Estimating the effects of forward guidance in rational expectations models[*]

Richard Harrison[†]

November 6, 2014

Abstract

Simulations of forward guidance in rational expectations models should be assessed using the "modest interventions" framework introduced by Eric Leeper and Tao Zha. That is, the estimated effects of a policy intervention should be considered reliable only if that intervention is unlikely to trigger a revision in private sector beliefs about the way that policy will be conducted. I show how to constrain simulations of forward guidance to ensure that they are regarded as modest policy interventions and illustrate the technique using a medium-scale DSGE model estimated on US data. I find that, in many cases, experiments that generate the large responses of macroeconomic variables that many economists deem implausible – the so-called "forward guidance puzzle" – would *not* be viewed as modest policy interventions by the agents in the model. Those experiments should therefore be treated with caution, since they may prompt agents to believe that there has been a change in the monetary policy regime that is not accounted for within the model. More reliable results can be obtained by constraining the experiment to be a modest policy intervention. The quantitative effects on macroeconomic variables are more plausible in these cases.

[†]Bank of England and Centre for Macroeconomics. Email: richard.harrison@bankofengland.co.uk

1 Introduction

Simulations of announcements about the prospective path of monetary policy ('forward guidance') in rational expectations models should be assessed using the "modest interventions" approach of Leeper and Zha (2003). That is, the estimated effects of a policy intervention should be considered reliable only if that intervention is unlikely to trigger a revision in private sector beliefs about the way that policy will be conducted. In contrast, estimates obtained from experiments that are not modest policy interventions may be unreliable because the single-regime rational expectations models typically used for monetary policy analysis not allow for the effects of shifts in agents' beliefs about the prevailing monetary policy regime.

In this paper, I apply the Leeper-Zha framework to simulations of announced paths for the policy rate in single-regime dynamic stochastic general equilibrium (DSGE) models with rational expectations. Using an estimated DSGE model of the US economy, I show that constraining forward guidance experiments to be modest policy interventions influences the estimated effects on macroeconomic variables. In particular, constraining the experiments to be modest policy interventions eliminates the extremely large estimates of the macroeconomic effects reported in many recent papers. However, my model can reproduce these extreme results if I do not constrain the forward guidance experiment to be a modest policy intervention. This suggests that recent estimates may be unreliable because they are not modest policy interventions.

I focus on the task of a policy advisor asked to provide estimates of the macroeconomic effects of forward guidance. Specifically, the policy advisor estimates the effects of the central bank announcing a planned path for the policy instrument that is fully believed by private agents.[1] The policy advisor uses a "New Keynesian" DSGE model (featuring optimizing forward-looking households and firms, explicitly modeled nominal rigidities and rational expectations) of the type commonly used to study monetary policy by academics and policymakers.[2]

To assess whether the experiment represents a modest policy intervention, I extend the framework developed by Adolfson et al. (2005) to measure modest interventions in DSGE models. I show how to simulate forward guidance under the assumption that the experiment is modest policy intervention. In some cases, it may be impossible to implement a modest policy intervention that delivers precisely the desired path for the policy rate. In those cases, I show how to implement a modest intervention that produces a path for the instrument that is as close as possible to the desired path.

I illustrate my technique using a variant of the Smets and Wouters (2007) medium-scale DSGE model, extended to include 'policy news shocks' that capture anticipated future changes in the Fed funds rate. These shocks represent the mechanism through which policymakers can provide guidance about the likely future path of monetary policy. The model is estimated on US data using data from 1984Q1–2008Q4. The sample period is chosen to exclude the recent period during which the Fed funds rate has been at the zero bound.

I use the model to conduct policy experiments in which the Fed funds rate is

---

[1]As noted in Section 2, not all forward guidance is intended to communicate a particular path for the policy rate. So this experiment is not necessarily the most appropriate for estimating the effects all forms of forward guidance.

[2]Many central banks use at least one model of this type as part of their policy analysis and forecasting processes. See Tovar (2008) for a review of how DSGE models are used at central banks.

held lower than the model-based forecast. The results of these experiments generate macroeconomic effects that most policymakers (and their advisors) would regard as implausibly large. This result is consistent with several previous studies documenting that, in this class of models, the announcement of a fully credible path for the short-term nominal interest rate can generate macroeconomic effects that most economists would regard as implausibly large (see, for example, Weale (2013)). This result is labeled the "forward guidance puzzle" by del Negro et al. (2012). I show that, from the perspective of agents in the model, these policy experiments do *not* represent modest policy interventions. So the predictions of extremely large effects of forward guidance are likely to be unreliable.

When I constrain the policy experiments so that agents in the model would regard them as modest policy interventions, the macroeconomic effects are greatly reduced. In many cases, the path for the Fed funds rate that delivers a modest policy intervention is very similar to the desired path that the policy advisor is asked to simulate.

The remainder of this paper is organized as follows. In Section 2, I briefly explain the nature of the 'forward guidance puzzle'. In Section 3, I introduce my methodology for measuring the 'modesty' of a forward guidance experiment and restricting an immodest policy intervention to be modest. In Section 4, I present an empirical exercise using the medium-scale estimated DSGE model. In Section 5, I examine the simulations in more detail to shed light on the underlying mechanisms that determine the results. Section 6 examines the robustness of my findings to two important assumptions: the way that the modesty of the policy experiment is measured; and the extent to which the estimated policy rule captures behavior from a stable monetary policy regime.

2   Forward guidance and estimates of its effects

Over the past decade, many central banks have explored ways to provide more information about how their assessments of the economy and of the appropriate way to achieve their policy objectives are likely to affect the future path for their monetary policy instruments. Such 'forward guidance' comes in many forms, ranging from qualitative descriptions of the key judgments underpinning policy discussions, to explicit projections of the policy instrument under alternative assumptions about the nature of the shocks hitting the economy.[3]

As noted by Woodford (2012), interest in the use of explicit forward guidance has increased in recent years. In part this is because many central banks reduced their policy rates to their effective lower bounds in response to the financial crisis, limiting the scope for further cuts in the policy rate.[4] The objectives of recent forward guidance (and the methods by which they have been communicated) are varied. In some cases, the guidance has been intended to clarify the stance of monetary policy that policymakers think is appropriate.[5] In other cases, the purpose of the guidance has been to

---

[3]See the Appendix in Monetary Policy Committee (2013) for a review of alternative forms of forward guidance and den Haan (2013) for commentary and analysis of the forward guidance policies implemented by the major central banks.

[4]Of course, many central banks have also responded by increasing the range of policy instruments to include so-called unconventional monetary policy tools. See Borio and Disyatat (2010) for a review of these policies.

[5]See, for example, Woodford (2012, Section 1.2).

clarify the nature of the monetary policy reaction function.[6]

2.1 Existing estimates and the "forward guidance puzzle"

Alongside the continual development of forward guidance strategies by central banks, economists have studied the effects of forward guidance in a variety of rational expectations models. Recent papers have assessed the effects of forward guidance using linearized New Keynesian DSGE models. These papers have studied experiments in which the policymaker announces that the policy rate will follow a particular path for a finite number of periods, thereafter being set in accordance with the monetary policy rule embedded in the model.

Specifically, the experiments assume that the monetary policy reaction function takes the following form:

$$r_t = f\left(x_t, \mathbb{E}_t x_{t+1}, x_{t-1}\right) + \epsilon_t^r \tag{1}$$

where r denotes the log-deviation of the (gross) nominal interest rate from steady state and f is a linear function of the vector of endogenous variables in the model (x) which may enter contemporaneously, as lags or as expected future values, where $\mathbb{E}$ denotes the expectations operator.[7] The term ϵ_t^r represents an exogenous shock to the policy rule. The experiment proceeds by computing a sequence of shocks $\{\epsilon_{t+i}^r\}$, $i = 0, \ldots, K - 1$ that, *when fully anticipated by agents in the model*, ensure that the policy instrument will follow the desired path for K periods (in the absence of the arrival of other shocks).[8]

This experiment has been conducted in both small-scale calibrated New Keynesian models and larger-scale estimated models of the type commonly used for forecasting and policy analysis at central banks. The results of these experiments in both sets of models have been striking: for example, for a common calibration of the prototypical New Keynesian model, Carlstrom et al. (2012) show than an eight quarter reduction in the policy rate by 4% can generate an immediate rise in annualized inflation of around 50%.[9] Carlstrom et al. (2012) and Laséen and Svensson (2011) document very large responses in, respectively, the Smets and Wouters (2007) model and the RAMSES DSGE model used at the Riksbank.[10] Indeed, these papers also demonstrate that both of these models can generate (large) *falls* in output and inflation in response to an anticipated reduction in the policy rate.[11] So the effects of these experiments on

---

[6]See, for example, Yellen (2012) and Bean (2013).

[7]The form of the policy rule (1) is not restrictive for the argument presented here: longer leads and lags can be included in the rule without altering the conclusions.

[8]For linear models, the solution is identical to taking a 'stacked time' approach in which the structural equations for the first K periods (with the policy rate treated as exogenous) are stacked together with the rational expectations solution of the model for period $K + 1$ onwards. See Carlstrom et al. (2012) for an example of this type of approach.

[9]Blake (2012) and Levin et al. (2010) have also shown that this experiment generates implausible results in the same prototypical New Keynesian model (as developed by, for example, Woodford (2003) and Galí (2009)).

[10]The RAMSES model is described by Adolfson et al. (2007).

[11]Carlstrom et al. (2012) show that one reason for this result is the presence of lagged state variables. For example, they demonstrate that the simple three-equation New Keynesian model can exhibit this behavior if price setting includes indexation to past inflation, so that a lag of inflation appears in the Phillips curve.

macroeconomic variables are thus typically regarded as implausibly large and sometimes counter-intuitive in their sign.

del Negro et al. (2012, p16–17) label the implausible responses to such simulations as the "forward guidance puzzle":

> [T]he apparently straightforward experiment "let us fix the short term interest rate to x percent for K periods" has implications for the short term rate that go well beyond the K-th period in medium scale DSGE models. As a consequence, these counterfactuals appear to have an over-sized effect on the macroeconomy.

Broadly speaking, the emergence of the puzzle has led to two lines of inquiry. In the context of the prototypical New Keynesian model it is possible to characterize the behavior of the model analytically.[12] This has enabled analysis of the features of the model structure that give rise to the implausible effects. For example, Levin et al. (2010) show that, for this simple model, the behavior of output and inflation when the interest rate is held fixed is determined by the size of the unstable eigenvalue of the transition matrix mapping the vector of current inflation and output gap to the vector of next period's inflation and output gap. The size of the unstable eigenvalue is determined by the product of the slope of the New Keynesian Phillips curve and the interest elasticity of demand. This type of analytical investigation has led to suggestions of how the underlying microfoundations of the model could be modified to eliminate the puzzling results. For example, Kiley (2014) argues that the 'sticky information' assumption delivers more plausible results than the standard New Keynesian assumption of sticky prices.

The second line of inquiry is to investigate whether the nature of the experiment can be modified to deliver more plausible results. del Negro et al. (2012) argue that the magnitudes of movements in long-term interest rates generated in their experiments are somewhat larger than the daily moves in US long-term rates on the days on which the FOMC issued significant forward guidance announcements. A persistently lower path for the nominal interest rate combined with a large initial rise in inflation creates a significant initial fall in the long-term real interest rate, which stimulates demand and validates the short-term rise in inflation. del Negro et al. (2012) therefore suggest a modification to the experiment in which the shocks to the monetary policy rule are used to deliver a particular change in the long-term interest rate as well as to influence the path of the policy rate over the near term. Haberis et al. (2014) allow the announcement that the policy rate will follow a particular path to be imperfectly credible. In this case, the magnitude of the responses to the experiment can be substantially dampened, even if the degree of imperfect credibility is very small.

A common feature of all of the experiments documented above is that they are interpreted as anticipated deviations from the systematic component of monetary policy (ie the function f in equation (1)). Indeed, this type of deviation from usual behavior has been advocated when the policy rate hits the lower bound. Forward guidance could in this case be interpreted as a commitment by the policymaker to hold the policy rate at the zero bound for longer than would be implied by adherence to the monetary policy reaction function that it typically follows (sometimes called a 'lower for

---

[12]See for, example, Blake (2012), Carlstrom et al. (2012) and Levin et al. (2010).

longer' strategy).[13] Committing to a deliberate deviation from normal policy behavior has been labeled "Odyssean forward guidance" by Campbell et al. (2012).

In this paper, I propose an alternative interpretation.

2.2 Modest policy interventions

The experiments documented in the previous section are interpreted a commitment by the policymaker to temporarily behave in an abnormal manner before returning to the normal conduct of monetary policy. Leeper and Zha (2003) show, however, that experiments of this nature may generate unreliable results in many rational expectations models. That is because a significant change in policy behavior is likely to shift agents' beliefs about the way that policy will be conducted in the future. In the experiments discussed in Section 2.1, such effects are ruled out by assumption because monetary policy is described by a single regime (the policy rule). As Leeper and Zha (2003, p1676) note: "Treating regime changes as surprises that will never occur again ascribes to the public beliefs about policy that are inconsistent with actual behavior – the government takes actions that the public thought were impossible." This critique can be applied to the experiments reviewed in Section 2.1 because, in these cases, forward guidance represents a temporary regime change (during the period that policy behaves 'abnormally').

Of course, a broader range of policy experiments can be contemplated in a richer setup. For example, Cooley et al. (1984) argue in favor of a model in which systematic policy behavior may change over time, with 'regime shifts' governed by a well-defined probability model. This expands the probability space over which agents form expectations, enabling rational expectations to be defined in a way that supports a broader range of policy experiments. This approach can be extended to environments of imperfect information. If agents do not perfectly observe the prevailing regime (only a noisy signal conferred by policy actions and statements), then they will apply a Bayesian learning procedure to form their view of the prevailing regime. Indeed, Leeper and Zha (2003) use a model of this type to illustrate the concept of "modest policy interventions" that I use in this paper.

Explicitly including the possibility of regime changes increases the size and complexity of any given model and, in practice, the scale of models typically used for policy analysis and forecasting in central banks precludes this.[14] As an alternative, Leeper and Zha (2003) argue that even a single-regime model can be used to examine "modest policy interventions" if it is estimated using a sample of data in which the policy regime did not change. Such a model is misspecified because it does not allow agents to revise their beliefs about the prevailing monetary policy regime.

However, Leeper and Zha (2003) define a "modest intervention" as a change in the policy instrument that does not prompt a change in agents' beliefs about the way that

---

[13]The motivation is that even if the short-term interest rate is constrained by the lower bound, forward-looking agents anticipate that future monetary policy will be looser than would usually be expected, reducing longer-term interest rates and stimulating demand in the near term. Arguments in favor of such a strategy are often based on the observation that, in simple New Keynesian models, the policy rate associated with the optimal commitment policy stays at the lower bound for longer than the optimal path under discretion, delivering better stabilization of the output gap and inflation. See, for example, Eggertsson and Woodford (2003).

[14]Encouragingly, recent work by Bianchi and Melosi (2013) suggests that building and estimating relatively large-scale models under these assumptions may be feasible in the near future.

policy is conducted. Given the assumption that agents expect policy to be conducted in line with recent behavior, such a policy experiment is likely to provide a good estimate of its effects. In contrast, policy experiments that are not regarded as modest interventions may generate "expectation-formation effects" as agents revise their view of the prevailing policy regime. These expectation-formation effects are not captured in the policy advisor's misspecified (single-regime) model, but can play a material role in the determination of the true rational expectations responses to a policy intervention. This means that policymakers (and their advisors) should be cautious of the results of policy experiments that would not be interpreted as modest policy interventions by agents in the model.

Leeper and Zha (2003) show how to test whether a policy experiment is likely to be regarded as a modest policy intervention by agents in the model. Loosely speaking, an intervention is modest if agents judge that the effects of that intervention are sufficiently similar to those typically observed under the prevailing policy regime.

The interpretation of forward guidance as a modest policy intervention chimes with some policymaker's descriptions of the rationale for their policies. For example, when discussing the forward guidance policy introduced in August 2013 by the Bank of England's Monetary Policy Committee, Bean (2013) notes that:

> This guidance is intended primarily to clarify our reaction function and thus make policy more effective, rather than to inject additional stimulus by pre-committing to a time-inconsistent lower for longer policy path in the manner of Woodford (2012).

This implies that this forward guidance policy was regarded as consistent with previous policy behavior rather than as a temporary period of abnormal behavior consistent with an Odyssean interpretation of the forward guidance strategy.

3   Forward guidance as a modest policy intervention

In this section I describe my approach for simulating forward guidance as a modest policy intervention. I consider the case in which a policy advisor is asked to estimated the effects of providing a signal to private agents about the future path of the policy instrument. I assume that the policy advisor uses a DSGE model estimated over a stable policy regime. I also assume that policy advisor uses a DSGE model that is linear in terms of variables measured as log-deviations from steady state.[15] Although the model is estimated over a period in which the policy regime has not changed, it is misspecified. That is because, as described previously, the model incorporates the assumption that agents believe that the probability of a policy regime change is zero. Since agents do not place any probability of policy behavior shifting to a new regime, only policy experiments that are consistent with agents' beliefs about policy behavior during the estimation period are likely to generate reliable estimates.

---

[15]I restrict attention to linearized models because most DSGE models used for forecasting and policy analysis at central banks are linearized, presumably because this facilitates their use for such experiments, given their size.

3.1 The monetary policy rule

A fundamental difference between my analysis and the experiments discussed in Section 2.1 is the interpretation of the monetary policy rule within the policy advisor's model. In the studies discussed previously, the shock ϵ_t^r to the monetary policy rule (1) is interpreted (explicitly or otherwise) as a deviation from the policymaker's systematic response (the function $f(\cdot)$) to developments in the economy. In contrast, I regard the monetary policy rule (1) as the policy advisor's model of private agents' beliefs about *systematic* monetary policy.

I assume that the monetary policy rule has the same generic form as equation (1). However, the disturbance to the rule is assumed to follow a particular stochastic process

$$\epsilon_t^r = \rho_r \epsilon_{t-1}^r + \sum_{j=0}^{J-1} \sigma_{\nu,j} \nu_{j,t-j}^r \tag{2}$$

This formulation of policy behavior means that, in each period, information about the future setting of the policy rate is revealed, which is the mechanism through which policymakers may influence expectations of the future path of the policy instrument. This mechanism is captured within the policy advisor's model of private agents beliefs in the form of the disturbances $\{\nu_{j,t}^r\}_{j=0}^{J-1}$. Here $\nu_{j,t-j}^r$ represents a j period ahead disturbance to the policy rule that is revealed in period $t-j$. Thus signals about policy made in period $t-j$ affect the policy rate in period t. I assume that all shocks ν^r are normally distributed with unit variance, so that $\sigma_{\nu,j} > 0$ measure the standard deviation of these disturbances.

Milani and Treadwell (2012) estimate a model with this type of policy disturbance, naming the $\nu_{j,t-j}^r$ terms as "policy news shocks".[16] Campbell et al. (2012) use a rule with disturbances of this form to analyze 'Odyssean' forward guidance over future policy in terms of the $\nu_{j,t-j}^r$ terms. However, as noted in Section 2.1, they interpret policy news shocks as anticipated deviations from the policymaker's usual behavior, interpreted as the function $f(\cdot)$ in equation (1). In contrast, I interpret $\{\nu_{j,t-j}^r\}_{j=0}^{J-1}$ in terms of the policy advisor's model of private agents' beliefs about monetary policy. The policy advisor's model of agents' true beliefs is imperfect (in particular, because it does not incorporate the possibility that agents may ascribe some probability to a regime change). This interpretation is crucial, because although agents in the model are assumed to treat $\{\nu_{j,t-j}^r\}_{j=0}^{J-1}$ as stochastic, they are in fact part of the systematic component of actual policy behavior. As such, they can be treated as choice variables by the policymaker and the policy advisor, as explained by Leeper and Zha (2003).

This interpretation may seem odd to some readers, but Sims (2011, p3) notes that it can be rationalized from a Bayesian perspective:

> From the perspective of a policy maker, her own choices are not "random", and confronting her with a model in which her past choices are treated as "random" and her available current choices are treated as draws from a probability distribution may confuse or annoy her. Indeed economists who provide policy advice and view probability from a frequentist perspective may themselves find this framework puzzling. A Bayesian perspective on inference makes no distinction between random and non-random objects. It distinguishes known or already observed objects from unknown

---

[16]Hirose and Kurozumi (2011) and del Negro et al. (2013) also augment their policy rule in this way.

objects. The latter have probability distributions, characterizing our uncertainty about them. There is therefore no paradox in supposing that econometricians and the public may have probability distributions over policy maker behavior, while policy makers themselves do not see their choices as random. The problem of econometric modeling for policy advice is to use the historically estimated joint distribution of policy behavior and economic outcomes to construct accurate probability distributions for outcomes conditional on contemplated policy actions not yet taken.

As noted, for the policy advisor, equations (1) and (2) describe an approximate model of agents' beliefs about monetary policy behavior. The quality of this approximation to the true probability model of agents in the actual economy will be important for the robustness of the results based on my procedure.

3.2 The model-based forecast

The set of equations in the model – including the policy rule (1) and the policy news process (2) – can be written as:

$$H_B\left(\theta\right) x_{t-1} + H_C\left(\theta\right) x_t + H_F\left(\theta\right) \mathbb{E}_t x_{t+1} = \Psi\left(\theta\right) z_t$$

where the vector z captures the shocks which are independently normally distributed, with unit variance: $z_t \sim N\left(0, \mathbb{I}\right)$. The vector x_t contains all of the endogenous variables in the model, including the policy instrument r_t and the forcing process ϵ_t^r that enters the monetary policy rule (1).

The matrices of coefficients (H_B, H_C and H_F) depend on a vector of parameters θ describing preferences and technology in the underlying model. From this point, I assume that the parameter vector θ is fixed and treated as known by the policy advisor. However, the analysis that follows could be easily extended to account for uncertainty about θ, for example by repeating the policy experiment using a set of draws from an estimated distribution of θ.

The rational expectations solution of the model can be written as:

$$x_t = B x_{t-1} + \Phi z_t \tag{3}$$

where the fact that the coefficient matrices B and Φ depend on the parameter vector θ has been suppressed for notational convenience.

