Robust Monetary Policy under Model Uncertainty: Incorporating Rational Expectations.

Alexei Onatski[1], Department of Economics, Columbia University

Abstract

One of the prominent ways to analyze the robustness of monetary policy under model uncertainty consists of the following three steps. First, choose a reference model of the economy. Next, define a set of perturbations around this model, where the set is structured so that the uncertainty is focused on potentially important weaknesses of the reference model. Finally, choose policy so that it works best for the worst model from the set. Previous applications of this approach allowed only for purely backward-looking models. This paper extends the analysis of robustness to models that may include forward-looking components. Empirical part of the paper studies simple policy rules under model, data and shock uncertainty in a small model of the US economy with rational expectations.

Key words: structured uncertainty, rational expectations, robustness.

JEL classification: E52, E58, C61,

I thank James Stock and Glenn Rudebusch for helpful discussions and suggestions.

[1]Department of Economics, Columbia University, 420 West, 118th Street, New York, NY, 10027. Tel.: (212) 854 3685. Fax: (212) 854-8059. E-mail: ao2027@columbia.edu.

1. Introduction

The question of robustness of monetary policy rules to model uncertainty has recently received much attention, both from practitioners and academic researchers. The recent Asian crisis, the steady decline of the natural unemployment rate in the US, and uncertainty about the workings of the economy in the new European environment have all contributed to this interest. Alan Greenspan (2000) recently acknowledged the challenges that model uncertainty introduces for monetary policy making. In particular he said:

> "Policymakers... inevitably construct working hypotheses or models of the way our economies work. Of necessity, these models are a major simplification of the many forces that govern the functioning of our system at any point in time. Obviously, to the extent that these constructs ... fail to capture critical factors driving economic expansion or contraction, conclusions drawn from their application will be off the mark."

How should a central bank conduct monetary policy in a situation when the existent models of the economy may be misspecified, imperfectly estimated and subject to structural breaks? This paper tries to answer this question. It follows the logic of analysis proposed recently by Onatski and Stock (1999). This analysis consists of the following three steps. First, choose a reference model that approximates the true workings of the economy. Next, define a non-parametric space of perturbations around this model, where the space is structured so that the uncertainty is focused on potentially important weaknesses of the reference model. So, for example, if a reference model states that the real interest rate affects the real output through a policy multiplier, a possible perturbation would add some "long and variable lags" to the effect. If, in addition, the reference model is known to result from a log-linearization procedure then the space of perturbations may be augmented to include non-linear perturbations consistent with the accuracy of the linearization of particular dynamic channels, etc. Finally, compute an index of robustness for a given policy rule. The index measures a distance between the reference model and those perturbations that result in the dynamic instability under the policy rule studied. The larger the index the more robust the rule.

Unfortunately, the practical application of this method was limited in an important way. Previous analysis allowed only for purely backward-looking reference models and uncertainty sets. This seriously impaired the analysis at several levels. Most fundamentally, restricting attention to backward-looking models rules out models with rational or partially rational expectations. A prominent example of such a model is the Clarida, Gali and Gertler (1999) New Keynesian model, which consists of *purely forward-looking* IS and Phillips-curve-type equations. Second, approximating model uncertainty by a strictly backward-looking version is not innocuous for certain policy rules. Thus, for example, the much discussed nominal income rules can result in dynamic instability in the purely backward-looking models of Ball (1999) and Svensson (1999). As shown by McCallum (1997), however, the instability disappears for a variety of alternative more forward-looking (and more plausible) model specifications. Finally, the backward-looking analysis is ill-suited for studying such phenomena as the indeterminacy of equilibrium and forecast-based rules. The latter has received a lot of attention in such inflation-targeting central banks as the Bank of England, the Bank of Canada, and the Reserve Bank of Australia. As was shown by Levin, Wieland, and Williams (1999), the equilibrium indeterminacy can be an important source of non-robustness for the forecast-based rules (see also Bernanke and Woodford (1997)).

The main methodological contribution of this paper is to solve the above problem, that is, to extend the Onatski-Stock approach to forward-looking reference models and uncertainty sets. Since forward-looking behavior is specific to social as opposed to physical processes, the extension cannot directly draw upon advances in the field of robust control in the engineering literature. Therefore, the ideas of robust stability analysis (see, for example, Dahleh and Diaz-Bobillo (1995)) must be considerably changed to use them for models incorporating rational expectations.[2] I implement this change and apply my theoretical results to analyze policy rules in the context of an empirical New Keynesian model of the US economy described in Rudebusch (2000).

There has been one previous attempt to generalize the approach to forward-looking models. Tetlow and von zur Muehlen (2000) propose to obtain a saddle path backward-looking representation of the reference model controlled by a particular

[2]As Hansen and Sargent (2000a) noted, it is natural to consider private agents sharing policy maker's uncertainty about the true model of economy. Thus, the philosophy of robust analysis suggests that the rational expectation hypothesis should be changed into something that could be called a "robust expectations" hypothesis. However in this paper I assume rational expectations, leaving the ideal of the robust expectations hypothesis for future research.

policy rule, then specify the uncertainty set around the controlled model, and finally compute the index of robustness as in the backward-looking case. One problem with this approach is that the uncertainty is introduced around the controlled model, so it is specific for each policy rule. Therefore, comparison of the degree of robustness for different policy rules is problematic. Another limitation of the approach is that it makes difficult to interpret the possible deviations from the reference model since the parameters of the model solved for rational expectations are highly non-linear functions of the parameters of the original model. Finally, the approach assumes that the private sector believes that the reference model of the economy is true which is not plausible.[3]

In this paper I show how to analyze model uncertainty specified directly in the original forward-looking model before it is controlled by a specific policy rule. This approach makes the economic interpretation of uncertainty easier. It also results in a measure of robustness consistent across different policy rules which is convenient for policy analysis. The private sector is assumed to believe that the right model of the economy is a particular (unknown to a policy maker) model from the model uncertainty set.

