mean growth rate of the investment-specific technology. In equilib-

⁶The parameter Φ is chosen so that intermediate good producers' profits are zero in steady state.

rium investment has a stochastic trend with log growth rate $q_t = \nu_t + \omega_t / (1 - \alpha)$.

3.3. Labor Markets

We adopt [Smets and Wouters \(2007\)](#)'s approach to introducing sticky wages when preferences are non-separable in consumption and labor. This approach involves a *wage mark-up shock* λ_t^w which follow a stationary process similar to λ_t^p :

$$\ln \lambda_t^w = (1 - \rho_w) \ln \lambda_*^w + \rho_w \ln \lambda_{t-1}^w - \phi_w \epsilon_{t-1}^w + \eta_t^w, \eta_t^w \sim N(0, \sigma_w^2).$$

3.4. Central Bank and Government

The central bank sets its policy rate, communicates with the public, and the public uses these communications to make its decisions in the way described in [Section 2](#). Our parametric specification of the monetary policy rule in [\(1\)](#) is

$$g_t(R_{t-1}, \pi_t^{gap}, y_t^{gap}) = \rho_R \ln R_{t-1} + (1 - \rho_R) \ln R_t^n,$$

where $\rho_R \in [0, 1]$ governs the degree of interest rate smoothing and R_t^n is the *notional* target interest rate given by

$$\ln R_t^n = \ln r_* + \ln \pi_t^* + \psi_1 \pi_t^{gap} + \psi_2 y_t^{gap}.$$

The constant r_* corresponds to the steady state real interest rate on government bonds and π_t^* is an exogenous *inflation drift shock* that could be interpreted as the central bank's intermediate target for inflation. The parameters $\psi_1, \psi_2 \geq 0$ correspond to the elasticities of the notional rate with respect to the gaps.

The drift term is included in the rule to address inflation's low-frequency movements during our sample.⁷ Since we have assumed the policy rule is perfectly communicated to the public so is π_t^* . Clearly this is not an innocuous assumption since the drift influences agents' inflation expectations and the Fed's communications about its inflation objectives were almost surely imperfect over much of the sample period we study. We leave an investigation

⁷See [Smets and Wouters \(2003\)](#) for an early example of a NK model with an inflation drift term in the monetary policy rule.

about inflation target communications to future research. The *inflation drift shock* evolves according to the stationary process

$$\ln \pi_t^* = (1 - \rho_\pi)\pi_* + \rho_\pi \ln \pi_{t-1}^* + \eta_t^\pi, \eta_t^\pi \sim N(0, \sigma_\pi^2),$$

where π_* is steady state inflation.

The gaps in the policy rule are measured as four quarter moving averages of variables observed by private agents. We assume the central bank measures the inflation gap as the average of the difference between twice and once lagged, current, and expected one-period-ahead log inflation and the contemporaneous value of the drift term:

$$\pi_t^{gap} = \frac{1}{4} E_t \sum_{j=-2}^1 (\ln \pi_{t+j} - \ln \pi_t^*).$$

The central bank measures the output gap as the average of the difference between the twice and once lagged, current, and expected one-period-ahead log level of aggregate output and its stochastic trend:

$$y_t^{gap} = \frac{1}{4} E_t \sum_{j=-2}^1 (\ln Y_{t+j} - \ln y^* - \ln A_{t+j}).$$

The constant y_* denotes steady state output. It appears in the output gap to ensure that the gaps are closed in steady state with the steady state nominal interest rate on government bonds $R_* = r_* \pi_*$.

The government issues bonds B_{t+1} and collects lump sum taxes T_t to pay for government spending $G_t = A_t g_t$ purchased in the final goods market. Therefore its one-period budget constraint is

$$G_t + B_t = T_t + \frac{B_{t+1}}{R_t}.$$

We assume the government balances its budget every period and has no legacy debt, so

government bonds are in zero net supply.⁸ The *government spending shock* g_t evolves as

$$\ln g_t = (1 - \rho_g) \ln s_*^g + \rho_g \ln g_{t-1} + \eta_t^g, \eta_t^g \sim N(0, \sigma_g^2),$$

where s_*^g is the government’s share of output in steady state.

3.5. Equilibrium

Equilibrium is defined in the usual way and is described in more detail in [Campbell et al. \(2017\)](#). We study the solution to the log linearized equilibrium conditions of the detrended economy and apply econometric techniques that rely on linearity to estimate a subset of the parameters and to conduct our study of central bank communications. One may question how such an approach can be squared with the ELB. Without forward guidance it is possible that at some dates agents’ expectations of future policy rates would violate the ELB constraint even if the contemporaneous rate did not. We use data on expected future funds rates, which of course do not violate the ELB, in our list of observables when we estimate our model. Forward guidance gives our model the flexibility to fit these data and thereby respect the ELB.

4. Estimation

[Chahrour and Jurado \(2018\)](#) show that models like ours can be represented in two observationally equivalent ways. We refer to our description of the model thus far as the *signal representation*.⁹ The model has an observationally equivalent *news representation* in which each period the central bank perfectly communicates news about future policy deviations. The news representation is easier to solve compared to the signal representation. Therefore, we estimate the news representation and then we use the mapping from the news representation to the signal representation to identify the parameters of the central bank’s

⁸As noted in [Campbell et al. \(2017\)](#) by including preferences for government bonds it is natural to extend the model to include a positive supply of government bonds. Doing so would be a step toward an environment where quantitative easing could be studied alongside forward guidance. We leave this avenue of inquiry to future work.

⁹[Chahrour and Jurado \(2018\)](#) use the terminology “noise representation.”

communication technology, Ξ_θ and Ξ_v . This approach is similar to [Blanchard, L’Huillier, and Lorenzoni \(2013\)](#).

This section begins by describing the mapping from the news representation to the signal representation. Then we describe how we estimate the news representation. An important feature of our estimation is that we use a rich array of macroeconomic data and data on the term structure of overnight interest rate futures. Our approach follows [Campbell et al. \(2017\)](#) quite closely and so our discussion is brief and emphasizes the key differences. The most important difference is that unlike in [Campbell et al. \(2017\)](#) we confront evidence that suggests both economic growth and the steady state interest rate on government bonds have declined over our sample period. Finally, we discuss briefly the properties of the estimated model.

4.1. Mapping the communications technology into news

Let $\varepsilon_{R,t}^h$ denote news about the policy deviation h periods hence revealed to the public through the policy signals at date t , s_t . By the definition of news

$$\varepsilon_{R,t}^h = E_t \theta_{t+h} - E_{t-1} \theta_{t+h}.$$

Collect news about policy deviations up to H periods ahead in the vector $\varepsilon_{R,t}$ so that we have

$$\varepsilon_{R,t} = E_t \boldsymbol{\theta}_t - E_{t-1} \boldsymbol{\theta}_t. \tag{6}$$

In the model of central bank communication, the date t revisions of private sector expectations about future policy deviations are given by equation (3). Combining equations (3) and (6) leads to the critical link between our model of central bank communications and news:

$$\kappa (\boldsymbol{\theta}_t + v_t - E_{t-1} \boldsymbol{\theta}_t) = \varepsilon_{R,t}. \tag{7}$$

We now show how to map the stochastic properties of news to the key parameters of the communications technology, Ξ_θ and Ξ_v .

The forecast errors of policy deviations for each $h = 0, 1, 0, \dots, H$ periods ahead are

related to news as follows:

$$\theta_{t+h} - E_{t-1}\theta_{t+h} = \sum_{j=0}^h \varepsilon_{R,t+h-j}^j. \quad (8)$$

This equation states that the error in the forecast of θ_{t+h} made in date $t - 1$ equals the sum of all news revealed from date t to date $t + h$. The variance-covariance matrix of the $H + 1$ random variables $\theta_{t+h} - E_{t-1}\theta_{t+h}$ corresponds to Ξ_θ . Therefore given the stochastic process of news we can calculate the variance-covariance matrix of the random variables $\sum_{j=0}^h \varepsilon_{R,t+h-j}^j$ and thereby obtain Ξ_θ using the relationship between news and the forecast errors of policy deviations given by (8).

Using (7) the variance-covariance matrices of the communications technology are related to the variance-covariance matrix of the vector of news, denoted by Ψ , according to

$$\Xi_\theta [\Xi_\theta + \Xi_v]^{-1} \Xi_\theta = \Psi.$$

With Ξ_θ in hand we can use this equation to solve for the variance-covariance matrix of noise:

$$\Xi_v = [\Xi_\theta^{-1} \Psi \Xi_\theta^{-1}]^{-1} - \Xi_\theta.$$

The foregoing demonstrates that if we have the stochastic process of news in hand we can obtain the parameters of the communications technology. We obtain the stochastic process for news by estimating the news representation of the model. In the news representation we suppose that instead of communicating noisy signals about future policy deviations the central bank communicates news. Specifically, each period t the central bank perfectly communicates the $(H + 1) \times 1$ vector $\varepsilon_{R,t}$. At the time this information is communicated agents believe it represents credible commitments to deviate from the rule up to H periods in the future with the full knowledge that more news may come along that changes expectations about future deviations. Any given period's policy deviation is communicated through news for up to H periods before the policy deviation is realized. It follows that the policy deviation

at date t , θ_t , is related to news according to

$$\theta_t = \sum_{j=0}^H \varepsilon_{R,t-j}^j.$$

The news representation thus involves replacing θ_t in equation (1) with $\sum_{j=0}^H \varepsilon_{R,t-j}^j$.

To complete the description of the news representation we need to specify the stochastic process for news. For this we adopt the factor structure used in [Campbell et al. \(2017\)](#) and described in the appendix. This factor structure allows for the possibility of serially correlated policy deviations and so is consistent with the signal representation. With this specification the model we estimate is essentially identical to the model estimated in [Campbell et al. \(2017\)](#).

