Columbia University

Department of Economics
Discussion Paper Series

Modeling Model Uncertainty

Alexei Onatski
Noah Williams

Discussion Paper #:0203-05

Department of Economics
Columbia University
New York, NY 10027

September 2002

Modeling Model Uncertainty

Alexei Onatski
Department of Economics
Columbia University
e-mail: ao2027@columbia.edu

Noah Williams*
Department of Economics
Princeton University
e-mail: noahw@princeton.edu

September 18, 2002

Abstract

Recently there has been a great deal of interest in studying monetary policy under model uncertainty. We develop new methods to analyze different sources of uncertainty in one coherent structure, which is useful for policy decisions. We show how to estimate the size of the uncertainty based on time series data, and how to incorporate this uncertainty in choosing policy. In particular, we develop a new approach for modeling uncertainty called model error modeling. The approach imposes additional structure on the errors of an estimated model, and builds a statistical description of the uncertainty around the model. We develop both parametric and nonparametric specifications of this approach, and use them to estimate uncertainty in a small model of the US economy. We then use our estimates to compute Bayesian and minimax robust policy rules, which are designed to perform well in the face of uncertainty.

1 Introduction

Uncertainty is pervasive in economics, and this uncertainty must be faced continually by policymakers. Poor quality of data, unpredictable shocks hitting the economy, econometric errors in estimation, and a lack of understanding of the fundamental economic mechanisms are among many different factors causing the uncertainty. Often, the uncertainty is so large that the effects of policy decisions on the economy are thought to be ambiguous. Under such

*Prepared for the 2002 ISOM in Frankfurt. We thank the participants in the seminar, especially our discussants Glenn Rudebusch and Ulf Söderström for detailed and insightful discussions and Chris Sims for useful comments. We are extremely grateful to Jim Stock for inviting us to participate and for providing helpful comments. We also thank Mario Milanese for extensive discussions which helped us initiate this project, and thank Glenn Rudebusch and Athanasios Orphanides for providing us with data. Finally, we thank the editor, Roberto Perotti, and two anonymous referees for comments and suggestions that greatly improved the substance and presentation of the paper.

an extreme uncertainty, any knowledge about the structure of uncertainty, as scarce as it might be, is very informative and must be useful. In this paper we therefore consider the structural modeling of the uncertainty relevant for policymaking.

We start by supposing that through some process of theorizing and data analysis, policymakers have arrived at a reference model of the economy. They want to use this model to set policy, but are concerned about potential uncertain deviations from it. There are three component blocks of our uncertainty model: first, uncertainty about the parameters of the reference model (including uncertainty about the model's order); second, uncertainty about the serial correlation properties of shocks; and third, uncertainty about data quality. Our analysis is focused on a simple, empirically-based macroeconomic model developed by Rudebusch and Svensson (1999) (henceforth, the RS model). In order to specify and measure the uncertainty about the RS model, we use a Model Error Modeling approach which draws upon recent advances in the control system identification literature due to Ljung (1999). We then apply both Bayesian and minimax techniques to develop policy rules which are robust to the uncertainty that we describe. Only Taylor-type policy rules, in which the interest rate is set in response to inflation and the output gap, are considered. The RS model and Taylor-type policy rules were chosen because they provide an empirically relevant, but technically simple, laboratory to illustrate the important features of our analysis.

Recently there has been a great deal of research activity on monetary policy making under uncertainty. Unfortunately, the practical implications of this research turn out to be very sensitive to different assumptions about uncertainty. For example, the classic analysis of Brainard (1967) showed that uncertainty about the parameters of a model may lead to cautious policy. More recently, Sargent (1999) showed that the introduction of extreme uncertainty about the shocks in the Ball (1999) model implies that very aggressive policy rules may be optimal. On the contrary, Rudebusch (2001) shows that focusing on the real time data uncertainty in the conceptually similar RS model leads to the attenuation of the optimal policy rule. Further, Craine (1979) and Söderström (2002) show that uncertainty about the dynamics of inflation leads to aggressive policy rules. Finally, Onatski and Stock (2002) find that uncertainty about the lag structure of the RS model requires a cautious reaction to inflation, but an aggressive response to variation in the output gap.

The fact that the robust policy rules are so fragile with respect to different assumptions about the structure of uncertainty is not surprising by itself. In fact, fragility is a general feature of optimizing models. Standard stochastic control methods are robust to realizations of shocks, as long as the shocks come from the assumed distributions and feed through the model in the specified way. But the optimal rules may perform poorly when faced with a different shock distribution, or slight variation in the model. The policy rules discussed above are each designed to be robust to a particular type of uncertainty, but may perform poorly when faced with uncertainty of a different nature. In our view, the most important message of the fragility of the robust rules found in the literature is that to design a robust policy rule in practice, it is necessary to combine different sources of uncertainty in a coherent structure and carefully estimate or calibrate the size of the uncertainty. In other words, it is necessary to structurally *model* uncertainty.

As we described above, we assume that policymakers start with a reference model of the economy. At a general level, model uncertainty can be adequately represented by suitable special restrictions on the reference model's shocks. For example, if one is uncertain about the parameters of the reference model or whether all relevant variables were included in the model, one should suspect that the reference shocks might actually be correlated with the explanatory variables in the model. That is, the reference model's shocks would now include "true" exogenous shocks and modeling errors. The model uncertainty can now be formulated by defining a set of potentially true models for these errors, or in other words by "Model Error Modeling".

One popular way to describe restrictions on the reference shocks (see for example Hansen and Sargent (2002)) is to assume that the shocks must be of bounded size, but arbitrary otherwise. We argue that a much more structured model of the shocks must be used to describe uncertainty relevant to monetary policymaking. In particular, we develop an example showing that the Hansen and Sargent (2002) approach may lead to the design of robust policy rules that can be destabilized by small parametric perturbations. Thus while the robust rule may resist shocks of a certain size, small variations in the underlying model can result in disastrous policy performance.

We then turn to the task of formulating an empirical description of uncertainty by model error modeling. In particular, we discuss and implement both parametric and nonparametric specifications for the RS model errors. The parametric specification imposes more structure and results in a probabilistic description of uncertainty. We estimate these parameters using Bayesian methods, obtaining a posterior distribution which characterizes the uncertainty. The nonparametric specification imposes fewer restrictions, and results in a deterministic specification of the uncertainty. This allows us to calibrate the size of the uncertainty set, but as it is a deterministic description, we cannot evaluate the likelihood of alternative models in the set.

After we estimate or calibrate the uncertainty, we use our results to formulate robust policy rules which are designed to work well for the measured uncertainty. From the parametric specification, we have a distribution over possible models. Therefore for this specification we find robust optimal rules which minimize the Bayesian risk. From the nonparametric specification, we have bounds on the uncertainty set. Therefore for this specification we find robust optimal rules which minimize the worst possible loss for the models in the set. This minimax approach follows much of the recent literature on robust control, and provides a tractable way of using our most general uncertainty descriptions. While there is the possibility that minimax results may be driven by unlikely models, we focus solely on empirically plausible model perturbations. Further, for many of our specifications the Bayesian and minimax results are quite similar. This suggests both that the stronger restrictions in the Bayesian framework do not greatly affect results, and that the minimax results are not driven by implausible worst case scenarios. It is worth noting that in all of our results we assume that policy makers commit to a rule once-and-for-all. Although this approach is common in the literature, it is clearly an oversimplification. This should be kept in mind, particularly when considering some of the bad outcomes we find for certain policy rules.

Both our Bayesian and minimax analysis indicate that without imposing much prior structure on the model perturbations, dynamic instability is a very likely outcome. This suggests that very large losses and very poor economic performance are to be expected when policy is conducted using simple interest rate rules. However when we tighten prior beliefs so that instability is deemed unlikely, our results change dramatically. In this case, the optimal rules under both our parametric-Bayesian and nonparametric-minimax specifications are quite close to the optimal rule in the absence of model uncertainty. This shows that the conventional optimal Taylor-type rule possesses strong robustness properties.

As a result of our minimax analysis, we find that the most damaging model perturbations come from very low frequency changes. Correspondingly, many of the robust policy rules that we find are relatively aggressive, stemming from policymakers' fears of particularly bad long-run deviations from the RS model. In particular, we impose a vertical long-run Phillips curve. Thus increases in the output gap would lead to very persistent increases in inflation in the absence of a relatively aggressive interest rate rule. The size of this persistent component is poorly measured, but has a huge impact on the losses sustained by the policy maker. However, the RS model is essentially model of short-run fluctuations, and is not designed to capture long-run phenomena. By asking such a simple model to accommodate very low frequency perturbations, we feel that we are pushing the model too far. A more fully developed model would be necessary to capture low frequency behavior.

We believe that for practical purposes, it is prudent to downweight the importance of the low frequency movements. To tailor our uncertainty description to more relevant worst case scenarios, we reconsider our minimax results when restricting attention to business cycle frequencies (corresponding to periods from 6 to 32 quarters). Interestingly, we find that, in this case, the robust optimal Taylor rules are less aggressive than the optimal Taylor rules under no uncertainty. This result can be interpreted as follows. Under no uncertainty, the optimal Taylor-type rules balance the performance of policy at all frequencies. When we introduce uncertainty at business cycle frequencies only, then the worst case scenarios occur at these frequencies, which makes policy very responsive to these frequencies. This comes at the cost of downweighting low frequency movements. Instead of fighting off any incipient inflation, policy becomes less aggressive, and focuses more on counter-cyclical stabilization. This contrasts with policymakers worried about low frequency perturbations, who may be reluctant to try to stimulate the economy in a recession.

One of the main benefits of our approach is that it allows us to treat many different forms of uncertainty in a unified framework. However it is interesting to also consider the different sources independently. This allows us to see how the uncertainty channels affect policy rules, and to determine which channels have the largest effects on losses. These results can provide useful information for users of similar models, by pointing out the most important parts of the model specification. When we consider uncertainty at all frequencies, we find that the uncertainty about the slope of the IS curve is the most important channel for worst-case losses, but that the real-time data uncertainty (in particular revisions in inflation data) has the largest effect on policy rules. When we restrict our attention to business-cycle frequencies, the shocks become more important and most of the policy rules become attenuated.