The empirical part of the paper addresses three sets of questions. The first set of questions concerns degree of policy activism appropriate under model uncertainty. A number of recent studies (see Sargent (1999), Stock (1999), Giannoni (1999)) found

[3]Tetlow and von zur Muehlen (2000) avoid the problems by considering a special structure of uncertainty that is invariant to solving out the expectations from the reference model. This assumption is made for technical convenience, however, and begs the question of robustness with respect to this assumed in formation structure.

that policy rules robust to model uncertainty must be very aggressive. In contrast, Brainard's (1967) classic analysis states that a policy maker facing model uncertainty must do less than in the certain environment. Previous studies of model uncertainty of the type considered in this paper produced results in between: relatively more robust policy turned out to be more responsive to some economic variables and less responsive to others. How will these results change for forward-looking models?

The second group of questions concerns the recent debate on the stabilization properties of nominal income rules. The nominal income rules are attractive from several perspectives. In particular, such rules imply automatic reaction to two variables of major interest for central bankers: prices and real output. A target for nominal output can serve as a nominal anchor for monetary policy. Reacting to changes in nominal income is similar to reacting to changes in money (under a relatively stable velocity). Finally, the rules do not rely upon output gap estimates, hence they are robust to real-time data uncertainty that has been a focus of many recent studies (see, for example, Orphanides (1999)).[4]

As I briefly mentioned before, stabilization properties of nominal income rules were seriously questioned by Ball (1999) and Svensson (1999). Because the unfavorable results of these authors turn out to be model-specific, McCallum (1999) suggested that the relative quality of the nominal income rules and other monetary policy rules be assessed by a wide-scale analysis, including, in particular, a study of robustness to model specification. Rudebusch (2000) conducts such a study. He

[4]The attractive features of the nominal income rules mentioned above are reviewed in greater detail in Rudebusch (2000).

compares the performance of nominal income targeting with that of a benchmark Taylor rule for three different parameter specifications of his model. He finds that though nominal income growth rules do not destabilize the economy, they perform much worse than Taylor type rules. Do his results hold under the more general model uncertainty analysis provided here?

The final issue addressed is that of robustness of forecast-based rules. This issue could not in principle be addressed by the Onatski and Stock approach in its backward-looking form. I compare the robustness of forecast-based rules with that of benchmark Taylor-type rules and nominal income rules. I am particularly interested in the question of whether, as suggested by studies of Levin, Wieland and Williams (1999) and Bernanke and Woodford (1997), the forecast-based rules can easily result in indeterminacy of the equilibrium.[5]

I consider several sources of uncertainty about the Rudebusch model. First, it is econometric uncertainty about point estimates of the model's parameters. Second, the reference model may be misspecified so that the true model may include different lags or leads of endogenous variables. Finally, there exist uncertainty about the quality of data available in the real time.

I find that, for the majority of the policy rules studied, statistically small perturbations of the Rudebusch model may lead to dynamic instability. For example, it is enough to change parameters of the model inside the 70% asymptotic confidence ellipsoid to face the dynamic instability under the famous Taylor rule (interest rate

[5]In contrast to Levin, Wieland and Williams (1999) paper, I consider policy rules based on private forecasts of inflation and the output gap. The private sector is assumed to know the true model of the economy so the forecasts are based on the perfect information.

reacts to inflation and the output gap with coefficients 1.5 and 0.5 respectively).

Robustness ranking of different policy rules strongly depends on the particular specification of the uncertainty chosen. For econometric error uncertainty, those rules that are relatively more aggressive are relatively more robust. Just the opposite is true for more broadly specified uncertainty: those rules that are relatively less aggressive are relatively more robust.

I find the nominal income growth rules to be much less robust than the Taylor-type rules for majority of the uncertainty specifications studied. However, for those uncertainty specifications that include real-time data uncertainty the robustness of nominal income policy rules may be comparable to that of the Taylor-type rules.

The indeterminacy in the equilibrium triggered by the model uncertainty is not a real threat for the forecast based rules. So, for example, reacting to 2 years ahead forecasts of inflation and the current output gap does not lead to indeterminacy for reasonable perturbations of the reference model. In general, the robustness of the forecast based rules seem to be higher than that of the rules based on current or past data.

Most of the above results depend on the particular choice of the reference model made. I illustrate this dependence by trying to replicate some of the above results for the Clarida, Gali, and Gertler (1999) model.

The rest of the paper is organized as follows. Section 2 introduces notions of a reference model and a perturbation space and defines a distance between a perturbation and the reference model. In Section 3, I define the index of robustness and develop an algorithm for its computation. Section 4 is devoted to an application of the developed techniques to the analysis of robustness in the New Keynesian model of economy described in Rudebusch (2000). Section 5 concludes. Some technical details of the paper are given in the Appendix.

2. Modeling uncertainty

In this section I define the basic components of the structured non-parametric approach to model uncertainty described in Onatski and Stock (1999). These components are: the reference model, the perturbation space, and the distance between a perturbation and the reference model.