The shocks z_t are partitioned as

$$z_t \equiv \left[\begin{array}{c} \eta_t \\ \nu_t \end{array} \right]$$

where η denotes the vector of non-policy shocks and ν collects the vector of policy news shocks:

$$\nu_t \equiv \left[\begin{array}{c} \nu_{0,t}^r \\ \nu_{1,t}^r \\ \vdots \\ \nu_{J-1,t}^r \end{array} \right]$$

The matrix Φ can be partitioned in a conformable manner and and the rational expectations solution (3) written as:

$$x_t = Bx_{t-1} + \Phi_\eta \eta_t + \Phi_\nu \nu_t \tag{4}$$

The policy advisor uses data for a set of observable variables over the sample $t = 1, \ldots, T$ to form a baseline forecast from the model over the period $T + 1, \ldots, T + H$. To produce this baseline forecast, the transition equation equation (4) is projected forward from an estimate of the current state vector, denoted $x_{T|T}$. This estimate can be produced using the Kalman filter together with a set of measurement equations linking the state vector x to a set of observable variables.[17] The projection at horizon $h \in \{1, \ldots, H\}$ is given by:

$$\mathbb{E}_T x_{T+h} = B^h x_{T|T} \tag{5}$$

3.3 Policy experiments

The model-based forecast is based on the assumption that agents expect future shocks to be zero. In particular, the policy advisor's model of agents' beliefs implies that agents expect the policy news shocks at date $T + 1$ to be zero ($\mathbb{E}_T \nu_{j,T+1} = 0, \forall j = 0, ..., J - 1$).[18] To study the implications of a particular path for the instrument, the policy advisor chooses a particular vector of policy news shocks $\bar{\nu}_{j,T+1}, j = 0, ..., J - 1$. The vector is chosen to implement a particular projection for the policy instrument, computed using the methods shown below. This experiment is interpreted as an announcement of future policy intentions that is communicated to agents at the point at which they make their decisions in period $T + 1$.

The vector of policy news shocks chosen by the policy advisor generates a new forecast given by

$$\bar{x}_{T+h} = B^{h-1} \left(Bx_{T|T} + \Phi_\nu \bar{\nu}_{T+1} \right)$$

The difference between the forecast under forward guidance and the model-based projection is:

$$\bar{x}_{T+h} - \mathbb{E}_T x_{T+h} = B^{h-1} \Phi_\nu \bar{\nu}_{T+1}$$

The policy advisor wishes to assess whether the result of the policy intervention $\bar{\nu}_{T+1}$ would be regarded by agents in the model as consistent with a modest policy intervention. To do this, the policy advisor considers the effects on a particular set of variables. Specifically, suppose that the variables of interest are given by

$$y_t = Qx_t + A\omega_t \tag{6}$$

where $\omega_t \sim N(0, \mathbb{I})$ is a vector of iid measurement errors. Then the effect of the policy intervention on the forecasts for these variables is given by:[19]

$$\bar{y}_{T+h} - \mathbb{E}_T y_{T+h} = Q \left[\bar{x}_{T+h} - \mathbb{E}_T x_{T+h} \right] \tag{7}$$

---

[17]The details of the measurement equation are unimportant for the exposition of the technique, but in the empirical exercises presented in Section 4, the measurement equations are made explicit.

[18]Of course, the expected path for the policy rate in the model-based forecast will be affected by policy news shocks that arrived within the sample period. For example, $\nu_{J-1,T-1}$ is part of agents' information set at date T and will affect date-T expectations of the policy rate in period $T + J - 2$.

[19]Forecasts of measurement errors satisfy $\mathbb{E}_T \omega_{T+h} = 0, \forall h \geq 1$.

To assess whether agents are likely to regard this effect as the result of a modest policy intervention, I use the multivariate modesty statistic developed by Adolfson et al. (2005). They show how to test whether the change in the forecast is statistically likely with respect to the distribution of outcomes implied by the assumed probability model for ν_{T+1}.[20] Adolfson et al. (2005) propose that if y contains n_y variables, then the following test statistic should be compared against critical values from a $\chi^2(n_y)$ distribution:

$$M_T^h = [\bar{y}_{T+h} - \mathbb{E}_T y_{T+h}]' \Omega_{T+h}^{-1} [\bar{y}_{T+h} - \mathbb{E}_T y_{T+h}] \tag{8}$$

where $\Omega_{t+h} = QP_{T+h|T}Q' + AA'$ and P is computed using the following recursion:[21]

$$P_{T+h|T} = BP_{T+h-1|T}B' \tag{9}$$

with $P_{T+1|T} = P_{T|T} + \Phi_\nu \Phi_\nu'$.

Adolfson et al. (2005) note that if the structure of the model is such that the history of observable variables is sufficient to uniquely identify the state vector $x_{T|T}$, then the iterations for the P matrix will start from $P_{T|T} = 0$. However, the inclusion of the policy news shocks means that the conditions under which this is true will generally not apply in my case. That is, the "poor man's invertibility condition' discussed by Fernández-Villaverde et al. (2007) is violated if the number of shocks exceeds the number of observable variables, which is very likely in applications of my approach. In this case, the initialization for $P_{T|T}$ can be computed from the Kalman smoother. The inclusion of policy news shocks also means that the multivariate modesty statistic can typically be computed for a larger set of variables than in Adolfson et al. (2005), because the rank of P is equal to the number of policy news shocks (that is, the dimension of ν_{T+1}).

Written in terms of the shocks used to implement the policy intervention, the modesty statistic is:

$$M_T^h = \left[QB^{h-1}\Phi_\nu \bar{\nu}_{T+1}\right]' \Omega_{T+h}^{-1} \left[QB^{h-1}\Phi_\nu \bar{\nu}_{T+1}\right] \tag{10}$$

or

$$M_T^h = \bar{\nu}_{T+1}' W \bar{\nu}_{T+1} \tag{11}$$

where $W \equiv \Phi_\nu' (B')^{h-1} Q' \Omega_{T+h}^{-1} QB^{h-1} \Phi_\nu$. The weighting matrix W is a function only of the parameters of the model and does not depend on $\bar{\nu}_{T+1}$

The preceding analysis has shown how to test whether a particular shock vector $\bar{\nu}_{T+1}$ would be regarded as a modest policy intervention by agents in the model. I now turn to the task of computing a shock vector that delivers a desired path for the policy instrument. Specifically, I assume that the policy advisor is asked to estimate the effects of an announcement that ensures that private agents' expectations of the policy rate follow a particular path for periods $T+1, \ldots, T+K$, with $K \leq J$.

I assume that the policy advisor uses the first $j = 1, \ldots, n$ policy news shocks (with $K \leq n \leq J$) to implement the policy experiment.[22] This enables me to compare

---

[20]Adolfson et al. (2005) focus on 'constant interest rate projections' generated by a model-based forecast constrained so that the projection of policy-relevant variables meet particular target criteria under the assumption that the policy rate remains fixed over some horizon. Those forecasts are different from the policy simulations considered in my paper because they use the target criteria for policy-relevant variables to compute the required (constant) path for the policy rate.

[21]The recursion presented by Adolfson et al. (2005) contains an additional term that captures the effect of *unanticipated* shocks over the forecast horizon. This term is absent from my formulation because the shocks used to implement the policy experiment are revealed in period $T+1$.

[22]I assume $\bar{\nu}_{T+1,j} = 0$ for $j = n+1, \ldots, J$.

alternative approaches for implementing the policy experiment, including those used in almost all of the previous studies discussed in Section 2.1. From the perspective of the model, the most natural approach would be to set $n = J$. But most existing analyses of such policy experiments choose $n = K$.

Let the $K \times 1$ vector for the desired path of the policy rate be denoted $\bar{r}$ and let S_r be the selector matrix that isolates the row of x corresponding to the policy rate. The shocks that deliver the desired path for the policy instrument must satisfy:

$$\bar{r} = R\bar{\nu}_{T+1} + c$$

where

$$R = \begin{bmatrix} S_r \Phi_\nu \\ S_r B \Phi_\nu \\ \vdots \\ S_r B^{h-1} \Phi_\nu \\ \vdots \\ S_r B^{K-1} \Phi_\nu \end{bmatrix}$$

and c represents the model-based projection for the instrument, given by

$$c = \begin{bmatrix} S_r B x_{T|T} \\ S_r B^2 x_{T|T} \\ \vdots \\ S_r B^h x_{T|T} \\ \vdots \\ S_r B^K x_{T|T} \end{bmatrix}$$

For the case in which $n > K$ there are many vectors $\bar{\nu}_{T+1}$ that can be chosen to impose the desired path for the policy rate. I consider three approaches to finding a vector $\bar{\nu}_{T+1}$.

Method 1: least squares solution

The first method chooses the shocks to minimize the sum of their squared values. The minimization problem is:

$$\min \frac{1}{2} \bar{\nu}'_{T+1} \bar{\nu}_{T+1} - \lambda' \left(R\bar{\nu}_{T+1} + c - \bar{r} \right)$$

which has the solution:

$$\bar{\nu}^* = R' \left(RR' \right)^{-1} \left(\bar{r} - c \right) \tag{12}$$

This vector may or may not imply that the policy intervention is modest. This can be assessed by evaluating the test statistic (11) using the shock vector computed using (12) with the relevant χ^2 critical value. Methods 2 and 3 allow the policy advisor to constrain the experiment to be modest.

Method 2: the minimum distance solution

If the value of M_T^h evaluated using ν^* implies that the experiment conducted using Method 1 does *not* represent a modest policy intervention, then the new projection for the variables of interest may be unreliable. In the case when $n = K$, ν^* is the unique vector of shocks that ensures that the policy rate is expected to follow the desired path. In this case, the policy advisor can find a shock vector that delivers a projection for the policy rate that is as close to the desired path as possible, but would still be considered to be a modest policy intervention from the perspective of agents in the model.

The minimization problem in this case is:

$$\min \left(R\bar{\nu}_{T+1} + c - \bar{r}\right)' \left(R\bar{\nu}_{T+1} + c - \bar{r}\right)$$

subject to:

$$\bar{\nu}'_{T+1} W \bar{\nu}_{T+1} - \bar{M} = 0$$

where $\bar{M}$ is the target value of the modesty statistic imposed by the policy advisor. This value can be chosen by computing the critical value of the relevant χ^2 distribution used to evaluate the modesty statistic for an 'acceptable' p-value. For example, the advisor may judge that a p-value of 0.05 is sufficient to ensure that the intervention is considered modest by the agents in the model.

The minimization problem is a quadratically constrained quadratic programming (QCQP) problem that can be solved using a variety of numerical methods. A common issue with numerical search procedures of this type is that they may converge to local minima. A heuristic approach to guard against this is to start the optimizer from a number of distinct initial conditions.

This method can also be applied to the case in which the the test statistic M_T^h evaluated at ν^* is sufficiently small that the policy experiment implemented using Method 1 is judged to be a modest policy intervention. Then the policy advisor can use Method 2 to deliver an expected path for the policy instrument that is as close to the desired path as possible, but generates macroeconomic effects that are as *unlikely* as possible, subject to the constraint that agents in the model would still regarded the experiment as a modest policy intervention. We would expect an experiment implemented in this way to generate *larger* effects on the variables of interest, so this may be a useful approach if a policy advisor wants to explore the range of possible macroeconomic effects of the experiment.[23]

Method 3: the modesty-constrained solution

When $n \geq K$, there will be many shock vectors that implement the desired path for the policy instrument. This means that the policy advisor can find a set of shocks that delivers the desired path *and* ensures that the experiment is considered modest by agents in the model. In practice, this means incorporating a constraint to ensure that the modesty statistic attains a particular critical value of the relevant χ^2 distribution, denoted $\bar{M}$.

---

[23]However, in general, there is not a direct correlation between the size of the effects on the variables of interest and the extent to which the policy experiment is considered to be a modest policy intervention. This is demonstrated in Section 4.4.

To implement this method, I specify another QCQP problem, given by:

$$\min S\left(\bar{\nu}_{T+1}\right) \equiv \left(D\bar{\nu}_{T+1} - D\bar{\nu}^{*}\right)' \Sigma^{-1} \left(\bar{D}\nu_{T+1} - D\bar{\nu}^{*}\right) \tag{13}$$

subject to:

$$R\bar{\nu}_{T+1} + (c - \bar{r}) = 0$$
$$\bar{\nu}'_{T+1} W \bar{\nu}_{T+1} - \bar{M} = 0$$

where, again, $\bar{M}$ is the target value of the modesty statistic imposed by the policy advisor.[24] The matrices D and Σ are: $D = QB^{h-1}\Phi_{\nu}$ and $\Sigma = DD' - DR'\left(RR'\right)^{-1} RD'$.

Appendix A shows that $D\bar{\nu}^{*}$ and Σ are the mean and variance of the distribution of effects on the variables of interest, *conditional* on the shocks delivering the desired path for the policy instrument (i.e., $R\nu = \bar{r} - c$). This objective function is chosen because it provides a metric for considering the plausibility of the experiment from the perspective of the distribution of shocks that generate the desired path for the policy instrument. Section 5 provides an example of how $S\left(\bar{\nu}_{T+1}\right)$ can be used as a diagnostic on experiments constructed using this method.[25]

3.4 Comparison with existing approaches

Most of the existing studies documented in Section 2.1 construct their policy experiments using Method 1 with $n = K$ policy news shocks. In this case, the vector of policy news shocks that implements the experiment, ν^{*}, is unique. Existing approaches do not inspect the properties of ν^{*} or whether the effects on the rest of the model would be regarded as a modest policy intervention.[26] These studies effectively assume that the policy experiment is reliable and that agents within the model would regard policy behavior as consistent with some underlying belief about the systematic conduct of policy. My approach treats this question as an empirical matter. The extent to which ν^{*} can be considered as consistent with agents' beliefs about the distribution of ν_{T+1} determines whether the results from the policy experiment are likely to be reliable. In constrast, very few existing studies characterize a model for agents' beliefs about the properties of ν_{T+1}.[27]

---

[24] Again, a heuristic approach of starting from many initial conditions is recommended to ensure that a global minimum is found.

[25] To gain some intuition, note that, because $n \geq K$, it will in general be possible to deliver the desired path for the policy rate and ensure that the intervention is modest at any significance level. In the limit, the policy advisor could set $\bar{M} = 0$. Inspection of equation (10) reveals that this limiting case would require that the variables of interest (y) are unchanged from their values in the baseline forecast. This is likely to generate a somewhat contorted path for the policy instrument, since it is constrained to follow the desired path for the first K periods and must subsequently move in a way that completely offsets the effects on the variables of interest. The subsequent movements in the policy rate may be considered implausible and the objective function for Method 3 is intended to provide a way to measure that.

[26] Some studies (for example, Carlstrom et al. (2012)) use a piecewise linear, or stacked-time, solution approaches to compute the effects of (anticipated) temporary deviations from a monetary policy rule. This means that the shocks ν^{*} are not computed directly. But, as noted in Section 2.1, the results computed using those methods are identical to the use of $n = K$ policy news shocks to compute $\nu*$ using Method 1.

[27] Indeed, del Negro et al. (2012) argue that (because ν^{*} is unique) the variances of the policy news shocks are not important. Those studies that do make explicit assumptions about the distribution of the policy news shocks (for example, Hirose and Kurozumi (2011) and Milani and Treadwell (2012)) do not examine the type of policy simulations that I study in this paper.

Method 2 has some similarities to the approach proposed by del Negro et al. (2012): in both cases the effects of the policy experiment may be regarded as more plausible than when the experiment is implemented using Method 1; and the path for the policy instrument that agents expect may not exactly match the desired path. The motivation for the approach proposed in del Negro et al. (2012) is that the responses of longer-term interest rates to experiments implemented using Method 1 are too large.[28] To generate more plausible responses, del Negro et al. (2012) use anticipated shocks to the monetary policy rule to deliver a particular change in the long-term interest rate (specified by the policy advisor conducting the experiment). The influence of the shocks on the path of the policy rate in the short term is only indirect, through the use of a weighting function that penalizes the implied changes in the expected policy rate over a ten year horizon.[29]

A key difference between the approaches is that my method uses a test for assessing the reliability of the estimated macroeconomic effects based on the beliefs of the agents that inhabit the estimated model, rather than a penalty function designed by the policy advisor constructing the experiment. Of course, my method still requires some judgment on the part of the policy advisor, both in terms of the variables which are used to construct the modesty statistic, the forecast horizon at which the statistic is tested and the probability level that constitutes whether or not a particular policy experiment is deemed to be "modest".[30] Nevertheless, my procedure provides the policy advisor with a framework in which the sensitivity of her conclusions may be examined very easily.

4   Empirical exercise

In this section, I use the approach outlined in Section 3 to conduct policy experiments in a medium-scale DSGE model. Before presenting the policy experiments, I first outline the model and the estimation approach.

4.1   The model

I use a variant of the Smets and Wouters (2007) model, which has formed the blueprint for a number of models in active use at central banks.[31] The model structure is attractive to policymakers because it contains similar frictions to those used by Christiano et al. (2005) to demonstrate that a DSGE model can replicate empirical estimates of

---

[28]del Negro et al. (2012) observe that a simulation that temporarily lowers the expected path of the policy instrument lowers the expected path for the policy instrument for a very prolonged period. They assess the plausibility of the size of the model's predictions by comparing them to typical changes in long-term interest rates observed for dates on which the FOMC made forward guidance announcements.

[29]del Negro et al. (2012) argue that, conditional on imposing a particular change in the ten-year spot rate, the macroeconomic effects are relatively unaffected by the path for the short-term interest rate that delivers that change. The weighting function can therefore be used to influence the path of the policy rate over the shorter term, while delivering plausible macroeconomic responses.

[30]As shown in Section 6.1, these choices may have important effects regardless of whether the estimated macroeconomic effects of a particular policy experiment are deemed modest.

[31]For example, the Riksbank's RAMSES model (Adolfson et al., 2007); the ECB's NAWM (Christoffel et al., 2008); the EDO model of the Federal Reserve Board of Governors (Edge et al., 2007; Chung et al., 2010); and the Bank of England's COMPASS model (Burgess et al., 2013).

the effects of a monetary policy shock. The model is also attractive because Smets and Wouters (2003) and Smets and Wouters (2007) have shown that, when estimated using Bayesian methods, the model can compete with VARs in its ability to fit (and forecast) US and euro area data.

The model is very close to the specification of Smets and Wouters (2007). I make two small modifications to the monetary policy rule. First I include "policy news" shocks in the specification of the monetary policy rule, as specified in equation (1), following Milani and Treadwell (2012), Campbell et al. (2012), Hirose and Kurozumi (2011) and del Negro et al. (2013) among others. As noted in Section 3, these shocks are the mechanism through which forward guidance experiments are implemented. The second modification is to exclude the term in the *change* in the output gap from the monetary policy rule. In preliminary estimation work, I found that the coefficient on this term was quite difficult to identify alongside the coefficient on the output gap itself.[32]

I also include equations for the expectations theory of the term structure and link them to data on longer term bond yields, as in De Graeve et al. (2009) and Hirose and Kurozumi (2011). While this addition does not change the macroeconomic structure of the model, it does provide additional information with which to identify the policy news shocks.

Since the Smets and Wouters (2007) model is well known, I focus on the key differences. The policy rule is:

$$r_t = \rho_r r_{t-1} + (1 - \rho_r) \left[r_\pi \pi_t + r_y \left(y_t - \tilde{y}_t \right) \right] + \epsilon_t^r \tag{14}$$

where r denotes the log-deviation of the (gross) nominal interest rate from steady state and π is quarterly (consumer price) inflation. The output gap is defined as the log deviation of output, y, from the level of output that would prevail if prices and wages were fully flexible and price and wage markups were constant, denoted $\tilde{y}$. The disturbance to the policy rule ϵ_t^r is given by equation (2). I assume that news about future policy actions can be communicated over a three year horizon (so $J = 12$ in equation (2)).

The equations for longer-term nominal interest rates are simple representations of the expectations theory of the term structure:

$$r_t^N = \frac{1}{N} \sum_{i=0}^{N-1} \mathbb{E}_t r_{t+i} \tag{15}$$

and I include equations for the one and five year spot nominal interest rates ($N = 4, 20$). Given my assumption that policy news shocks extend for three years, these yields are included to provide a balance between providing information to help identify the shocks containing news about the near future and a more general indication of the expected path of the policy rate over the longer term.

The measurement equations for the macroeconomic variables are identical to those used by Smets and Wouters (2007). For the long-term bond yields I use the following measurement equations to map to the raw data for spot rates (denoted R^N):

$$100 \frac{R_t^N}{4} = r_t^N + 100 \left(\bar{\pi} \bar{\beta}^{-1} - 1 \right) + \tau_N + \sigma_{me,N} \omega_t^N \tag{16}$$

---

[32]Hirose and Kurozumi (2011) also exclude the change in the output gap from their policy rule specification (though they use a production function measure of the output gap rather than a model-consistent measure of potential output).

which states that the bond yield is equal to the steady state (short-term) interest rate, plus a constant term premium (τ_N) and a measurement error $\omega_t^N \sim N(0,1)$.

4.2   Data and estimation

For the macroeconomic variables, I use an updated version of the dataset constructed by Smets and Wouters (2007).[33] The yield curve data is from the Federal Reserve Board of Governors website, computed using the methodology of Gurkaynak et al. (2007). The daily yield curve estimates are converted to quarterly frequency by taking simple averages. Given the discussion in Section 3, it is important to estimate the model over a period that can be regarded as a stable monetary policy regime. I estimate the model over the period 1984Q1–2008Q4 in order to ensure that the sample size is reasonably large.[34] Given that the results in Hirose and Kurozumi (2011) suggest that the properties of policy news shocks may not have been stable across this period, I estimate the model over alternative subsamples to investigate robustness in Section 6.2.

I estimate the model using Bayesian likelihood estimation.[35] Following Smets and Wouters (2007), five parameters are held fixed during the estimation process. The depreciation rate δ is set at 0.025. The steady state ratio of government spending to GDP (g_y) is set at 0.18 and the steady state wage markup (λ_w) is calibrated to 1.5. Finally, the curvature of the Kimball aggregators in the goods and labor markets are set to $\varepsilon_p = \varepsilon_w = 10$.

The estimation results are shown Tables 1 and 2. The prior moments for the structural parameters and AR(1) coefficients for the forcing processes are set identically to Smets and Wouters (2007).[36] The priors for the steady state term premia and measurement errors for the long-term yields are intended to be diffuse.

To estimate the posterior parameter distribution, I numerically maximized the posterior density and used a numerical approximation of the Hessian at the mode to produce simulated draws from the posterior using a Random Walk Metropolis Hastings algorithm.[37] I generated 250,000 simulated draws from four chains of length 87,500 (where the first 25,000 observations in each chain are burned). The acceptance rates for each chain were between 25% and 28% and MCMC convergence diagnostics indicated that the chains had converged.