In the next section of the paper we describe the framework for our analysis at a general level. In Section 3 we present an example highlighting the importance of the model of uncertainty, and show that parametric and shock uncertainty must be considered separately. Section 4 describes our application of the Model Error Modeling approach to find both parametric and nonparametric measures of the uncertainty associated with the Rudebusch-Svensson model. Section 5 formulates robust monetary policy rules based on our uncertainty descriptions. Section 6 concludes.

2 General Framework

The general issue that we consider in this paper is decision making under model uncertainty. In particular, we focus on the policy-relevant problem of choosing interest rate rules when the true model of the economy is unknown and may be subject to different sources of uncertainty. The goal of the paper is to provide characterizations of the empirically relevant sources of uncertainty, and to design policy rules which account for that uncertainty.

The starting point of our analysis is a reference model of the economy:

$$x_{t+1} = A(L)x_t + B_1(L)u_t + B_2(L)\varepsilon_t \quad (1)$$

$$y_t = C(L)x_t + D(L)\varepsilon_t, \quad (2)$$

where x_t is a vector of macroeconomic indicators, u_t is a vector of controls such as taxes, money, or interest rates, y_t is a vector of variables observed in real time, ε_t is a vector of white noise shocks, and $A(L), B_i(L), C(L)$, and $D(L)$ are matrix lag polynomials. Note that the majority of purely backward-looking models of the economy can be represented in the above form. In fact, by defining the state appropriately, this system of equations has a standard state-space form. We consider this form of the reference model because, as will soon be clear, it accords with our description of the uncertainty.

As mentioned in the introduction, we assume that through some unmodeled process of trial and error policy makers have arrived at a reference model of the economy. In this paper, we do not address an important question of how to choose a reference model. Instead, we assume that the reference model is given, and policy makers are concerned about small deviations of the true model from the reference one. This is also the starting point of much of the literature on robustness in economics, as described for example in Hansen and Sargent (2002). A more ambitious question of what policy a central bank should follow under vast disagreement about the true model of the economy is addressed, for example in Levin, Wieland, and Williams (1999).

We assume that policymakers have a time-additively separable quadratic loss function:

$$L_t = E_t \sum_{i=0}^{\infty} \beta^i x'_{t+i} \Lambda x_{t+i}.$$

They seek to minimize losses by choosing a policy rule from an admissible class:

$$u_t = f(y_t, y_{t-1}, ..., u_{t-1}, u_{t-2}, ...).$$

The admissible class does not necessarily include the optimal control because the optimality of a rule may be traded off with its other characteristics, such as simplicity. In some cases it is more convenient to discuss policymakers maximizing a utility function, which is simply the negative of the loss function.

Equations (1) and (2) can be estimated for a time period in the past for which both real-time data y_t, u_t and the final data x_t are available. The obtained estimates can then be used to compute the best policy rule from the admissible class. The quality of the policy rule obtained in this way will depend on the accuracy of the reference model. In general, this model will not be completely accurate. The reference model is likely to be a stylized macroeconomic model, which for tractability may leave out certain variables or focus only on the first few lags of the relevant variables. While these simplifications may be justified for both practical and statistical reasons, we will show that they can have a large impact on policy decisions.

We assume that a more accurate model of the economy encompasses the reference model as follows:

$$x_{t+1} = \left(A(L) + \tilde{A}(L)\right) x_t + \left(B_1(L) + \tilde{B}_1(L)\right) u_t + \left(B_2(L) + \tilde{B}_2(L)\right) \varepsilon_t \quad (3)$$

$$y_t = \left(C(L) + \tilde{C}(L)\right) x_t + \left(D(L) + \tilde{D}(L)\right) \varepsilon_t, \quad (4)$$

where $\tilde{A}(L), \tilde{B}_i(L), \tilde{C}(L)$ and $\tilde{D}(L)$ are relatively unconstrained matrix lag polynomials of potentially infinite order. Our assumption allows for a rich variety of potential defects in the reference model. Econometric errors in the estimation of the reference parameters, misspecifications of the lag structure of the reference equations, and misinterpretations of the real-time data are all considered as distinct possibilities.

We assume that the central bank wants to design a policy rule that works well not only for the reference model but also for statistically plausible deviations from the reference model having form (3,4). Formally, such a set can be defined by a number of restrictions $\mathcal{R}$ on the matrix lag polynomials $\tilde{A}(L), \tilde{B}_i(L), \tilde{C}(L)$ and $\tilde{D}(L)$. The restrictions $\mathcal{R}$ may be deterministic if sets of the admissible matrix lag polynomials are specified, or stochastic if distributions of the polynomials' parameters are given.

We formalize policy makers' desire for robustness by assuming that they use Bayesian or minimax strategy for choosing the policy, depending on whether $\mathcal{R}$ is stochastic or deterministic. That is, in the stochastic case policy makers solve the Bayes problem:

$$\min_{\{u_t = f(.)\}} E_{\mathcal{R}} L_t \quad (5)$$

where the expectation is taken with respect to distributions of the potential deviations from the reference model specified by $\mathcal{R}$. In the deterministic case, they solve the minimax problem:

$$\min_{\{u_t = f(.)\}} \max_{\mathcal{R}} L_t \quad (6)$$

where the maximum is taken over all matrix lag polynomials $\tilde{A}(L), \tilde{B}_i(L), \tilde{C}(L)$ and $\tilde{D}(L)$ satisfying the deterministic restrictions $\mathcal{R}$.[1]

It is needless to say that, at least in principle, the particular structure of the restrictions $\mathcal{R}$ will strongly affect solutions to the above problems. In the next section, we illustrate importance of this structure through a simple example.

3 Consequences of Different Uncertainty Models

It is useful to re-write (3)-(4) to represent the model uncertainty in the form:

$$\begin{aligned} x_{t+1} &= A(L)x_t + B_1(L)u_t + w_t \\ y_t &= C(L)x_t + s_t, \end{aligned}$$

where we define the "model errors" as:

$$\begin{aligned} w_t &= \tilde{A}(L)x_t + \tilde{B}_1(L)u_t + \left(B_2(L) + \tilde{B}_2(L)\right)\varepsilon_t, \\ s_t &= \tilde{C}(L)x_t + \left(D(L) + \tilde{D}(L)\right)\varepsilon_t, \end{aligned} \tag{7}$$

and $\tilde{A}(L), \tilde{B}_i(L), \tilde{C}(L)$ and $\tilde{D}(L)$ comply with $\mathcal{R}$. This representation shows that, the uncertainty may be described by restrictions (7) on the model errors w_t and s_t.

One approach to model uncertainty, similar in spirit to that developed by Hansen and Sargent (2002), does not impose any special structure on w_t and s_t. Instead, the approach considers all errors subject to the restriction:

$$E\sum_{t=0}^{\infty} \beta^t (w_t'\Phi_1 w_t + s_t'\Phi_2 s_t) < \eta. \tag{8}$$

The parameter η in the above inequality regulates the size of uncertainty, and it may be calibrated so that the corresponding deviations from the reference model are statistically plausible. While this approach seems quite general and unrestrictive, not taking into account the particular structure of w_t and s_t may seriously mislead decision makers. Below we develop an example illustrating this fact. In the example, we are trying to be as close as possible to a practically important situation.

We consider a two-equation purely backward-looking model of the economy proposed and estimated by Rudebusch and Svensson (1999). This model will be the benchmark for the rest of the paper as well, and is given by:

$$\begin{aligned} \pi_{t+1} &= \underset{(.08)}{.70}\pi_t - \underset{(.10)}{.10}\pi_{t-1} + \underset{(.10)}{.28}\pi_{t-2} + \underset{(.08)}{.12}\pi_{t-3} + \underset{(.03)}{.14}y_t + \varepsilon_{\pi,t+1} \\ y_{t+1} &= \underset{(.08)}{1.16}y_t - \underset{(.08)}{.25}y_{t-1} - \underset{(.03)}{.10}\left(\bar{\imath}_t - \bar{\pi}_t\right) + \varepsilon_{y,t+1} \end{aligned} \tag{9}$$

[1]Note that in our formulation, the model uncertainty takes form of a one-time uncertain shift in the parameters or specification of the reference model. For an analysis of uncertainty interpreted as a stochastic process in the space of models see Rudebusch (2000).

The standard errors of the parameter estimates are given in parentheses. Here the variable y stands for the gap between output and potential output, π is inflation and i is the federal funds rate. All the variables are quarterly, measured in percentage points at an annual rate and demeaned prior to estimation, so there are no constants in the equations. The variables $\bar{\pi}$ and $\bar{i}$ stand for four-quarter averages of inflation and the federal funds rate respectively.

The first equation is a simple version of the Phillips curve, relating the output gap and inflation. The coefficients on the lags of inflation in the right hand side of the equation sum to one, so that the Phillips curve is vertical in the long run. The second equation is a variant of the IS curve, relating the real interest rate to the output gap. A policymaker can control the federal funds rate and wants to do so in order to keep y and π close to their target values (zero in this case). For the present, we ignore the real-time data issues so that our reference model does not include equations describing real-time data generating process.

In general, the policy maker's control policy may take the form of a contingency plan for her future settings of the federal funds rate. We, however, restrict attention to the Taylor-type rules for the interest rate. As emphasized by McCallum (1988) and Taylor (1993), simple rules have the advantage of being easy for policymakers to follow and easy to interpret. In this section, we assume that the policymaker chooses among the following rules:

$$i_t = g_\pi \bar{\pi}_{t-1} + g_y y_{t-2} \tag{10}$$

Here, the interest rate reacts to both inflation and the output gap with delay. The delay in the reaction to the output gap is longer than that in the reaction to the inflation because it takes more time to accurately estimate the gap. The timing in the above policy rule is unorthodox, and is made here to sharpen our results. In later sections we use the more conventional timing, in which the interest rate responds contemporaneously to inflation and the output gap.

Following Rudebusch and Svensson, we assume that a policy maker has the quadratic loss function:[2]

$$L_t = \bar{\pi}_t^2 + y_t^2 + \frac{1}{2}(i_t - i_{t-1})^2. \tag{11}$$

If the policy maker were sure that the model is correctly specified, she could use standard methods to estimate the expected loss for any given policy rule (10). Then she could find the optimal rule numerically. Instead, we assume that the policy maker has some doubts about the model. She wants therefore to design her control so that it works well for reasonable deviations from the original specification.