4.2. Estimating the news representation

Our estimation of the model’s news representation proceeds in two steps. The first step is to calibrate the parameters that the model has in common with the neoclassical growth model to match sample averages from the U.S. economy.¹⁰ With the exceptions noted below to address low frequency movements in growth and interest rates these parameters are held fixed throughout the analysis. The second step takes the calibrated parameters as given and applies standard Bayesian methods to estimate the remainder of the model plus some auxiliary parameters which are used to map the model into the data.

Our calibration strategy is essentially the same as in [Campbell et al. \(2017\)](#) and so we do not elaborate upon it much here. The only difference is that we address the well known evidence of secular declines in the growth potential of the economy and rates of return on nominally risk free assets. We address these developments by imposing a change in steady state in 2008q4 (the choice of this date is motivated below.) The model’s potential GDP growth is governed by the growth rates of the neutral and investment-specific technologies. We adjust the growth rate of investment-specific technology down to account for the slower decline in the relative price of investment since 2008q4. Given this change we then lower the

¹⁰This “first-moments-first” approach ensures that any success in replicating second moments does not come at the cost of counterfactual long-run predictions.

steady state growth of the neutral technology so that potential growth is reduced to 2%. We increase the steady state marginal utility of government bonds using ε_*^s to match a lower real risk-free rate of 1%.¹¹ These adjustments leave the other parameters of tastes and technology unchanged but do change the steady state values of the endogenous variables and therefore the point at which the economy is log-linearized.¹²

Our Bayesian estimation uses the same split-sample strategy as in [Campbell et al. \(2017\)](#) except that we incorporate the change in steady state described above and one other change noted below. As in [Campbell et al. \(2017\)](#) our sample begins in 1993q1. This date is based on the availability and reliability of the overnight interest rate futures data. The sample period ends in 2016q4 but we impose a sample break in 2008q4. Our choice of this latter date is motivated by four considerations. First, it seems clear that the horizon over which forward guidance was communicated by the Fed lengthened substantially during the ELB period. Second, there is the evidence that points to lower interest rates and productivity growth later in the sample. Third, during the ELB period there is little information about the parameters of the policy rule. Finally, the downward trends in inflation and inflation expectations from the early 1990s appear to come to an end in the mid-2000s. Splitting the sample in 2008q4 and assuming certain model parameters change at that date is our way of striking a balance between parsimony and addressing the multiple structural changes that occur around the same time.

We estimate the full suite of non-calibrated structural parameters in the first sub-sample under the assumption that forward guidance extends for $H = 4$ quarters. Starting in 2008q4 we assume the model environment changes in three ways. First we assume the change in the steady state described above. Second, forward guidance lengthens to $H = 10$ quarters. Third, the time-varying inflation target from the first sample becomes a constant equal to the steady state rate of inflation, 2% at an annual rate. All three changes are assumed to be unanticipated and permanent.

¹¹The targets potential growth and the risk-free rate reflect a variety of evidence including the Fed's Summary of Economic Projections.

¹²Our re-calibration changes the return on private assets by a little. This small change is consistent with [Yi and Zhang \(2017\)](#) who show that rates of return on private capital stay roughly constant in the face of declines in risk free rates.

The measurement equations for the first sub-sample estimation are as follows:

$$\begin{aligned}
\Delta \ln Q_t^{obs} &= f\left(\hat{c}_t, \hat{c}_{t-1}, \hat{l}_t, \hat{l}_{t-1}, \hat{g}_t, \hat{\omega}_t, \hat{\pi}_t^{g,obs}\right) \\
\Delta \ln C_t^{obs} &= z_* + \Delta \hat{c}_t + \hat{z}_t \\
\Delta \ln I_t^{obs} &= z_* + \omega_* + \Delta \hat{l}_t + \hat{z}_t + \hat{\omega}_t \\
\log H_t^{obs} &= \hat{H}_t \\
\pi_t^{i,obs} &= \omega_* + \hat{\omega}_t + \varepsilon_t^i \\
R_t^{obs} &= R_* + \hat{R}_t \\
R_t^{j,obs} &= R_* + E_t \hat{R}_{t+j}, j = 1, 2, \dots, H \\
\pi_t^{l,j,obs} &= \pi_* + \pi_*^{l,j} + \frac{\beta^{l,j}}{l} \sum_{i=1}^l E_t \hat{\pi}_{t+i} + \varepsilon_t^{l,j,\pi}, j = 1, 2, l = 1, 40 \\
\pi_t^{j,obs} &= \pi_* + \pi_*^j + \beta^{\pi,j} \hat{\pi}_t + \gamma^{\pi,j} \pi_t^{d,obs} + \varepsilon_t^{j,p}, \text{ with } \beta^{\pi,1} = 1, j = 1, 2, 3 \\
\Delta \ln w_t^{j,obs} &= z_* + w_*^j + \beta^{w,j} (\hat{w}_t - \hat{w}_{t-1} + \hat{z}_t) + \varepsilon_t^{j,w}, \text{ with } \beta^{w,1} = 1, j = 1, 2 \\
\pi_t^{d,obs} &= \pi_*^d + \beta_{1,1} \pi_{t-1}^{d,obs} + \beta_{1,2} \pi_{t-2}^{d,obs} + \varepsilon_t^d \\
\pi_t^{g,obs} &= \pi_*^g + \beta_{2,1} \pi_{t-1}^{g,obs} + \beta_{2,2} \pi_{t-2}^{g,obs} + \varepsilon_t^g
\end{aligned}$$

In these equations the “hatted” variables represent log deviations from steady state, lower case letters correspond to de-trended counterparts of the upper case variables described in Section 3, and “ Δ ” is the first difference operator. The left hand side variables represent data (Q denotes chain-weighted GDP). These data are described in Campbell et al. (2017).¹³ The function f in the first equation represents the linear approximation to the chain-weighted GDP formula discussed in Campbell et al. (2017). Two variables are included to complete the mapping from model to data but are not endogenous to the model. Specifically, the consumption price of government consumption plus net exports, $\pi_t^{g,obs}$, is used to construct our model-consistent measure of chain-weighted GDP, and inflation in the consumption price of consumer durable goods, $\pi_t^{d,obs}$, is used to complete the mapping from model inflation to measured inflation.¹⁴

¹³We use three additional time series, all measures of expected inflation from the Survey of Professional Forecasters: PCE expected inflation over the next 10 years and both CPI and PCE expected inflation one quarter out.

¹⁴The measurement equations introduce some additional notation: $\pi_t^{j,obs}$ and $w_t^{j,obs}$ represent the inflation

These equations indicate we use 21 time series to estimate the model in the first sub-sample. In addition to the real quantities and federal funds rate that are standard in the literature our estimation includes multiple measures of wage and consumer price inflation, two measures each of average inflation expected over the next ten years and over one quarter, and $H = 4$ quarters of interest rate futures. Our second sub-sample estimation is restricted to estimating the parameters of the stochastic process for forward guidance news with $H = 10$ plus the processes driving $\pi_t^{g,obs}$ and $\pi_t^{d,obs}$. This estimation uses the measurement equations involving the current federal funds rate and 10 quarters of expected future policy rates plus the last two equations. We take into account the change in steady state but otherwise hold fixed the parameters at their first sub-sample values.

It is worth reiterating that our estimation respects the ELB. This is because we measure expected future rates in the model, the $E_t \hat{R}_{t+j}$, using the corresponding empirical futures rates, $R_t^{j,obs}$, and we use futures rates extending out 10 quarters. Because our estimation forces data on real activity, wages and prices to coexist with the interest rate futures data, we expect the estimation to mitigate the forward guidance puzzle.

4.3. The estimated model

The differences in the estimation strategy compared to [Campbell et al. \(2017\)](#) do not change the desirable properties of the estimated model. The vast majority of business cycle variation in real variables is driven by five shocks: shocks to the two technologies, liquidity preferences, the discount factor, and investment demand. Markup shocks contribute very little to real fluctuations and in the first sample monetary policy shocks contribute little as well. The model also continues to provide a plausible interpretation of the two recessions in our sample. The 2001 recession is explained by tighter financial conditions, i.e. an increase in demand for government bonds, and to a lesser extent weaker technology growth. During the recession the sharp drop in the funds rate is larger than stipulated by the policy rule and this lifts

and wage indicators discussed in [Campbell et al. \(2017\)](#); ε_t^i , $\varepsilon_t^{l,j,\pi}$, $\varepsilon_t^{j,p}$ and $\varepsilon_t^{j,w}$ denote classical measurement errors; ε_t^d and ε_t^g denote regression residuals; π_*^j and $\pi_*^{l,j}$ are constants that account for the average differences between the observable measures of inflation and inflation expectations and steady state inflation; $\beta^{\pi,j}$ and $\gamma^{\pi,j}$ denote the factor loadings relating observable inflation to model inflation and observed consumer durable inflation; $\beta^{l,j}$ are factor loadings relating observable inflation expectations to their model counterparts; and $\beta^{i,j}$ denote regression coefficients.

aggregate activity and inflation. Tighter financial conditions are the main driver of the Great Recession and the constraint of the ELB led to monetary policy being a substantial drag on real activity. Finally, the estimated model does not exhibit a forward guidance puzzle. In particular, when we conduct the experiment in [Del Negro et al. \(2015\)](#) in which the low interest rate peg anticipated by markets in 2012q3 is extended by one quarter we find small effects on real activity and inflation. These results bring credibility to the forward guidance we estimate.