Perhaps unsurprisingly, the posterior parameter estimates are generally very similar to those presented by Smets and Wouters (2007, Table 5) for the 1984–2004 subsample. Some notable exceptions are households' inverse elasticity of substitution and habit formation parameters (σ_c and λ respectively), which I estimate to be somewhat lower, perhaps reflecting an increase in consumption volatility over the financial crisis. The wage stickiness and labor supply elasticity parameters (ζ_w and σ_l) are both estimated to be somewhat lower than the sub-sample results in Smets and Wouters

---

[33]These data were sourced from the Federal Reserve Economic Data database maintained by the St Louis Fed.

[34]Data from 1980–1983 is used as a training sample for the Kalman filter.

[35]See An and Schorfheide (2007) for an excellent review of Bayesian DSGE model estimation.

[36]The distributions chosen for the parameters are also the same with one exception: I use a Gamma distribution to ensure that the inflation response parameter in the Taylor rule satisfies the Taylor principle.

[37]The estimation was performed using the MAPS toolkit described in Burgess et al. (2013).

(2007). Together with the higher variance attributed to the wage markup shock process (given that I estimate higher values for ρ_w and σ_w), this points to a more flexible and volatile wage setting block than estimated by Smets and Wouters (2007). The parameters defining the forcing processes for the shocks are generally quite similar, though I estimate the process for the risk premium shock to be very persistent (ρ_b is close to unity), whereas Smets and Wouters (2007) find little persistence.

For the policy rule parameters, the response of the Fed funds rate to inflation is estimated to be slightly higher and the responsiveness to the output gap slightly lower than in Smets and Wouters (2007). The standard deviation of the contemporaneous monetary policy shock $\sigma_{\nu,0}$ is estimated to be slightly smaller, but the overall variance of all monetary policy shocks (including the policy news components $\sum_j \sigma_{\nu,j}$) is larger. The shock process is also estimated to be more persistent than the results presented in Smets and Wouters (2007).

There are several interpretations for these differences. It is possible that the additional data in my sample has provided valuable information for the identification of the parameters in the model. Alternatively, the difference could reflect a structural break in the data brought about by the emergence of the financial crisis at the end of my sample period. In that case, the assumption that I have estimated the model over a stable policy regime may be cast into doubt, since policy behaviour may have changed in light of the financial crisis.

Table 1: Priors and posteriors for estimated parameters

Parameter	Description	Dist'n	Mean	Std[a]	Mean	Std	5%	95%
			Prior		Posterior			
φ	Capital adjustment cost elasticity	Normal	4	1.5	6.514	1.139	4.694	8.448
σ_c	Inverse elasticity of substitution	Normal	1.5	0.375	0.9763	0.1621	0.7355	1.267
λ	Consumption habits	Beta	0.7	0.1	0.5809	0.07134	0.4617	0.6962
ξ_w	Calvo probability for wages	Beta	0.5	0.1	0.4324	0.1032	0.284	0.6211
σ_l	Labor supply elasticity	Normal	2	0.75	1.436	0.5614	0.584	2.419
ξ_p	Calvo probability for prices	Beta	0.5	0.1	0.7252	0.08279	0.6004	0.8531
ι_w	Wage indexation parameter	Beta	0.5	0.15	0.4517	0.1457	0.2196	0.701
ι_p	Inflation indexation parameter	Beta	0.5	0.15	0.1517	0.06201	0.06387	0.2651
ψ	Elasticity of capital utilization	Beta	0.5	0.15	0.6619	0.1115	0.4675	0.8337
Φ	Production fixed cost	Normal	1.25	0.125	1.528	0.08673	1.388	1.673
$r_\pi - 1$	Policy rule feedback on inflation	Gamma	0.5	0.25	1.187	0.3525	0.6704	1.814
ρ_R	Policy rule response to lagged rate	Beta	0.75	0.1	0.8434	0.0396	0.7727	0.8996
r_y	Policy rule response to output gap	Normal	0.125	0.05	0.01518	0.02904	-0.02132	0.07113
$\bar{\pi}$	Inflation target	Gamma	0.625	0.1	0.6379	0.09544	0.4894	0.802
$100\left(\beta^{-1}-1\right)$	Discount rate	Gamma	0.25	0.1	0.2296	0.08718	0.1062	0.3901
$\bar{L}$	Hours measurement equation constant	Normal	0	2	-0.5214	1.856	-3.579	2.547
$\bar{\gamma}$	Quarterly steady state growth rate	Normal	0.4	0.1	0.4568	0.02559	0.4133	0.4958
ρ_{ga}	Government and productivity shock correlation	Normal	0.5	0.25	0.3962	0.09867	0.2347	0.5589
α	Capital share in production	Normal	0.3	0.05	0.1717	0.02293	0.1341	0.2096
ρ_a	Persistence of producivity process	Beta	0.5	0.2	0.916	0.02493	0.8732	0.9549
ρ_b	Persistence of risk premium process	Beta	0.5	0.2	0.9842	0.009121	0.9675	0.996
ρ_g	Persistence of government spending process	Beta	0.5	0.2	0.9756	0.01163	0.9542	0.9917
ρ_I	Persistence of investment cost process	Beta	0.5	0.2	0.636	0.1086	0.4574	0.8176
ρ_r	Persistence of monetary policy process	Beta	0.5	0.2	0.6232	0.05214	0.5336	0.7042
ρ_p	Persistence of price markup process	Beta	0.5	0.2	0.927	0.06305	0.8115	0.9922
ρ_w	Persistence of wage markup process	Beta	0.5	0.2	0.9898	0.004745	0.9812	0.9965

[a]Degrees of freedom for inverse gamma distributions.

Table 2: Priors and posteriors for estimated parameters (continued)

Parameter	Description	Prior			Posterior			
		Dist'n	Mean	Std[a]	Mean	Std	5%	95%
μ_p	Price markup moving average parameter	Beta	0.5	0.2	0.6838	0.1047	0.4893	0.8325
μ_w	Wage markup moving average parameter	Beta	0.5	0.2	0.6258	0.1804	0.3183	0.8934
σ_a	Std dev of productivity shock	Invgamma	0.1	2	0.3804	0.03018	0.3336	0.4323
σ_b	Std dev of risk premium shock	Invgamma	0.1	2	0.02464	0.006533	0.01529	0.03657
σ_g	Std dev of govt spending shock	Invgamma	0.1	2	0.3866	0.02871	0.343	0.4367
σ_I	Std dev of investment price shock	Invgamma	0.1	2	0.3396	0.06564	0.2409	0.4551
σ_p	Std dev of price markup shock	Invgamma	0.1	2	0.09304	0.01197	0.07388	0.1133
σ_w	Std dev of wage markup shock	Invgamma	0.1	2	0.4039	0.06514	0.3163	0.528
$100 \times \sigma_{\nu,0}$	Std dev of monetary policy shock	Invgamma	2.752	2.001	9.723	1.066	8.142	11.6
$100 \times \sigma_{\nu,1}$	Std dev of 1 quarter ahead policy news shock	Invgamma	2.752	2.001	2.7	1.324	0.7446	4.926
$100 \times \sigma_{\nu,2}$	Std dev of 2 quarters ahead policy news shock	Invgamma	2.752	2.001	5.024	1.402	2.219	6.982
$100 \times \sigma_{\nu,3}$	Std dev of 3 quarters ahead policy news shock	Invgamma	2.752	2.001	3.562	1.565	0.8822	5.905
$100 \times \sigma_{\nu,4}$	Std dev of 4 quarters ahead policy news shock	Invgamma	2.752	2.001	1.432	0.767	0.5317	2.965
$100 \times \sigma_{\nu,5}$	Std dev of 5 quarters ahead policy news shock	Invgamma	2.752	2.001	1.309	0.6326	0.5181	2.514
$100 \times \sigma_{\nu,6}$	Std dev of 6 quarters ahead policy news shock	Invgamma	2.752	2.001	1.392	0.683	0.5393	2.706
$100 \times \sigma_{\nu,7}$	Std dev of 7 quarters ahead policy news shock	Invgamma	2.752	2.001	1.507	0.7798	0.5564	3.045
$100 \times \sigma_{\nu,8}$	Std dev of 8 quarters ahead policy news shock	Invgamma	2.752	2.001	1.316	0.6542	0.5114	2.6
$100 \times \sigma_{\nu,9}$	Std dev of 9 quarters ahead policy news shock	Invgamma	2.752	2.001	1.326	0.6631	0.5197	2.619
$100 \times \sigma_{\nu,10}$	Std dev of 10 quarters ahead policy news shock	Invgamma	2.752	2.001	1.346	0.6509	0.5301	2.616
$100 \times \sigma_{\nu,11}$	Std dev of 11 quarters ahead policy news shock	Invgamma	2.752	2.001	1.778	0.8739	0.61	3.374
$\sigma_{me,4}$	Std dev of measurement error, 1 year spot rate	Invgamma	0.2	2	0.02317	0.004896	0.01588	0.03192
$\sigma_{me,20}$	Std dev of measurement error, 5 year spot rate	Invgamma	0.2	2	0.02918	0.006633	0.01933	0.04093
τ_4	Steady state term premium on 1 year bond	Normal	0	1	0.04965	0.01482	0.02479	0.07335
τ_{20}	Steady state term premium on 5 year bond	Normal	0	1	0.2872	0.05999	0.185	0.3827

[a]Degrees of freedom for inverse gamma distributions.

4.3 Model-based forecast distributions

In this section I briefly summarize the model's forecasts for key variables over the forecast horizon 2009Q1–2013Q2 and compare them to the actual realizations of the data over that period.

Figure 1: Model forecasts and actual outturns for key variables
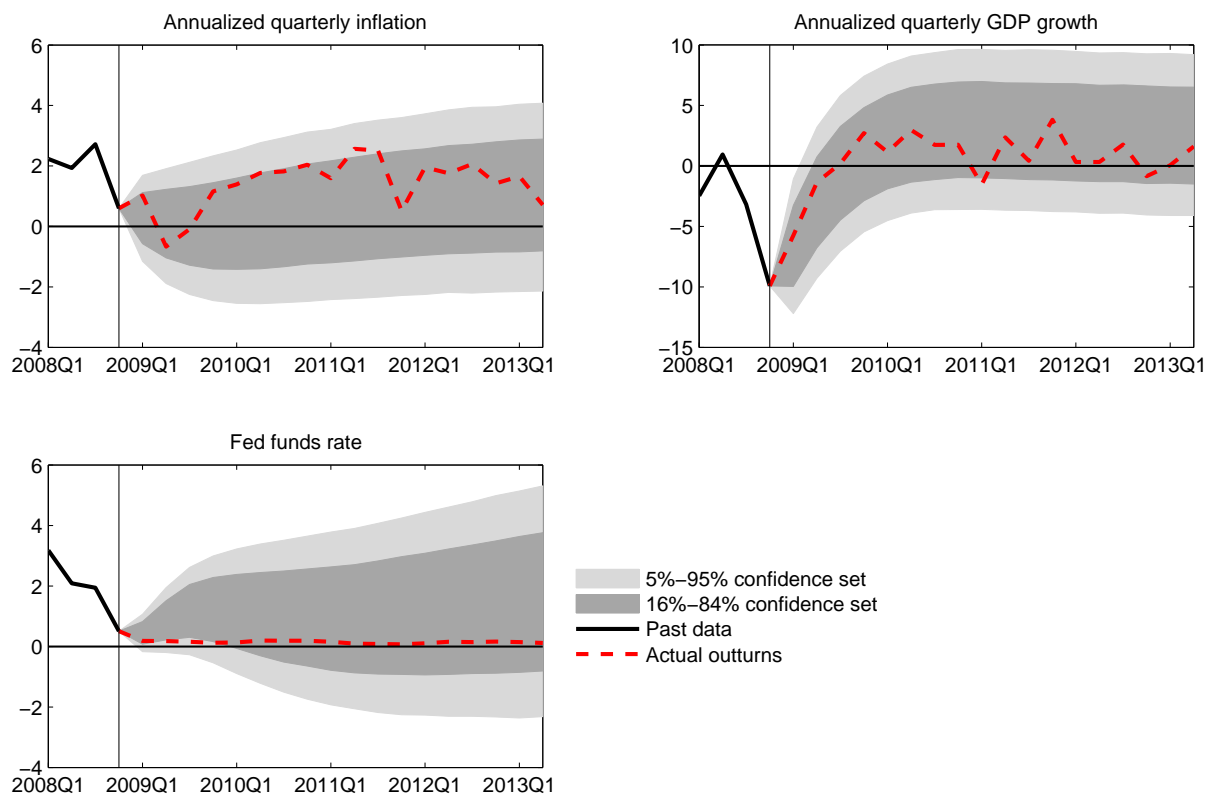
Notes: The panels show confidence sets of forecasts generated by the DSGE model, obtained by drawing from the posterior distribution of the estimated parameters. The past data are shown by the solid black lines and the actual realizations over the forecast horizon are shown by the dashed red lines.

Figure 1 shows that the model's forecasts of inflation and GDP growth were broadly in line with the subsequent realizations of the data over the first year or so of the forecast horizon. Thereafter, the forecast densities diverge from the realized data somewhat, with the model tending to under-predict inflation and over-predict GDP growth. Perhaps most striking is the divergence between the forecast density for the Fed funds rate, which has significantly more mass above the realized path than below it for the entire forecast horizon.

However, the forecast densities in Figure 1 do not impose the zero bound on the Fed funds rate. As shown by Coenen and Warne (2013), imposing the zero bound on such simulations will increase deflation risks considerably, because the policy rate cannot be reduced in response to deflationary shocks. Given that even the *unconstrained* densities in Figure 1 exhibit a sizable deflation risk, policymakers faced with these forecasts would likely be concerned about the limited ability to respond by further reducing the Fed funds rate. Indeed, the minutes of the FOMC meeting on 15–16 December 2008 reveal that some Committee members "expressed concern that inflation might decline below levels consistent with price stability in the medium term".[38]

---

[38]See Federal Open Market Committee (2008, p8).

Such concerns motivate the consideration of policies that might mitigate the risk of deflation. One such experiment is to consider holding the Fed funds rate at a lower level than expected by agents in the baseline forecast. At their meeting on 15–16 December 2008, the FOMC discussed the effects of forward guidance over the path for the Fed funds rate: "participants judged that communicating the Committee's expectation that short-term interest rates were likely to stay exceptionally low for some time could be useful because it would lead to pricing of longer term interest rates consistent with the path of monetary policy that policymakers saw as most likely".[39]

4.4 Forward guidance experiments

In this section, I consider the task of a policy advisor who is asked to estimate the macroeconomic effects of alternative paths for the Fed funds rate. The policy advisor uses the model described in Section 4.1 to estimate the consequences of holding the policy rate temporarily lower than the baseline forecast produced by the model. These experiments are similar to those analyzed by del Negro et al. (2012).

Of course, in the midst of the financial crisis, it may have been entirely sensible for agents to have expected the central bank to take extraordinary actions in response to extraordinary events. However, a policy advisor with a single-regime rational expectations model cannot reliably estimate the effects of such extraordinary policies. As explained in Section 2.2, the best that she can do is estimate the effects of policies that are consistent with agents' beliefs about the conduct of policy in the recent past. This is what I estimate in the following experiments. Within the limitations of my approach, I apply a relatively loose constraint when testing whether an experiment constitutes a modest policy intervention. Specifically, I use the 5% critical values of the relevant χ^2 distributions, which means that I allow experiments that agents would consider relatively unlikely *ex ante* to be classed as modest policy interventions. This captures at least the spirit of the idea that quite radical policy experiments may have been contemplated at this stage of the financial crisis.

A key choice for the policy advisor is the specification of the modesty statistic used to assess the effects of the policy experiment. The advisor must choose the set of variables for which the effects of the experiment will be tested and the horizon at which the responses of those variables are tested. In the context of the analysis in Section 3, this corresponds to the choice of the matrix Q in the mapping $y_t = Qx_t$ and the horizon, h, at which the effects $[\bar{y}_{T+h} - \mathbb{E}_T y_{T+h}]$ are tested. Given that a key part of the forward guidance puzzle is the fact that the effect on macroeconomic variables is so front-loaded, I focus on the impact effects of the policy announcement by setting the horizon equal to $h = 1$.[40] I include the responses of inflation and GDP growth in the set of variables for which modesty is tested, given the focus on inflation and economic activity in del Negro et al. (2012) and Yellen (2012). I assess the robustness of the analysis to these choices in Section 6.1.
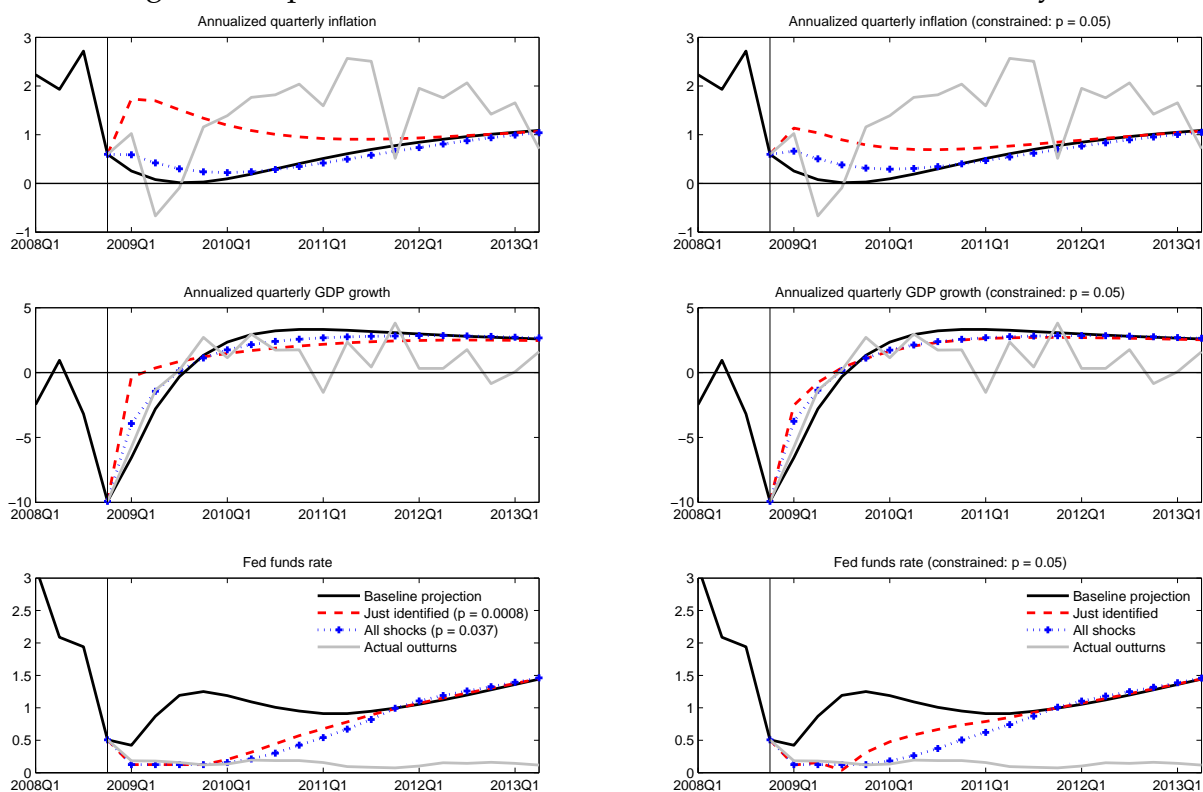
Figure 2 shows the forecasts for GDP growth, inflation and the Fed funds rate under different assumptions about policy. In all charts, the black solid lines depict the baseline forecast from the estimated DSGE model.[41] The gray lines show the realized

---

[39] Federal Open Market Committee (2008, p8).

[40] This means that I test whether the responses are plausible in period $T + 1$, using the modesty statistic (10) with $h = 1$.

[41] The 'baseline projection' is generated using the posterior mean of the parameter vector.

Figure 2: Experiment in which Fed funds rate is held low for one year

Notes: The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 4$ periods is generated using $n = K = 4$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 4$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

outcomes over the forecast horizon. The remaining lines plot the forecasts corresponding to policy experiments in which the Fed funds rate is held fixed at 12.5bps for four quarters.[42] The left and right columns of charts impose this path for policy using the methods introduced in Section 3.

In the left column, the red dashed lines are generated under the assumption that $n = 4$ policy news shocks are used to hold the Fed funds rate at 12.5bps for $K = 4$ quarters: this corresponds to a version of Method 1 that is 'just-identified' (since $n = K$). This implementation generates a sizable increase in both GDP growth and inflation: signs of the "forward guidance puzzle". The p-value of the modesty statistic for this simulation is less than 0.01%, suggesting that agents would not regard this policy experiment as a modest intervention. The dotted blue lines in the left column show the case in which the policy experiment is implemented using $n = J = 12$ policy news shocks. This represents the case in which the intention to hold the Fed funds rate at 12.5bps for one year is signaled using the full set of policy news shocks that appear in equation (1). The macroeconomic effects of this simulation are considerably smaller

---

[42]I choose 12.5bps as this is the midpoint of the targeted range for the funds rate since 2009Q1 (and corresponds roughly to the level of the realized funds rate depicted in the gray line).

than for the 'just identified' case depicted in the red dashed lines.[43]

From a very mechanical perspective, this result follows from the fact that the procedure uses a larger number of monetary policy news shocks. The ability to control the path for the Fed funds rate beyond the four quarters for which the rate follows the desired path makes it possible to support an equilibrium in which inflation and the output gap respond by less and so the size of the policy news shocks required to hold the Fed funds rate on its desired path are smaller. This equilibrium is achieved by selecting policy news shocks that produce a trajectory for the Feds fund rate that lies *above* the baseline path further out in the forecast horizon.[44] Despite the markedly smaller macroeconomic responses, this intervention would still not be considered modest using a significance level of 5%.

The right hand column of Figure 2 show the cases in which the policy experiments are constrained to ensure that the intervention is considered to be modest by agents in the model. A critical value of 5% for the modesty statistic is used to make this assessment. Once again, the dashed red lines are 'just identified' and are implemented using $n = K = 4$ shocks (using Method 2). As expected, in this case the macroeconomic effects are smaller than the case in which the intervention is not constrained to be modest (left hand column). But this is achieved by implementing a path for the Fed funds rate that misses its target level over the first four quarters of the forecast horizon: in particular, the funds rate tightens noticeably in the fourth quarter.

The dotted blue lines in the right hand column of Figure 2 are implemented using the full set of policy news shocks ($n = J = 12$), but using Method 3 so that the intervention is constrained to be modest at the 5% significance level. The responses in this case are very similar to those in the left panel (since in that case, the p-value of the intervention was almost 4%).