One of the most straightforward ways to represent her doubts is to assume that the model parameters may deviate from their point estimates as, for example, is assumed in Brainard (1967). It is also likely, that the policy maker would not rule out misspecifications of the model's lag structure. As Blinder (1997) states, "Failure to take proper account of lags is, I believe, one of the main sources of central bank error."

[2] The inclusion of the interest-smoothing term $(i_t - i_{t-1})^2$ in the loss function is somewhat controversial. Our results will not depend on whether this term is included in the loss function or not.

For the sake of illustration, we assume that the policy maker contemplates the possibility that one extra lag of the output gap in the Phillips curve and IS equations and one extra lag of the real interest rate in the IS equation were wrongfully omitted from the original model. She therefore re-estimates Rudebusch-Svensson model with the additional lags. The re-estimated model has the following form:

$$\begin{aligned} \pi_{t+1} &= \underset{(.08)}{.70}\pi_t - \underset{(.10)}{.10}\pi_{t-1} + \underset{(.10)}{.28}\pi_{t-2} + \underset{(.09)}{.12}\pi_{t-3} + \underset{(.10)}{.14}y_t + \underset{(.10)}{.00}y_{t-1} + \varepsilon_{\pi,t+1} \\ y_{t+1} &= \underset{(.08)}{1.13}y_t - \underset{(.12)}{.08}y_{t-1} - \underset{(.08)}{.14}y_{t-2} - \underset{(.14)}{.32}\left(\bar{\imath}_t - \bar{\pi}_t\right) + \underset{(.14)}{.24}\left(\bar{\imath}_{t-1} - \bar{\pi}_{t-1}\right) + \varepsilon_{y,t+1} \end{aligned} \tag{12}$$

Then she obtains the covariance matrix of the above point estimates and tries to design her control so that it works best for the worst reasonable deviation of the parameters from the point estimates. For example, she may consider all parameter values inside the 50% confidence ellipsoid around the point estimates.[3]

We will soon return to this problem, but for now let us give an alternative, less structured, description of the uncertainty. In general, we can represent uncertainty by modeling the errors w_{1t}, w_{2t} of the Phillips curve and the IS equations as any processes satisfying:

$$E\sum_{t=0}^{\infty}\beta^t\left(\frac{w_{1t}^2}{\mathrm{Var}(\varepsilon_{\pi t})} + \frac{w_{2t}^2}{\mathrm{Var}(\varepsilon_{yt})}\right) < \eta.$$

Here we will consider the case $\beta \to 1$. The special choice of the weights on errors to the Phillips curve and the IS equations was made to accommodate the MATLAB codes that we use in our calculations.

In the extreme case when η tends to infinity, our uncertainty will be very large, so the corresponding robust (minimax) rule must insure the policy maker against a large variety of deviations from the reference model. It can be shown that such an "extremely robust" policy rule minimizes the so-called H_∞ norm of the closed loop system transforming the noise $\varepsilon_t = \left(\varepsilon_{\pi t}/\sqrt{\mathrm{Var}(\varepsilon_{\pi t})}, \varepsilon_{yt}/\sqrt{\mathrm{Var}(\varepsilon_{yt})}\right)$ into the target variables $z_t = \left(\pi_t, y_t, (i_t - i_{t-1})/\sqrt{2}\right)'$ (see Hansen and Sargent (2002)). It is therefore easy to find such a rule numerically using, for example, commercially available MATLAB codes to compute the H_∞ norm. Our computations give the following rule:

$$i_t = 3.10\bar{\pi}_{t-1} + 1.41y_{t-2}. \tag{13}$$

Now let us return to our initial formulation of the problem. Recall that originally we wanted to find a policy rule that works well for all deviations of the parameters of the re-estimated model (12) inside a 50% confidence ellipsoid around the point estimates. Somewhat surprisingly, the above "extremely robust" rule does not satisfy our original criterion for robustness. In fact, it destabilizes the economy for deviations from the parameters' point estimates inside as small as a 20% confidence ellipsoid. More precisely, the policy rule (13)

[3] In the later sections of the paper we will discuss a more systematic way of representing and estimating the model uncertainty. We also will not restrict our attention to the max-min setting as we do in this section.

results in infinite expected loss for the following perturbation of the Rudebusch-Svensson (RS) model:

$$\begin{aligned} \pi_{t+1} &= .68\pi_t - .13\pi_{t-1} + .35\pi_{t-2} + .10\pi_{t-3} + .30y_t - .15y_{t-1} + \varepsilon_{\pi,t+1} \\ y_{t+1} &= 1.15y_t - .07y_{t-1} - .18y_{t-2} - .51\left(\bar{\imath}_t - \bar{\pi}_t\right) + .41\left(\bar{\imath}_{t-1} - \bar{\pi}_{t-1}\right) + \varepsilon_{y,t+1}. \end{aligned} \tag{14}$$

Let us denote the independent coefficients of the above model, the re-estimated RS model (12), and the original RS model as c, c_1, and c_0 respectively.[4] Also, let V be the covariance matrix of the coefficients in the re-estimated model (12). Then we have:

$$\begin{aligned} (c - c_1)'V^{-1}(c - c_1) &= 6.15 \\ (c_0 - c_1)'V^{-1}(c_0 - c_1) &= 5.34. \end{aligned}$$

Both numbers are smaller than the 20% critical value of the chi-squared distribution with 10 degrees of freedom. This may be interpreted as saying that both the original Rudebusch-Svensson model and the perturbed model are statistically close to the encompassing re-estimated model. In spite of this, the robust rule leads to disastrous outcomes.

Why does our "extremely robust" rule perform so poorly? It is not because other rules work even worse. For example, we checked that the famous Taylor rule $i_t = 1.5\bar{\pi}_{t-1} + 0.5y_{t-2}$ guarantees stability of the economy at least for all deviations inside 60% confidence ellipsoid. Rule (13) works so bad simply because it was not designed to work well in such a situation. To see this, note that our original description of the model uncertainty allowed deviations of the slope of the IS curve from its point estimate. Therefore our ignorance about this parameter would be particularly influential if we were to use a very aggressive control rule. It may even be consistent with instability under such an aggressive rule. However no effects of this kind are allowed under the unstructured description of model uncertainty. The specific interaction between the aggressiveness of policy rules and uncertainty about the slope of the IS curve is not taken into account. This lack of structure in the uncertainty description turns out to be dangerous because the resulting robust rule happens to be quite aggressive.

The example just considered should not be interpreted in favor of a particular description of uncertainty. Instead, it illustrates that when designing robust policy rules, we must carefully specify and thoroughly understand the model uncertainty that we are trying to deal with. Robust policy rules may be fragile with respect to reasonable changes in the model uncertainty specification. In the next sections, we therefore use a systematic approach based on model error modeling to estimate the uncertainty about the Rudebusch-Svensson model introduced above. Then we use our estimates of the model uncertainty to find interest rate rules which perform well in the face of this uncertainty.

[4] Recall that the sum of coefficients on inflation in the Phillips curve is restricted to be equal to 1. We therefore exclude the coefficient on the first lag of inflation from the vector of independent coefficients. Collecting our estimates, these are: $c = (-.13, .35, .10, .30, -.15, 1.15, -.07, -.18, -.51, .41)'$, $c_1 = (-.10, .28, .12, .14, .00, 1.13, -.08, -.14, -.32, .24)'$, $c_0 = (-.10, .28, .12, .14, 0, 1.16, -.25, 0, -.10, 0)'$.

4 Model Error Modeling

As was shown in the previous section, model uncertainty can be reformulated in terms of restrictions (7) on the errors of the reference model. Hence, to form an empirically relevant description of the uncertainty, one should find a set of models having the form (7) which are consistent with available data and prior beliefs. We now begin specifying the model uncertainty model for our application.

4.1 Specifying the Uncertainty Models

We start by adding equations describing the real-time data to the Rudebusch and Svensson's reference model of the economy described in the previous section. Such an extension of the reference model is important because the central bank's policy must feedback on the information available in real time. As emphasized by Orphanides (2001), there is a substantial amount of uncertainty in such information. Initial estimates of GDP, and hence the deflator and output gap, are typically revised repeatedly and the revisions may be substantial.

Our reference assumption is that the real-time data on inflation, π_t^*, and the output gap, y^*, are equal to noisy lagged actual data, and the noise has AR(1) structure. That is:

$$\begin{aligned}
\pi_t^* &= \pi_{t-1} + \eta_t^\pi, \text{ where } \eta_t^\pi = \rho^\pi \eta_{t-1}^\pi + v_t^\pi \\
y_t^* &= y_{t-1} + \eta_t^y, \text{ where } \eta_t^y = \rho^y \eta_{t-1}^y + v_t^y.
\end{aligned}$$

The assumption of the AR(1) noise in the real-time data accords with previous studies (see for example Orphanides (2001) and Rudebusch (2001)). The choice of timing in the above equations is consistent with the fact that lagged final data predicts the real-time data better than the current final data do. This is true at least for the sample of the real-time data on the output gap and inflation for the period from 1987:1 to 1993:04 that we use, which was kindly provided to us by Athanasios Orphanides from his 2001 paper.

Using the Rudebusch-Svensson data set kindly provided to us (some time ago) by Glenn Rudebusch, we compute the errors of the RS Phillips curve, e_{t+1}^π, and the IS curve, e_{t+1}^y. Using Athanasios Orphanides' data, we compute the errors of our reference model for the real-time data on inflation, $e_t^{d,\pi}$, and the output gap $e_t^{d,y}$.[5] We then model the reference model's errors as follows:

$$\begin{aligned}
e_{t+1}^\pi &= a(L)(\pi_t - \pi_{t-1}) + b(L)y_t + \varepsilon_{t+1}^\pi, \text{ where } \varepsilon_{t+1}^\pi = c(L)\varepsilon_t^\pi + u_{t+1}^\pi \\
e_{t+1}^y &= d(L)y_t + f(L)\pi_t + g(L)i_t + \varepsilon_{t+1}^y, \text{ where } \varepsilon_{t+1}^y = h(L)\varepsilon_t^y + u_{t+1}^y \\
e_t^{d,\pi} &= k(L)\pi_t + \eta_t^\pi, \text{ where } \eta_t^\pi = m(L)\eta_{t-1}^\pi + v_t^\pi \\
e_t^{d,y} &= n(L)y_t + \eta_t^y, \text{ where } \eta_t^y = p(L)\eta_{t-1}^y + v_t^y.
\end{aligned}$$

Several structurally distinct misspecifications of the RS model are consistent with our model of the errors. First, non-zero functions a, b, d, f, and g imply errors in the coefficients

[5] In our terminology, the "errors" of the real-time data reference equations are simply $\pi_t^* - \pi_{t-1}$ and $y_t^* - y_{t-1}$.

and lag specifications in the reference Phillips curve and the IS equations. Note that the econometric errors in the point estimates of the reference parameters are thus taken into account. The misspecification of the reference lag structure may be interpreted literally (say, more distant lags of the real interest rate have a direct non-trivial effect on the output gap), or as indicating omission of important explanatory variables from the reference model. Second, our inclusion of both inflation and the nominal interest rate in the model of the IS equation error e_{t+1}^{y} allows for the separation of the effect of real and nominal interest rates on the output gap.[6] Finally, non-zero functions c and h allow for rich serial correlation structure of the shocks to the Phillips and IS curves.