2.1. Reference model

Assume that a policy-maker's model of the economy is

$$\sum_{j=0}^{m} E_{t-j} M^j(L) X_t = u_t, \tag{2.1}$$

where X_t is a $n \times 1$ vector of endogenous variables including a policy instrument, u_t is $n \times 1$ vector of shocks, $M^j(L)$ is a $n \times n$ matrix lag polynomial containing both positive and negative powers of the lag operator, L, and where E_{t-j} is expectation given information available at time $t-j$ including knowledge of the model structure. The information set at time t consists of current and all past values of shock u_t. The first $n-1$ equations of the model describe dynamics of the endogenous variables and the last equation represents policy.[6]

A policy-maker can affect the state of the economy by choosing the coefficients of the policy equation. I assume that the policy-maker is able to commit to her choice so that the policy equation can be viewed as a policy rule. The only rules I consider are linear responses to current, past and expected future values of the variables X_t. This choice of possible rules is not very restrictive. Indeed, most simple policy rules that receive much attention are linear rules (see Taylor (1999)). Besides, as is well known, optimal rules in the case of conventional linear quadratic control are linear. However, the set of linear rules is too restrictive in some important settings, such as one with lower bound on the nominal interest rate explicitly taken into account (see Orphanides and Wieland (1999)). In this paper I will not consider such settings.

In what follows I employ Whiteman's (1983) solution principle to solve (2.1). That is, first, I restrict attention to the case when expectations are formed linearly. Hence, E_tX must be read as optimal linear predictor of X given information available by time t. Second, the shock process u_t is a zero-mean regular stationary process. Third, solutions will be sought in the space spanned by time-invariant square-summable linear combinations of u_t. I also require that the coefficients of $M^j(L)$ are absolutely summable, which is satisfied automatically for standard models where $M^j(L)$ represents a finite order polynomial.

[6]Models (2.1) were studied, for example, in Broze, Gourieroux and Szafarz (1995) and include a wide variety of special linear rational expectations models studied in the literature.

2.2. Perturbation space

The policy-maker understands that her model, which I will call the reference model, is only an approximation to reality because, perhaps, not all relevant variables are included in X_t, or not all relevant lags are considered, or because linear equations of the model are imperfect substitute for the true nonlinear relations describing the economy, or, maybe, true economic relations are subject to structural breaks etc. She prefers, therefore, to use a policy rule that works well for all models from a neighborhood of the reference model:

$$\sum_{j=0}^{m} E_{t-j}\left(M^j(L) + W_1^j(L)\Delta^j W_2^j(L)\right) X_t = u_t. \tag{2.2}$$

Here Δ^j are in general $k \times k$ nonlinear, time-varying, and not necessarily causal block-diagonal operators from the space of $k \times 1$ stationary processes to itself, and $W_1^j(L)$ and $W_2^j(L)$ are $n \times k$ and $k \times n$ weighting matrix lag-lead polynomials. It is assumed that the private sector knows the true model so that the expectations in (2.2) are taken with the full knowledge of Δ^j.

The blocks on the diagonal of Δ^j and the weighting matrices can be structured so that the uncertainty is focused on potentially important weaknesses of the model. The following extremely stylized example illustrates the idea of structured perturbations. Assume that the reference model of the policy-maker is

$$\begin{aligned} x_t &= ap_{t-1} + s_t \\ p_t &= kx_t + v_t, \end{aligned} \tag{2.3}$$

where p_t is a policy instrument, s_t, v_t are exogenous shocks, and x_t is the variable of interest for the policy-maker. The model has form (2.1) with $m = 0, M^0 = \begin{pmatrix} 1 & -aL \\ -k & 1 \end{pmatrix}$, $u_t = [s_t, v_t]'$, and $X_t = [x_t, p_t]'$.

Now, suppose that the policy-maker suspects that the effect of policy on x_t has some weak but long and variable unmodeled lag structure. Besides, it is suspected that x_t might have some inertia of its own and the value of x_t today is somewhat affected by expectations of future x_t. The policy-maker then may believe that a better model of the economy has the following form

$$\begin{aligned} x_t &= w_1\Delta_1 x_{t-1} + w_2 E_t \Delta_2 x_{t+1} + (a + w_3\Delta_3)p_{t-1} + u_t \qquad (2.4) \\ p_t &= kx_t + v_t, \end{aligned}$$

where Δ_1 is a linear, time-invariant, causal operator represented by an infinite lag polynomial, Δ_2 is a linear, time-invariant anti-causal operator represented by infinite lead polynomial, and Δ_3 is a linear, causal, slowly time-varying operator represented by infinite lag polynomial with slowly time-varying coefficients. The weights w_i reflect relative importance of the model weaknesses corresponding to Δ_1, Δ_2, and Δ_3. The above model has form (2.2) with $\Delta^0 = diag\,(\Delta_1, \Delta_2, \Delta_3)$, $W_1^0 = -\begin{pmatrix} w_1 & w_2 & w_3 \\ 0 & 0 & 0 \end{pmatrix}$, and $W_2^0 = \begin{pmatrix} 0 & 0 & L \\ L & L^{-1} & 0 \end{pmatrix}'$.

In this paper I consider only linear time invariant perturbations Δ^j, such that their infinite lag-lead polynomial representations have absolutely summable coefficients. The simplest case of such perturbations is multiplication by a constant (all coefficients in the lag-lead polynomial representation are zero except the one on $L^0 \equiv 1$). Such static operators may be used in (2.2) to represent uncertainty about values of model parameters. In the above example, if uncertainty were about the size of coefficient a only, we would have

$$x_t = (a + \Delta)p_{t-1} + s_t, \tag{2.5}$$

with $\Delta^0 = \Delta$ which is simply a constant, $W_1^0 = (-1, 0)'$, and $W_2^0 = (0, L)$.

Note that in general the perturbed model (2.2) may be represented in the "shock uncertainty" form

$$\sum_{j=0}^{m} E_{t-j} M_j(L) X_t = u_t + \xi_t, \tag{2.6}$$

where $\xi_t = -\sum_{j=0}^{m} W_{j,1}(L)\Delta_j W_{j,2}(L) X_t$. The shock representation of uncertainty that does not specify particular structure of the additive shock ξ_t became recently a popular vehicle of research on robustness (see, for example, Hansen and Sargent (2000) and references therein).