An important result from Figure 2 is that the size of the responses of the individual variables are not sufficient for assessing whether the experiment is regarded as modest. For example, the macroeconomic effects of the 'modesty constrained' experiment using $n = K = 4$ policy news shocks using Method 2 (right column, dashed red lines) are larger than those for the case in which Method 1 is used with $n = J = 12$ shocks (*left* column, dotted blue lines), even though the former experiment is considered more likely to be a modest policy intervention than the latter. This result underscores the fact that the modesty statistic is a multivariate statistic and is determined by the co-movement of the variables considered as well as by the absolute size and sign of the effects on each variable.

Figure 3 presents a policy experiment in which the Fed funds rate is held at 12.5bp for two years ($K = 8$). The broad pattern of results in this case is similar to that described for the one-year experiment. In the left hand column, the 'just identified' ($n = K = 8$) implementation depicted in the red dashed lines generates a substantial rise in inflation and an enormous instantaneous boom in GDP growth. From the perspective of the modesty statistic, this is a zero probability event. When implemented

---

[43]As noted in Section 3.3, from the perspective of the model, using the entire set of policy news shocks ($n = J = 12$, generating the blue dotted lines) is the most natural way to implement the experiment, since beliefs over the behavior of monetary policy are affected by all of these shocks. The main motivation for examining the 'just identified' experiment ($n = K = 4$, red dashed lines) is that this is the way the experiment has typically been implemented in recent papers.

[44]This is barely noticeable in Figure 2, but has an important effect on the results. Figure 6 in Section 5 provides more detailed illustrations of this effect for a policy experiment in which the Fed funds rate is held lower for eight quarters.

Figure 3: Experiment in which Fed funds rate is held low for two years
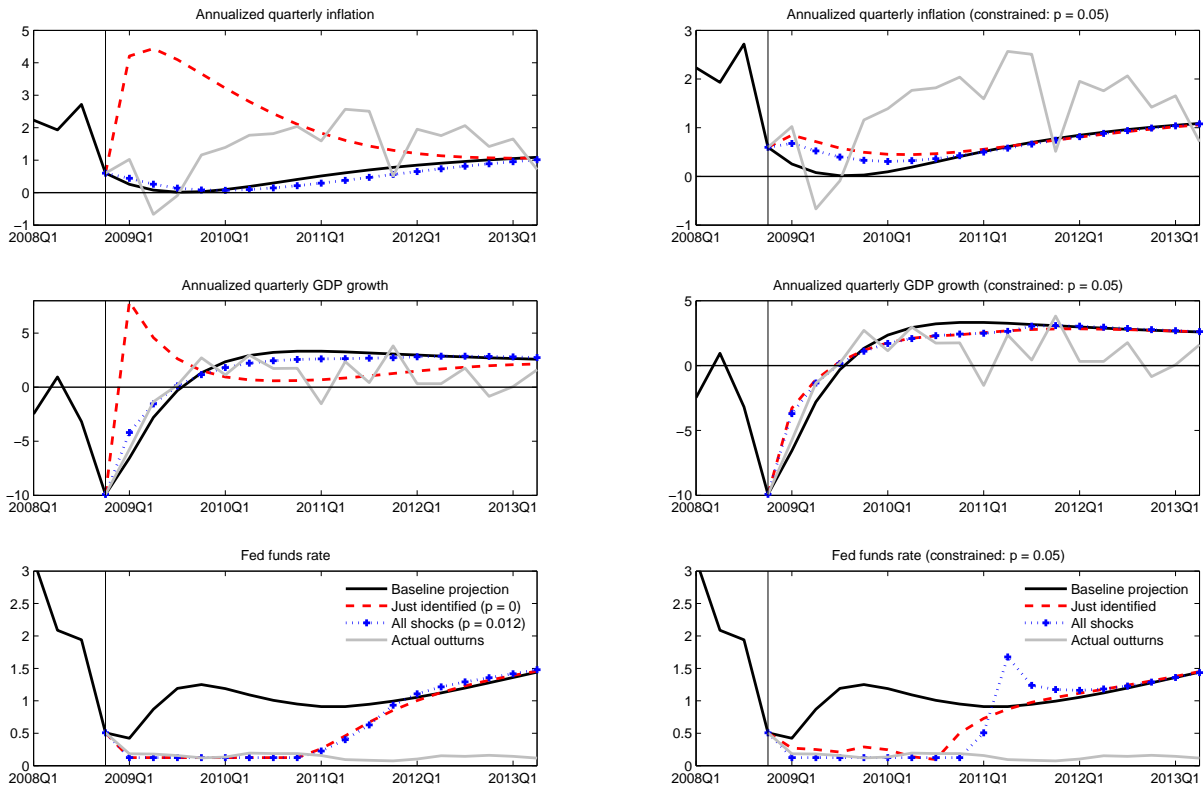
Notes: The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 8$ periods is generated using $n = K = 8$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 8$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

using the full set of available policy news shocks ($n = J = 12$), the macroeconomic effects are substantially smaller, though even so the intervention is not modest at the 5% significance level.

The right column of Figure 3 shows the case in which the policy intervention is constrained to be modest at the 5% significance level. When the experiment is implemented using $n = K = 8$ shocks (red dashed lines), the effects are substantially dampened, but once more this is achieved by missing the desired path for the Fed funds rate and raising the rate somewhat earlier than prescribed by the desired path.

Two interesting results emerge from Figure 3. First, when the experiment is implemented using $n = J = 12$ policy news shocks (dotted blue lines), the effect on inflation is *larger* than the responses plotted in the left hand column. Once again, the multivariate modesty statistic is influenced by the relative co-movement of the macroeconomic variables. Second, the right column of Figure 3 shows that the paths of GDP growth and inflation are almost identical for the cases in which the Fed funds rate path is imposed using $n = K$ and $n = J$ shocks, despite the fact that the paths for the Fed funds rate are quite different. The path for the Fed funds rate tightens earlier (but by less) for the $n = K$ implementation. Again, this emphasizes that it is the entire path for the

policy rate that matters for determining the macroeconomic effects. Both results are explored further in Section 5.

Figure 4: Experiment in which Fed funds rate is held low for three years
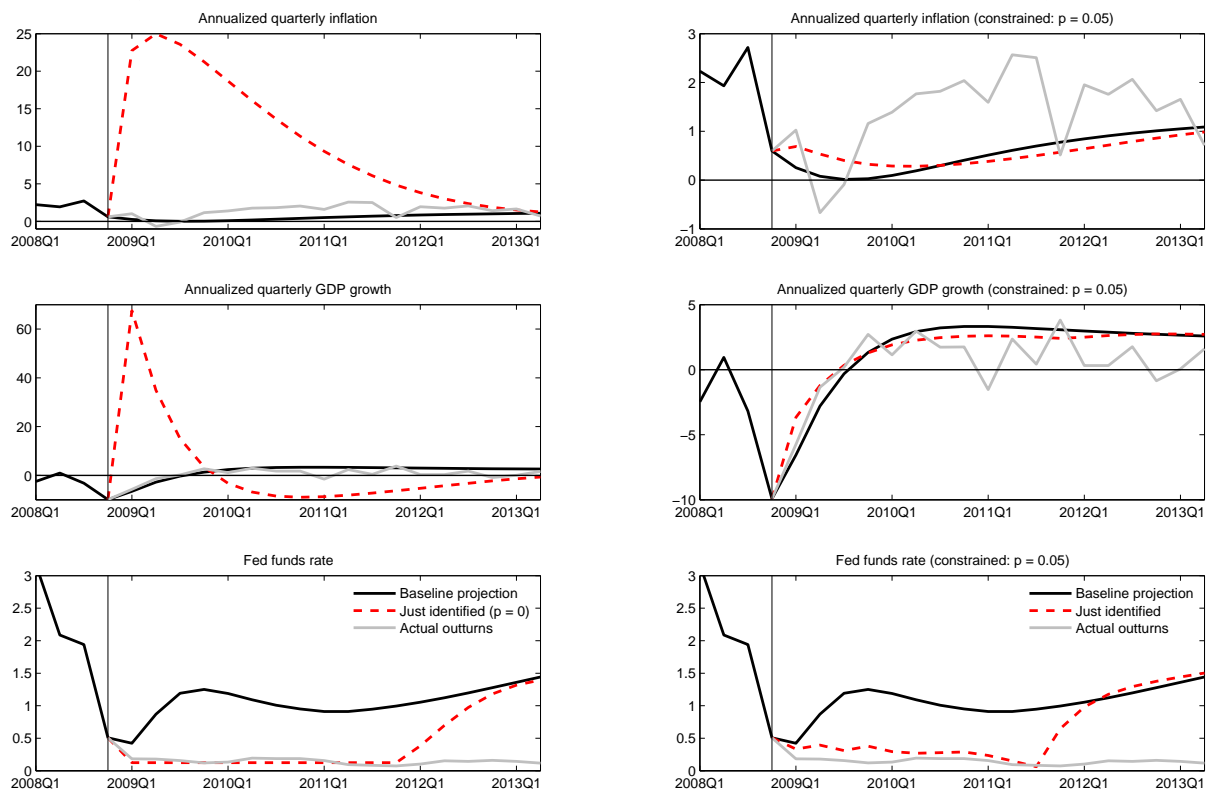
Notes: The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 12$ periods is generated using $n = K = 12$ policy news shocks ('just identified', red dashed lines), using Method 1. The p-value of the modesty statistics associated with this simulation is reported in the legend. The right column repeats the experiments, but imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$. Since $n = K$, Method 2 is used.

Figure 4 shows the case in which the Fed funds rate is held at 12.5bp for three years. Because the Fed funds rate is to be fixed for three years, this experiment requires the use of all 12 available policy news shocks. Implementing the experiment using Method 1 with $n = J = K = 12$ (left column) generates results that are truly incredible: and indeed the p-value of the modesty statistic is zero.[45]

The right column of Figure 4 shows the results when the experiment is constrained to be modest at the 5% significance level. By construction, the Fed funds rate does not follow the desired path exactly. In the twelfth quarter of the forecast horizon, the Fed funds rate rises sharply such that the trajectory of the policy rate lies above the baseline forecast from the fourteenth quarter onwards. This path for the funds rate generates outcomes for growth and inflation that would presumably be regarded as far more plausible by most policymakers and their advisors.
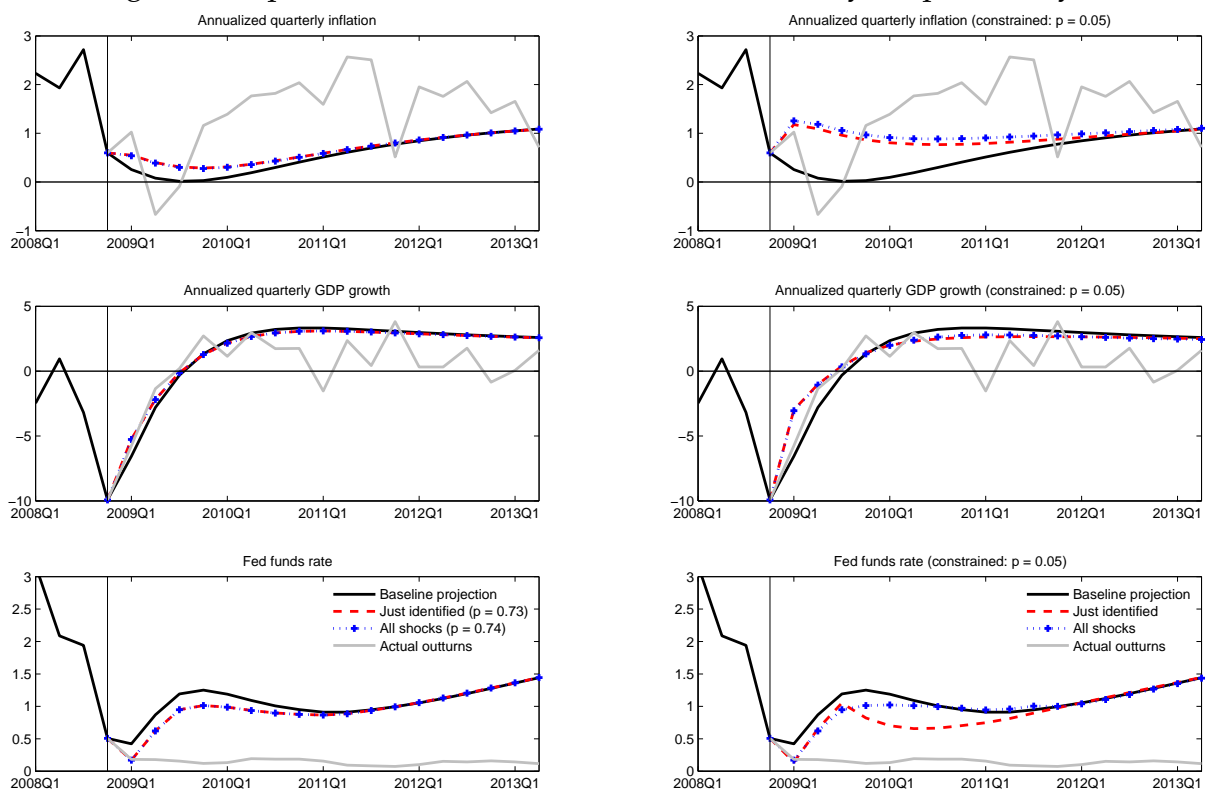
The final experiment in this section considers the case in which the Fed funds rate is reduced by a small amount for a short time. Specifically, I simulate the case in which the Fed funds rate is held 25bp lower than the baseline projection for one year. The

---

[45]The responses in Figure 4 are of a similar order of magnitude as those presented in Carlstrom et al. (2012).

objective is to demonstrate how Methods 2 and 3 can be applied to the case in which Method 1 generates a modest intervention. This allows the policy advisor to produce a simulation in which the effects of the intervention are as unlikely as possible, subject to the constraint that agents still believe that the intervention is modest at a particular significance level.

The results of this experiment are shown in Figure 5. The left column shows the results when the path for the Fed funds rate is implemented using Method 1. The two variants presented are 'just identified' ($n = K = 4$, red dashed lines) and using all policy news shocks ($n = J = 12$, blue dotted lines). Each variant generates virtually identical trajectories for the variables of interest. In both cases, the intervention would be judged modest by agents in the model, since the p-values of the modesty statistics exceed 0.7. The right-hand column shows the case in which the path for the Fed funds rate is implemented using Methods 2 and 3.[46]

Figure 5: Experiment in which Fed funds rate is cut by 25bp for one year

Notes: The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 4$ periods is generated using $n = K = 4$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 4$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

---

[46]Method 2 is used for the just-identified case ($n = K = 4$) and Method 3 when $n = J = 12$.

5 Inspecting the mechanism

In this section, I explore some of the policy experiments in more detail, to uncover the underlying mechanisms at work. I start with the results in the left-hand column of Figure 3, which shows results for the experiment in which the Fed funds rate is held low for two years. The left-hand column shows that, when the experiment is implemented using $n = J = 12$ policy news shocks, the effects on inflation and GDP are much smaller than when the experiment is implemented using $n = K = 8$ shocks. However, the paths for the Fed funds rate over the five year forecast horizon are very similar.

Figure 6: Diagnostics for experiment in which Fed funds rate is held low for two years
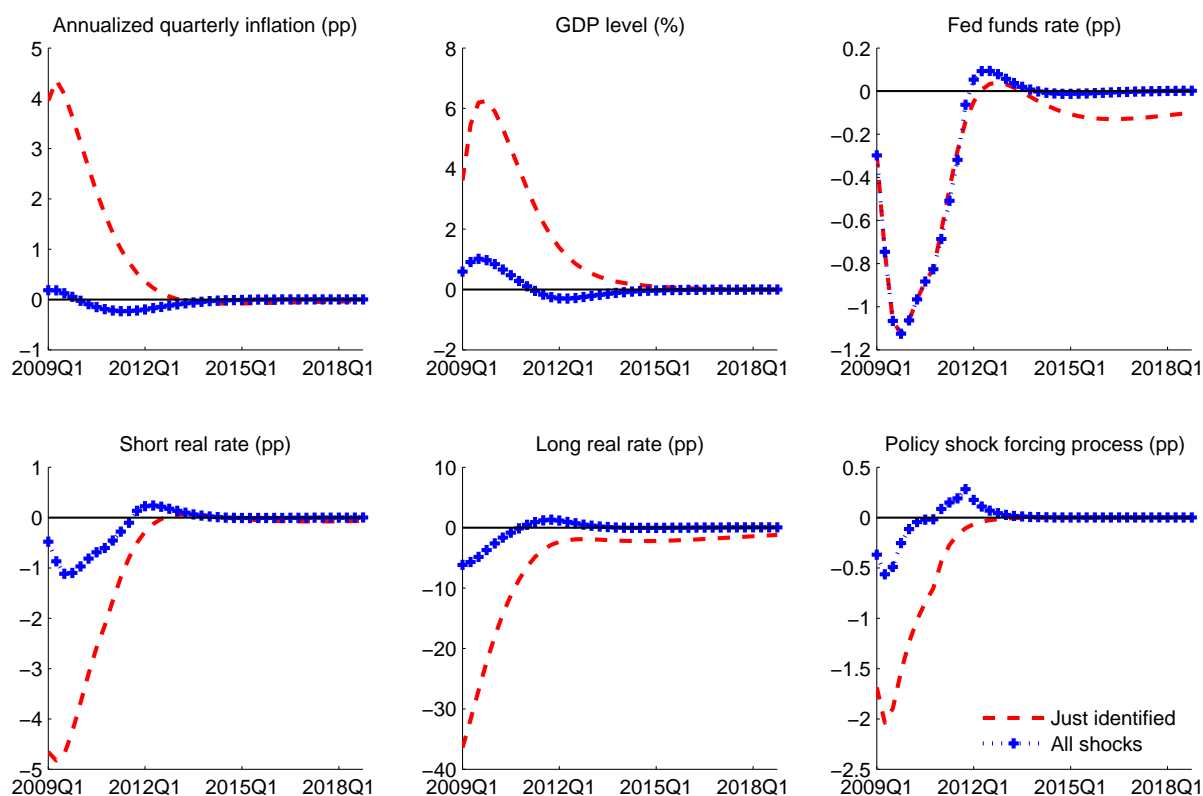
Notes: Each panel shows the responses of a particular variable to the policy experiment conducted in the left hand column of Figure 3. The responses are measured relative to the baseline forecast, so that each line represents the marginal effect of the policy experiment on that variable. The policy experiment implemented using $n = K = 8$ shocks is shown by the dashed red line. The experiment using $n = J = 12$ shocks is shown by the blue dotted line.

To explain this result, Figure 6 plots the *marginal* responses of key variables, measured as the difference between the projection for the variable when the Fed funds rate is held lower and the baseline projection. These responses are plotted over a ten-year horizon, to help illustrate the mechanisms at work.

The dashed red lines in Figure 6 show the responses when the Fed funds rate is held low for two years using $n = K = 8$ policy news shocks, so that the experiment is 'just identified'. The dotted blue lines show the effects when $n = J = 12$ shocks are used. In the 'just identified' case, lowering the Fed funds rate boosts inflation by 4pp and the peak effect on the level of output is around 6%. In contrast, when all policy news shocks are used, the peak effect on GDP is around 1% and the initial increase in inflation is small (indeed inflation is *weaker* after about a year). In the two experiments

the Fed funds rate paths are (by construction) identical for the first two years and are very similar during the third year of the forecast horizon. Later on, however, the paths diverge.

Indeed, the 'just identified' variant of the policy experiment exhibits the property that the path for the Fed funds rate lies below the baseline path for a prolonged period (it is still around 0.1 pp below the baseline path after 10 years). del Negro et al. (2012) emphasize the persistence of the effects of a similar experiment that temporarily lowers the expected path of the policy instrument. In their model, the expected path for the policy instrument also remains lower for a prolonged period. In both their experiment and mine, persistently lower nominal interest rates are determined by the monetary policy reaction function, which responds to a prolonged period in which inflation and/or the output gap are (slightly) *lower* than forecast in the baseline case. In the 'just identified' variant of my simulation (dashed red lines), inflation falls below the baseline projection from around 2013 and GDP falls below the baseline projection from the end of 2017.

A persistently lower nominal interest rate combined with a large initial rise in inflation creates a significant initial fall in the long-term real interest rate, which stimulates demand and validates the initial rise in inflation. The bottom row of Figure 6 shows that the short-term annualized real interest rate (defined as $rr_t \equiv r_t - \mathbb{E}_t \pi_{t+1}$) falls by almost 5 percentage points and a measure of the longer term rate (defined as $rr_t^L \equiv \sum_{j=0}^{\infty} rr_{t+j}$) falls by nearly 40 percentage points. Because forward-looking consumption and investment behavior are driven by these longer-term real interest rates, this represents an enormous stimulus to the economy.

The path for the policy shock forcing process (ϵ_t^r) that delivers these striking results is plotted as the dashed red line in the bottom right panel. The path of the forcing process implies a loosening of around 200bp initially, even though the Fed funds rate only falls by around 100bp. The disturbance to the policy rule must be larger than the fall in the Fed funds rate because inflation and GDP rise so much in response to the policy experiment. After its initial decline, the path for ϵ_t^r returns relatively smoothly back to zero, in line with (2).

When all of the policy news shocks are used, the path for the forcing process is rather different. As shown in the dotted blue lines, the forcing process initially falls by less and then becomes positive in the third year of the forecast horizon. The path is generated by positive policy news shocks during the third year. The net result is a path for the Fed funds rate that is temporarily higher than the baseline path in the third and fourth year of the forecast horizon. The anticipated future tightening of policy mitigates the effects of the policy experiment on longer term nominal interest rates. This variant of the experiment is therefore consistent with a smaller decline in the long-term real interest rate and a correspondingly smaller increase in GDP and inflation. The expected future tightening of policy (relative to the baseline forecast) leads to a small decline in GDP in the third year of the projection. Because inflation is very forward looking, inflation falls below the baseline forecast even earlier than this. The more muted responses of output and inflation are thus consistent with the smaller initial response of ϵ_t^r.