Similarly, for the reference real-time data equations, non-zero functions m and p extend the possible serial correlation structure of the noise η beyond the reference AR(1) assumption. As to the functions k and n, they model the "news component" of the data revision process.[7] To see this, note that the revisions $\pi_t^* - \pi_t$ and $y_t^* - y_t$ can be expressed in the form $(k(L) - 1 + L)\pi_t + \eta_t^{\pi}$ and $(n(L) - 1 + L)y_t + \eta_t^{y}$ respectively. The functions k and n are thus responsible for the structure of the correlation between the final data and the revisions, which defines the news component.

One possible extension of our analysis would be to include additional variables in the model errors. For example, it is not unreasonable to think that the true dynamics of the inflation and the output gap should depend on the real exchange rate. However our description of uncertainty does allow for such a relationship, albeit an implicit one. In this paper, we deal with reduced form models. Of course, uncertainty about the reduced form dynamics may correspond to a deeper uncertainty about a background structural model that includes much more variables than just inflation and the output gap. However, we could potentially sharpen our estimates of uncertainty by explicitly including "omitted" variables directly in the error model. We leave such important extensions of our analysis for future research.

4.2 Estimating the Models

We have structured the compound model uncertainty faced by policymakers via the lag polynomials $a(L)$ through $p(L)$ describing the dynamics of the model errors. The restrictions on these polynomials may either be parametric or nonparametric. In this section we describe one parametric and one nonparametric specification. We also describe a possible way of formulating empirically relevant constraints for each specification. The parametric specification imposes more structure, and allows us to determine a probability distribution over the class of alternative models. The nonparametric specification imposes significantly less structure, but only provides bounds on the class of feasible alternative models. Later when we use our measures of model uncertainty for policy decisions, these differences will be crucial.

First, for the parametric case, we assume that $a, b, c, d, f, g,$ and h (which affect the Phillips and IS curve errors) are fourth order lag polynomials, and $k, m, n,$ and p (which affect the real-time data errors) are second order lag polynomials. The choice of these

[6]We thank Glenn Rudebusch for suggesting this extension of the reference model.

[7]See Mankiw and Shapiro (1986) for a discussion of news versus noise in the revisions of real-time data.

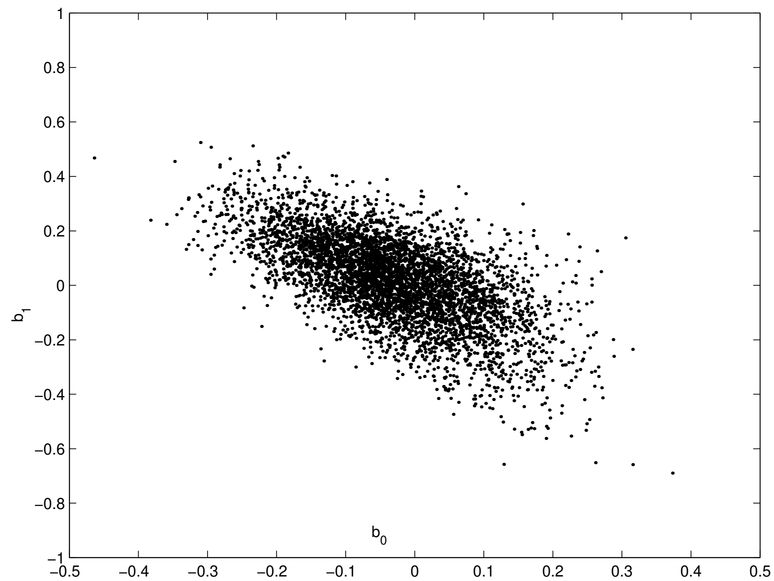

Figure 1: MCMC draws from the posterior distribution of b_0 and b_1.

particular orders of the polynomials is rather ad hoc. Looking ahead, we will estimate the error model using a relatively short sample of the real-time data errors and a longer sample of the RS errors. Therefore, the polynomials describing the dynamics of the real-time data errors are chosen to have smaller order than those for the RS model.

We estimate an empirically relevant "distribution of the uncertainty" using Bayesian estimation methods. In particular, we sample from the posterior distributions of the coefficients of $a, b, c, \ldots, p$ and the posterior distributions of the variances of the shocks u and v using the algorithm of Chib and Greenberg (1994) based on Markov Chain Monte Carlo simulations. We assume diffuse priors for all parameters and obtain six thousand draws from the posterior distribution, dropping the first thousand draws to ensure convergence. In Figure 1 we show the MCMC draws from the joint posterior density of the coefficients b_0 and b_1 on the zero's and the first degree of L respectively in the polynomial $b(L)$. These parameters can roughly be interpreted as measuring the error of the RS model's estimates of the effect of the output gap on inflation. The picture demonstrates that the RS model does a fairly good job in assessing the size of the effect of a one time change in the output gap on inflation (as most of the points on the graph are near the origin). However, there exist some chances that the effect is either more spread out over the time or, vice versa, that the initial response of inflation overshoots its long run level. Averaging the draws from the posterior, we can obtain the point estimates $\hat{a}, \hat{b}, \hat{c}, ..., \hat{p}$ of the parameters of our error model. We will need these estimates to calibrate the non-parametric uncertainty restrictions that we now discuss.

Clearly, restricting the polynomials to be of this specific order may rule out some plausible deviations from the reference model. Such an undesirable restrictiveness, together with the absence of clear rules for determining the orders of the lag polynomials, calls for an alternative, non-parametric description of the uncertainty. For such a description, we allow the polynomials $a(L),...,p(L)$ to be of infinite order. We interpret these polynomials as general causal linear filters having absolutely summable coefficients, that is we assume:

$$\begin{aligned} a(L) &= \sum_{j=0}^{\infty} a_j L^j, \text{ where } \sum_{j=0}^{\infty} |a_j| < \infty \\ b(L) &= \sum_{j=0}^{\infty} b_j L^j, \text{ where } \sum_{j=0}^{\infty} |b_j| < \infty, \end{aligned}$$

and so on.[8]

In general, any linear filter $x(L)$ with absolutely summable coefficients is uniquely determined by the Fourier transform of its coefficients, called the transfer function of the filter:

$$\Gamma_x(\omega) = \sum_{j=0}^{\infty} x_j e^{-i\omega j}. \tag{15}$$

We specify the model uncertainty restrictions in terms of restrictions on the transfer functions of the filters $a(L),...,p(L)$ as follows. For each frequency ω, we require that:

$$\left|\Gamma_a(\omega) - \hat{\Gamma}_a(\omega)\right| < W_a(\omega), \ldots, \left|\Gamma_p(\omega) - \hat{\Gamma}_p(\omega)\right| < W_p(\omega) \tag{16}$$

where $\hat{\Gamma}_i(\omega)$ and $W_i(\omega)$ are some complex-valued and positive real-valued functions of frequency, respectively. We interpret $\hat{\Gamma}_i(\omega)$ as our best guess about the value of the transfer function $\Gamma_i(\omega)$ and $W_i(\omega)$ as a frequency-dependent parameter regulating the size of our uncertainty about $\Gamma_i(\omega)$. Geometrically, the inequalities (16) restrict possible values of the transfer functions $\Gamma_i(\omega)$ to lie in circles in the complex plane centered at $\hat{\Gamma}_i(\omega)$ and having radius $W_i(\omega)$.

The model uncertainty described by the inequalities (16) takes a form of the deterministic set of models alternative to the reference model. Such a set can be made small if the weights W_i are chosen to be small. Indeed, the uncertainty set can be reduced to a singleton if $W_i = 0$. On the contrary, if the W_i are large, then the set is big, and therefore the amount of uncertainty about the reference model is large.

To calibrate our non-parametric description of the uncertainty to an empirically relevant size, we use the following strategy. At each frequency point ω, a rough idea about the possible values of the transfer functions $\Gamma_a(\omega), \ldots, \Gamma_p(\omega)$ can be obtained by plotting a cloud of the MCMC draws of the parametric versions of $a(z),...,p(z)$ evaluated at $e^{-i\omega}$. Therefore, we define our best guesses about the transfer functions at that frequency as:

$$\hat{\Gamma}_a(\omega) = \hat{a}(e^{-i\omega}), \ldots, \hat{\Gamma}_p(\omega) = \hat{p}(e^{-i\omega})$$

[8] The requirement of absolute summability of the filters' coefficients is not really necessary for our analysis. For the results in the rest of the paper to remain valid it is enough to assume that the linear filter preserve the stationarity of inputs. However, the absolute summability is a standard requirement (see, for example, Priestley (1981) ch. 4), and we keep it here.

where $\hat{a}, \ldots, \hat{p}$ are the point estimates of the parametric specification of the polynomials defined earlier in this section. Next, we calibrate $W_a(\omega), \ldots, W_p(\omega)$ so that the circles in the complex plane with centers $\hat{\Gamma}_a(\omega), \ldots, \hat{\Gamma}_p(\omega)$ and radiuses $W_a(\omega), \ldots, W_p(\omega)$ include 50% of the MCMC draws of (the parametric versions of) $a(e^{i\omega}), \ldots, p(e^{i\omega})$. The 50% cutoff value is arbitrary and can be adjusted, but we choose it to focus solely on plausible values of model uncertainty.