5. The information flows from Fed communications

In this section, we quantify the information content of the estimated signals emitted by the Fed. We do this in two ways. First, we measure the reduction in the uncertainty about the future path of monetary policy at a point in time due to the individual entries of the signal vector. Second, we study the rate at which agents' uncertainty falls as the time of the policy implementation approaches. These exercises offer evidence on the limits of forward guidance when the central bank tries to steer the private sector's expectations at long horizons.

5.1. Information about the policy path embedded in s_t

We measure the information about the policy path embedded in s_t starting from a situation in which agents have received signals up to $t-1$ and do not have any time t signals. We then suppose that at time t the central bank sends signals within the vector s_t sequentially starting with the longest horizon H . For each new signal we measure the reduction in uncertainty relative to not having received any time t signals. This allows us to quantify the role of the individual signals in reducing the agents' uncertainty about the monetary policy path.

In the signal representation, the monetary policy path viewed by the public is a random vector of unobserved future deviations from the rule so its uncertainty is encoded in its variance-covariance matrix. The concept of entropy, defined as the average uncertainty of random vectors, allows us to summarize the uncertainty about the monetary policy path with a scalar while preserving the estimated correlation structure of the signals. Reduction in entropy is a widely used metric in the engineering literature to measure information

flows, e.g. [Cover and Thomas \(1991\)](#). In economics, the rational inattention literature uses this measure to characterize the information-processing constraint, e.g. [Sims \(2003\)](#) and [Maćkowiak and Wiederholt \(2009\)](#). For a Gaussian distributed random vector, $x \sim N(\mu, \Sigma)$, the entropy is given by

$$\xi(x) = \frac{1}{2} \log_2 |\Sigma| + \frac{n}{2} (\log_2 2\pi e)$$

where n is the dimension of x .

We focus on the future path of monetary policy, $\boldsymbol{\theta}_t^p$, the $H \times 1$ vector of future policy deviations that excludes the perfectly revealing contemporaneous signal $s_t^0 = \theta_t$. The *information gains*, or reduction in uncertainty, induced by the revelation of signals can be measured by taking the difference between the posterior entropy of receiving the signals of policy deviations H to $H - h$ quarters out, $h = 0, 1, \dots, H - 1$, minus the prior entropy of not receiving any new signal at all. Usually, entropy is measured in ‘bits’. We express the information gains as fractions. Specifically, we define the information gains of receiving the $h + 1$ signals for horizon $H - h$ to H as

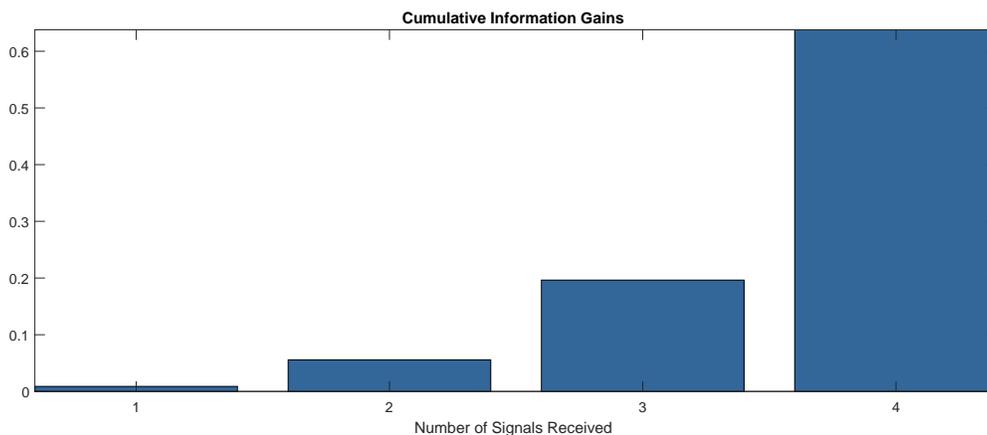
$$\begin{aligned} \mathcal{G}(h + 1) &= 1 - \exp \left[\xi(\boldsymbol{\theta}_t^p \mid s^{t-1}, s_t^H, s_t^{H-1}, \dots, s_t^{H-h}) - \xi(\boldsymbol{\theta}_t^p \mid s^{t-1}) \right] \\ &= 1 - \exp \left[\frac{1}{2} \log_2 |\Xi_{\theta, h}^p| - \frac{1}{2} \log_2 |\Xi_{\theta}^p| \right]. \end{aligned}$$

Here s^{t-1} denotes the history of all signals received up to and including period $t - 1$; Ξ_{θ}^p and $\Xi_{\theta, h}^p$ correspond to the variance-covariance matrices of the posterior and prior distributions of $\boldsymbol{\theta}_t$, respectively, and are derived in the Appendix.

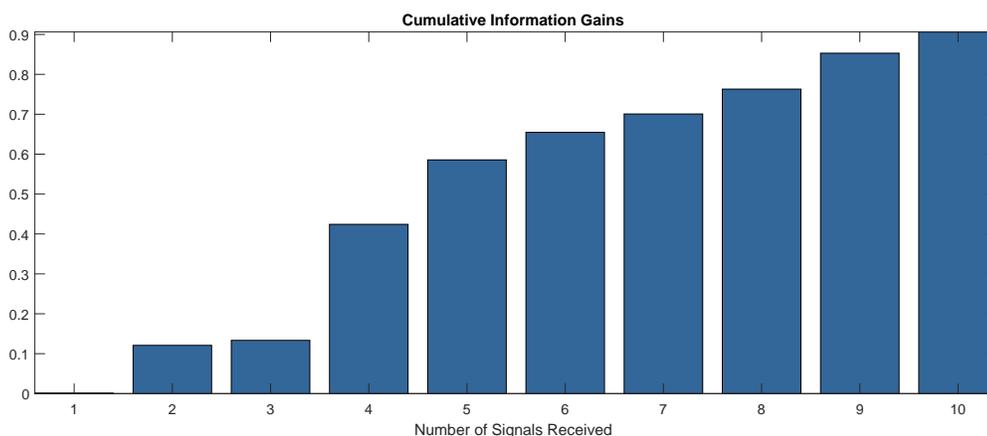
The information gains $\mathcal{G}(h + 1)$ induced by signals $H - h$ to H indicate the fraction of the reduction in the entropy of the monetary policy path $\boldsymbol{\theta}_t^p$ that is attributable to these signals alone. When the difference between the posterior and prior entropy, the argument in the exponential, is close to zero, the reduction in uncertainty is small and hence so are the information gains. Conversely, when the reduction in the posterior entropy is sizable, the argument in the exponential is negative and information gains become large.

Figure 1 reports the information gains in the first and second samples. The information gains show the accumulated reduction in uncertainty in the monetary policy path as the

Figure 1: Cumulative information gains



(a) First Sample



(b) Second Sample

Note: Each bar represents $\mathcal{G}(h + 1)$ where the x -axis indicates the number of signals, $h + 1$, received.

central bank progressively reveals deviations from the policy rule at different horizons. We start from the horizon furthest out and progressively add signals at shorter horizons. In the first sample agents receive the first signal four quarters before its actual implementation in the policy rule, whereas in the second sample signals about future deviations are emitted as early as ten quarters in advance.

The first sample plot shows close to zero reduction in uncertainty four quarters out. This means that uncertainty of the monetary policy path when the central bank announces only the deviation from the rule one year out hardly changes the agents' uncertainty about future policy deviations. Agents barely update their beliefs. The reduction in uncertainty is still less than 50% when agents have received three signals already. It is only when they

receive the signal about the next quarter policy rate that we see a substantial reduction in uncertainty but even then it is only about 60%. So in normal times, it appears the Fed has a very limited ability to use forward guidance to influence private sector expectations.

The second sample plot shows virtually zero reduction in uncertainty 10 quarters out. There is a discrete jump 7 quarters out and another 6 quarters out that lifts the information gain above 50%. By 2 quarters out about 80% of uncertainty is reduced. These results cast doubt about the power of forward guidance to affect agents' uncertainty about future deviations at horizons very far from the policy implementation. However the larger information gains at longer horizons in the second sample compared to the first sample suggest that the tool of forward guidance may have some bite during periods of extreme economic distress. Nevertheless there are limits to the power of this tool even then. It seems very difficult to convince the public that an interest rate peg can be extended for very long. From this perspective there is no forward guidance puzzle. In particular, [Del Negro et al. \(2015\)](#) frame the forward guidance puzzle as implausibly large effects on aggregate activity from extending the 10 quarter low interest rate peg expected by markets in 2012q3 to 11 quarters. [Figure 1](#) suggests it was not possible for the Fed to communicate such an extension.

We hypothesize that the relatively large information gains at longer horizons in the second sample arise from two factors. First, it's possible that the extreme economic conditions during the second sample may have led agents to pay more attention to the Fed than they did during more normal times. This should be reflected in the Kalman gain matrix.¹⁵ Second, it may be easier to communicate when policy is constrained by the ELB. Communicating policy deviations while being constrained by the ELB boils down to explaining how long the policy rate will remain near zero which seems relatively easy to do.¹⁶

¹⁵We have not framed our analysis in terms of the optimal allocation of attention. However there is a close connection between the Kalman gain and this allocation problem. See for example [Maćkowiak and Wiederholt \(2009\)](#) and [Melosi \(2014\)](#).

¹⁶Our analysis abstracts from the Fed's large scale asset purchases. To the extent that the macroeconomic effects of these policies were through signalling or establishing a commitment to keep rates lower for longer they can be viewed through the lens of forward guidance.