2.3. Distance between perturbations and the reference model

To define a neighborhood of the reference model in the perturbation space we need a notion of distance between the reference model and the alternatives. I define the distance between the reference model, M, and perturbation (2.2), M_Δ, as

$$d(M, M_\Delta) = \max_j \left\| \Delta^j \right\|.$$

Here the norm of Δ, $\|\Delta\|$, is taken to be L_∞ norm of the function $\Delta(e^{i\omega})$, that is[7]

$$\|\Delta\| = \left\{ \sup_{\omega} \text{maxeval} \left[\Delta'(e^{-i\omega})\Delta(e^{i\omega}) \right] \right\}^{1/2},$$

where maxeval denotes maximum eigenvalue.

If the uncertainty exists only about values of parameters of the reference model so that Δ is a diagonal matrix of constants, the L_∞ norm of $\Delta(e^{i\omega})$ is simply the maximum of absolute values of its diagonal elements. So if in example (2.3) we are uncertain only about value of a then the distance between perturbed model (2.5) and the reference model is equal to the absolute value of the difference between the perturbed parameter, $a + \Delta$, and the reference parameter, a.

General linear time invariant perturbations Δ may be viewed as linear filters acting in the space of stationary random processes with finite variances. Then L_∞ norm of $\Delta(e^{i\omega})$ is the maximal gain of filter Δ across different frequencies. That is, for any stationary input of the filter, ξ_t, with variance 1 the variance of the output will be less than or equal to $\|\Delta\|$. Moreover, there exists input ξ_t with variance 1 such that the output has variance arbitrarily close to $\|\Delta\|$.

Indeed, let ξ_t be a stationary input of the filter. Then, for the variance of the

[7] In what follows I will ignore indices j as if all expectations in the model were taken as of time t. This will simplify my notations. I will use the indices whenever they are needed for understanding of the material.

output we have:

$$
\begin{aligned}
& tr\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta(e^{i\omega})F_\xi(\omega)\Delta'(e^{-i\omega})d\omega \\
= \ & tr\frac{1}{2\pi}\int_{-\pi}^{\pi}\Delta'(e^{-i\omega})\Delta(e^{i\omega})F_\xi(\omega)d\omega \qquad (2.7)
\end{aligned}
$$

where F_ξ is the spectral density matrix of the process ξ_t. Let f be the eigenvector corresponding to $\sup_\omega \text{maxeval}[\Delta'(e^{-i\omega})\Delta(e^{i\omega})]$ and ω_0 is the frequency where the supremum is attained. Then the spectral density matrix can be chosen so that it is proportional to $f * f'$ with huge coefficient of proportionality at ω_0 and it is small for other ω.

Indeed, let ζ_t be a real-valued stationary process with spectral density function equal to $f * f^{'}$ at frequency ω_0. Consider a sequence of filters with positive Fourier transforms $g_n(\omega)$ such that $\frac{1}{2\pi}\int_{-\pi}^{\pi} g_n(\omega)d\omega = 1$, $g_n(\omega) = g_n(-\omega)$, and g_n converges to zero uniformly outside any open set containing $\pm\omega_0$. Apply these filters to ζ_t and denote the resulting processes (scaled so as to have unit variance) as ξ_{nt}. Expression (2.7) can be made arbitrarily close to $\|\Delta\|$ by choosing $\xi_t = \xi_{nt}$ for large enough n. On the other hand, this expression is obviously no larger than $\|\Delta\|$. It seems reasonable to consider perturbation operators that cannot increase the variance of an input process too much.

Interpretation of the distance between models depends on the reference model, the weighting matrices and assumptions made about operators on the diagonal of Δ. Depending on these factors there may or may not exist a monotone relationship between the distance and some statistical measure of closeness of models. More on this will be said in the application part of the paper.

3. Index of robustness

The policy-maker is assumed to have a quadratic loss function

$$L_t = (1 - \beta) E_t \sum \beta^i \left(X_{t+i}' \Lambda X_{t+i} \right), \qquad (3.1)$$

where Λ is some positive-definite weighting matrix, and β is between zero and one. Below I consider the case $\beta \to 1$, so that, for stationary X_t, the loss is equal to the variance of a linear combination of variables from X_t.

$$L = EX_t' \Lambda X_t.$$

The policy maker's problem is to choose a rule for the policy instrument so that the loss for all models from the vicinity of the reference model is not too high. The two most popular formalizations of this problem lead to the Bayesian and the minimax criterion of optimality. Let $\mathcal{F}$ be the set of feasible policy rules rules, f. Then a Bayesian policy-maker will choose the policy rule

$$f = \arg\min_{f \in \mathcal{F}} \int \sup_{X_t \in S} EX_t' \Lambda X_t dF(\Delta), \qquad (3.2)$$

where the inner supremum is taken over all X_t from the set S of all stationary solutions to a particular model and F is a probability measure over the perturbation space.

Note that if there is indeterminacy situation, when S consists of more than one element, the inner supremum is infinite. Indeed, let X_{1t} and X_{2t} be two different stationary solutions. Then any linear combination $Z_t = \lambda X_{1t} + (1-\lambda)X_{2t}$ is also a solution. It is then possible to choose λ so as to make loss associated with Z_t as large as one wants.

A "minimax policy maker" will choose

$$f = \arg\min_{f\in\mathcal{F}} \sup_{||\Delta||\leq r} \sup_{X_t\in S} EX_t'\Lambda X_t, \tag{3.3}$$

where the outer supremum is taken over all perturbations from the ball of radius r. The radius can be chosen in many different ways. One way would be to choose r large enough for uncertainty set to include some particular alternative to the reference model. In the example given in the previous section suppose that a prominent alternative to the reference model was

$$\begin{aligned} x_t &= bx_{t-1} + ap_{t-1} + s_t \\ p_t &= kx_t + v_t. \end{aligned}$$

Then for the uncertainty set to include this model $\|\Delta\|$ must be at least b/w_1. Hence, we can choose $r = b/w_1$.