Note that a specific choice of transfer functions satisfying (16) may be very different from the sampled (parametric) transfer functions. In particular, although the frequency-by-frequency analysis has a cutoff value of 50%, any resulting filter pieced together across frequencies may have a much smaller likelihood of being observed. Therefore by allowing these remote possibilities, there is a clear chance that we will overestimate the amount of uncertainty. However this method does provide a tractable, implementable way of capturing model uncertainty without imposing a great deal of *a priori* structure on the dynamics of the possible models. This generality is a benefit of the approach, and is absent from the parametric case we considered above.

To summarize, the greater generality of the above non-parametric description of the uncertainty comes with two big costs. First, a probabilistic description of uncertainty is substituted by a deterministic description. Second, the deterministic uncertainty set may include some irrelevant models because the calibration procedure proposed above is too crude. The latter cost can be reduced by introducing more careful calibration techniques which is an important topic left for the future research. In the next section, we show how to the use of our measures of uncertainty to set policy.

5 The Robustness of Simple Policy Rules

In the previous section we constructed both parametric and nonparametric model uncertainty sets for the RS model. We now use Bayesian and minimax techniques to analyze the robustness of Taylor-type rules, and we develop Taylor-type rules which are optimal in presence of the estimated uncertainty.

5.1 Bayesian Analysis

In this section, we numerically solve the Bayesian problem (5), using our estimates of the parametric uncertainty. Before proceeding, we must address the issue of the loss function. Since we do not put restrictions on our priors, the posterior distribution of the coefficients does not have finite support. Moreover, in our estimates we typically find non-negligible probability that the system will by dynamically unstable. Therefore if we use the typical quadratic loss (as in RS), non-zero mass will be assigned to infinite loss and any rule will correspond to infinite Bayesian risk.

One solution to this problem is to restrict our priors to rule out instability and infinite losses. Another solution is to make the loss function bounded from above. We choose the second solution. Clearly, the standard quadratic loss functions are only justified as a local approximation of the true, non-quadratic loss (see Woodford (2002) for example, who however derives a slightly different loss function). Thus there is danger in extrapolating too far away from the mean, and it is not clear that the same loss functions are relevant in extremely bad times. Moreover, bounded utility functions and losses help to avoid the so-called St. Petersburg paradox in which individuals would risk all of their wealth on a repeated coin toss lottery (see Mas-Colell, Whinston, and Green (1995) for a discussion).

We choose the loss to be:

$$\overline{L}_t = \min\left(|\pi_t|, 25\right)^2 + \min\left(|y_t|, 25\right)^2 + \min\left(|i_t - i_{t-1}|/\sqrt{2}, 25\right)^2.$$

This states that all situations in which the absolute value of inflation or the output gap are greater than 25% or interest rates change by more than $25\sqrt{2}\% \approx 35\%$ are ranked equally. This choice, which gives an upper bound on the losses of $3 \times 625 = 1875$, is clearly arbitrary. However our results did not depend much on the precise values chosen.

First, we compute the Bayesian risk for Taylor type policy rules:

$$i_t = g_\pi \bar{\pi}_t^* + g_y y_t^*$$

where $\bar{\pi}_t^*$ is a four quarter average of the real-time data on inflation and y_t^* is the real time data on the output gap.[9] We make our computations on a grid for g_π going from 1.25 to 4 in increments of 0.25 and for g_y running from 0.25 to 3 in increments of 0.25. By experimenting with the grid size, we found that this region contains most of the solutions. We refer to different policy rules by the ordered pairs (g_π, g_y).

A surface plot of the estimated Bayesian risk is shown in Figure 2. For all policy rules studied, the risk is extremely high, ranging from about 800 to 1800. Note that, as explained earlier, the maximum possible risk is 1875. Such a high level of risk is caused by the fact that many deviations from the RS model drawn by Monte Carlo method turn out to be dynamically unstable. For example, under the famous Taylor rule of (1.5,0.5), which belongs to our grid, 50% of the MCMC draws result in dynamic instability!

The optimal Bayesian rule is outside the region covered by our grid. The rule corresponding to the smallest Bayesian risk among those rules included is $(1.25, 0.5)$. This finding is consistent with Brainard (1967) intuition that the introduction of uncertainty should make policy makers cautious, as the optimal rule under no uncertainty is $(2.5, 1.4)$. Thus uncertainty results in attenuation. From a practical point of view, however, these results look depressing. It seems that the amount of uncertainty that policy makers have to deal with is so large that even the best solutions remain extremely risky. The lower bound of the risk for the rules on our grid is 800. This may, for example, correspond to expected standard deviations of inflation and the output gap close to 20%. So at the very "best", we can expect that "every-day events" will not look much worse than the Great Depression!

[9]For each MCMC draw, we check whether the corresponding deviation from the reference model is stable or not. If it is unstable, then we associate maximum loss of 1875 with such a deviation. In cases when the deviation is stable, we compute the variance-covariance matrix of the stationary normal distribution for $(\pi_t, y_t, (i_t - i_{t-1})/\sqrt{2})'$. Then we simulate 10,000 draws from this stationary distribution and compute the average loss over these draws. We take this as our estimate for the risk.

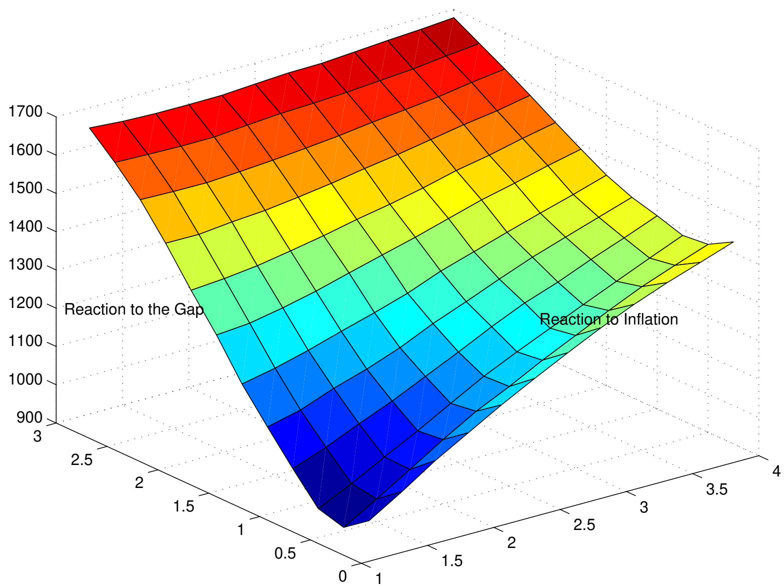

Figure 2: Estimated Bayesian risk for different policy rules under a diffuse prior.

Our results may, however, be interpreted in a less apocalyptic way. They may indicate, for example, that mechanically following a simple Taylor rule is a very dangerous exercise. Instead, potentially much more complex rules should be followed in an adaptive and cautious way. Policy makers do not commit to a rule once-and-for-all, but instead would be likely to abandon a rule leading to bad outcomes. Another reason to be less pessimistic is that policymakers may have informative priors about the uncertainty they face. Our estimates imply that without imposing stronger prior beliefs, the data suggests that instability is a likely outcome. But policymakers likely know much more than attribute to them. For example, one may *a priori* believe that most of the plausible deviations from the RS model will *not* result in instability if policymakers follow a rule which closely approximates their observed historical behavior.

To explore this possibility, we compute another sample of MCMC draws assuming informative priors on the coefficients of the polynomials a, b, d, f, and g.[10] Recall that these polynomials correspond to the effects of the macroeconomic variables in the Phillips curve and IS equation errors. Later we refer to these polynomials as the "slope" terms. The priors were calibrated so that about 90% of the draws from these prior distributions resulted in dynamic stability for the economy governed by the famous Taylor rule. Such a calibration of the prior distributions changes our posterior distribution drastically. This is clearly illus-

[10] Changing the prior for the polynomials c, h, m, and p, which recall relate to the serial correlation of the driving shocks, has little effect on the stability of deviations. We choose not to impose informative priors on the polynomials k and n describing the news content of the real-time data. Doing so, however, does not change our results much.

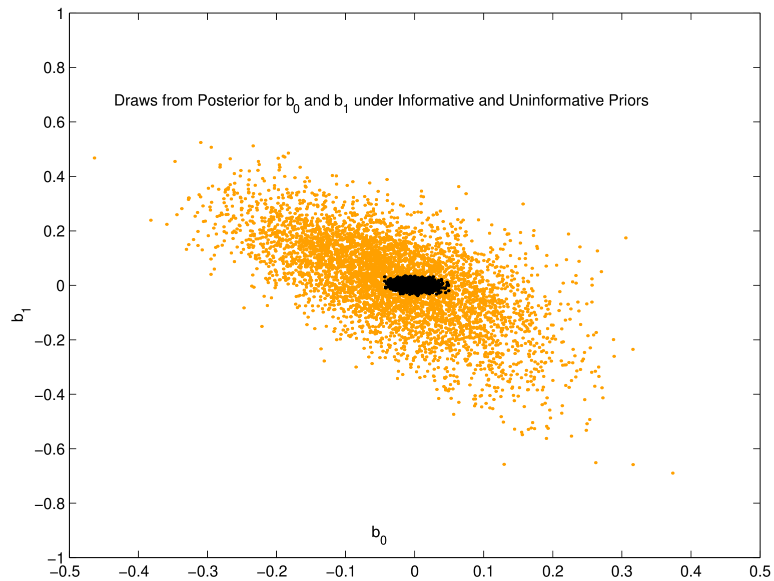

Figure 3: Draws from the posterior distribution under a diffuse (orange points) and informative (black points) prior.

trated in Figure 3 which superimposes the MCMC draws from the posterior distribution of the coefficients b_0 and b_1 corresponding to the uninformative and informative priors. The informative priors lead to enormous shrinkage in the posterior distribution, as the draws are now in a much tighter cloud around zero.

Under the informative priors, Figure 4 shows a surface plot of the *inverse* of the Bayesian risk for the policy rules in our grid. We report the inverted risk because a few rules in the grid produce extremely large risk, whereas the majority of the rules correspond to small risk. Such an unbalanced situation distorts the graph so that it is easier interpreted when inverted. The minimal risk of 16.6 is attained by the rule $(3.25, 1.5)$. However the risk is nearly flat over a wide range, resulting in a large region of rules with comparable risk. For example, the rule $(2.5, 1.5)$ that is the closest on our grid to the optimal rule under no uncertainty corresponds to risk of 18.2, only a 10% degradation in performance from the minimum. Our findings therefore lend partial support to the results in Rudebusch (2001), who shows that robustness to many different kinds of uncertainty does not result in a substantial attenuation of the policy response. In fact, our robust optimal rule with informative priors is more aggressive than the optimal rule in the absence of uncertainty, but the difference in losses are slight.