5.2. Accumulation of knowledge about a future policy deviation

We now examine the rate at which the agents' uncertainty about the deviation from the monetary policy rule *at a specific date* resolves as time goes by. We call this the *accumulation of knowledge* about a future policy deviation. We quantify the accumulation of knowledge by looking at the reduction in entropy of the deviation from the policy rule as we approach the policy implementation. Specifically, the reduction in entropy after receiving $h + 1$ vectors of signals about the policy deviation at $t + H$ is given by

$$\mathcal{K}(h + 1) = 1 - \exp [\xi(\theta_{t+H}|s^{t+h}) - \xi(\theta_{t+H}|s^{t-1})], \quad h = 0, 1, 2, \dots, H.$$

Here $\xi(\theta_{t+H}|s^{t-1})$ is the prior entropy of θ_{t+H} , which measures agents' uncertainty about θ_{t+H} before having received any signals about it, and $\xi(\theta_{t+H}|s^{t+h})$ is the posterior entropy conditional on receiving $h + 1$ vectors of signals.¹⁷

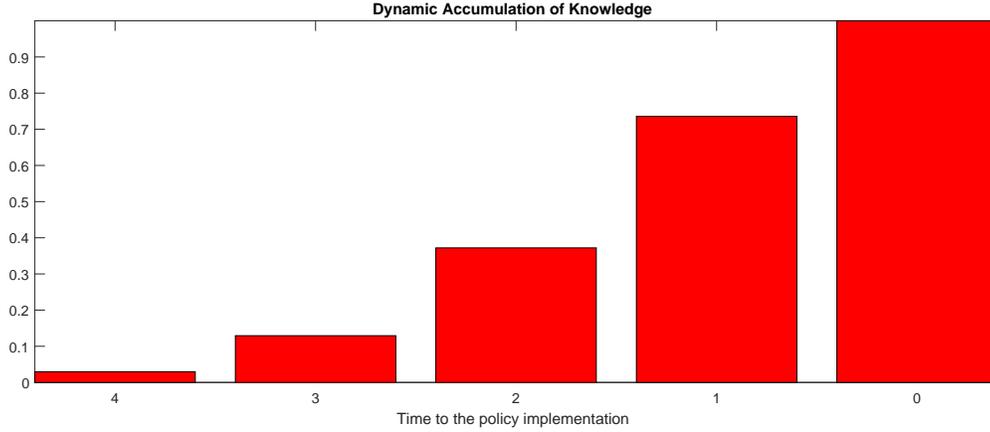
Figure 2 reports the accumulation of knowledge, $\mathcal{K}(h + 1)$, in the first and second sub-samples. The x -axis in the plots indicates the quarters before the policy deviation is implemented. For example in the first sub-sample the left-most bar corresponds to the case $h = 0$ in which only one vector of signals about the policy deviation at $t + 4$ has been communicated by the central bank. Knowledge flows at a fairly steady pace both in the first and second sample. In the first sample the reduction in uncertainty is sizable only two periods away from the policy implementation, while in the second sample the size of uncertainty is reduced by half more than a year before the policy is implemented. Figure 2 reinforces the main lessons from Figure 1. In the first sub-sample it was hard to communicate much more than 1 quarter before a policy decision. In the second sub-sample the communication horizon was substantially longer, but limited.

6. The dynamic effects of a forward guidance shock

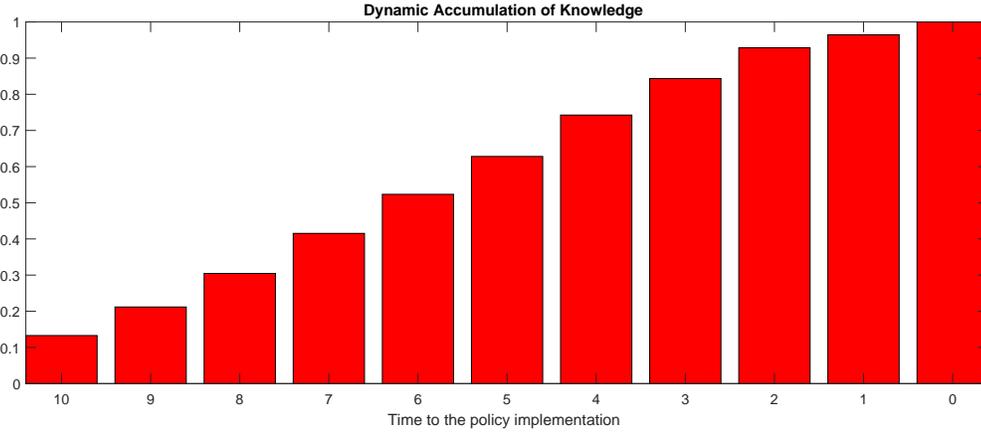
In this section we explore the role of imperfect communication in the propagation of forward guidance to the economy. To this end we explore the dynamic response of the economy to a

¹⁷The variance used in the posterior entropy $\xi_{t+j}(\theta_{t+H}|s^{t+h})$ corresponds to the $(H + 1 - h) \times (H + 1 - h)$ element of the variance-covariance matrix of the posterior distribution of the entire state $\theta_t|s^{t+h}$.

Figure 2: Dynamic accumulation of knowledge



(a) First Sample



(b) Second Sample

Note: Dynamic accumulation of knowledge in the first (top panel) and in the second sample (bottom panel). Horizontal axis reports the time to the implementation of the deviation from the policy rule. Vertical axis reports the fraction of the reduction in uncertainty about the intercept in the policy rule.

hypothetical *forward guidance shock*. We define a forward guidance shock as an orthogonalized shock to \mathbf{s}_t that implies a zero signal about the current deviation and non-zero signals about all future deviations up to H periods ahead.

The forward guidance shock is constructed by first decomposing the signal equation (2) as follows:

$$\mathbf{s}_t = \boldsymbol{\theta}_t + v_t = \Phi u_t \quad (9)$$

where $u_t \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is a $(H + 1) \times 1$ random vector of shocks and the matrix Φ is lower triangular with $E(\Phi u_t u_t' \Phi') = E_{t-1}(\mathbf{s}_t \mathbf{s}_t') = \Xi_\theta + \Xi_\nu$, which is the variance and covariance

matrix of time t signals that embodies information about the cross-correlations of the deviations θ_t and the cross-correlations of noise v_t . We then exploit the triangular structure of the matrix Φ to characterize the forward guidance shock. Specifically, we define a forward guidance shock as a unit innovation to the second element of the vector u_t fixing all other elements equal to zero. This represents a forward guidance shock since it consists of an orthogonalized vector of signals that implies no change in the intercept contemporaneously but a change in the path going forward.¹⁸

6.1. Propagation of forward guidance

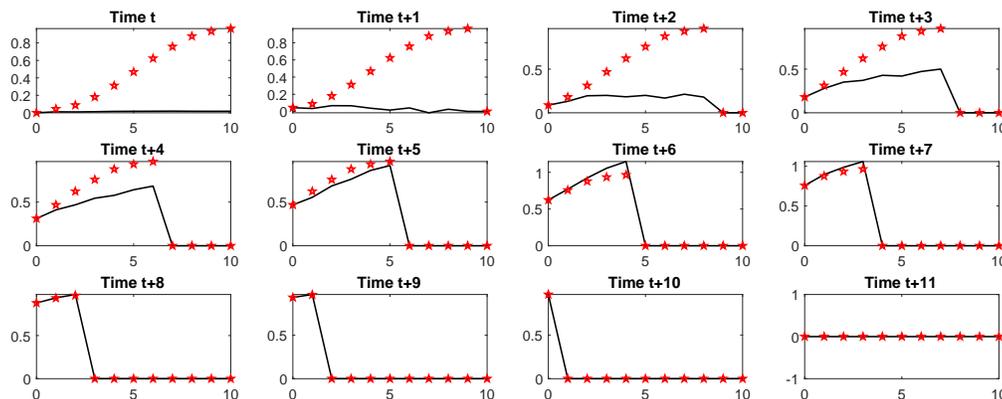
To study the effects of a forward guidance shock we consider the following scenario. The central bank announces a vector of *noiseless* signals in an arbitrary period t that comprise its actual policy deviations from period $t + 1$ through period $t + H$. The announcement comes when the economy is initially at steady state and so before it is made agents expect that all future deviations are zero. However, to be consistent with our learning environment, we also assume that before the announcement agents have received $H - h$ signals equal to zero about the h^{th} quarter ahead deviation from the rule and this information is embedded in their prior uncertainty. Further, we assume that over the period $t + 1$ to $t + H$ the central bank continues to send truthful signals about its policy deviations but does not send signals beyond the horizon of the initial forward guidance. Since the signals are the only source of information for the private sector to learn about future policy deviations, agents' expectations about the deviations from time $t + H$ onward equal zero, which is the unconditional mean of such deviations. No more deviations will be carried out after period $t + H$. These two assumptions imply that agents correctly anticipate the model economy is not hit by any shock after period $t + H$ and will transition back to the steady state.

Given this setup we consider two ways in which the forward guidance is communicated to private agents. In the *perfect communication* case we assume agent's learn with the Kalman gain matrix κ set equal to the identity matrix. In *imperfect communication* case agent's

¹⁸Other shocks are potentially interesting to study. The h^{th} shock in u_t is the shock that does not affect signals concerning deviations that will occur from period t through period $t + h - 1$. While these shocks certainly belong to the forward guidance class, they do not affect *all* the signals concerning future deviations from the rule. Therefore, the second shock seems like the most natural candidate to study.

learn with our estimated Kalman gain matrix. These two cases allow us to measure how imperfect communication influences the response to forward guidance.

Figure 3: Expectations about future policy deviations after a forward guidance shock in the second sample



Note: Response of private sector’s expectations about future deviations from the policy rule to a forward guidance shock. The black solid line denotes the case of imperfect communication and the red stars mark the case of imperfect communication. The units are percentage points of annualized rates.