Another way to choose r would be to include in the model uncertainty set only those models that are statistically close to the reference one. For example, if a point estimate of parameter a in the reference model $x_t = ap_{t-1} + \varepsilon_t$ is, say, 0.5 and its standard deviation is 0.1 we can represent the uncertainty by a set of models $x_t = (a + \Delta)p_{t-1} + \varepsilon_t$ with $|\Delta| < 0.1$. Still another possibility would be not to choose r in advance but vary it from zero to infinity to get a whole family of the minimax rules robust to the uncertainty of different size.[8]

At present no numerical algorithms solving either the Bayesian or the minimax problem as they are stated above are known.[9] However, as I show below, it is possible to compute the set of rules that do not result in economic instability for each perturbation from the set

$$D_r = \{\Delta : \Delta \text{ has particular block diagonal structure specified at the stage} \tag{3.4}$$

$$\text{of formulating the model uncertainty and } ||\Delta|| < r\}.$$

Similar to Christiano and Gust (1999), I focus attention on the extreme economic instability that results either in non-stationarity or in indeterminacy in the equilibrium. Obviously, the set of stabilizing rules contains the minimax rule. It also contains the optimal Bayesian rule if the support of measure F on perturbations Δ includes D_r. In general, the larger the size of possible perturbations, r, the smaller the set of stabilizing rules. Hence, for sufficiently large r the set of stabilizing rules

[8]Note also that I consider a situation when the model uncertainty represented by the probability distribution $F(\Delta)$ in (3.2) and by the perturbation set $\|\Delta\| \leq r$ in (3.3) is not changing over time. There is no learning and rule f is chosen once and for all given the model uncertainty prevailing at the time of the choice.

[9]Paganini (1996) gives an algorithm for computation of the minimax rules when the model and uncertainty operators are backward looking. Application of these techniques for the robustness analysis in the Rudebusch-Svensson model can be found in Onatski (2000).

is very narrow and, therefore, it characterizes the optimal rules fairly precisely. On the other hand, when r is small, the set of stabilizing rules can be large and it is relatively uninformative about the nature of the optimal rules that solve (3.2) and (3.3).

One way to summarize stabilization properties of a policy rule is to compute the maximum r such that the rule still results in the finite loss for any model from the ball D_r. I call such maximum the index of robustness for the rule. More formally,

Definition 1. *I define the index of robustness for a rule f as supremum r such that f results in unique stationary solution for any model from D_r except, maybe, a degenerate set of models such that it does not include any open subset.*[10]

Since it is natural to assume that the precise autocorrelation structure of the noise is not known, by the existence of a solution to the model I mean existence of a solution for any stationary noise process with the correlation structure arbitrarily close to that assumed for the reference noise. This definition avoids some pathological situations when the solution to the model exists only for a particular correlation structure of the noise. For example, consider a model

$$\frac{1}{2^{i+1}} \sum_{i=-1}^{\infty} E_t X_{t+i} = u_t$$

[10] *The "degenerate set" qualification simplifies computation of the index. It is tempting to say that it also makes the index less sensitive to extremely improbable destabilizing perturbations. However, if there exists an open neighborhood of a degenerate set of models where the policy-maker's loss is not necessarily infinite but simply very high then sensitivity of the index to degenerate sets of perturbations that literally destabilize the model of the economy may be desirable because it indicates existence of a non-degenerate set that leaves the model stable but makes it extremely volatile.*

Let $u_t = \sum_{k=0}^{\infty} c_k \varepsilon_{t-k}$ be Wold representation of u_t. The model has solution if and only if $2c_0 = c_1$. This condition cannot be granted if second moments of the noise are known imprecisely.

One should interpret the index of robustness with caution. If the size of uncertainty, r, is known to the policy-maker, then an optimally robust rule in terms of either (3.2) or (3.3) must have the index greater than r. It would be wrong, however, to recommend the rule with the largest index. Indeed, such a rule may trade off extreme stability robustness with poor conventional loss. An example of such a situation will be given in the application section.

3.1. Computation of the index

In this section I explain how to compute the index of robustness for a given policy rule. To get this computation done one needs to have a criterion for existence and uniqueness of a stationary solution for any given perturbation around the reference model. Conditions for existence and uniqueness of solution to forward-looking models are well known for the case when the model has only finite number of leads and lags (see, for example, Whiteman (1983)). For example, if the model has form

$$P(L)x_t + E_t Q(L^{-1})x_t = \varepsilon_t$$

where x_t is one-dimensional variable and P and Q are polynomials, then a stationary solution exists and is unique if and only if the number of zeros of $P(z)+Q(z^{-1})$ lying outside (inside) the unit circle is exactly equal to the degree of P (respectively degree of Q).

However, because I do not restrict attention to perturbations with finite number of lags or leads I need a criterion of existence and uniqueness that will work for models having infinite lag-lead structure. Below I formulate such a criterion that was developed in a separate paper (see Onatski (2001)).

Define a winding number of a complex-valued function $f(e^{i\omega})$, wind$f$, as the number of times the graph of f rotates around zero counter-clockwise in the complex plane when ω goes from 0 to 2π.[11] Define a function $M(e^{i\omega}) = \sum_{j=0}^{m} e^{-im} M^j(e^{i\omega})$. Then the following criterion of existence and uniqueness of solution holds except for a degenerate set of models

Criterion. *Model (2.1) has a unique solution, multiple solutions, or, no solutions if and only if the winding number of* $\det M(e^{i\omega})$ *is equal to zero, less than zero, or greater than zero respectively.*

The first step in computing the index of robustness for a given rule is to check whether the reference model has unique stationary solution. I assume that the reference (not perturbed) model has only finite lags and leads so that one can use standard criteria described in Whiteman (1983). If the reference model does not have unique solution then the index of robustness for the rule is zero. Otherwise, the radius is positive and according to the above criterion the winding number of $\det M(e^{i\omega})$ is zero.