Figure 7 presents a similar analysis of the results in Figure 3 when $n = J = 12$ policy shocks are used. The objective in this case is to study how the constraint on the modesty of outcomes affects the results.[47] The results in Figure 7 show that, to satisfy

---

[47]The solid green lines in Figure 3 correspond to the dotted blue lines in the right hand column of

Figure 7: Diagnostics for experiment in which Fed funds rate is held low for two years
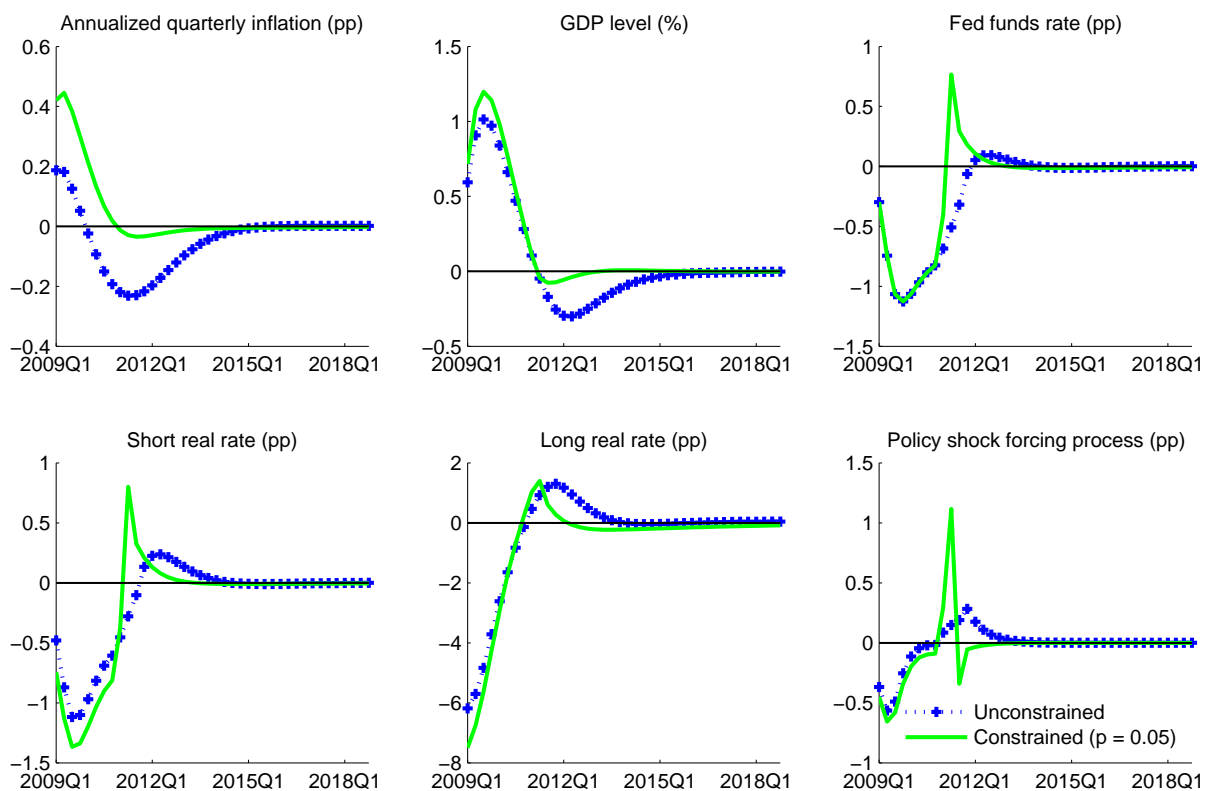
Notes: Each panel shows the responses of a particular variable to the policy experiment conducted in Figure 3. The responses are measured relative to the baseline forecast, so that each line represents the marginal effect of the policy experiment on that variable. The policy experiment implemented using $n = J = 12$ shocks is shown by the blue dotted line. The experiment using $n = J = 12$ shocks that is also constrained to be modest at the 5% significance level is shown by the solid green line.

the restriction that the intervention is modest, the policy experiment must generate an impact effect on GDP is slightly larger and an impact effect on inflation is considerably larger (around twice the size of the case in which no modesty constraint is imposed).

The joint behavior of the variables plotted in Figure 7 are, of course, the result of the general equilibrium effect of the policy shocks used to implement the experiment. The larger initial effect on inflation is generated by a larger near-term increase in GDP and a smaller subsequent undershoot of the baseline projection in 2011-12. This requires a long-term real rate that returns more quickly to the baseline forecast. To deliver this time profile for the long-term real interest rate, the Fed funds rate must rise more sharply above the baseline path after the initial $K = 8$ periods for which it is held low. That is because, in the solid green lines, the behavior of real interest rates after the first year or so of the simulation are driven by the behavior of the nominal rate path rather than by expected inflation.[48] Generating this path for the Fed funds rate requires a volatile profile for the policy shock forcing process (bottom right panel).[49]

Figure 3. The dotted blue lines in Figure 3 correspond to the dotted blue lines in the left hand column of Figure 3 (ie without a constraint on the modesty of the policy experiment).

[48]The converse is true in the blue dotted lines.

[49]These results demonstrate again that the entire path for the policy instrument is important in determining the macroeconomic effects of the policy intervention. In particular, the effects depend on how the policy instrument behaves *after* the period for which it is constrained to follow the desired path. This result echoes the findings of Blake (2013) who finds that experiments in which the policy rate

Figure 8: Joint distributions of effects of policy interventions on GDP and inflation
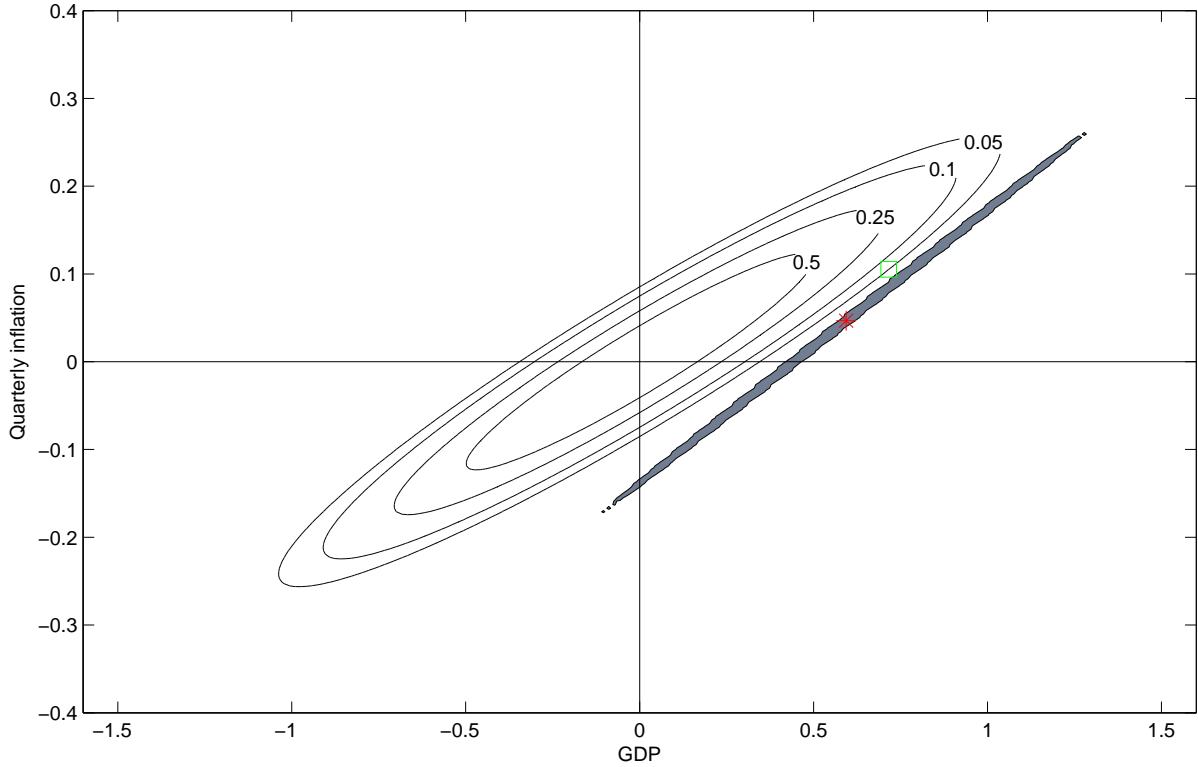
Notes: The solid gray lines are contours of the ex ante joint distribution of the effects on GDP and infla-tion at date $T + 1$ generated by policy announcements ν_{T+1}. The labels denote the mass of probability *outside* each contour line. These contours therefore encompass the set of effects that would be regarded as modest at the labeled significance level. The gray shaded area denotes the area encompassed by 99% of the joint distribution of effects on GDP and inflation of interventions that deliver a lower path for the Fed funds rate for two years ($R\nu_{T+1} = \bar{r} - c$). The red asterisk is the mean of the distribution. The green square denotes the effects on GDP and inflation when the path for the Fed funds rate is constrained to satisfy the restriction $R\nu_{T+1} = \bar{r} - c$ and the policy intervention is constrained to be modest at the 5% significance level, using Method 3. Appendix A provides details of how these densities are computed.

To see why imposing the modesty constraint generates a larger initial response of inflation, Figure 8 presents a two dimensional representation of the χ^2 distribution used to evaluate the modesty test statistic. The x-axis measures the effect of the policy intervention on the level of GDP in the first quarter of the simulation, measured as a percentage deviation from the baseline forecast. The y-axis measures the percentage point effect on quarterly inflation in the first quarter of the simulation, again mea-sured relative to the baseline forecast. Each axis therefore corresponds to one row of $[\bar{y}_{T+h} - \mathbb{E}_T y_{T+h}]$ in equation (8). The solid lines correspond to the contours of the χ^2 distribution for M_T^h. The labels refer to the probability of observations lying outside the region defined by the contour.

Recall that the distribution used to evaluate the modesty of the policy intervention is the distribution of outcomes for inflation and GDP that agents expect to be generated by policy interventions ν_{T+1}. The solid contour lines are centered on zero, indicating that the most likely (and expected) effect on GDP and inflation is zero, as would be generated by the most likely (and expected) intervention $\nu_{T+1} = 0$. The contours are also symmetric around zero because, prior to a policy intervention, agents believe that

---

tightens immediately after a prolonged (anticipated) loosening may have very small effects on growth and inflation.

it is equally likely that the intervention could tighten or loosen policy. The upward slope in the contours indicates that, in general, a policy intervention that generates an increase in GDP will also generate an increase in inflation.

The red asterisk shows the marginal effect of the policy intervention that holds the Fed funds rate at 12.5bp for $K = 8$ quarters when it is implemented using $n = J = 12$ policy news shocks.[50] The asterisk lies outside the 5% contour of the distribution of M_T^h and is therefore not regarded as modest at this significance level.

The gray disk that is centered on the red asterisk is based on the distribution of outcomes for GDP and inflation, conditional on the policy intervention delivering the desired path for the Fed funds rate (equal to 12.5bp for 8 quarters). Specifically, 99% of outcomes lie within the disk. This distribution is used for the test statistic $S\left(\bar{\nu}_{T+1}\right)$ in equation (13). Again, there is an upward slope in this distribution indicating holding the Fed funds rate at 25bp for two years in a way that generates a larger (smaller) effect on GDP would also tend to generate a larger (smaller) effect on inflation. Compared with the distribution of M_T^h, the gray disk has relatively more mass in the region in which both GDP and inflation increase, reflecting the fact that the distribution is conditional on the intervention generating a prolonged fall in the Fed funds rate.

Applying Method 3 to impose that the policy intervention is modest at the 5% significance level involves finding a point on the 5% contour of the distribution for M_T^h that is most likely from the perspective of the distribution represented by the gray disk. That point is the green square in Figure 8. The green square lies to the North East of the red asterisk, so the initial effects on GDP and inflation are larger when the policy experiment is constrained to be a modest policy intervention.

The fact that the green square lies outside the gray disk – which encompasses 99% of the outcomes conditional on the policy intervention delivering the desired path for the Fed funds rate – indicates that it is not a very likely outcome from the perspective of this distribution.

The test statistic $S\left(\bar{\nu}_{T+1}\right)$ can therefore be used to assess the plausibility of policy experiments from the perspective of the distribution of outcomes *conditional* on the policy rate following the desired path. Figure 9 shows three alternative policy experiments. In each case, the Fed funds rate is held low for two years, but the experiments differ in terms of the extent to which they are constrained to be modest. Forcing the experiment to be more modest (increasing the p-value of constraint on M_T^h from 0.05 to 0.25 to 0.5) reduces the initial effect on growth and inflation. In terms of Figure 8 this represents a movement towards the origin, so that the policy intervention is more likely from the perspective of the distribution of M_T^h. However, from the perspective of test statistic S in equation (13), the outcomes are increasingly unlikely. The values of S are reported in Figure 9 and represent zero probability events for the relevant χ^2 distributions.

Figure 9 shows that, to deliver interventions that are increasingly modest requires increasingly large overshoots of the path for the Fed funds rate for the period after the policy rate is constrained to follow the desired path. Larger overshoots mitigate the stimulus provided by the reduction in the Fed funds rate over the first two years of the projection (so that long-term real interest rates fall by less). But the subsequent movements in long term real interest rates generate arguably implausible secondary cycles in growth and inflation.

These observations suggest that a policy advisor may find it useful to inspect the

---

[50]The experiment is shown in the left column of Figure 3 in the blue dotted lines.

## Figure 9: Policy experiments imposing alternative modesty constraints
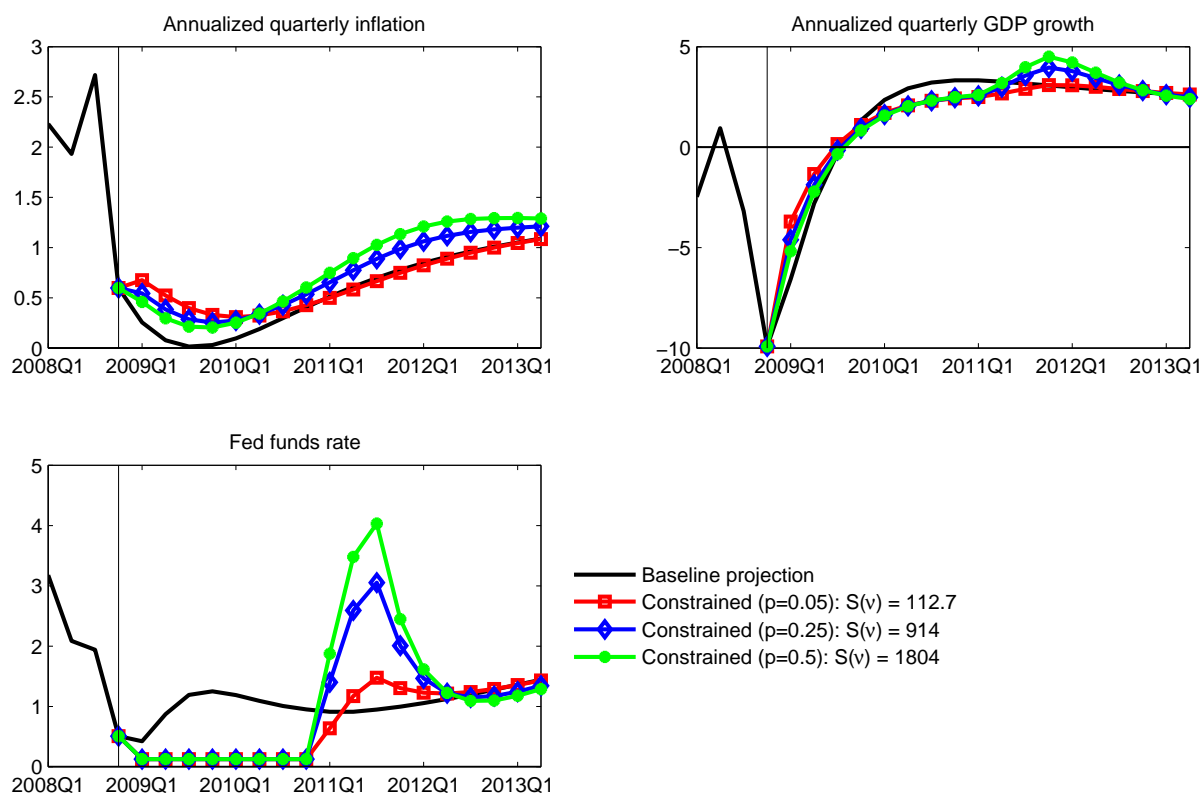
Notes: The figures show baseline forecasts (solid black lines) together with three policy simulations. In each case, the Fed funds rate is held low (at 12.5bp) for $K = 8$ quarters and is implemented using using $n = J = 12$ policy news shocks using Method 3. Each experiment is constrained to be a modest policy intervention at a different significance level (indicated by the legend specifying the value of p). Also reported is the value of $S(\nu)$, the minimand of the problem solved by Method 3. In the absence of the modesty constraint, $S(\nu) \sim \chi^2(2)$.

distribution of outcomes that agents would expect to be generated by policy interventions (rather than the arrival of other shocks). The analysis of Section 3.3 shows that the distribution of the variables of interest in period $T + h$ generated by policy interventions in period $T + 1$ is a multivariate normal densities with mean $QB^h x_{T|T}$ and covariance Ω_{T+h}.[51] Once again, I abstract from the zero bound on the Fed funds rate.

Figure 10 presents these distributions for the variables of interest and the Fed funds rate. The dashed line in the bottom left panel indicates the 12.5bp level of the Fed funds rate that has been used in the policy simulations. Beyond the first quarter of the forecast horizon, the dashed line lies outside the 5%–95% confidence set for the path of the Fed funds rate. So the probability of observing a policy intervention that holds the Fed funds rate at this level for a prolonged period is very low and this may prompt the policy advisor to suspect that such an intervention is unlikely to be regarded as modest by agents in the model.

---

[51]The solid contour lines in Figure 8 uses this result to plot the joint distribution of the effects on GDP and inflation in the period that the policy intervention is announced.

Figure 10: Distribution of outcomes generated by policy interventions
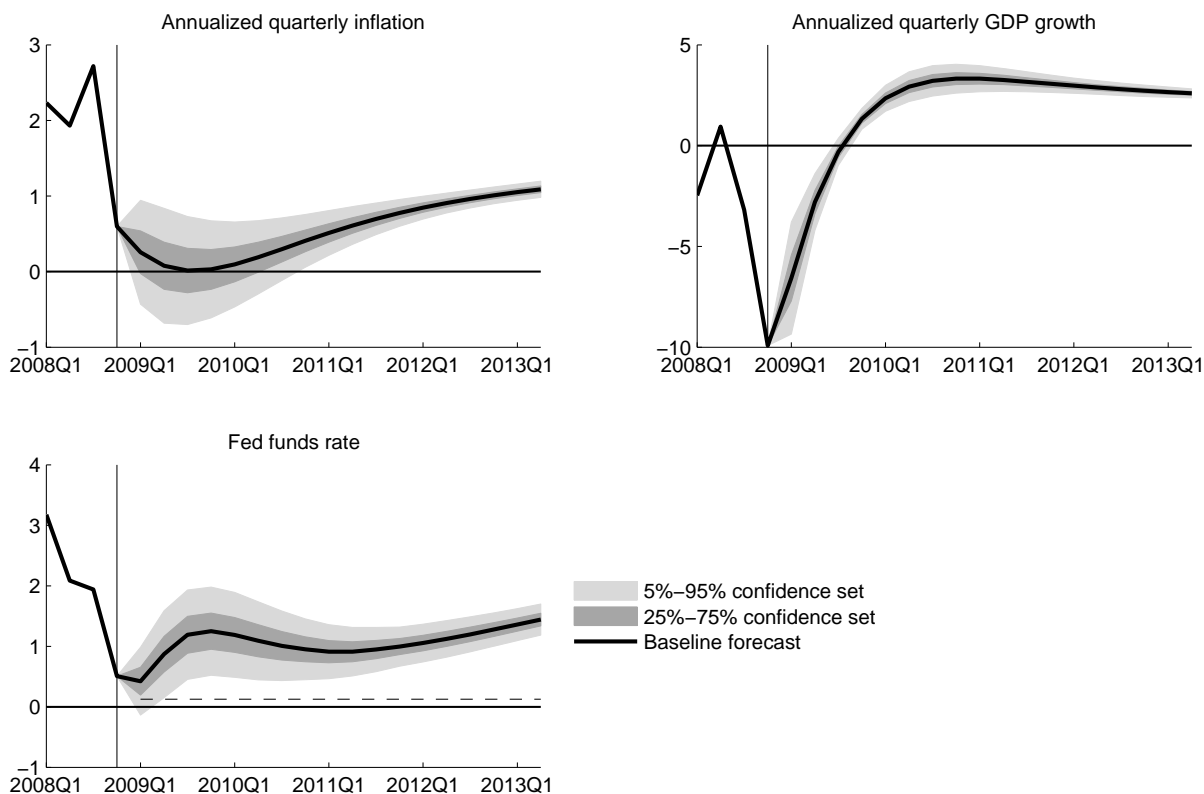
Notes: Each panel shows the baseline forecast (solid black line) together with confidence sets for the outcomes generated by policy interventions ν_{T+1}. Appendix A provides details of how these densities are computed.

6   Robustness analysis

In this section, I explore the sensitivity of the results in two important dimensions. First, I examine whether the results are robust to the definition of the modesty statistic. The second question is whether the estimated model used in Section 4.4 contains a reliable characterization of agent's beliefs about monetary policy behavior. I examine these questions by estimating the model over different samples, according to alternative (narrative) definitions of what might constitute a stable monetary regime and by estimating the model using different specifications for the monetary policy reaction function.

6.1   Specification of the modesty statistic

As noted in Section 4.4, the extent to which a policy experiment is viewed as modest by agents in the model is determined by the specification of the modesty statistic that will be tested. Specifically, the researcher must make two choices: which variables are to be included in the assessment of whether the change in the forecast is modest; and the horizon, h, at which that assessment is made. In Section 4.4, I focused on the effect in the period of announcement, $T + 1$, (i.e., I set $h = 1$) on inflation and GDP growth, given the key importance of these variables for policy analysis and the observation that the responses to forward guidance simulations are likely to be regarded by many economists as excessively front-loaded.