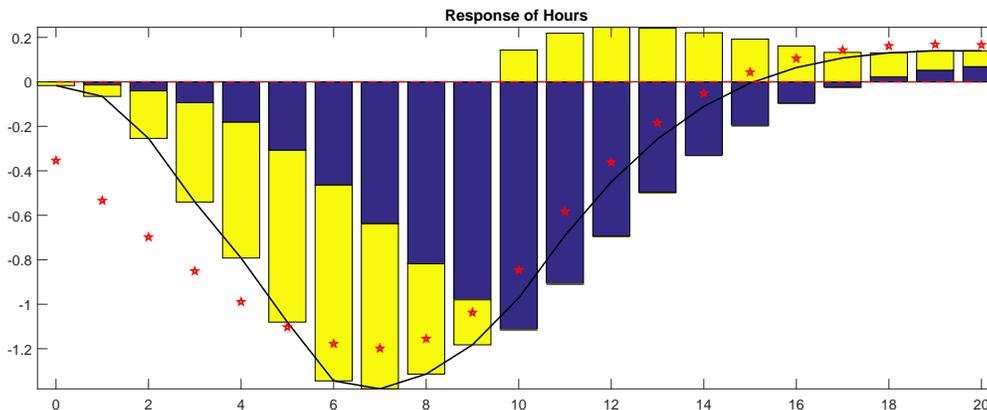
We begin by considering the most recent sub-sample. For this we normalize the forward guidance shock so that it leads the annualized interest rate to deviate from its rule-consistent value by 100 basis points after 10 quarters.¹⁹ The red stars in Figure 3 correspond to the actual future deviations from the rule looking ahead from each date following the forward guidance shock. Because the signals are noiseless the stars also correspond to the signals sent by the central bank. Recall that the signals perfectly reveal the current deviation and hence the stars associated with horizon 0 on the x -axis of each plot correspond to the actual deviation in the indicated period following the shock. The black lines of Figure 3 show the private sector’s expectations about policy deviations over the next 10 quarters, $E_t\theta_t$, at time t , which is when the first announcement is made, and in the following periods.

At time t , the central bank spectacularly fails to communicate its future deviations from the rule. Expectations about future deviations hardly budge. The second announcement at time $t + 1$ does not seem to materially move the private sector’s expectations either. At time $t + 2$, the third announcement has some impact and lifts private expectations toward the truth (the red stars) a bit. From period $t + 3$ on, agents’ expectations quickly rise, overshoot

¹⁹This requires us to re-scale the forward guidance shock by 4.

the true deviations in periods $t + 6$ and $t + 7$, and finally line up with the truth in period $t + 8$ and in subsequent periods. These patterns seem very much in line with the cumulative information gains shown in Figure 1, suggesting that it takes the first four or five signals for agents to have gathered enough information to start adjusting their expectations in line with what the future deviations will actually be.

Figure 4: Response of hours to a forward guidance shock



Note: Response of hours to a forward guidance shock that causes the interest rate to deviate from its rule-consistent value by 25 basis points in 10 quarters. The black solid line denotes the case of imperfect communication, the red stars mark the case of imperfect communication, and the blue circles refer to the case where the forward guidance shock does not lead to any actual deviations from the rule (pure beliefs). The units are percentage deviations from steady state.

The finding in the top left plot is striking and calls for caution in using the news representation of a model to study the macroeconomic effects of forward guidance. In the news representation future deviations are perfectly communicated. This plot shows that the central bank’s ability to steer expectations about future monetary policy is extremely limited. Therefore, using the news representation is likely to lead to inaccurate predictions or to puzzles (e.g., the forward guidance puzzle) that would not arise if one takes into account the imperfect ability of the central bank to communicate. But to do this, one needs to consider the signal representation.

Figure 4 shows the response of hours to the forward guidance shock.²⁰ We discuss the

²⁰We use the news representation to obtain these impulse response functions. The case of perfect communication is constructed by imposing that the news shocks in the first period are equal to the actual future deviations and zero in the following periods. The case of imperfect information is constructed by premultiplying the actual future deviations by the estimated Kalman gain matrix and then by using the resulting revision of expectations about future deviations as surprise and news shocks to simulate the news

yellow and blue bars below and focus here on the black line and red stars. The black line illustrates the case in which the central bank communicates using the imperfect communication technology. The red stars show the response of hours to a forward guidance shock in the counterfactual scenario in which the central bank can perfectly communicate with the private sector. Hence, the difference between these two lines captures the effects of the central bank not being able to perfectly communicate the future course of policy. Imperfect communication delays the response of real activity to forward guidance and magnifies it in the medium term. At the time of the first forward guidance (time t), hours barely adjust because private sector expectations fail react to the announcement, as shown in Figure 3. As time goes by and more and more guidance is provided, private sector expectations adjust and consequently economic activity quickly deteriorates and contracts more than that in the case of perfect information.

There are two reasons why imperfect communication triggers a deeper recession compared to perfect communication. First, as shown in Figure 3, the overshooting of expectations about the path of contractionary monetary shocks in period $t + 6$ and $t + 7$ contributes to deepen the recession. Second, the delayed revisions of agent's expectations contribute to lower hours. In the perfect communication case, agents never revise their expectations about future deviations after period t , when the first announcement is made. Under imperfect communications agents largely fail to anticipate the path of policy in period t and then revise their expectations slowly over time as shown in Figure 3. Consequently, when agents finally learn the future policy deviations, they have less time to adjust to the consequences of them relative to the case of perfect communication. A shorter anticipation time for future deviations boosts the response of real activity. This is a feature of all standard NK models.²¹ When firms anticipate a monetary shock long in advance, the effects of this anticipated shock will be relatively small because less firms are constrained by the Calvo lottery before the anticipated policy deviation is realized.²² More flexible prices imply a smaller response of real

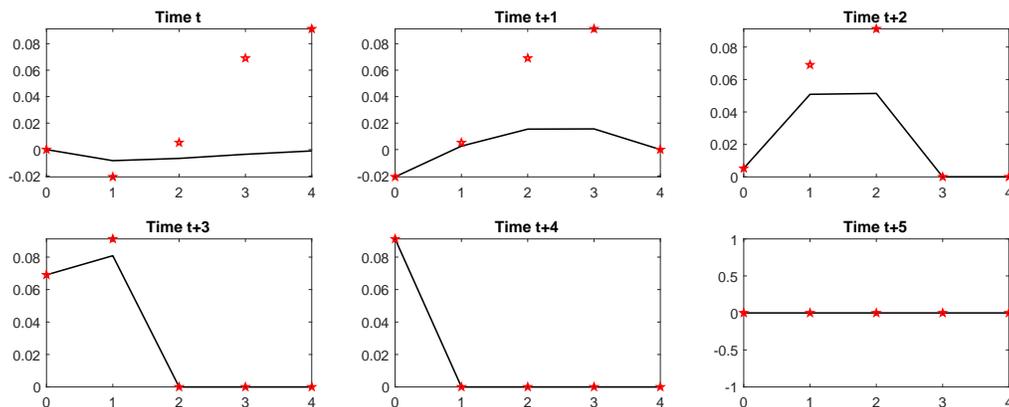
representation.

²¹Here we are not referring to the effects of extending an interest rate peg which grow with the horizon of the peg.

²²This result also holds in presence of Rotemberg-style price adjustment. A long anticipation horizon allows firms to smooth out the price changes over a longer period and hence lowers the cost for firms to change their price relative to a surprise monetary shock.

activity. Thus, by slowing down the information flow from the central bank to the private sector, imperfect communication magnifies the effects of forward guidance on real activity.

Figure 5: Expectations about future policy deviations after a forward guidance shock in the first sample



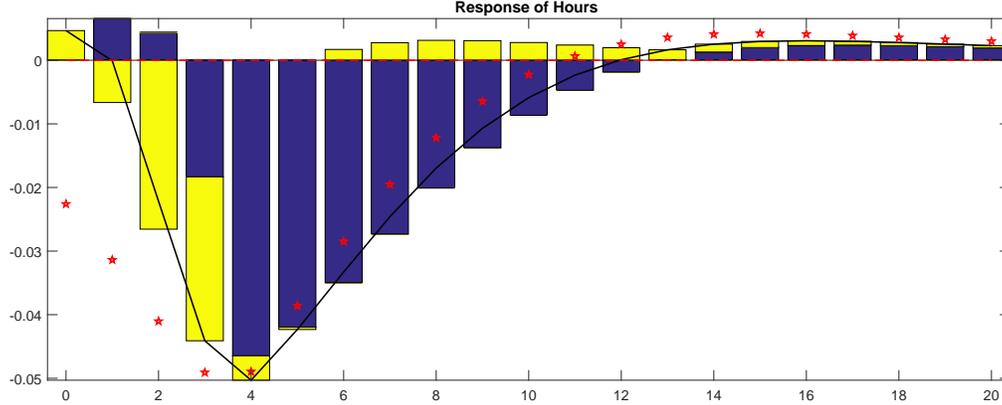
Note: Response of private sector’s expectations about future deviations from the policy rule to a forward guidance shock. The black solid line denotes the case of imperfect communication and the red stars mark the case of imperfect communication. The units are percentage points of annualized rates.

The results in Figure 4 suggest that a typical forward guidance shock takes a while to fully unravel its effects on real activity. This is mainly due to imperfect communication that delays the effects of forward guidance. This finding warns about the practical use of forward guidance away from the ELB as a tool to stabilize the economy in response to changes in the risk environment which might lead the central bank to want to deviate from its established rule.