5.2 Minimax Analysis

The Bayesian analysis in the previous section is limited to the parametric model of uncertainty. We now analyze the robustness of policy rules under the much less restrictive, nonparametric description of uncertainty we discussed above. However, as we noted there,

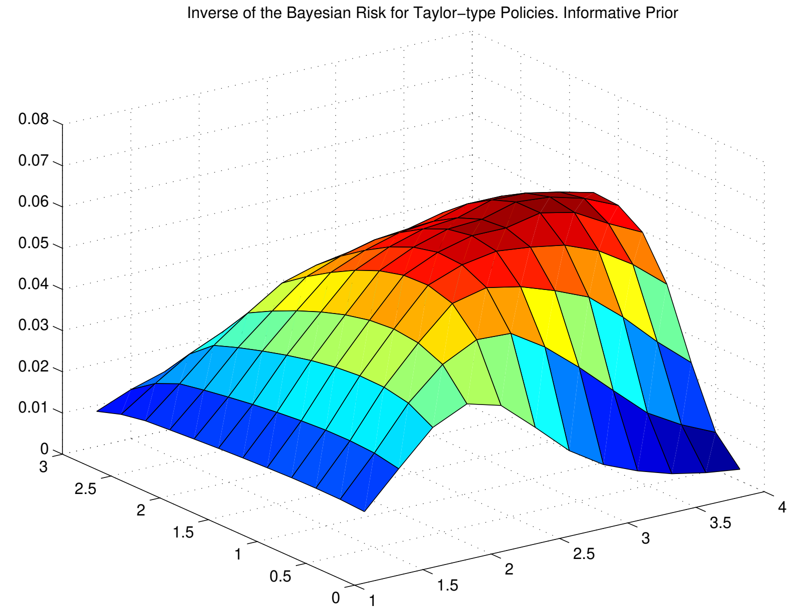

Figure 4: Inverse of the estimated Bayesian risk for different policy rules under an informative prior.

we do not have a probability distribution over this nonparametric set. Therefore in this section we use a minimax approach (6), minimizing the worst case loss.[11] We estimate the worst case loss for each policy rule from above using the algorithm described in Chapter 6 of Paganini (1996). Unfortunately, there are no theoretical guarantees that the upper bound on the worst possible loss that we compute is tight. However, our experience with relatively simple uncertainty descriptions suggests that the gap between the upper bound and the actual worst possible loss is not very large. Besides, the bound has an appealing interpretation of the exact worst possible loss under slowly time-varying uncertainty and a special noise structure (see Paganini (1996)).

5.2.1 Uncertainty at All Frequencies

From the preceding analysis, it is clear that the nonparametric description of uncertainty must use an informative prior. Nonparametric model uncertainty calibrated using MCMC draws corresponding to an uninformative prior will simply be too large to produce interesting results. Since so many of the draws result in instability, the worst case loss will clearly be maximal. We checked that for such a calibration, all the policy rules on our grid did indeed correspond to dynamic instability in the worst case. We therefore use the MCMC draws corresponding to the informative prior to calibrate the uncertainty.

Figure 5 shows the *inverse* of the worst possible losses for the Taylor-type rules on our grid. Qualitatively, the graph is similar to one that we obtained using Bayesian analysis in

[11]For the minimax analysis that follows we assume the untruncated quadratic loss function (11).

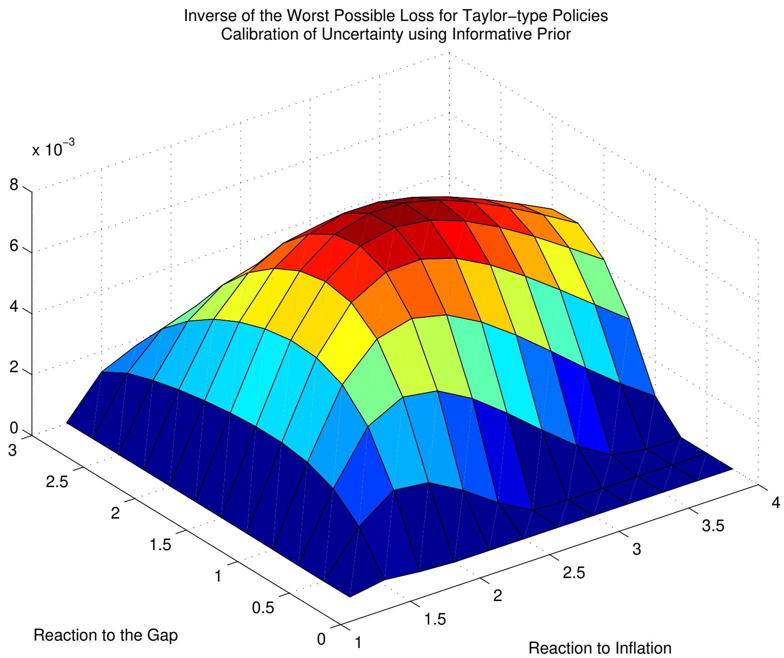

Figure 5: Inverse of the worst case losses for different policy rules.

Figure (4), with a slightly different peak location. The minimal worst possible loss is 130.6 and it is attained by the policy rule (2.5,1.5), which is indistinguishable on our grid from the optimal rule under no uncertainty. This shows that the conventional optimal Taylor-type rule, formulated in the absence of model uncertainty, possesses strong robustness properties. Even though we incorporate an informative prior, limiting perturbations which result in instability, we still allow for a broad range of perturbations from the reference model. The optimal rule under no uncertainty effectively deals with these perturbations, and results in good performance under both the reference model (as it was designed to do) and under the worst case model. Further, note that we do not find any attenuation of policy in response to uncertainty.

However, these results should be treated with some caution. In the next section, we discuss why the nonparametric description of uncertainty we employ here may not be capturing the uncertainty relevant for policy.

5.2.2 Uncertainty at Business Cycle Frequencies

There are two reasons why we may want to limit our analysis of uncertainty to business cycle frequencies. One is theoretical, reflecting the nature of our reference model as a model of business cycle fluctuations. The other is more technical, relating to how our parametric and nonparametric uncertainty descriptions differ in their treatment of high and low frequency perturbations. We now describe each of these in turn.

One of the features of minimax analysis is that it provides a simple method of diagnosing

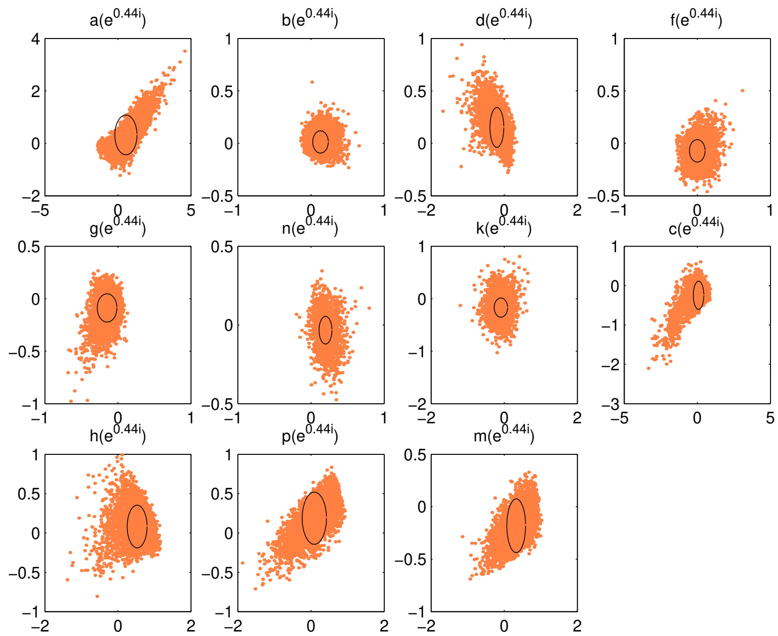

Figure 6: MCMC draws (points) and our nonparametric uncertainty bound (circles) at a business cycle frequency.

possible defects in the model. By inspecting the worst case deviations from the RS model under different policy rules, we found that the biggest losses result from the deviations at very low frequencies. More precisely, from (16) the biggest losses are inflicted by larger differences between $\Gamma_i(\omega)$ and $\hat{\Gamma}_i(\omega)$ at frequencies ω close to zero. Thus the most damaging perturbations represent deviations in some of the very long-run properties of the model.

However, we feel that changing the low frequency properties of the RS model is pushing the model too far. After all, the model is designed to explain business cycle frequency fluctuations and not to describe long-run phenomena. The model is estimated based on demeaned quarterly data, and we make no effort to model the means or any possible changes in the means over time. Additionally, just as the loss function is best viewed as a quadratic approximation, the reference model is best viewed as a linear approximation to a nonlinear true model. The linearization is much more appropriate for business cycle fluctuations than for deviations which may push the model away from its mean for extended periods of time. A more fully developed model, for example incorporating growth or explicitly modeling time variation in the data, would be necessary to seriously consider long-run issues.