Now consider perturbations (2.2) to the reference model. Define function $M_\Delta(e^{i\omega})$

[11] A clockwise rotation of the graph around zero is counted with the negative sign.

as

$$M_\Delta(e^{i\omega}) = e^{-im} \sum_{j=0}^{m} \left(M^j(e^{i\omega}) + W_1^j(e^{i\omega}) \Delta^j(e^{i\omega}) W_2^j(e^{i\omega}) \right).$$

It is convenient to rewrite $M_\Delta(e^{i\omega})$ in the form

$$M_\Delta(e^{i\omega}) = e^{-im} \left(M(e^{i\omega}) + W_1(e^{i\omega}) \Delta(e^{i\omega}) W_2(e^{i\omega}) \right),$$

where $M = \sum_{j=0}^{m} M^j$, $W_1 = [W_1^0, W_1^1, ..., W_1^m]$, $W_2 = [W_2^0, W_2^1, ..., W_2^m]'$, and $\Delta = diag\,(\Delta^0, \Delta^1, ..., \Delta^m)$. For each point on the unit circle, e^{iw}, the value of $\det M_\Delta(e^{i\omega})$ is a continuous function of $\Delta(e^{iw})$. Recall that the size of the perturbation operator, Δ, is measured by the L_∞ norm of $\Delta(e^{i\omega})$. Hence, the graph of $\det M_\Delta(e^{i\omega})$ changes continuously with respect to small (in L_∞ sense) changes in the perturbation operator, Δ. Therefore, wind($\det M_\Delta$) can become different from zero only after the perturbation Δ becomes large enough for $\det M_\Delta(e^{i\omega})$ to hit zero for some $w \in [0, 2\pi)$.

Suppose that the graph of $\det M_\Delta$ hits zero for some $\Delta = \Delta_0$ of size r but not for smaller Δ. Then, the index of robustness is larger than or equal to r. Indeed, for perturbations Δ such that $||\Delta|| < r$ the winding number of $\det M_\Delta$ is equal to zero, so according to the criterion the perturbed model has a unique stationary solution unless it belongs to a degenerate set of models that we exclude from consideration. On the other hand, the index must be less than or equal to r because it is possible to change Δ_0 marginally so that the graph of $\det M_\Delta$ will cross zero and the winding number of $\det M_\Delta$ becomes different from zero. Thus, as the criterion implies, there exists a perturbation operator, $\bar{\Delta}_0$, of the size marginally larger than r such that the model either has multiple or no solutions.[12]

To summarize, to get the index of robustness one needs to find minimum $\|\Delta\|$ such that matrix $M(z)+W_1(z)\Delta(z)W_2(z)$ is singular for some $z:|z|=1$. Note that M is invertible on the unit circle because the reference model has unique solution under the policy rule. Therefore, on the unit circle we have

$$\begin{aligned}\det\left(M+W_1\Delta W_2\right) &= \det\left(M\right)\det\left(I_n+M^{-1}W_1\Delta W_2\right)\\ &= \det\left(M\right)\det\left(I_k+W_2M^{-1}W_1\Delta\right)\end{aligned}$$

where I_n denotes $n\times n$ unity matrix. Denote $-W_2M^{-1}W_1$ as S. Then we are looking for minimum $||\Delta||$, having a particular block-diagonal structure, that makes matrix $I_m-S(z)\Delta(z)$ singular at some point on the unit circle.

This problem is known in engineering literature as the problem of computing structured norm of operator S (see, for example, Dahleh and Diaz-Bobillo (1995)). In the next section I implement numerical algorithms[13] for computing the structured norm to analyze robustness of simple policy rules under model uncertainty in a small empirical New Keynesian model of the economy studied in Rudebusch (2000b).

[12] Note that small deviations from $\bar{\Delta}_0$ leave the winding number of $\det M_\Delta$ different from zero so that the perturbed model corresponding to $\bar{\Delta}_0$ is not from the degenerate set mentioned in Definition 3.1.

[13] Computer codes I use are based on the programs available in Mu Analysis and Synthesis Toolbox in Matlab.

4. Application

In this section, I use the above results to study three sets of questions. The questions concern policy activism appropriate under model uncertainty, the robustness of nominal income rules, and the robustness of forecast-based rules. As was explained in the introduction I am particularly interested to know whether the extension of Onatski and Stock (1999) to forward-looking models recommends extreme activism or not, how robust nominal income rules are relative to benchmark Taylor-type rules, and whether forecast-based rules can easily lead to indeterminacy of equilibrium.

4.1. Reference model

To perform the analysis I first need to choose a reference model. As John Taylor (2000) notes, despite a lot of differences in models now used for normative policy analysis, there is a common general framework. It consists of three basic equations: a Phillips-curve-type equation, an IS-type equation relating real GDP and the real interest rate, and an equation for monetary policy rule. On one end of the spectrum of models having the above form there are New Keynesian forward-looking models of the economy such as those described in Woodford (1999) and Clarida, Gali and Gertler (1999). These models have solid micro foundations and based on general equilibrium analysis but fail to fit data well. On the other end of the spectrum there are empirical purely backward-looking models such as Rudebusch and Svensson's (1999) and Ball's (1999) models. These models fit data surprisingly well but have obscure foundations.