Figure 5 shows the expected deviations from the rule following a forward guidance shock in the first sub-sample which we construct using the same method as the forward guidance shock in the second sample. The figure is constructed analogously to Figure 5.²³ Recall that forward guidance in the first sub-sample extends out only $H = 4$ quarters. The central bank fails to accurately communicate its future policy deviations in the first three periods. This is reflected in the black lines being far from the red stars in period t , when the first

²³The scale of Figure 5 is comparable to the one of Figure 5 meaning that the typical forward guidance shock in the first sample is not only very small ($2 = 8/4$ basis points but also only 8% of the typical shock in the second sample. The extremely small shocks in the first sample indicate two important features of our estimation. First, because the deviations are small the estimated policy rule is an excellent summary of the Fed’s behavior then. Second, the Fed did not do much forward guidance in the first sample. These features are likely influenced by our assumption that the inflation drift shock is perfectly communicated.

Figure 6: Response of hours to a forward guidance shock in the first sample



Note: Response of hours to a forward guidance shock that causes the interest rate to deviate from its rule-consistent value by 25 basis points in 10 quarters. The black solid line denotes the case of imperfect communication, the red stars mark the case of imperfect communication, and the blue circles refer to the case where the forward guidance shock does not lead to any actual deviations from the rule (pure beliefs). The units are percentage deviations from steady state.

announcement is made, and in the following two periods. Consistent with the cumulative information gains in Figure 1, only when the central bank has sent four signals and has carried out three deviations from the rule, do private sector expectations become very close to the true deviations.

The responses of hours to a forward guidance shock in the first sub-sample under perfect and imperfect communication are shown in Figure 6. The responses of hours are qualitatively similar to those in the second sub-sample. Imperfect communication seems to initially delay and later magnify the volatility of hours. Nevertheless, forward guidance shocks have smaller and less persistent effects on real activity than in the second sub-sample primarily due to the shorter duration of the guidance.

6.2. The expectation channel of forward guidance

The signal representation allows us to decompose the macroeconomic effects of forward guidance into two additive components. The first component represents the sole effects of the change in expectations triggered by forward guidance. The second component captures the effects of implementing the deviations by taking the private sector by surprise in every period. We call the first component the *expectation channel* of forward guidance. A proof

of this result is in the Appendix.

To evaluate the expectation channel of forward guidance, we construct a policy experiment in which the path of signals implied by the forward guidance shock turns out to be just noise. We still assume that the central bank has an imperfect communication technology as in the baseline case of truthful forward guidance. There are two things that should be noticed about the simulation with noise-driven forward guidance. First, expectations about the future deviations of monetary policy are the same as those in the baseline case in every period. This can be seen by inspecting equation (3). Second, when forward guidance is driven by noise only, the perfectly revealing signal concerning the current deviation must be always equal to zero. This experiment captures the expectations channel of forward guidance because there are no actual policy deviations and so the only effects of forward guidance arise due to expectations that deviations will occur.

In Figure 4, the yellow bars capture the effects of forward guidance on hours in this experiment. The expectation channel of forward guidance heavily affects the propagation of forward guidance shocks. It is quite potent in the first year and a half after the shock and its intensity quickly diminishes thereafter. The response of hours during the first year of forward guidance is almost entirely explained by expectations. The expectation channel makes the effects of monetary shocks on hours more front-loaded and generally more effective.

The case of noise-driven forward guidance is not only helpful to evaluate the expectation channel of forward guidance but it also allows us to study the implications of forward guidance entirely driven by noise. This interpretation of the counterfactual exercise captures situations in which either the central bank's forward guidance is misinterpreted by the private sector or the central bank changes its mind about future deviations from the rule after it has announced them. The magnitude of the yellow bars suggests that such communications may feed macroeconomic volatility and can challenge the central bank's ability of stabilizing real activity and inflation (the latter is not shown).

The blue bars in Figure 4 equal the vertical distance between the response of hours in the baseline case of truthful forward guidance and the counterfactual case of noise-driven forward guidance. As shown in the Appendix, the blue bars are the response of hours when the central bank opts for not providing forward guidance (no communication). To put

it differently, the central bank knows how it will deviate from the rule in the future but does not reveal anything about these deviations and keeps surprising the private sector in every period. Comparing the case of no communication (blue bars) to the case of truthful forward guidance (the black line) sheds light on how communication alters the propagation of monetary shocks. From this perspective, our main finding is that communication raises the volatility of hours and inflation (the latter not shown) during the implementation of the policy.

The effects on hours of the expectation channel and the absence of forward guidance on hours for the first sub-sample are shown in Figure 6. They are broadly similar to those in the second sub-sample.

7. Conclusion

This paper measured the imperfect ability of the Fed to communicate forward guidance and the influence of its imperfect communications on private sector expectations and macroeconomic outcomes. We found that the Fed’s ability to affect expectations at horizons that are sufficiently long to give rise to the forward guidance puzzle is substantially limited. The difficulties inherent in communicating complicated decisions about the future path of interest rates may be too great for forward guidance to be the powerful instrument predicted by standard NK models.

We also found that imperfect communications heavily influences the propagation of monetary shocks. First, the dynamic response of the economy to a shock that changes the policy rate path is substantially different under imperfect communications compared to the case of perfect communications. Second, unintended communication that contains no information about future policy deviations leads to sizable macroeconomic volatility. This finding suggests that communication away from the ELB is a fairly delicate instrument to handle. Finally, we show that forward guidance raises macroeconomic volatility compared to no communication.

While our findings show clearly that the Fed has a limited ability to shape expectations about the path of the Federal Funds rate it is not so limited as to suggest the Fed should eliminate forward guidance from its policy toolbox. We did find the power of forward guidance

to be extremely limited in the period when the funds rate was far from the ELB. However we also found that in the period after 2008 the horizon over which the Fed could influence expectations grew substantially. We conjectured that this may be due to two factors. First, its possible that the extreme economic conditions during after 2008 may have led the public to pay more attention to the Fed than they did during more normal times. Second, it just may be easier to communicate when policy is constrained by the ELB.

Imperfect central bank communications seems ripe for further study. More work can be done in isolating instances in which the Fed was best able to communicate its future intentions. This might guide central banks in developing better communications strategies. Using the methodology of this paper to investigate central bank communications in other economies would shed further light on this issue. One area not addressed in our analysis is how well the Fed was able to use its communications to shape inflation and inflation expectations during their gradual decline over the 1990s and early 2000s. Our methodology seems straightforward to apply to this question. More generally it is worth exploring other approaches to understand the ability of the central bank to communicate its future intentions. Recently [Coibion, Gorodnichenko, and Weber \(2019\)](#) have opened up a new and exciting area of research which uses an experimental approach to understand the effects of central bank communication. Finally, we have not taken a stand on the welfare implications of trying to communicate forward guidance. Are such communications advisable? Some of our findings hint at an answer to this question but clearly there is much more to be done here as well.

References

- Andrade, P., G. Gaballo, E. Mengus, and B. Mojon (2018). Forward guidance and heterogeneous beliefs. *American Economic Journal: Macroeconomics*. Forthcoming.
- Angeletos, G.-M. and C. Lian (2018, September). Forward guidance without common knowledge. *American Economic Review* 108(9), 2477–2512.
- Angeletos, G.-M. and A. Pavan (2007). Efficient use of information and social value of information. *Econometrica* 75(4), 1103–1142.
- Auclert, A., M. Rognlie, and L. Straub (2018). The intertemporal Keynesian cross. Harvard manuscript.
- Bianchi, F. and L. Melosi (2017). Constrained discretion and central bank transparency. *Review of Economics and Statistics*. Forthcoming.
- Blanchard, O., J.-P. L’Huillier, and G. Lorenzoni (2013). News, noise and fluctuations: An empirical exploration. *American Economic Review* 103(7), 3045–3070.
- Campbell, J., J. Fisher, A. Justiniano, and L. Melosi (2017). Forward guidance and macroeconomic outcomes since the financial crisis. *NBER Macroeconomics Annual 2016* 31, 283–357.
- Campbell, J. R., C. L. Evans, J. D. Fisher, and A. Justiniano (2012). Macroeconomic effects of Federal Reserve forward guidance. *Brookings Papers on Economic Activity Spring*, 1–54.
- Campbell, J. R. and J. P. Weber (2019). Discretion rather than rules: Equilibrium uniqueness and forward guidance with inconsistent optimal plans. Federal Reserve Bank of Chicago Working Paper 2018-14.
- Chahrour, R. and K. Jurado (2018). News or noise? The missing link. *American Economic Review* 108(7), 1702–1736.
- Christiano, L., M. Eichenbaum, and C. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113(1), 1–45.
- Coibion, O., Y. Gorodnichenko, and M. Weber (2019, January). Monetary policy communications and their effects on household inflation expectations. Working Paper 25482, National Bureau of Economic Research.
- Cover, T. M. and J. A. Thomas (1991). *Elements of Information Theory*. New York: Wiley.
- Del Negro, M., M. P. Giannoni, and C. Patterson (2015). The forward guidance puzzle. New York Fed working paper.
- Eggertsson, G. B. and M. Woodford (2003). The zero bound on interest rates and optimal monetary policy. *Brookings Papers on Economic Activity 2003*(1), 139–211.