Furthermore, there are some features of our nonparametric methods which increase our measurement of uncertainty at very low and very high frequencies. Recall that (16) defines bounds on the transfer functions which can be viewed as describing a circle in the complex plane. We calibrate the size of uncertainty by insuring that 50% of the MCMC draws lie within each circle. Thus the circles provide an approximation of a level set of the empirical distribution of the MCMC draws. This approximation is good if the empirical distribution

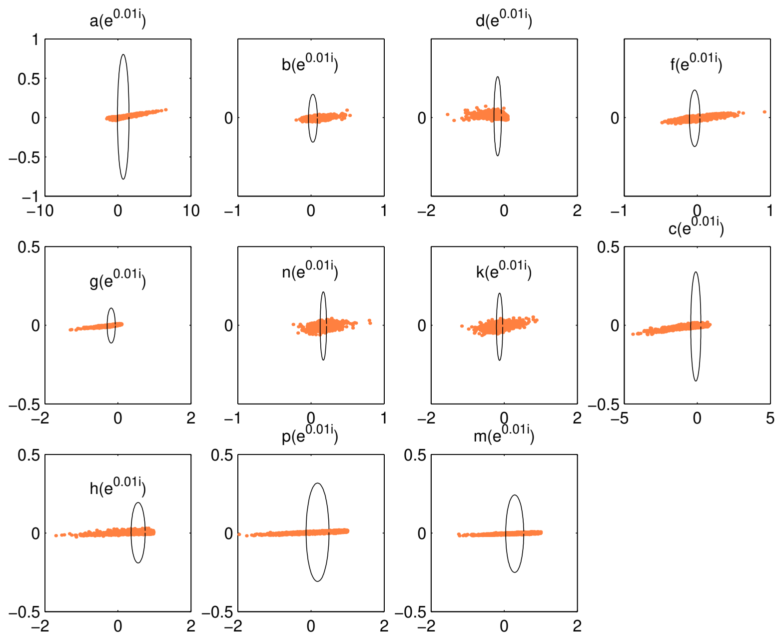

Figure 7: MCMC draws (points) and our nonparametric uncertainty bound (circles) at a low frequency.

is nearly "circular". However the quality of this approximation decreases substantially if the empirical distribution of MCMC draws is not circular. For intermediate frequencies, the approximation seems to be quite good. An illustration of this for a business cycle frequency is shown in Figure 6. The figure plots the cloud of MCMC draws for each $\Gamma_i(\omega)$, with $i = a, \ldots, p$, associated with the parametric description, along with the circle which contains the possible $\Gamma_i(\omega)$ for the nonparametric description. The nonparametric descriptions seem appropriate in this case.

However, if we look at very low frequencies the correspondence breaks down. Recall that our MCMC draws are based on low (second or fourth) order lag polynomials. At very low (and very high) frequencies the imaginary parts of the transfer functions as in (15) disappear. For example, $p(L)$ is a second order polynomial, so its transfer function is:

$$\begin{aligned} p(\omega) &= p_0 + p_1 e^{-i\omega} + p_2 e^{-2i\omega} \\ &= p_0 + p_1(\cos\omega - i\sin\omega) + p_2(\cos 2\omega - i\sin 2\omega). \end{aligned}$$

Clearly for ω very near zero, both $\sin\omega$ and $\sin 2\omega$ will be also be very near zero, so the imaginary part will be negligible. Only very high order polynomials have significant imaginary parts at low frequencies. An illustration of this is shown in Figure 7, which is similar to Figure 6, except now for a frequency near zero. This clearly shows that in this case, our approximation of the clouds of MCMC draws by a circle in the complex plane is not accurate. Exactly the same logic applies for very high frequencies. Thus our calibration of uncertainty is most accurate for intermediate, business cycle frequencies.

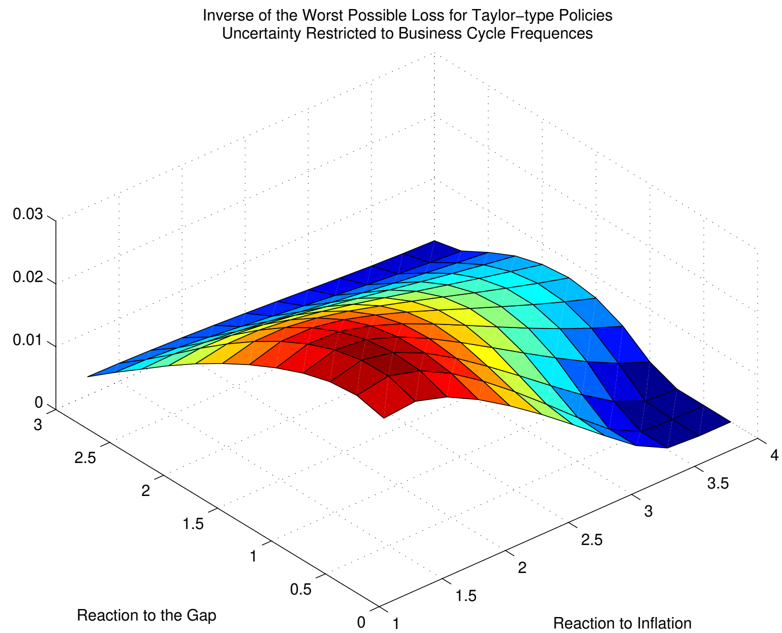

Figure 8: Inverse of the worst case losses for business cycle frequency uncertainty.

Based on these theoretical and technical reasons, we therefore extend our analysis by truncating uncertainty to be zero except at the business cycle frequencies. Technically, this means that we require the weights $W_a, \ldots, W_p$ from (16) to be equal to zero for any frequency outside the range $\left[\frac{2\pi}{32}, \frac{2\pi}{6}\right]$. Figure 8 reports our results, which differs dramatically from our results with uncertainty at all frequencies as in Figure 5. Now, the best rules are those with a relatively small reaction to both inflation and the output gap. The minimal worst possible loss is attained at (1.5,0.5) and equal to 37.0, which is substantially less than the value of 130.6 for uncertainty at all frequencies. This result is intriguing because the optimal minimax rule coincides with the famous Taylor rule.

One possible explanation for the attenuation of the optimal interest rate reaction to inflation and the output gap under the "business cycle uncertainty" is as follows. Robustness against bad scenarios at the business cycle frequencies comes at the cost of downweighting low frequency movements, even relative to the no uncertainty case. As we saw above, the optimal rule under no uncertainty effectively deals with low frequency perturbations. Thus, it is aggressive enough to stop inflation before it takes hold. However with business cycles uncertainty, policy makers become more concerned with the performance of policy over the business cycle horizon. Concerns about long-term increases in inflation become less important than counter-cyclical stabilization. This leads policymakers to choose attenuated policy rules.

Uncertainty Channel	Optimal Bayesian Rule	Optimal Minimax Rule	Bayesian Risk	Worst Loss
No uncertainty	(2.4,1.5)	(2.4,1.5)	13.1	13.1
Slope of Phillips curve	(2.4,1.2)	(2.7,1.8)	17.5	72.6
Shock to Phillips curve	(2.4,1.3)	(2.3,1.4)	14.0	31.9
Slope of IS curve	(1.5,1.0)	(2.5,1.4)	13.0	103.2
Shock to IS curve	(2.6,1.5)	(2.5,1.6)	17.1	40.9
News in revisions to π^*	(2.9,1.3)	(3.4,1.8)	14.8	47.4
Noise in revisions to π^*	(2.4,1.3)	(2.3,1.3)	14.9	31.6
News in revisions to y^*	(2.4,1.5)	(2.3,1.4)	14.7	37.1
Noise in revisions to y^*	(2.4,1.3)	(2.3,1.1)	15.0	32.8

Table 1: The coefficients of the robust optimal Bayesian and minimax Taylor-type rules and corresponding Bayesian risk and worst possible losses. Diffuse priors on specific uncertainty channels.

5.3 Analysis of Distinct Uncertainty Channels

One of the purposes of this paper was to bring together different studies on robustness which focus on special forms of uncertainty. Therefore, in previous sections, we combined many different sources of uncertainty into one encompassing structure. In this section, we analyze the different components of our uncertainty description taken separately. A goal of this analysis is to find out which components of the uncertainty have the largest effects on policy and on losses. This can be useful for researchers working with similar Phillips-curve-IS-type models, by showing which parts of the specification require the most attention.

In this section, we analyze uncertainty about the "slope" and shock of the Phillips curve, the "slope" and shock of the IS curve, the news and noise components of inflation revisions, and the news and noise components of the output gap revisions, each taken separately. To do this, we simulate different MCMC samples, each 6000 draws long, corresponding to a zero prior on all uncertainty except the chosen uncertainty channel. Thus, for example, the MCMC sample corresponding to uncertainty about the "slope" of the Phillips curve is taken under a zero prior on all the parameters of our error model except the coefficients of $b(L)$. The MCMC sample corresponding to uncertainty about the "slope" of the IS curve is taken under a zero prior on all of the polynomials but $f(L)$ and $g(L)$ and so on. The prior on the chosen uncertainty channel is taken to be diffuse.

After simulating the MCMC samples, we calibrate the nonparametric description of uncertainty as outlined in Section 4, and then perform both Bayesian and minimax analysis of the robustness of Taylor-type policy rules. Instead of analyzing rules on a grid as we previously did, we now use standard MATLAB optimization package routines to compute optimally robust Bayesian and minimax rules. Our results are reported in Table 1.

As one would expect from our previous analysis, the majority of the robust optimal rules are very close to the benchmark optimal rule under no uncertainty. For most of the uncertainty channels, the robust response to the output gap is slightly less than the no uncertainty response. Additionally, for the most of the uncertainty channels the optimal minimax rules are more aggressive than the optimal Bayesian ones. A couple of individual channels stand out. First, the optimal Bayesian response to uncertainty about the slope of the IS curve leads to a considerably attenuated policy rule. Additionally, the optimal minimax rule corresponding to the uncertainty about the news component in the revisions of the real-time inflation data is quite aggressive. The optimal Bayesian rule for the same uncertainty channel also leads to a relatively aggressive response to inflation. We discuss these findings below.

The effect of the uncertainty on the losses differs significantly in the Bayesian and minimax analysis. The Bayesian risk does not vary by more than 35% relative to the benchmark loss in the absence of uncertainty. Perturbations of the reference model carrying the most losses correspond to the uncertainty about the slope of the Phillips curve and the shock to the IS curve. Interestingly, the Bayesian risk corresponding to the uncertainty about the slope of the IS curve is slightly *less* than the benchmark loss. This happens because most of the empirically plausible parametric perturbations of the reference model associated with this uncertainty turn out to be favorable to policymakers, meaning that the perturbed model can be more efficiently controlled than the reference model. This also describes why the robust Bayesian policy rule is attenuated, as not much variation in interest rates is needed to ensure good economic performance.

For the minimax analysis, the worst possible losses vary substantially across different uncertainty channels. The least harmful uncertainty is that about the shock to the Phillips curve and about noise in the revisions of the real-time data. The associated worst possible losses for these channels are about 2.5 times larger than the benchmark loss. The most damaging worst possible loss, which is about eight times larger than the benchmark, corresponds to the nonparametric uncertainty about the slope of the IS curve. This contrasts with the small Bayesian risk under the parametric uncertainty of this same channel. This suggests that the difference may be driven by an unlikely, but extremely damaging, perturbation to the model. As was discussed earlier, at very low and very high frequencies our nonparametric description overstates the amount of uncertainty relative to the parametric case. This is due to the fact that the nonparametric description does not restrict the number of lags in the reference model's perturbations. Therefore, a potential explanation of the contrast is that the worst possible perturbations of the slope of the IS curve are structured so that some distant lags of the interest rate and/or inflation are added to the IS equation and the perturbations are particularly damaging at very low or very high frequencies. This hypothesis finds strong support from the analysis of particular uncertainty channels truncated to the business cycles frequencies that we carry out below. Table 2 reports the results of the analysis.