34

Figure 11: Modesty test statistics for alternative combinations of variables
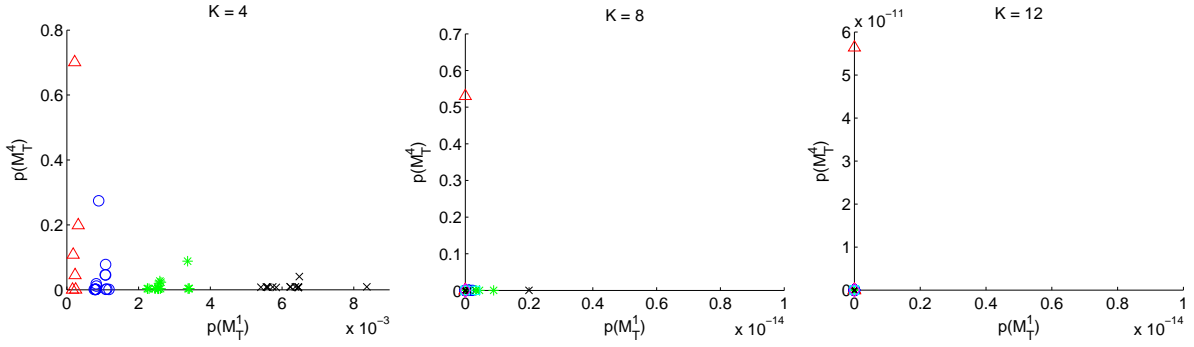
Notes: Each panel shows a scatter plot of the p-values associated with modesty tests at different horizons. The x-axis measures the p-value of the test statistic at horizon $h = 1$ and the y-axis measures the p-value of the test statistic for horizon $h = 4$. Each marker type shows the results when a particular subset of observable variables are chosen for the test: red triangles = 1 variable; blue circles = 2 variables; green asterixes = 3 variables; black crosses = 4 variables; cyan squares = 5 variables; magenta diamonds = 6 variables. The panels contain results for the implementation of experiments in Figures 2, 3 and 4 using Method 1 with $n = K$ (i.e., just-identified).

In this section, I re-examine the extent to which the policy experiments conducted in Section 4.4 would be viewed as modest under alternative specifications of the test statistic. I consider the use of different combinations of variables at different horizons. I consider all possible combinations of the macroeconomic variables in the set of observable variables used to estimate the model. The set of macroeconomic variables is: GDP growth; consumption growth; investment growth; quarterly inflation; total hours worked; and quarterly nominal wage growth.[52] This set of six variables generates 63 distinct combinations of variables that can be used to specify the modesty statistic.[53] To examine the sensitivity of the results to the forecast horizon at which the modesty statistic is computed, I compare the results under my baseline assumption of $h = 1$ with the results for $h = 4$, which correspond to the four-quarter-ahead projections.

The results are presented in the form of scatter plots in Figure 11. Each panel shows the results based on the cases in which the Fed funds rate is held fixed for $K = 4, 8, 12$ quarters respectively. The calculations are based on the cases in which the policy experiment is just identified (that is, using Method 1, with $n = K$).[54] So the results are based on the forecasts that generate the red dashed lines in the left-hand columns of Figures 2, 3 and 4.

Figure 11 shows that, for $h = 1$, none of the three policy experiments (corresponding to $K = 4, 8, 12$) would be considered modest for the commonly used critical significance level of 5% (or even the 1% significance level), regardless of the specification of the modesty statistic.

However, the distribution of the points shows that a there are a number of defini-

---

[52]This means that I exclude the Fed funds rate and the long-term interest rate measures, following Adolfson et al. (2005) who exclude the policy instrument from the set of variables used to define the modesty statistics that they test.

[53]The number of combinations of j variables from a set of N variables is given by $C_j \equiv \frac{N!}{N!(N-j)!}$. The total number of combinations is therefore given by $\sum_{j=1}^{N} C_j$, which gives 63 combinations for $N = 6$.

[54]A technicality that arises here is that for the case of $K = 4$ quarters, I cannot compute modesty statistics for more than 4 observable variables because the rank of the covariance matrix used to compute M^h is equal to K. In this case, I simply plot the results from modesty statistics based on all combinations of 4, 3, 2 and 1 variables.

tions of the modesty statistic with $h = 4$ that would imply that the policy experiment is modest using conventional significance levels. This is particularly true for the left-most panel (in which $K = 4$), but there is also one instance in which the results shown in Figure 3 would also be classified as modest, despite the fact that the macroeconomic responses in this case would likely be considered incredible by most economists and policymakers.

There is a straightforward explanation for these findings. The dynamics of the model are very front loaded, so that inflation and output growth (and the other variables in the model) move a lot on the announcement of the policy experiment ($h = 1$) before converging back towards their baseline paths. Inspection of Figures 2, 3 and 4 reveals that GDP growth and inflation are relatively close to their baseline paths after 4 quarters, so the change in the projection at this horizon is more likely to be interpreted as modest. This logic is also consistent with the observation that modesty statistics based on a single variable (red triangles) are more likely to indicate that the experiment represents a modest policy intervention when the horizon is $h = 4$. That is because these specifications of the test statistic are defined using only the (squared) size of the change in a single variable and the co-movement of that variable with other variables is ignored.

As a final illustration of the importance of the specification of the modesty statistic, I consider variants that focus attention on the responses of expected future interest rates in the spirit of del Negro et al. (2012). Figure 12 plots the p-values associated with the modesty test statistics defined using forward interest rates at different horizons. Each p-value is computed by setting the Q matrix in equation (6) to select the expectation of the policy rate in a particular future period. Each panel is based on the case in which the shocks used to implement the path for the policy rate are just identified (so $n = K$ and Method 1 is used) for the one-year, two-year and three-year experiments depicted in Figures 2 , 3 and 4.[55]

Figure 12 shows that neither the one-year ($K = 4$) nor two-year ($K = 8$) experiments would cause concern for a policy advisor using conventional significance levels. Even the three-year experiment ($K = 12$) generates effects on forward interest rates between four and five years ahead (16–20 quarters) that would be judged statistically likely by the modesty statistic.

These results are easily rationalized using the discussion above: measuring interventions using the response of a single variable is more likely to deliver a 'modest' result. Given the 'front loaded' nature of the responses to anticipated reductions in the Fed funds rate, this is particularly true for variables at longer forecast horizons. Because the forward rates are computed using the expectations theory of the term structure, the n quarter forward rate corresponds to the forecast of the policy instrument n quarters ahead. Inspection of the results in Figures 2 , 3 and 4 reveals that the policy rate returns close to its baseline path relatively quickly in most cases.

I can also assess whether the experiments are likely to be regarded as modest policy interventions in terms of their effects on a long-term *spot* rate by setting the Q matrix in equation (6) to select a measure of the ten-year rate defined using equation (15), with $N = 40$. I construct modesty statistics using this measure of the ten-year rate for the 'just identified' simulations in the left hand panels of Figures 2 , 3 and 4. The p-values for these statistics are 0.17, 0.04 and 3.5×10^{-6} for the one-year, two-year and

---

[55]That is, the results are computed from the simulations shown by the dashed red lines in the left hand panels of those figures.

Figure 12: P-values for modesty statistics using forward interest rate expectations
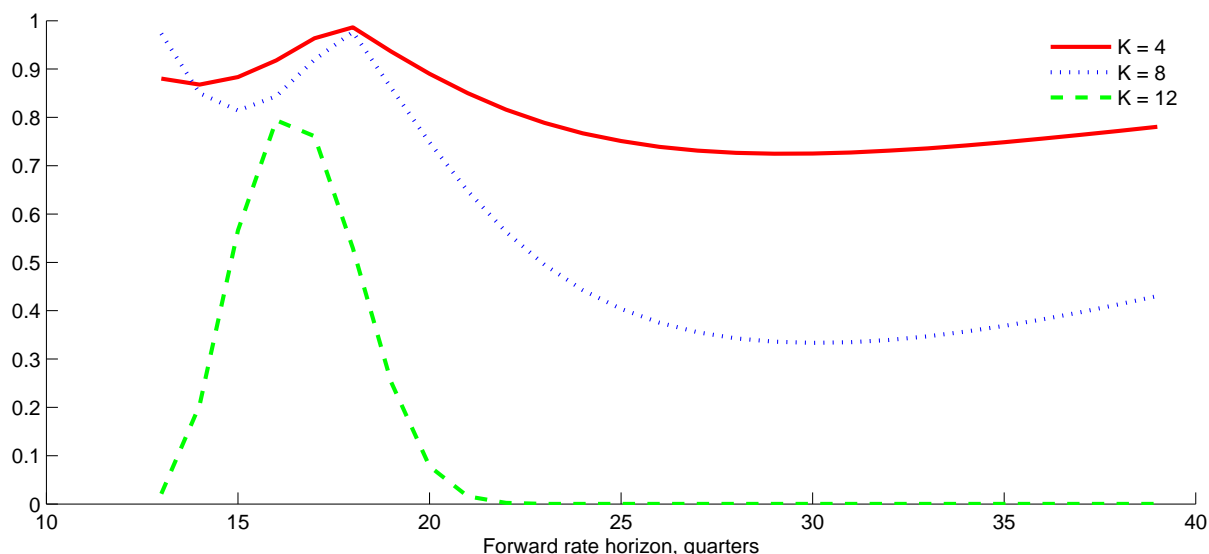
Notes: The chart plots the p-values associated with modesty statistics defined for forward interest rates at different maturities. The x-axis measures the forward rate maturity in quarters and the y-axis measures the p-value of the modesty statistic. The results are based on the case in which the desired path for the policy rate is implemented using shocks that are just-identified (ie Method 1 with $n = K$).

three-year experiments respectively.

6.2 Specification of agents' beliefs about monetary policy

My method uses the policy advisor's model of agent's beliefs about monetary policy to assess the plausibility of a policy experiment. For the results to be reliable, the policy advisor must be satisfied that the specification of agents' beliefs about monetary policy is adequate and that the sample period used to estimate the parameters governing agents' beliefs represents a stable monetary policy regime. In this section, I assess the extent to which my results may be sensitive to the baseline specification of agents' beliefs about the behavior of monetary policy. I explore two possible miss-specifications.

First, my choice of sample period may not represent a stable policy regime: for example, it is possible that policy behavior as changed systematically as the Chair of the FOMC has changed over time. Indeed, this may be a particularly relevant concern in the context of forward guidance. As noted in Section 2, the use of forward guidance and the methods by which it is implemented have changed over time, which suggests that it may be inappropriate to treat the post-1984 period as a single regime. To investigate this issue, I re-estimate the model over a shorter sample period and present the results in Section 6.2.1.

Second, I examine the sensitivity of the results to the specification of agents' beliefs about monetary policy behavior: the reaction function in equation (14) may not be an adequate description of agents' beliefs. Indeed, Cúrdia et al. (2011) show that the specification of the monetary policy reaction function can have a dramatic effect on the ability of New Keynesian models to fit the US data. To address this issue requires re-estimating the model with an alternative specification of the monetary policy reaction function. Results from doing so are presented in Section 6.2.2.

6.2.1 Sensitivity to estimation sample

To assess the sensitivity of the results to the estimation sample period, I estimate the model using a data sample starting in 1996, using the same priors for the parameters as in the baseline estimation.[56] Appendix B presents results from this exercise.

Tables B.1 and B.2 in Appendix B document summary statistics of the posterior distribution of the parameters. The posterior densities for most of the parameters describing private sector decision making are broadly in line with the baseline estimation over the full sample.[57] For the purposes of the exercise at hand it is useful to focus on the parameters describing agents' beliefs about monetary policy. Compared with the baseline results, estimation using the shorter sample suggests that the Fed funds rate responds more to the output gap, *less* to inflation and with (slightly) less inertia. Moreover, the estimated variances of the policy news shocks (at all horizons) and the measurement errors for one-year and five-year bond rates are larger when the shorter sample is used. In addition, the estimated 'term premia' on one-year and five-year bond rates (τ_4 and τ_{20}) are noticeably smaller when the shorter sample is used.

Figure 13: Forecasts from model estimated on shorter sample and actual outturns
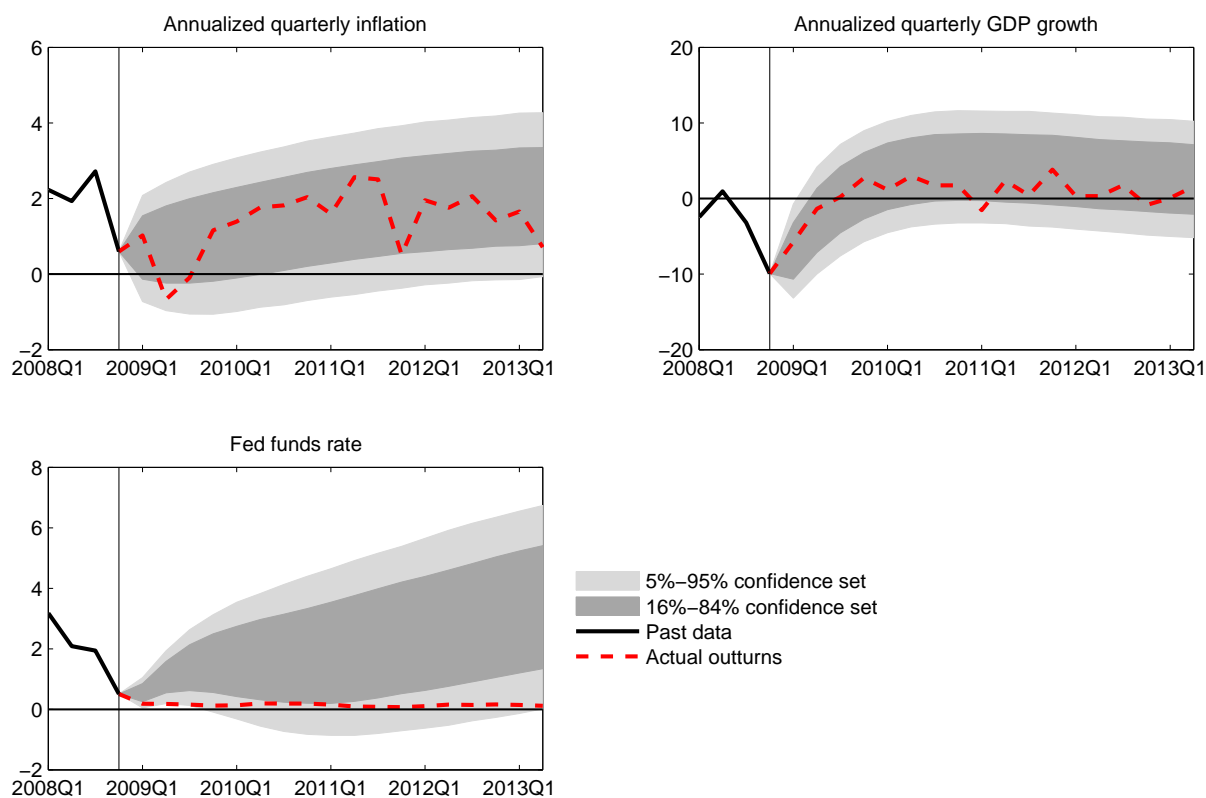
Notes: The panels show confidence sets of forecasts generated by the DSGE model estimated using the shorter data sample (1996–2008), obtained by drawing from the posterior distribution of the estimated parameters. The past data are shown by the solid black lines and the actual realizations of the data over the forecast horizon are shown by the dashed red lines.

Figure 13 presents the model-based forecasts using the estimation results based on the shorter sample. The projections for inflation and GDP growth are somewhat

---

[56]Data from 1992 are used as a training sample for the Kalman filter. The sample ends in 2008Q4 as in the baseline specification.

[57]There are, of course, some differences. For example, the shorter sample is characterized by higher wage stickiness and lower persistence of wage markup shocks.

stronger than the forecasts from the model estimated on the full sample (see Figure 1). In particular, most of the mass of the forecast density for GDP growth lies above the actual realizations for growth over the forecast horizon. The strong outlook for growth (and indeed inflation) means that the forecast density for the Fed funds rate is significantly higher (and more steeply sloped) than the forecasts based on the full sample estimation. The path for the Fed funds rate rises more steeply despite the fact that the information set on which the projection is based includes data for longer-term yields that are very low. This is consistent with the smaller estimated term premia and larger estimated variances of the measurement errors in the measurement equations for long-term yields.

Because the projection for the Fed funds rate is rather different than the projection using the model estimated on the longer sample (Figure 1), I adjust my experiments to make them more comparable with those in Section 4.4. To do so, I apply the marginal change in the Fed funds rate generated by the experiments in Section 4.4 to the baseline projection for the Fed funds rate from the model estimated on the shorter sample. I focus here on the results for the experiment in which the Fed funds rate is held low for three years, depicted in Figure 14.[58]

The left-hand column of Figure 14 shows that holding the Fed funds rate low for three years generates less extreme effects on growth and inflation than the experiment presented in Section 4.4 (see Figure 4). Nevertheless, the experiment does not constitute a modest policy intervention at the 5% significance level. The right-hand column shows that when the experiment is constrained to be modest at the 5% significance level, the macroeconomic effects are considerably smaller (and in fact resemble the results based on the model estimated on the full sample presented in Figure 4).

The results in this section illustrate that the extent to which a policy intervention is modest naturally depends on the policy advisor's model of private agents' beliefs about monetary policy. When the model is estimated over a shorter, more recent, sample, the parameters of describing private agents' beliefs about monetary policy do change somewhat. These changes influence both the baseline projection and the extent to which a given policy intervention is modest according to the policy advisor's model of private agents' beliefs. Based on my experiments, it seems that a given policy intervention is more likely to be modest using a variant of the model estimated using the more recent data. This seems consistent with the finding that, over this sample, the variances of policy news shocks are estimated to be larger. This is likely to increase the size of the elements in the Φ_ν matrix that determines the denominator of the modesty statistic (see equations (8) and (9)).

Nevertheless, many plausible policy interventions of the magnitude studied in recent papers investigating the effects of pegging the interest rate would not be deemed modest when the model is estimated on the shorter sample. And the results generated by constraining the policy experiment to be modest are very similar to those obtained in Section 4.4.

---

[58]Results for the one-year and two-year experiments are presented in Appendix B and are consistent with the discussion that follows. Those results show that the one-year experiment is modest at the 5% significance level, though the two-year experiment is only modest when all policy news shocks are used.

Figure 14: Experiment in which Fed funds rate is held low for three years using model estimated over shorter sample
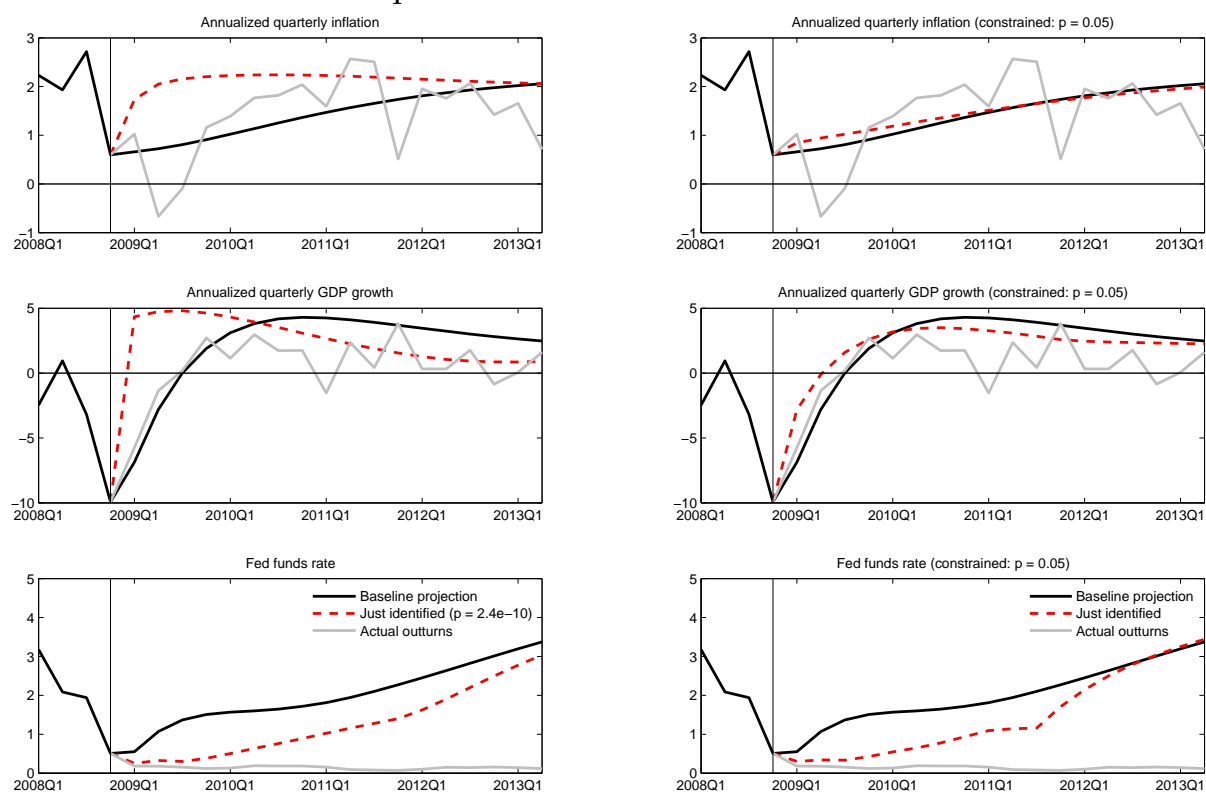
Notes: The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 12$ periods is generated using $n = K = 12$ policy news shocks ('just identified', red dashed lines), implemented using Method 1. The p-value of the modesty statistic associated with this simulation is reported in the legend. The right column repeats the experiment, but imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Since $n = K$, Method 3 is used.

6.2.2 Sensitivity to specification of monetary policy behavior

To assess the importance of the policy advisor's model of agents' beliefs about monetary policy, I estimate a variant of the model in which the interest rate smoothing parameter in the monetary policy reaction function (14) is calibrated to zero. Cúrdia et al. (2011) find that specifications of New Keynesian models in which the monetary policy rules features interest rate smoothing fit the US data better than rules without smoothing. Indeed, the marginal data density of the variant of the model with no interest rate smoothing is -364, compared with -230 for the baseline specification of the model.

Tables B.1 and B.2 in Appendix B document summary statistics of the posterior distribution of the parameters. The posterior densities for some of the key parameters describing private sector decision making differ somewhat those from the baseline version of the model with interest rate smoothing. In particular, the mean estimates of the intertemporal consumption elasticity and intratemporal labour supply elasticity are both much larger for the no-smoothing rule variant. Prices and wages are also much stickier in this variant. Focusing again on the parameters describing agents' beliefs about monetary policy, this variant suggests that the monetary policy response to

inflation is extremely weak (barely satisfying the Taylor principle). Moreover, the estimated variances of the policy news shocks are substantially larger at several horizons (contemporaneous, eight quarter ahead and twelve quarter ahead) when the shorter sample is used. The estimated 'term premia' on one-year and five-year bond rates (τ_4 and τ_{20}) are noticeably smaller for this variant of the model.

Figure 15 shows an experiment in which the Fed funds rate is held lower for three years. Again, to make the policy experiments more comparable with those for the baseline specification of the model, I apply the marginal change in the Fed funds rate generated by the experiments in Section 4.4 to the baseline projection of the Fed funds rate from the variant of the model without interest rate smoothing. The results in the left panel are quite striking. A persistent reduction in the Fed funds rate generates a sizable and protracted *reduction* in both growth and inflation. The p-value for the modesty statistic for these responses is 0.31.