- Eusepi, S. and B. Preston (2010, July). Central bank communication and expectations stabilization. *American Economic Journal: Macroeconomics* 2(3), 235–71.
- Evans, C., J. Fisher, F. Gourio, and S. Krane (2015). Risk management for monetary policy near the zero lower bound. Forthcoming, Brookings Papers on Economic Activity.
- Farhi, E. and I. Werning (2017, March). Monetary policy, bounded rationality, and incomplete markets. Working Paper 23281, National Bureau of Economic Research.
- Fisher, J. (2015, March-April). On the structural interpretation of the smets-wouters “risk premium” shock. *Journal of Money, Credit and Banking* 47(2-3), 511–516.
- Gabaix, X. (2016). A behavioral New Keynesian model. NYU-Stern manuscript.
- Gürkaynak, R., B. Sack, and E. Swanson (2005). Do actions speak louder than words? *International Journal of Central Banking* 1(1), 55–93.
- Hagedorn, M., J. Luo, I. Manovskii, and K. Mitman (2019). Forward guidance. *Journal of Monetary Economics*.
- Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010a, March). Investment shocks and business cycles. *Journal of Monetary Economics* 57(2), 132–145.
- Justiniano, A., G. E. Primiceri, and A. Tambalotti (2010b). Investment shocks and business cycles. *Journal of Monetary Economics* 57(2), 132 – 145.
- Kaplan, G., B. Moll, and G. L. Violante (2018, March). Monetary policy according to hank. *American Economic Review* 108(3), 697–743.
- Krishnamurthy, A. and A. Vissing-Jorgensen (2011). The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity Fall*, 215–265.
- Krugman, P. R. (1998). It’s baaack: Japan’s slump and the return of the liquidity trap. *Brookings Papers on Economic Activity Fall*, 137–187.
- Kuttner, K. N. (2001). Monetary policy surprises and interest rates: evidence rom the federal funds futures market. *Journal of Monetary Economics* 47, 523–544.
- Laséen, S. and L. E. Svensson (2011). Anticipated alternative policy-rate paths in policy simulations. *International Journal of Central Banking*.
- Leeper, E. and T. Zha (2003). Modest policy interventions. *Journal of Monetary Economics* 50(8), 1673–1700.
- Maćkowiak, B. and M. Wiederholt (2009). Optimal sticky prices under rational inattention. *American Economic Review* 99(3), 769–803.
- McKay, A., E. Nakamura, and J. Steinsson (2015). The power of forward guidance revisited. Columbia University manuscript.

- Melosi, L. (2014). Estimating models with dispersed information. *American Economic Journal: Macroeconomics* 6(1), 1–31.
- Melosi, L. (2016). Signalling Effects of Monetary Policy. *The Review of Economic Studies* 84(2), 853–884.
- Michaillat, P. and E. Saez (2018). A New Keynesian model with wealth in the utility function. NBER Working Paper 24971.
- Morris, S. and H. S. Shin (2002, December). Social value of public information. *American Economic Review* 92(5), 1521–1534.
- Nakamura, E. and J. Steinsson (2018, 01). High-Frequency Identification of Monetary Non-Neutrality: The Information Effect*. *The Quarterly Journal of Economics* 133(3), 1283–1330.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics* 50(3), 665–690.
- Smets, F. and R. Wouters (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European Economic Association* 1(5), 1123–1175.
- Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: A Bayesian DSGE approach. *The American Economic Review* 97(3), 586–606.
- Woodford, M. (2003). *Interest and Prices*. Princeton, NJ: Princeton University Press.
- Yi, K.-M. and J. Zhang (2017). Understanding global trends in long-run real interest rates. *Economic Perspectives* 41(2).

Appendix

A. Constructing the prior matrix of the signals

As explained in the main text, we can use equations (8) along with the estimated statistical properties of the news shocks $\{\varepsilon_{R,t}^j\}_{j=0}^H$ to construct the prior matrix Ξ_θ . Here are some examples to show how one can construct the matrix Ξ_θ (with $H = 10$).

Since the deviations from the rule θ_t are assumed to be correlated up to the H -th lag in our model, we allow news shocks in the observationally equivalent news representation to be correlated. To capture this, we follow [Campbell et al. \(2017\)](#) and assume that a factor structure determines the cross-correlation among news shocks. In symbols:

$$\varepsilon_{R,t}^i = \alpha_i f_t^\alpha + \beta_i f_t^\beta + \psi_i \eta_t. \quad (10)$$

where the factors f_t^α and f_t^β are scalar, η_t are an $H \times 1$ column vector of shocks, and ψ_i is $1 \times H$ vector of coefficients that depend on the model's structural parameters. Let us denote the diagonal variance and covariance matrix $\Sigma_\eta \equiv E(\eta_t \eta_t')$.

Example 1 Characterize $\Xi_\theta(10, 9) = E_{t-1}(\theta_{t+10} \theta_{t+9})$

$$\begin{aligned} \Xi_\theta(10, 9) = & [\alpha_0, \dots, \alpha_9] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{10} \end{bmatrix} \text{var}(f_t^\alpha) + \\ & [\beta_0, \dots, \beta_9] \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_{10} \end{bmatrix} \text{var}(f_t^\beta) + \\ & [\psi_0, \dots, \psi_9] \cdot [\mathbf{I}_{10} \otimes \Sigma_\eta] \cdot \begin{bmatrix} \psi'_1 \\ \psi'_2 \\ \vdots \\ \psi'_{10} \end{bmatrix} \end{aligned}$$

where \mathbf{I}_n denotes the $n \times n$ identity matrix.

Example 2 Characterize $\Xi_\theta(10, 1) = E_{t-1}(\theta_{t+10} \theta_{t+1})$

$$\begin{aligned} \Xi_\theta(10, 1) = & [\alpha_9, \alpha_{10}] \cdot \begin{bmatrix} \alpha_0 \\ \alpha_1 \end{bmatrix} \text{var}(f_t^\alpha) + \\ & [\beta_9, \beta_{10}] \cdot \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix} \text{var}(f_t^\beta) + \end{aligned}$$

$$[\psi_9, \psi_{10}] \cdot [\mathbf{I}_2 \otimes \Sigma_\eta] \cdot \begin{bmatrix} \psi'_0 \\ \psi'_1 \end{bmatrix}$$

Example 3 Characterize $\Xi_\theta(1, 1) = E_{t-1}(\theta_{t+1}\theta_{t+1})$

$$\Xi_\theta(1, 1) = \alpha_0^2 \text{var}(f_t^\alpha) + \beta_0^2 \text{var}(f_t^\beta) + \psi_0 \Sigma_\eta \psi'_0$$

Example 4 Characterize $\Xi_\theta(5, 5) = E_{t-1}(\theta_{t+5}\theta_{t+5})$

$$\begin{aligned} \Xi_\theta(5, 5) &= [\alpha_0, \dots, \alpha_5] \cdot \begin{bmatrix} \alpha_0 \\ \vdots \\ \alpha_5 \end{bmatrix} \text{var}(f_t^\alpha) + \\ &[\beta_0, \dots, \beta_5] \cdot \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_5 \end{bmatrix} \text{var}(f_t^\beta) + \\ &+ [\psi_0, \dots, \psi_5] [\mathbf{I}_6 \otimes \Sigma_\eta] \begin{bmatrix} \psi_0 \\ \vdots \\ \psi_5 \end{bmatrix} \end{aligned}$$

B. Calculating the information gains

In this section we provide the derivations to compute the first and second moments of $\boldsymbol{\theta}_t$, $\boldsymbol{\theta}_t|_{s_{t-1}}$ and $\boldsymbol{\theta}_t|_{s_t}$. To this aim, it is useful to write the vector of signals in matrix notation as follows

$$\boldsymbol{\theta}_t = \begin{bmatrix} \theta_t \\ \theta_{t+1} \\ \vdots \\ \theta_{t+H} \end{bmatrix} = \begin{bmatrix} \varepsilon_{R,t-H}^H + & \varepsilon_{R,t-H+1}^{H-1} + \dots & + \varepsilon_{R,t-1}^1 + & \varepsilon_{R,t}^0 & 0 + & \dots & + 0 \\ 0 + & \varepsilon_{R,t-H+1}^H + \dots & + \varepsilon_{R,t-1}^2 + & \varepsilon_{R,t}^1 & \varepsilon_{R,t+1}^0 + & \dots & + 0 \\ & & & \vdots & & & \\ 0 + & \dots & + \varepsilon_{R,t-1}^H & \varepsilon_{R,t}^{H-1} + & \varepsilon_{R,t+1}^{H-2} + & \dots & + 0 \\ 0 + & \dots & + 0 & \varepsilon_{R,t}^H + & \varepsilon_{R,t+1}^{H-1} + & \dots & + \varepsilon_{R,t+H}^0 \end{bmatrix}$$

or equivalently

$$\begin{aligned} \boldsymbol{\theta}_t &= J_H \begin{bmatrix} \varepsilon_{t+H}^0 \\ \vdots \\ \varepsilon_{t+H}^H \end{bmatrix} + \dots + J_1 \begin{bmatrix} \varepsilon_{t+1}^0 \\ \vdots \\ \varepsilon_{t+1}^H \end{bmatrix} + J_0 \begin{bmatrix} \varepsilon_t^0 \\ \vdots \\ \varepsilon_t^H \end{bmatrix} + J'_1 \begin{bmatrix} \varepsilon_{t-1}^0 \\ \vdots \\ \varepsilon_{t-1}^H \end{bmatrix} + \dots + J'_H \begin{bmatrix} \varepsilon_{t-H}^0 \\ \vdots \\ \varepsilon_{t-H}^H \end{bmatrix} \\ &= J_H \boldsymbol{\varepsilon}_{t+H} + \dots + J_1 \boldsymbol{\varepsilon}_{t+1} + \boldsymbol{\varepsilon}_t + J'_1 \boldsymbol{\varepsilon}_{t-1} + \dots + J'_H \boldsymbol{\varepsilon}_{t-H} \end{aligned}$$

where J_k is a $((H + 1) \times (H + 1))$ matrix of zeros with ones on the k lower diagonal; J_0 coincides with the identity matrix. For example, for when $H = 4$ and $k = 2$

$$J_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Define $\Sigma_\varepsilon = E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t')$; we have an estimated Σ_ε from the news representation . We have that the first moments are

$$E(\boldsymbol{\theta}_t) = 0$$

$$E_{t-1}(\boldsymbol{\theta}_t) = J_1' \boldsymbol{\varepsilon}_{t-1} + \dots + J_H' \boldsymbol{\varepsilon}_{t-H}$$

$$\begin{aligned} E_t(\boldsymbol{\theta}_t) &= E_{t-1}(\boldsymbol{\theta}_t) + \kappa(s_t - E_{t-1}(\boldsymbol{\theta}_t)) \\ &= E_{t-1}(\boldsymbol{\theta}_t) + \kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t)) \end{aligned}$$