For convenience of comparison, we include in the table the minimax results for uncertainty at all frequencies and for uncertainty truncated to business cycle frequencies. We see that for the majority of the uncertainty channels, the optimal minimax rules associated with the business-cycle uncertainty are significantly less aggressive than those associated with all-frequencies uncertainty. These optimal minimax rules resemble the Bayesian rules reported in Table 1, and in many cases they are less aggressive than these rules. Now we find that the most attenuated minimax rule corresponds to the uncertainty about the slope of the IS

Uncertainty Channel	All frequencies	BC frequencies only	Worst Loss all uncertainty	Worst Loss BC frequencies
No uncertainty	(2.4,1.5)	(2.4,1.5)	13.1	13.1
Slope of Phillips curve	(2.7,1.8)	(2.2,1.1)	72.6	31.9
Shock to Phillips curve	(2.3,1.4)	(1.3,1.2)	31.9	27.9
Slope of IS curve	(2.5,1.4)	(1.1,0.8)	103.2	28.4
Shock to IS curve	(2.5,1.6)	(2.4,1.5)	40.9	36
News in revisions of π^*	(3.4,1.8)	(2.6,1.2)	47.4	30.3
Noise in revisions of π^*	(2.3,1.3)	(2.3,1.3)	31.6	30.5
News in revisions of y^*	(2.3,1.4)	(2.3,1.4)	37.1	30.1
Noise in revisions of y^*	(2.3,1.1)	(2.4,1.3)	32.8	30

Table 2: The coefficients of the robust minimax Taylor-type rules and corresponding worst possible losses. Uncertainty at all frequencies vs. uncertainty at business cycle frequencies.

curve, in accordance with the Bayesian result.

Also analogous to the Bayesian results, we find than now the worst possible losses for the different uncertainty channels do not vary much. In particular, some of the channels experience a dramatic reduction in the worst possible losses relative to uncertainty at all frequencies. For example, the worst possible loss corresponding to the uncertainty about the slope of the IS curve is 3.5 times smaller than before. This supports our proposed explanation for the discrepancy in the minimax and Bayesian results. In fact, the discrepancy disappears when the uncertainty is truncated to the business cycle frequencies. In correspondence with the Bayesian analysis, the most significant worst case losses are associated with uncertainty about the slope of the Phillips curve and the uncertainty about the shock to the IS curve. The least troubling perturbations are those corresponding to uncertainty about the shock to the Phillips curve and the slope of the IS curve. Interestingly, the losses corresponding to the real-time data uncertainty channels are relatively damaging, and do not change much across frequencies. This suggests that real-time data uncertainty is important both for monitoring cyclical fluctuations and for more long-run issues, such as detecting a change in the trend rate of productivity growth.

An interpretation of the effect of truncation of the uncertainty to business cycle frequencies was discussed earlier in the paper. The worst case scenarios among all frequencies often correspond to very low frequency perturbations in which inflation grows out of control. The aggressiveness of policy rules essentially varies with how much weight they put on low frequencies. The optimal rule weights all frequencies equally, while the robust rule considering uncertainty at all frequencies pays a lot of attention to the potentially damaging low frequencies. However the rule robust to business cycle uncertainty increases the attention paid to cyclical fluctuations, and so downweights the low frequencies. The robust rules at business cycle frequencies give a large role to providing counter-cyclical stabilization, and are thus less aggressive in their responses.

6 Conclusion

In this paper we analyzed the effects of uncertainty on monetary policy decisions. We considered three different types of uncertainty: uncertainty about the specification of a reference model, uncertainty about the serial correlation of noise, and uncertainty about data quality. We argued that different specifications of uncertainty may have significantly different implications for monetary policy. Further, uncertainty which enters at different frequencies may have substantially different effects. It is therefore necessary to model the uncertainty itself and try to carefully estimate or calibrate the uncertainty model.

We introduced a systematic approach to the formulation of uncertainty relevant for policy making based on the Model Error Modeling literature. As the name suggests, this approach describes the uncertainty about an estimated reference model by building models of the model's errors. Throughout the paper we focused on a small macroeconometric model of the US economy proposed and estimated by Rudebusch and Svensson (1999). We formulated models for the errors of the RS model, focusing on the aspects of uncertainty that are relevant for that model. We then implemented both parametric and nonparametric descriptions of uncertainty for the model, and used them to design robust monetary policy rules of the Taylor type.

Our parametric description of uncertainty assumed that the model errors could be fit by simple low-order lag polynomials. We estimated these models using Bayesian methods, obtaining a distribution over the potential alternative models. For use in policy, this led naturally to Bayesian optimization methods to determine policy rules. The robust rule was defined as the policy rule that minimizes Bayesian risk over the distribution of potentially true models.

For the nonparametric description of uncertainty, we did not restrict the order of the error models, but instead calibrated the size of the uncertainty set in the frequency domain. In particular, we used our parametric estimates so that at each frequency half of the draws from the posterior distribution were in our chosen sets. Without having a distribution over this large (but empirically plausible) class of alternative models, for policy purposes we focused on minimax optimization methods. In this case, the robust rule was defined as the policy rule that minimizes losses under the worst possible scenario consistent with the uncertainty description.

Without imposing much prior knowledge, we found that the amount of uncertainty in both our parametric and nonparametric specifications was too large to produce sensible recommendations for policymakers. With uninformative prior beliefs, dynamic instability is a likely outcome for many policy rules. However this result may not be empirically plausible, particularly for policy which does not strongly deviate from the past. Therefore we then imposed stronger prior beliefs to downweight the likelihood of instability. We found that the resulting policy rules in both the Bayesian and minimax cases were close to the optimal rule in the absence of uncertainty. All of these rules are relatively aggressive, especially in comparison to directly estimated policy rules.

Our analysis also showed that very low frequency perturbations have the most impact on policy. The aggressiveness we found in policy rules is driven largely by the low frequency properties of the model. Since our baseline model is essentially a model of short-run fluctuations, we felt that it was extreme to ask it to accommodate very low frequency perturbations. Therefore we recalculated our results by restricting attention to business cycle frequencies. In these cases we found that instead of reacting aggressively, our policy rules were more attenuated than in the absence of uncertainty. Under some specifications, our results were quite close to the policy rules that have been directly estimated. We also analyzed separately the effects uncertainty of each of the different channels. We found that in the minimax case low frequency perturbations can lead to large changes in losses and policy rules across channels. But focusing on business cycle frequencies or weighting across models using Bayesian methods leads to much less variation.

Many important issues are left for future research. For example, in this paper we left open the question of how to choose a reference model. Additionally, the baseline model that we used was intentionally simple, and completely backward looking. Much more could be done to extend both the baseline model that we analyze and therefore the methods we use, for example to consider forward-looking behavior or unrestricted optimal policy rules. This paper is only a first step in the analysis, but even by focusing on a simple case we find some interesting results. More work remains to be done to accurately measure the uncertainty relevant for policy. This requires even more careful modeling of model uncertainty.

References

Ball, L. (1999). Policy rules for open economies. In J. Taylor (Ed.), *Monetary Policy Rules*, pp. 127–144. Chicago: University of Chicago Press for the NBER.

Blinder, A. (1997). Distinguished lecture on economics in government: What central bankers could learn from academics - and vice versa. *Journal of Economic Perspectives 11*, 3–19.

Brainard, W. (1967). Uncertainty and the effectiveness of policy. *American Economic Review 57*, 411–425.

Chib, S. and E. Greenberg (1994). Bayes inference in regression models with ARMA(p,q) errors. *Journal of Econometrics 64*, 183–206.

Craine, R. (1979). Optimal monetary policy with uncertainty. *Journal of Economic Dynamics and Control 1*, 59–83.

Hansen, L. P. and T. J. Sargent (2002). *Robust Control and Model Uncertainty in Macroeconomics*. Unpublished book manuscript.

Levin, D., V. Wieland, and J. Williams (1999). Robustness of simple monetary policy rules under model uncertainty. In J. Taylor (Ed.), *Monetary Policy Rules*, pp. xxx. Chicago: University of Chicago Press for the NBER.

Ljung, L. (1999). Model validation and model error modeling. Control and Communcations Group Technical Report LiTH-ISY-R-2125, Linkoping University.

Mankiw, N. G. and M. D. Shapiro (1986). News or noise: An analysis of GNP revisions. *Survey of Current-Business 66*, 20–25.

Mas-Colell, A., M. D. Whinston, and J. R. Green (1995). *Microeconomic Theory*. Oxford University Press.

McCallum, B. T. (1988). Robustness properties of a rule for monetary policy. *Carnegie-Rochester Conference Series on Public Policy 29*, 175–203.

Onatski, A. and J. H. Stock (2002, February). Robust monetary policy under model uncertainty in a small model of the US economy. *Macroeconomic Dynamics 6*, 85–110.

Orphanides, A. (2001). Monetary policy rules based on real-time data. *American Economic Review 91*, 964–985.

Paganini, F. (1996). Sets and constraints in the analysis of uncertain systems. Ph.D. thesis, California Institute of Technology.

Priestley, M. B. (1981). *Spectral Analysis and Time Series*. Academic Press.

Rudebusch, G. and L. E. O. Svensson (1999). Policy rules for inflation targeting. In J. Taylor (Ed.), *Monetary Policy Rules*, pp. 203–246. Chicago: University of Chicago Press for the NBER.

Rudebusch, G. D. (2001, May). Is the Fed too timid? Monetary policy in an uncertain world. *Review of Economics and Statistics 83*, 203–217.

Sargent, T. J. (1999). Discussion of "Policy rules for open economies," by Laurence Ball. In J. Taylor (Ed.), *Monetary Policy Rules*, pp. 145–158. Chicago: University of Chicago Press for the NBER.

Söderström, U. (2002). Monetary policy with uncertain parameters. *Scandinavian Journal of Economics 104*, 125–145.

Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy 39*, 195–214.

Woodford, M. D. (2002). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton: Princeton University Press. Forthcoming.