These results demonstrate that the modesty test statistic is just one way to judge the plausibility of responses to forward guidance simulations. Even though the simulation is very likely to be a modest policy intervention (at a 30% significance level), many economists, policy advisors and policymakers would find the results in Figure 15 implausible because the sign of the responses are counterintuitive.[59] Mechanically, it is straightforward to explain these results. Table B.1 shows that the estimated persistence of the policy shock process ρ_r is extremely high (the mean is almost 0.98). The most likely explanation for this parameter estimate is that it is helps the model to match the persistence of the nominal interest rate in the data, given the absence of an interest rate smoothing coefficient in the monetary policy rule. The presence of a highly persistent shock to the monetary policy rule can imply that a *tightening* in policy (positive policy shock) leads to a *fall* in the nominal interest rate as an endogenous reaction to the reduction in activity and inflation induced by the policy tightening.[60]

In addition, the large variances of the estimated policy shocks (see Table B.2) are likely to imply that the elements in the Φ_ν matrix are also large. As discussed in Section 6.2.1, Φ_ν determines the denominator of the modesty statistic (see equations (8) and (9)), so that a "larger" Φ_ν will tend to generate smaller values of the test statistic, for a given $\bar{\nu}_T$ vector. Finally, the fact that the estimated variances of policy news shocks are particularly large at specific horizons implies that the 'size' of the Φ_ν matrix changes with the number of policy shocks used to implement the experiment. An implication is that, for this variant of the model, experiments in which the Fed funds rate is fixed for a shorter period are *less* likely to be deemed modest: see Figures B.5 and B.6 in in Appendix B. As Figure B.3 demonstrates, this implies that the results for this variant of the model are particularly sensitive to the specification of the modesty statistic.

Once again, the results of this section highlight the importance of using an adequate specification for private agents' beliefs about monetary policy. The variant of the model used in this section fits the data poorly in comparison to the baseline specification. The deterioration of fit is one indication that the specification of private agents' beliefs about policy behavior may be inadequate. So in this case, modesty test statistics may not provide a reliable assessment of the plausibility of policy simulations.

---

[59]Quantitatively, the effects of the policy experiment are very persistent, though they are not unambiguously "oversized".

[60]See Galí (2009, page 51) for an analytical investigation of this result in a prototypical New Keynesian model.

Figure 15: Experiment in which Fed funds rate is held low for three years using model with no interest rate smoothing
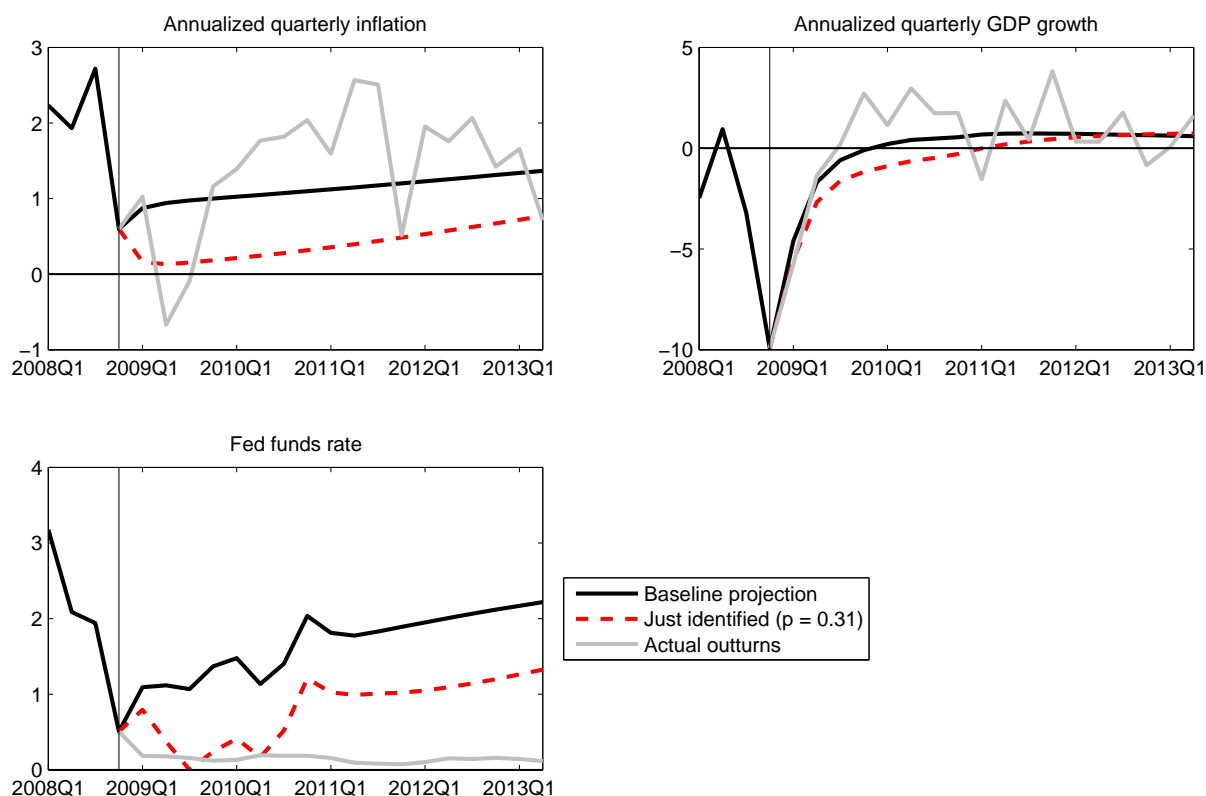
Notes: The baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. A policy experiments in which the path for the Fed funds rate for $K = 12$ periods is implemented using $n = J = 12$ policy news shocks is shown in the dashed red line. The p-value of the modesty statistic associated with this simulations is reported in the legend.

7 Conclusion

In this paper, I argue that simulations of forward guidance should be assessed using the "modest interventions" approach of Leeper and Zha (2003). I extend work by Adolfson et al. (2005) that shows how to test whether policy experiments are modest policy interventions in DSGE models. In particular, I show how to constrain policy experiment so that it is likely to be regarded as a modest policy intervention. In some cases it is possible to deliver the desired path for the policy instrument exactly. In other cases, this may not be feasible (for example when the number of shocks used to implement the experiment equals the number of periods for which the desired interest rate path is defined). In those cases, I show how to construct a modest policy intervention in which the path for the policy instrument is as close as possible to the desired path.

When the implementation of a desired future path for the short-term policy rate is constrained in this way, the macroeconomic effects are arguably much more plausible. One interpretation of this implementation is that it constrains the effects of the policy experiment to be consistent with agents' beliefs about past monetary policy behavior. This chimes with statements by some central bankers that their forward guidance policies are not intended to signal a deviation from their past behavior.

My analysis applies to single regime DSGE models in which monetary policy is assumed to be conducted according to a particular monetary policy rule. Models that incorporate multiple policy regimes (and an articulation of agents' probability

model over the regimes) could provide reliable estimates to policy experiments without recourse to my approach. Until recently, estimation of large scale regime-switching DSGE models was infeasible. But recent advances by Bianchi and Melosi (2013) among others suggests that building and estimating such models may be feasible in the near future.

Bibliography

ADOLFSON, M., S. LASÉEN, J. LINDÉ, AND M. VILLANI (2005): "Are Constant Interest Rate Forecasts Modest Policy Interventions? Evidence from a Dynamic Open-Economy Model," *International Finance*, 8, 509–544.

——— (2007): "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through," *Journal of International Economics*, 72, 481–511.

AN, S. AND F. SCHORFHEIDE (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26, 113–172.

BEAN, C. (2013): "Global aspects of unconventional monetary policies," Remarks at the Federal Reserve Bank of Kansas City Economic Policy Symposium, Jackson Hole, Wyoming, Bank of England.

BIANCHI, F. AND L. MELOSI (2013): "Modeling the Evolution of Expectations and Uncertainty in General Equilibrium," Working Papers 13-14, Duke University, Department of Economics.

BLAKE, A. (2012): "Fixed interest rates over finite horizons," Bank of England Working Paper 454, Bank of England.

——— (2013): "A cumulated expected target and an interest rate peg," Mimeo, Bank of England.

BLINDER, A. S., M. EHRMANN, M. FRATZSCHER, J. DE HAAN, AND D.-J. JANSEN (2008): "Central Bank Communication and Monetary Policy: A Survey of Theory and Evidence," *Journal of Economic Literature*, 46, 910–45.

BORIO, C. AND P. DISYATAT (2010): "Unconventional monetary policies: an appraisal," *The Manchester School*, 78, 53–89.

BURGESS, S., E. FERNANDEZ-CORUGEDO, C. GROTH, R. HARRISON, F. MONTI, K. THEODORIDIS, AND M. WALDRON (2013): "The Bank of England's forecasting platform: COMPASS, MAPS, EASE and the suite of models," Bank of England Working Paper 471, Bank of England.

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*, 44, 1–80.

CARLSTROM, C., T. FUERST, AND M. PAUSTIAN (2012): "Inflation and output in New Keynesian models with a transient interest rate peg," Bank of England Working Paper 459, Bank of England.

CHRISTIANO, L., G. EVANS, AND M. EICHENBAUM (2005): "Nominal rigidities and the dynamic effect of a shock to monetary policy," *Journal of Political Economy*, 113, 1–45.

CHRISTOFFEL, K., G. COENEN, AND A. WARNE (2008): "The new area-wide model of the euro area - a micro-founded open-economy model for forecasting and policy analysis," Working Paper 944, European Central Bank.

CHUNG, H. T., M. T. KILEY, AND J.-P. LAFORTE (2010): "Documentation of the Estimated, Dynamic, Optimization-based (EDO) model of the U.S. economy: 2010 version," Finance and Economics Discussion Series Working Paper 2010-29, Board of Governors of the Federal Reserve System.

COENEN, G. AND A. WARNE (2013): "Risks to price stability, the zero lower bound and forward guidance: a real-time assessment," Working Paper Series 1582, European Central Bank.

COOLEY, T. F., S. F. LEROY, AND N. RAYMON (1984): "Econometric Policy Evaluation: Note," *American Economic Review*, 74, 467–70.

CÚRDIA, V., A. FERRERO, G. CEE NG, AND A. TAMBALOTTI (2011): "Evaluating interest rate rules in an estimated DSGE model," Tech. rep., Staff Report, Federal Reserve Bank of New York.

DE GRAEVE, F., M. EMIRIS, AND R. WOUTERS (2009): "A structural decomposition of the US yield curve," *Journal of Monetary Economics*, 56, 545–559.

DEL NEGRO, M., S. EUSEPI, M. GIANNONI, A. SBORDONE, A. TAMBALOTTI, M. COCCI, R. HASEGAWA, AND M. H. LINDER (2013): "The FRBNY DSGE model," Staff Reports 647, Federal Reserve Bank of New York.

DEL NEGRO, M., M. GIANNONI, AND C. PATTERSON (2012): "The forward guidance puzzle," Staff Reports 574, Federal Reserve Bank of New York.

DEN HAAN, W., ed. (2013): *Forward Guidance: Perspectives from Central Bankers, Scholars and Market Participants*, VoxEU.

EDGE, R. M., M. T. KILEY, AND J.-P. LAFORTE (2007): "Documentation of the Research and Statistics Division's estimated DSGE model of the U.S. economy: 2006 version," *Finance*.

EGGERTSSON, G. B. AND M. WOODFORD (2003): "Zero bound on interest rates and optimal monetary policy," *Brookings Papers on Economic Activity*, 2003, 139–233.

FEDERAL OPEN MARKET COMMITTEE (2008): "Minutes of the Federal Open Market Committee, December 15–16," Federal Reserve Board of Governors.

FERNÁNDEZ-VILLAVERDE, J., J. F. RUBIO-RAMÍREZ, T. J. SARGENT, AND M. W. WATSON (2007): "ABCs (and Ds) of Understanding VARs," *The American Economic Review*, 97, 1021–1026.

GALÍ, J. (2009): *Monetary policy, inflation, and the business cycle: an introduction to the New Keynesian framework*, Princeton University Press.

GURKAYNAK, R. S., B. SACK, AND J. H. WRIGHT (2007): "The U.S. Treasury yield curve: 1961 to the present," *Journal of Monetary Economics*, 54, 2291–2304.

HABERIS, A., R. HARRISON, AND M. WALDRON (2014): "Transitory interest rate pegs under imperfect credibility," Tech. rep., Bank of England.

HIROSE, Y. AND T. KUROZUMI (2011): "Changes in the Federal Reserve communication strategy: a structural investigation," Bank of Japan Working Paper 11-E-2, Bank of Japan.

KILEY, M. T. (2014): "Policy Paradoxes in the New Keynesian Model," Finance and Economics Discussion Series 2014-29, Board of Governors of the Federal Reserve System (U.S.).

LASÉEN, S. AND L. E. SVENSSON (2011): "Anticipated Alternative policy Rate Paths in Policy Simulations," *International Journal of Central Banking*, 7, 1–35.

LEEPER, E. M. AND T. ZHA (2003): "Modest policy interventions," *Journal of Monetary Economics*, 50, 1673–1700.

LEVIN, A., D. LÓPEZ-SALIDO, E. NELSON, AND T. YUN (2010): "Limitations on the Effectiveness of Forward Guidance at the Zero Lower Bound," *International Journal of Central Banking*, 6, 143–189.

MILANI, F. AND J. TREADWELL (2012): "The Effects of Monetary Policy News? and Surprises?" *Journal of Money, Credit and Banking*, 44, 1667–1692.

MONETARY POLICY COMMITTEE (2013): "Monetary policy trade-offs and forward guidance," Bank of England.

SIMS, C. A. (2011): "Statistical modeling of monetary policy and its effects," *Nobel Prize Lecture*.

SMETS, F. AND R. WOUTERS (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1, 1123–1175.

——— (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, 586–606.

TOVAR, C. E. (2008): "DSGE Models and Central Banks," Economics Discussion Papers 2008-30, Kiel Institute for the World Economy.

WEALE, M. (2013): "Forward guidance and its effects," Speech given at the National Institute of Economic and Social Research.

WOODFORD, M. (2003): *Interest and prices*, Princeton Univ. Press.

——— (2012): "Methods of policy accommodation at the interest-rate lower bound," in *Jackson Hole symposium, August, Federal Reserve Bank of Kansas City*.

YELLEN, J. (2012): "Perspectives on monetary policy," remarks given at the Boston Economic Club Dinner, Federal Reserve Bank of Boston, Board of Governors of the Federal Reserve System.

A   Distributions conditional on an interest rate path

Ex ante, agents expect $\nu_{T+1} \sim N(0, \mathbb{I})$. From Section 3.3, the effect of ν_{T+1} on the variables of interest is given by:

$$\tilde{y} \equiv y_{T+h} - \mathbb{E}_T y_{t+h} = D\nu_{T+1}$$

where $D = QB^{h-1}\Phi_\nu$.

The effect of the shocks on the policy instrument are given by

$$\tilde{r} \equiv R\bar{\nu}_{T+1}$$

If

$$\tilde{z} \equiv \left[\begin{array}{c} \tilde{y} \\ \tilde{r} \end{array} \right]$$

then the properties of the multivariate normal distribution mean that $\tilde{z} \sim N(0, \Sigma_{\tilde{z}})$, where

$$\Sigma_{\tilde{z}} = \left[\begin{array}{cc} DD' & DR' \\ RD' & RR' \end{array} \right]$$

This result leads directly to the distributions of the variables of interest conditional on a choice of shocks that delivers a particular interest rate path, $\bar{r}$. Standard properties of multivariate normal distributions mean that:

$$\tilde{y} | (R\nu = \bar{r} - c) \sim N\left(DR'(RR')^{-1}(\bar{r} - c), DD' - DR'(RR')^{-1}RD' \right)$$

Note that the mean of this distribution can be written as $D\nu^*$, where ν^* is the solution to the minimization problem solved by Method 1 (unsurprisingly as ν^* is the shock vector that is most likely ex ante to deliver the path $\bar{r} - c$).

B   Results from alternative variants

This Appendix presents results from the model estimated on a shorter sample and under the restriction that the interest rate smoothing parameter is calibrated to a value of zero.

Tables B.1 and B.2 present summary statistics from estimation of the two variants of the model considered in Section 6. The tables present posterior means and 5-th and 95-th percentiles for each variant together with the baseline estimation results (also presented in the main text).

Figures B.1 and B.2 present estimates of policy experiments using the variants of the model estimated over the shorter sample. These figures show the effects of holding the Fed funds rate lower for 4 and 8 quarters respectively. Figure B.3 presents analysis of the sensitivity of inference based on the specification of the modesty statistic.

Figures B.5 and B.6 present estimates of policy experiments using the variants of the model estimated using a specification of the monetary policy rule without interest rate smoothing. These figures show the effects of holding the Fed funds rate lower for 4 and 8 quarters respectively. Figure B.7 presents analysis of the sensitivity of inference based on the specification of the modesty statistic.
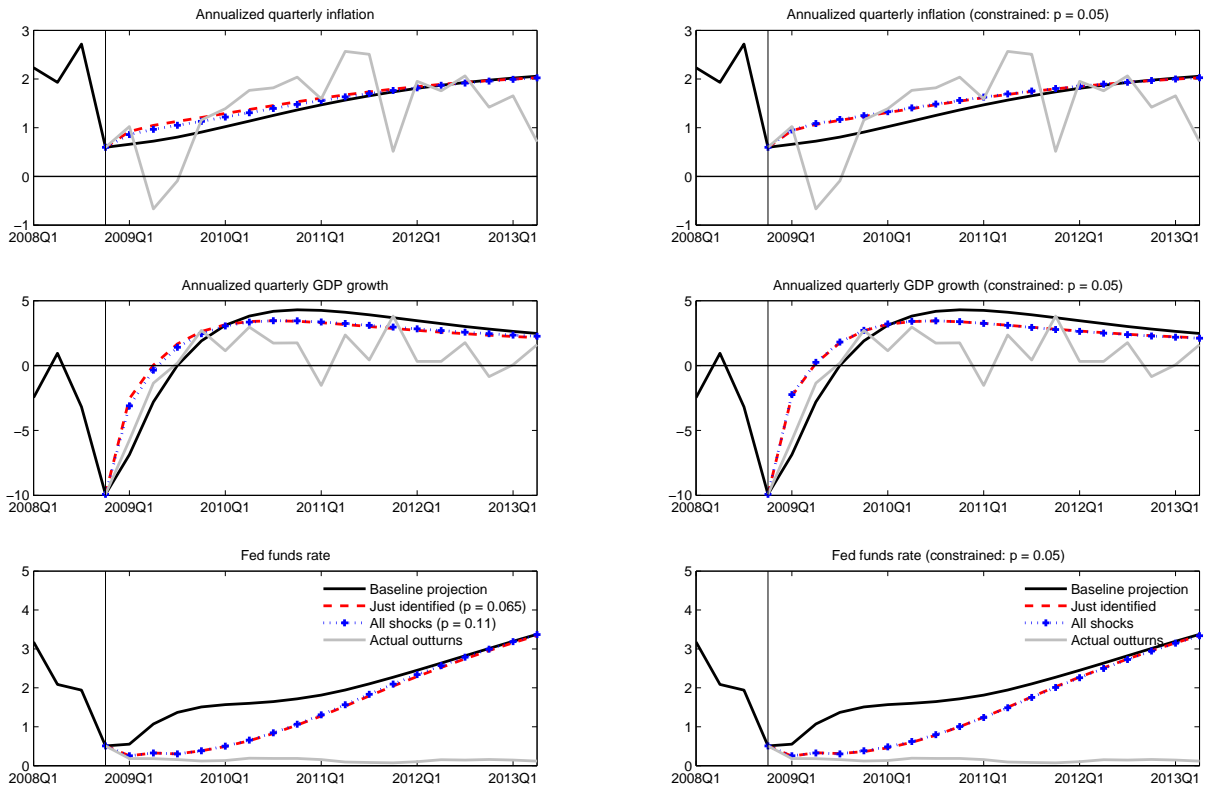
Table B.1: Comparison of posterior densities of estimated parameters for three variants

Parameter	Full sample			Shorter sample			No int rate smoothing		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
φ	6.514	4.694	8.448	5.935	4.204	7.838	6.304	4.506	8.281
σ_c	0.9763	0.7355	1.267	1.293	1.019	1.649	1.701	1.467	1.929
λ	0.5809	0.4617	0.6962	0.58	0.4617	0.6926	0.6812	0.6154	0.7453
ξ_w	0.4324	0.284	0.6211	0.8234	0.7535	0.8792	0.744	0.6767	0.8027
σ_l	1.436	0.584	2.419	1.638	0.682	2.72	2.201	1.342	3.134
ξ_p	0.7252	0.6004	0.8531	0.7704	0.6659	0.8571	0.9368	0.9187	0.9508
ι_w	0.4517	0.2196	0.701	0.3659	0.1639	0.6064	0.4426	0.2178	0.6822
ι_p	0.1517	0.06387	0.2651	0.2608	0.11	0.4674	0.1381	0.04962	0.2641
ψ	0.6619	0.4675	0.8337	0.7144	0.5416	0.8642	0.9105	0.8406	0.9626
Φ	1.528	1.388	1.673	1.624	1.471	1.78	1.306	1.204	1.414
$r_\pi - 1$	1.187	0.6704	1.814	0.3939	0.1335	0.7661	0.04512	0.01561	0.0875
ρ_R	0.8434	0.7727	0.8996	0.8806	0.8285	0.9235	0	–	–
r_y	0.01518	-0.02132	0.07113	0.1701	0.09897	0.2438	0.1318	0.1037	0.1623
$\bar{\pi}$	0.6379	0.4894	0.802	0.5575	0.4403	0.6887	0.5988	0.4552	0.7553
$100\left(\beta^{-1} - 1\right)$	0.2296	0.1062	0.3901	0.2199	0.1004	0.3746	0.1448	0.06926	0.2369
$\bar{L}$	-0.5214	-3.579	2.547	0.1052	-1.744	1.93	-0.4821	-2.147	1.215
$\bar{\gamma}$	0.4568	0.4133	0.4958	0.429	0.3653	0.487	0.06742	-0.0161	0.1495
ρ_{ga}	0.3962	0.2347	0.5589	0.5456	0.3418	0.7496	0.4389	0.2786	0.601
α	0.1717	0.1341	0.2096	0.2078	0.1637	0.2518	0.2025	0.1632	0.2412
ρ_a	0.916	0.8732	0.9549	0.8773	0.8096	0.935	0.979	0.9699	0.9857
ρ_b	0.9842	0.9675	0.996	0.9127	0.8816	0.9405	0.2779	0.09852	0.5573
ρ_g	0.9756	0.9542	0.9917	0.9566	0.9197	0.9844	0.9145	0.8776	0.9472
ρ_I	0.636	0.4574	0.8176	0.6441	0.4302	0.8452	0.6493	0.5533	0.7301
ρ_r	0.6232	0.5336	0.7042	0.7464	0.6221	0.8514	0.9778	0.9706	0.9846
ρ_p	0.927	0.8115	0.9922	0.728	0.4641	0.8939	0.7379	0.6575	0.8053
ρ_w	0.9898	0.9812	0.9965	0.4395	0.1911	0.7043	0.6704	0.5101	0.8078