The second moments are respectively

$$\Sigma_\theta = E(\boldsymbol{\theta}_t \boldsymbol{\theta}_t') = J_H \Sigma_\varepsilon J_H' + \dots + J_1 \Sigma_\varepsilon J_1' + \Sigma_\varepsilon + J_1' \Sigma_\varepsilon J_1 + \dots + J_H' \Sigma_\varepsilon J_H$$

$$\Xi_\theta = E(\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t))(\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t))' = J_H \Sigma_\varepsilon J_H' + \dots + J_1 \Sigma_\varepsilon J_1' + \Sigma_\varepsilon$$

$$\begin{aligned} \Xi_{\theta,1} &= E(\boldsymbol{\theta}_t - E_t(\boldsymbol{\theta}_t))(\boldsymbol{\theta}_t - E_t(\boldsymbol{\theta}_t))' \\ &= E[\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t) - \kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t))][\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t) - \kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t))] \\ &= E[\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t)][\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t)]' + E[\kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t))][\kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t))] \\ &\quad - 2 \times E[\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t)][\kappa(\boldsymbol{\theta}_t + v_t - E_{t-1}(\boldsymbol{\theta}_t))] \\ &= \Xi_\theta + \kappa(\Xi_\theta + \Xi_v) \kappa' - 2 \Xi_\theta \kappa' \\ &= \Xi_\theta + \Xi_\theta (\Xi_\theta + \Xi_v)^{-1} (\Xi_\theta + \Xi_v) (\Xi_\theta + \Xi_v)^{-1} \Xi_\theta' - 2 \Xi_\theta (\Xi_\theta + \Xi_v)^{-1} \Xi_\theta' \\ &= \Xi_\theta - \Xi_\theta (\Xi_\theta + \Xi_v)^{-1} \Xi_\theta' \end{aligned}$$

where by construction we have $E[\boldsymbol{\theta}_t - E_{t-1}(\boldsymbol{\theta}_t)]v_t' = 0$. The derivations of the implied moments of the path of monetary policy, $\boldsymbol{\theta}_t^p = [\theta_{t+1}, \dots, \theta_{t+H}]'$, are quite straightforward as we need to apply the previous calculation to a linear transformation of $\boldsymbol{\theta}_t$. In particular, we have the future path of monetary policy is given by

$$\boldsymbol{\theta}_t^p = P \boldsymbol{\theta}_t$$

where P is a $H \times (H + 1)$ matrix that removes the first element of the vector $\boldsymbol{\theta}_t$, i.e.

$$P = \begin{bmatrix} 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ & & \ddots & & \\ 0 & 0 & \dots & 0 & 1 \end{bmatrix}$$

First and second moments of $\boldsymbol{\theta}_t^p$ are then easy to derive. In particular,

$$\Xi_\theta^p = E(P\boldsymbol{\theta}_t - E_{t-1}(P\boldsymbol{\theta}_t))(P\boldsymbol{\theta}_t - E_{t-1}(P\boldsymbol{\theta}_t))' = P\Xi_\theta P'$$

The expected value of $\boldsymbol{\theta}_t^p$ conditional of having received the signals about the deviations from the policy rule from H and until $H - h$ quarters out is given by

$$\begin{aligned} E(\boldsymbol{\theta}_t^p | s^{t-1}, s_t^{H-h}, \dots, s_t^H) &= E_{t-1}(\boldsymbol{\theta}_t^p) + \kappa_h S_h (s_t - E_{t-1}(\boldsymbol{\theta}_t^p)) \\ &= E_{t-1}(\boldsymbol{\theta}_t^p) + \kappa_h S_h (\boldsymbol{\theta}_t^p + v_t - E_{t-1}(\boldsymbol{\theta}_t^p)) \end{aligned}$$

where $\kappa_h = \Xi_\theta S_h' (S_h (\Xi_\theta + \Xi_v) S_h')^{-1}$ and S_h is a $(h + 1) \times H$ selection matrix that has ones on the right most diagonal and zeros elsewhere. In other words, we add signals at different horizon sequentially, starting from the signal at the longest horizon. For example, at the longest horizon when $h = 0$, $S_0 = [0, 0, \dots, 1]$. When $h = 1$,

$$S_1 = \begin{bmatrix} 0 & \dots & 1 & 0 \\ 0 & \dots & 0 & 1 \end{bmatrix}$$

As before we can compute the covariance matrix of $\boldsymbol{\theta}_t^p$ conditional of having received a subset of signal signals, which is given by

$$\Xi_{\theta,h}^p = Var(\boldsymbol{\theta}_t^p | s^{t-1}, s_t^{H-h}, \dots, s_t^H) = \Xi_\theta^p - \Xi_\theta^p S_h' (S_h (\Xi_\theta^p + \Xi_v) S_h')^{-1} S_h \Xi_\theta^p$$

C. Decomposing Forward Guidance

In the main text, we claim that the signal representation allows us to decompose the macroeconomic effects of forward guidance into two additive interesting components. The first component represents the sole effects of the change in expectations triggered by forward guidance (the yellow bars in Figures 4 and 6). The second component captures the effects of implementing the deviations implied by the forward guidance shocks without announcing them in advance (the blue bars in Figures 4 and 6). In this section we sketch the proof of this claim.

It is easier to work with the news representation of the model. The surprise and news monetary shocks that realize from period $t - H$ through period t can be collected along the

columns of the matrix below

$$\begin{bmatrix} \varepsilon_t^0 & \varepsilon_{t-1}^0 & \cdots & \varepsilon_{t-H}^0 \\ \varepsilon_t^1 & \varepsilon_{t-1}^1 & \cdots & \varepsilon_{t-H}^1 \\ \varepsilon_t^2 & \varepsilon_{t-1}^2 & \cdots & \varepsilon_{t-H}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_t^H & \varepsilon_{t-1}^H & \cdots & \varepsilon_{t-H}^H \end{bmatrix}. \quad (11)$$

The elements above the diagonal are shocks that are past at time t . The diagonal elements are the shocks that hit the economy at time t . Indeed, the actual deviation from the rule at time t is given by summing all the diagonal elements of the matrix of shocks above; that is,

$$\theta_t = \sum_{j=0}^H \varepsilon_{t-j}^j \quad (12)$$

The elements below the diagonal are those shocks that are known at time t and that will hit the economy in the future.

The news and surprise shocks in this matrix are obtained from the signal representation by using equation (7). The response of hours in the baseline economy of truthful forward guidance (i.e., the black line in Figures 4 and 6) is obtained by assuming that the forward guidance shock is driven by actual future deviations from the rule (zero noise). In the text, we also study the response of hours in a *alternative scenario* (the yellow bars) in which all forward guidance shocks are driven by noise. This alternative scenario is obtained by simulating the news representation under the restrictions that the surprise shocks are such that

$$\varepsilon_t^0 = \tilde{\varepsilon}_t^0 \equiv \sum_{j=1}^H \varepsilon_{t-j}^j, \quad (13)$$

for every t . The other shocks, including the news shocks from the second to the H -th row of the matrix above, are the same in both the baseline case and in the alternative case.

Denote the infinite moving average representation of our model economy:

$$h_t = \Phi_h(L)^{-1} \xi_t \quad (14)$$

where h_t denotes hours and ξ_t is the column vector containing all the shocks realized at time t . In the baseline case, the vector ξ_t also includes all the surprise and news shocks in the first column of the matrix above. In the counterfactual case, the vector of shocks ξ_t is the same except that the surprise shocks are modified by equation (13).

The blue bars for hours in Figures 4 and 6 are constructed by taking the difference of the response of hours to forward guidance shocks in the baseline (h_t), the black line, and in the alternative scenarios of noise-driven forward guidance (\tilde{h}_t), the yellow bars. In symbols,

$$h_t - \tilde{h}_t = \Phi_h(L)^{-1} \begin{bmatrix} \mathbf{0} \\ (\varepsilon_t^0 - \tilde{\varepsilon}_t^0) \\ \mathbf{0} \end{bmatrix} \quad (15)$$

$$= \Phi_h(L)^{-1} \begin{bmatrix} \mathbf{0} \\ \varepsilon_t^0 - \sum_{j=1}^H \varepsilon_{t-j}^j \\ \mathbf{0} \end{bmatrix} \quad (16)$$

$$= \Phi_h(L)^{-1} \begin{bmatrix} \mathbf{0} \\ \theta_t^R \\ \mathbf{0} \end{bmatrix} \quad (17)$$

where in the first line we use the infinite average representation of hours, the second line is obtained by plugging equation (13), and the third line stems from equation (12). So the difference in hours $h_t - \tilde{h}_t$ between these two cases can be obtained by simulating the news representation using the actual deviations from the rule $\theta_t^R = \sum_{j=0}^H \varepsilon_{t-j}^j$ as surprise monetary policy shocks (ε_t^0). This scenario is tantamount to the case of no communication, in which the central bank does not say anything about its future deviations and it just implements them over time.