I chose to study an empirical New Keynesian model proposed and estimated in Rudebusch (2000). The model nests both theoretically appealing forward-looking models and empirically sound backward-looking models. It consists of two equations estimated using US quarterly data from 1968:Q3 to 1998:Q2

$$\begin{aligned}\pi_t &= \underset{(.08)}{.26} E_{t-1}\bar{\pi}_{t+3} + .74\left(\underset{(.14)}{.69}\pi_{t-1} - \underset{(.14)}{.15}\pi_{t-2} + \underset{(.14)}{.41}\pi_{t-3} + \underset{(.12)}{.07}\pi_{t-4}\right) + \\ &\quad \underset{(.05)}{.16} y_{t-1} + \varepsilon_t \end{aligned} \tag{4.1}$$

$$y_t = \underset{(.09)}{1.15} y_{t-1} - \underset{(.09)}{.27} y_{t-2} - \underset{(.03)}{.09}(i_{t-1} - E_{t-1}\bar{\pi}_{t+3}) + \eta_t. \tag{4.2}$$

Here $\bar{\pi}_t = \frac{1}{4}(\pi_t + \pi_{t-1} + \pi_{t-2} + \pi_{t-3})$ is the four-quarter average value of the annual percentage rate of inflation, y_t is the output gap[14], measured as 100 times the log ratio of actual real output to potential output, and i_t is the federal funds rate at an annual rate. The standard errors of the coefficient estimates are given in parenthesis.

The first equation is an accelerationist Phillips-curve-type relationship. It is a hybrid of a backward-looking Phillips curve and a forward-looking equation. Inclusion of the backward-looking terms substantially improves the equation's fit to data. Besides, the backward-looking terms can be theoretically justified by assuming sticky inflation as in Fuhrer and Moore (1995) or sticky information as in Mankiw and Reis (2001).

The second equation is an IS curve linking the output gap to the past output gap and ex ante real interest rate. Such a backward-looking own dynamics of the output

[14] Ken Kuttner kindly provided me with estimates of the gap obtained using the method outlined in his (1994) paper.

gap contrasts with the forward-looking theoretical behavior described by Woodford (1999) and Clarida, Gali and Gertler (1999). It is, however, possible to justify the backward-looking dynamic of the output gap theoretically by assuming consumers' habit formation or capital adjustment costs. The timing structure of the equations (4.1) and (4.2) reflects "real world recognition, processing, and adjustment lags" as discussed in Rudebusch (2000).

4.1.1. Policy Rules

I consider the policy rule equations of two different types. The first type is represented by the rules of the form

$$i_t = g_i i_{t-1} + (1 - g_i) i^* + g_\pi E_t(\bar{\pi}_{t+J} - \pi^*) + g_y E_t y_{t+K}. \tag{4.3}$$

that set nominal interest rate equal to a linear combination of inflation (lagged if $J < 0$, current if $J = 0$, or expected if $J > 0$), the output gap (lagged if $K < 0$, current if $K = 0$, or expected if $K > 0$), and the lagged interest rate. Constants π^* and i^* correspond to the inflation target and the unconditional mean of the nominal interest rate given that there is no inflation bias.[15] The second type is represented by the nominal income rules proposed by Orphanides (1999) and McCallum and Nelson

[15]The rules of the above type are equivalent to

$$i_t = \rho i_{t-1} + (1 - \rho)(r^* + E_t \bar{\pi}_{t+J}) + \alpha E_t(\bar{\pi}_{t+J} - \pi^*) + \beta E_t y_{t+K}$$

that might look more familiar (see, for example, Levin, Wieland, and Wiliams (1999b)). Here r^* denotes the unconditional mean of the equilibrium real interest rate. Parameters of (4.3) can be expressed in terms of ρ, α, and β as follows: $g_i = \rho, g_\pi = 1 + \alpha - \rho$, and $g_y = \beta$.

(1999):

$$i_t = i^* + (\bar{\pi}_t - \pi^*) + g_{n1}(\bar{\pi}_t - \pi^* + y_t - y_{t-4}) \quad (4.4)$$

$$i_t = g_{n2}(\pi_t - \pi^* + 4(y_t - y_{t-1})) + g_i i_{t-1}. \quad (4.5)$$

According to these rules a policy maker changes nominal interest rate in response to deviations in growth in the nominal income from the target growth. In what follows I use standard zero normalization (see Rudebusch and Svensson (1999)) for π^* and i^*.[16] Such a normalization does not affect "stability and uniqueness robustness" characteristics of the rules because adding constants to system equations (4.1,4.2) changes neither the system's stability nor its determinacy.[17]

Attractive properties of nominal income rules were briefly discussed in the introduction. The rules of type (4.3) were introduced by J. Taylor (1993) and since then were subject to active research. The famous Taylor rule corresponds to $g_\pi = 1.5$, $g_y = 0.5$, and $g_i = J = K = 0$. Such rules fit data quite well, at least since late 80's, and were shown to be near optimal in variety of models and relatively robust to different model specifications (see Taylor (1999)).

Smoothing of interest rate ($g_i \neq 0$) in (4.3) considerably improves data fit and robustness properties of the rules (see Levin, Wieland and Williams (1999a)). Letting $K, J < 0$ can account for data processing lags, importance of which was emphasized

[16] Note that the variables used to estimate (5.1,5.2) were demeaned prior to estimation so the normalization was imposed.

[17] Choosing incorrect value of i^* does , however, increase loss associated with a given policy rule because it implies inflation bias. Therefore, the normalization matters for more precise level of analysis than that considered in this paper.

by McCallum (1997). The forecast-based rules correspond to positive J, or K, or both and have many desirable properties. In particular, as argued by Svensson (1997), inflation forecast is a very good intermediate target for inflation targeting policy adopted by many central banks. Therefore, reacting to forecast of inflation is an "information-encompassing" strategy.