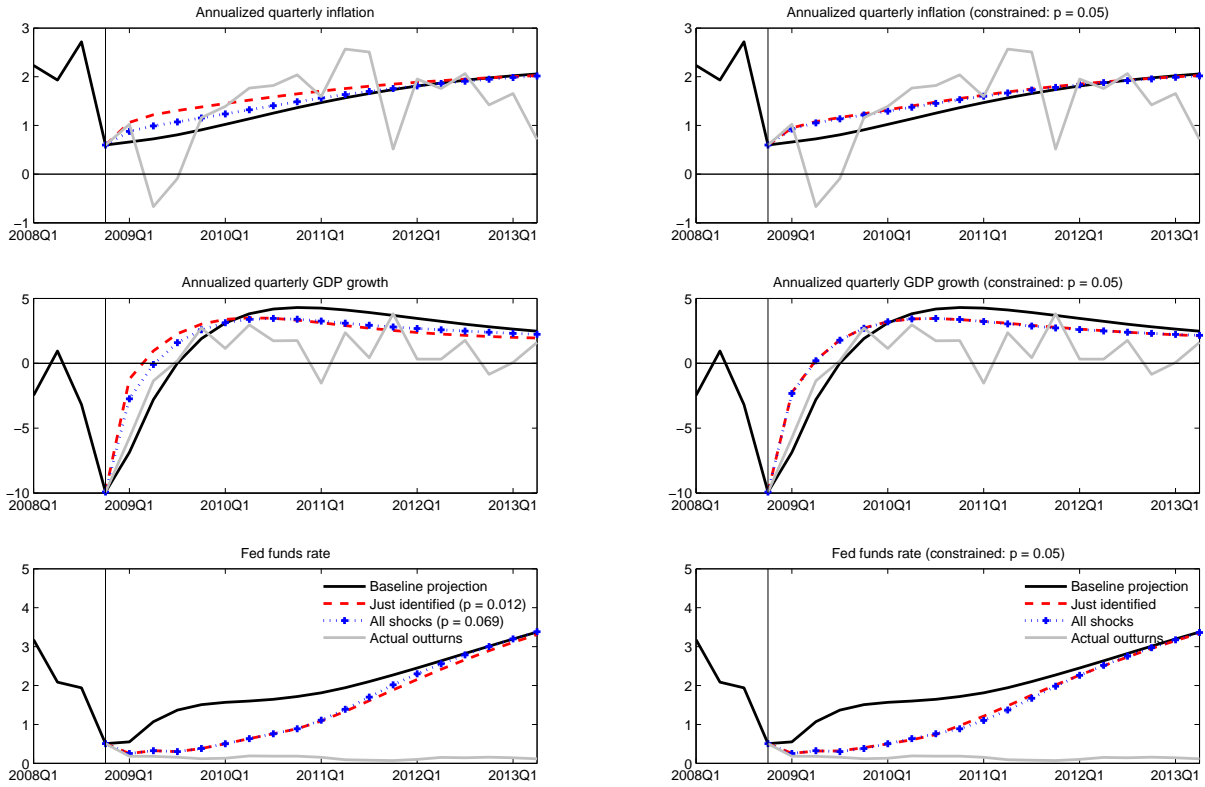Table B.2: Comparison of posterior densities of estimated parameters for three variants (continued)

Parameter	Full sample			Shorter sample			No int rate smoothing		
	Mean	5%	95%	Mean	5%	95%	Mean	5%	95%
μ_p	0.6838	0.4893	0.8325	0.5659	0.2647	0.8032	0.5703	0.3791	0.7183
μ_w	0.6258	0.3183	0.8934	0.438	0.1975	0.6829	0.398	0.1575	0.6348
σ_a	0.3804	0.3336	0.4323	0.4047	0.3398	0.4815	0.3804	0.3363	0.4291
σ_b	0.02464	0.01529	0.03657	0.05829	0.04092	0.07903	0.2008	0.148	0.2466
σ_g	0.3866	0.343	0.4367	0.3936	0.3319	0.4676	0.3936	0.3431	0.4526
σ_I	0.3396	0.2409	0.4551	0.3135	0.2093	0.4462	0.4198	0.3406	0.5115
σ_p	0.09304	0.07388	0.1133	0.1079	0.07817	0.1409	0.08798	0.06089	0.1131
σ_w	0.4039	0.3163	0.528	0.3762	0.2935	0.4709	0.2559	0.1996	0.3198
$100 \times \sigma_{\nu,0}$	9.723	8.142	11.6	6.642	5.306	8.2	18.46	16.27	19.85
$100 \times \sigma_{\nu,1}$	2.7	0.7446	4.926	3.67	0.8936	6.827	1.446	0.5284	3.054
$100 \times \sigma_{\nu,2}$	5.024	2.219	6.982	5.464	1.4	8.347	2.094	0.5758	5.403
$100 \times \sigma_{\nu,3}$	3.562	0.8822	5.905	2.127	0.5955	4.925	5.58	0.9686	9.687
$100 \times \sigma_{\nu,4}$	1.432	0.5317	2.965	1.792	0.5582	4.26	1.779	0.5572	4.37
$100 \times \sigma_{\nu,5}$	1.309	0.5181	2.514	1.392	0.5227	2.818	3.62	0.7345	8.355
$100 \times \sigma_{\nu,6}$	1.392	0.5393	2.706	1.475	0.5385	3.054	2.046	0.5797	4.864
$100 \times \sigma_{\nu,7}$	1.507	0.5564	3.045	1.356	0.5233	2.75	1.778	0.5616	4.158
$100 \times \sigma_{\nu,8}$	1.316	0.5114	2.6	1.485	0.5482	3.051	8.035	4.034	11.2
$100 \times \sigma_{\nu,9}$	1.326	0.5197	2.619	1.415	0.5375	2.866	1.676	0.5561	3.851
$100 \times \sigma_{\nu,10}$	1.346	0.5301	2.616	1.386	0.5307	2.757	1.916	0.5667	4.83
$100 \times \sigma_{\nu,11}$	1.778	0.61	3.374	1.419	0.5303	2.918	8.98	6.322	11.81
$\sigma_{me,4}$	0.02317	0.01588	0.03192	0.02539	0.01686	0.03571	0.02673	0.01727	0.03874
$\sigma_{me,20}$	0.02918	0.01933	0.04093	0.03881	0.02459	0.05616	0.05541	0.02462	0.1007
τ_4	0.04965	0.02479	0.07335	0.01828	-0.008663	0.04463	0.001588	-0.03694	0.03923
τ_{20}	0.2872	0.185	0.3827	0.043	-0.08068	0.161	0.1506	0.03376	0.2622

Figure B.1: Experiment in which Fed funds rate is held low for one year
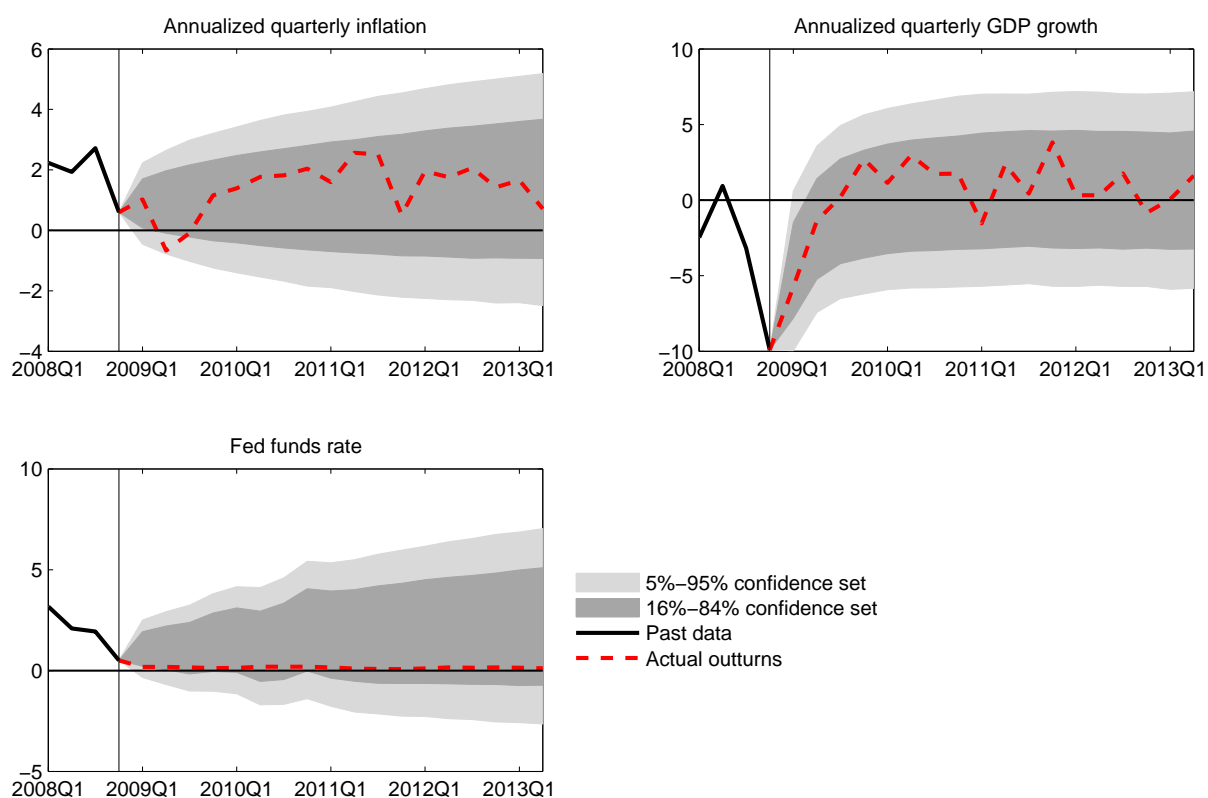

Notes: The left and right hand columns of charts show projections under alternative policy assumptions using a version of the model estimated over the sample period from 1996 to 2008. The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 4$ periods is generated using $n = K = 4$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 4$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

Figure B.2: Experiment in which Fed funds rate is held low for two years

Notes: The left and right hand columns of charts show projections under alternative policy assumptions using a version of the model estimated over the sample period from 1996 to 2008. The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 8$ periods is implemented using $n = K = 8$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 8$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

## Figure B.3: Modesty test statistics for alternative combinations of variables
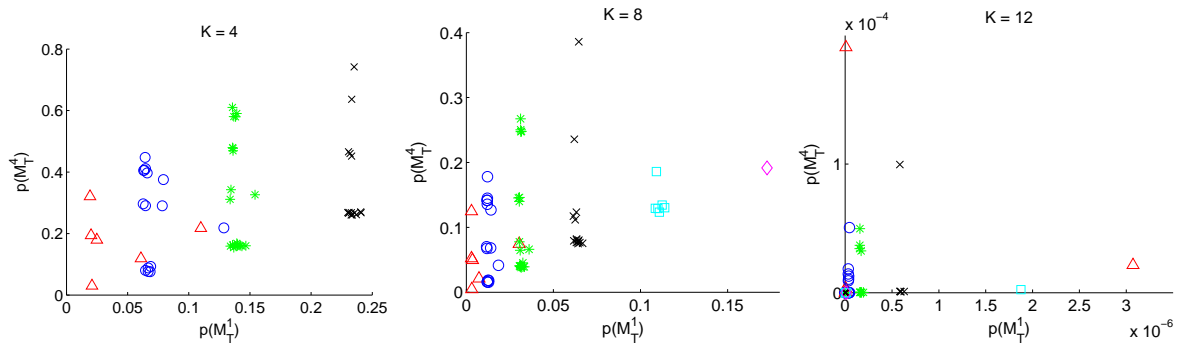
Notes: Each panel shows a scatter plot of the p-values associated with modesty tests at different horizons. The x-axis measures the p-value of the test statistic at horizon $h = 1$ and the y-axis measures the p-value of the test statistic for horizon $h = 4$. Each marker type shows the results when a particular subset of observable variables are chosen for the test: red triangles = 1 variable; blue circles = 2 variables; green asterixes = 3 variables; black crosses = 4 variables; cyan squares = 5 variables; magenta diamonds = 6 variables. The panels contain results for the just-identified (Method 1) implementation of experiments in Figures B.1, B.2 and 14.

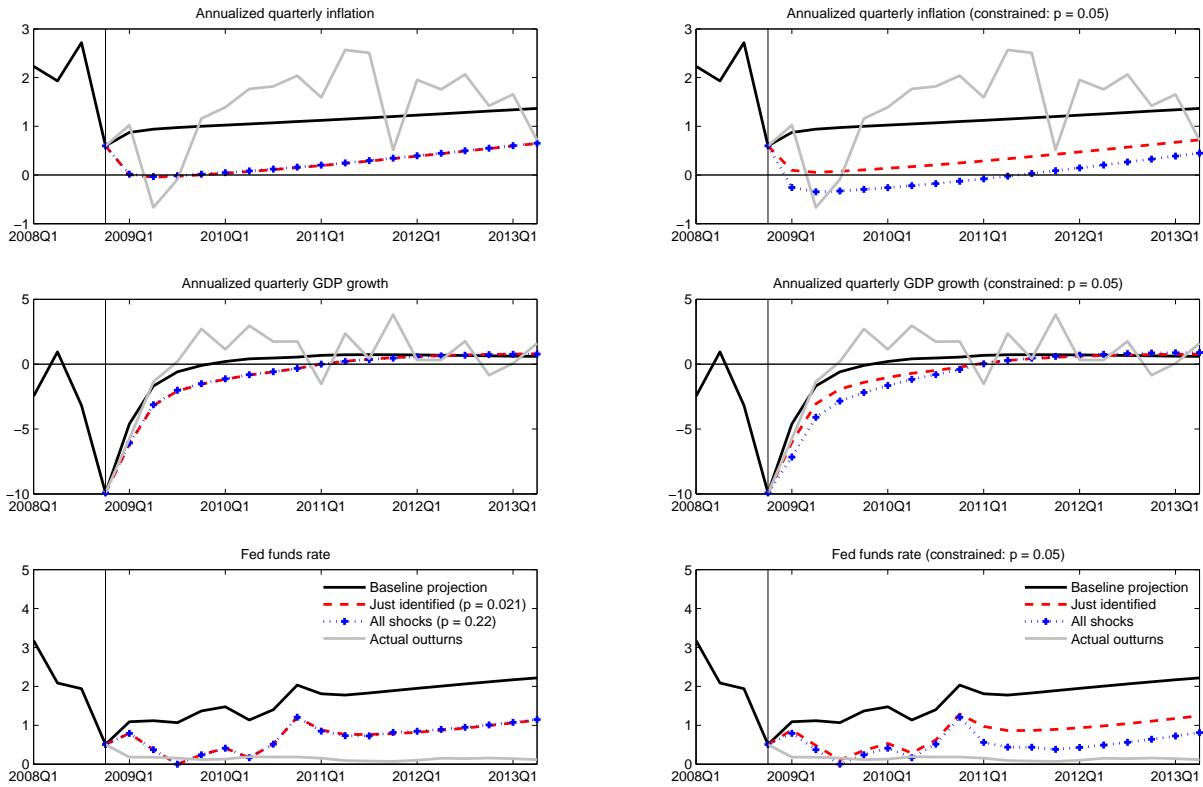Figure B.4: Forecasts from model without interest rate smoothing and actual outturns

Notes: The panels show confidence sets of forecasts generated by a variant of the model in which the interest rate smoothing parameter is fixed to zero. The sets are obtained by drawing from the posterior distribution of the estimated parameters. Actual outturns over the forecast horizon are shown by the dashed red lines.

# Figure B.5: Experiment in which Fed funds rate is held low for one year
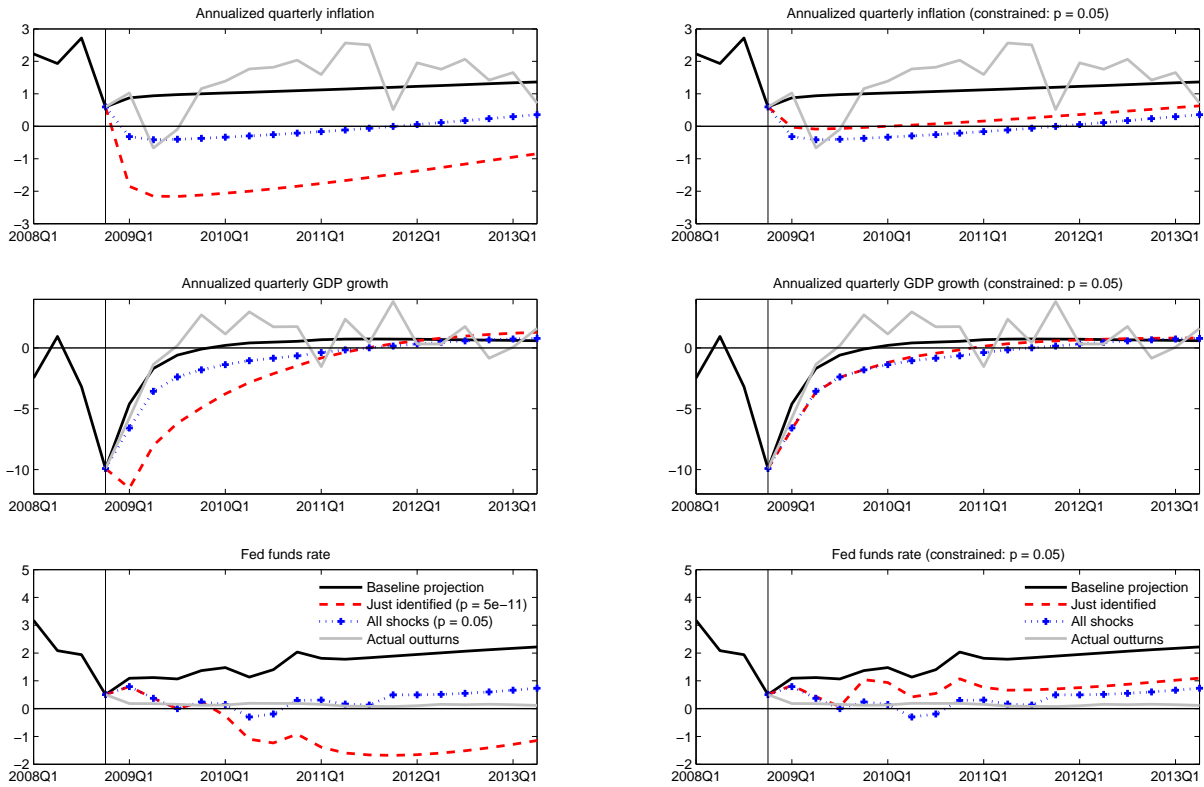
Notes: The left and right hand columns of charts show projections under alternative policy assumptions using a version of the model in which the interest rate smoothing parameter in the reaction function is calibrated to 0. The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 4$ periods is implemented using $n = K = 4$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 4$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

Figure B.6: Experiment in which Fed funds rate is held low for two years

Notes: The left and right hand columns of charts show projections under alternative policy assumptions using a version of the model in which the interest rate smoothing parameter in the reaction function is calibrated to 0. The left and right hand columns show projections under alternative policy assumptions. In both columns, the baseline forecast is plotted in the black solid line and the realized data is plotted in the solid gray line. The left column shows the results of policy experiments in which the path for the Fed funds rate for $K = 8$ periods is implemented using $n = K = 8$ policy news shocks ('just identified', red dashed lines) and $n = J = 12$ policy news shocks ('all shocks', blue dotted lines), both implemented using Method 1. The p-value of the modesty statistics associated with these simulations are reported in the legend. The right column repeats the experiments, but in each case imposes the additional restriction that the policy intervention be considered modest at the $p = 0.05$ significance level. Method 2 is used for $n = K = 8$ ('just identified', red dashed lines) and Method 3 for $n = J = 12$ ('all shocks', blue dotted lines).

Figure B.7: Modesty test statistics for alternative combinations of variables
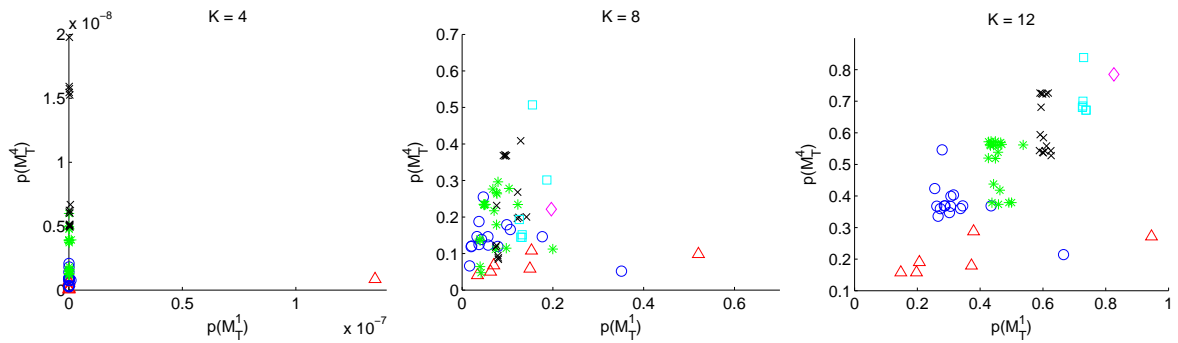
Notes: Each panel shows a scatter plot of the p-values associated with modesty tests at different horizons, for policy experiments using a version of the model in which the interest rate smoothing parameter is calibrated to zero. The x-axis measures the p-value of the test statistic at horizon $h = 1$ and the y-axis measures the p-value of the test statistic for horizon $h = 4$. Each marker type shows the results when a particular subset of observable variables are chosen for the test: red triangles = 1 variable; blue circles = 2 variables; green asterixes = 3 variables; black crosses = 4 variables; cyan squares = 5 variables; magenta diamonds = 6 variables. The panels contain results for the just-identified (Method 1) implementation of experiments in Figures B.5, B.6 and 15.