4.2. Uncertainty

I consider several sources of uncertainty about the reference model.

4.2.1. Econometric error.

The parameters of the reference model are estimated with econometric error so that the true values may differ from the point estimates in the range of a p% confidence ellipsoid. Unfortunately, it is impossible to represent such an uncertainty set in the form (3.4) that is convenient for computations. I therefore consider a set of models whose parameters differ from the point estimates in the range of a confidence parallelepiped that is a linear approximation of the confidence ellipsoid.

Precisely, denote vector of deviations of the perturbed model parameters from the point estimates[18] as $d = [d_1, ..., d_8]'$. And denote the variance-covariance matrix of the point estimates as V. Define $\delta = (\Delta_1, ..., \Delta_8)'$ as $\delta = V^{-1/2'}d$. I consider the set of models corresponding to all δ such that $|\Delta_i| < r$. Since δ is distributed approximately as 8-dimensional standard normal random variable, the confidence

[18] Since the long run verticality of the Phillips curve is maintained, there are 8 independently estimated parameters in the reference model.

level corresponding to such a cube (or parallelepiped in terms of d) is equal to $(2\Phi(r)-1)^8$ where Φ is a cumulative distribution function (cdf) for standard normal distribution. The parallelepiped is encircled by the confidence ellipsoid with the level of confidence $p = F_{\chi_2(8)}(8r^2)$ where $F_{\chi_2(8)}$ stands for cdf of chi-squared distribution with 8 degrees of freedom. As shown in Appendix, such a set can be represented in form (3.4).

The index of robustness for the econometric error uncertainty measures the smallest size of δ (defined as $\max_i |\Delta_i|$) such that there exist a "bad" combination of parameters inside the corresponding confidence parallelepiped that results in instability or indeterminacy in the economy. Clearly, for this particular description of uncertainty, the index of robustness can be rescaled to measure the confidence level of the smallest parallelepiped (or the corresponding encircling ellipsoid) that includes "bad" parameter values.

4.2.2. Specification error.

The reference model may be misspecified. At the simplest level, a few lags or leads of the endogenous variables may be wrongfully omitted from the reference equations. As an example, I reestimate the reference model with one additional lag of the output gap added to the Phillips curve and the IS equations and one additional lag of the real interest rate added to the IS equation. Then I form a set of models whose parameters differ from these point estimates in the range of a p% confidence parallelepiped.

The above treatment of possible misspecifications is obviously very limited. Therefore, to introduce less restricted specification errors, I perturb the reference equations so as to include potentially infinite number lags or leads of all variables in the right hand side. Denote the right hand side of equations (4.1) and (4.2) as ref_π and ref_y respectively. I consider the following perturbations:

$$\begin{aligned}
\pi_t &= ref_\pi + w_{\pi\pi} E_{t-1} \Delta_{\pi\pi}(\pi_{t+1} - \pi_t) + w_{\pi y} \Delta_{\pi y} y_{t-1} \\
y_t &= ref_y + w_{yy} E_{t-1} \Delta_{yy} y_{t+1} + w_{yr} E_{t-1} \Delta_{yr}(i_{t-1} - E_{t-1}\bar{\pi}_{t+3}),
\end{aligned}$$

where $\Delta_{\pi\pi}, \Delta_{\pi y}, \Delta_{yy}$, and Δ_{yr} represent uncertainty about the four dynamic channels of the model: inflation-inflation, inflation-output gap, output gap-output gap, and output gap-real interest rate. The weights w_{ij} are supposed to reflect relative importance of the uncertainties. I measure them by an average standard error in coefficient estimates corresponding to a particular channel.[19] Hence, $w_{\pi\pi} = 0.09, w_{\pi y} = 0.05, w_{yy} = 0.09$, and $w_{yr} = 0.03$. Obviously, such a choice of the weights is arbitrary. Therefore, I vary the weights in the numerical computations below to check robustness of my results with respect to the above choice.

Uncertainty operator $\Delta_{\pi\pi}$ is taken to be linear time invariant operator with absolutely summable coefficients. It acts on the first differences instead of the level of inflation because I want to keep sum of inflation coefficients in the right hand side of the Phillips curve equal to one. By adding $\Delta_{\pi\pi}$ I allow for deviations from the reference model that have different inflation lags and leads structure of the Phillips curve. In

[19]Since Δ_{ij} may include infinite number of lags/leads, this weighting scheme makes little sense from the econometrics point of view. Moreover, the index of robustness for this particular description of uncertainty has little connection with statistical size of the smallest "destabilizing uncertainty set". The robustness analysis under such perturbations is still useful. For example, it may suggest the structure of statistically relevant misspecifications that bring most harm to the policymaker.

particular, choosing large enough uncertainty size I can get a purely forward-looking form of the Phillips curve.

Similarly, I choose uncertainty operator Δ_{yy} to be a two-sided (mixed forward and backward-looking) linear time invariant operator. Thus, I can consider a deviation from IS curve (4.2) to a purely forward-looking theoretical IS curve. It would be enough to choose $\Delta_{yy} = \frac{1}{w_{yy}}(1 - 1.15L^2 + .27L^3)$.

Uncertainty Δ_{yr} is considered to be a forward-looking linear time invariant uncertainty. It captures a fact that the monetary policy affects the economy not only through the short-term interest rate but also through longer-term interest rates and precise specification of this transmission is uncertain. Finally, the uncertainty $\Delta_{\pi y}$ is taken to be mixed linear time invariant uncertainty. This captures uncertain lags/leads in the effect of a change in the output gap on inflation.

4.2.3. Potential output uncertainty

The next source of uncertainty that I consider is uncertainty about potential output. As discussed for example in Orphanides (1999), the real-time estimate of the output gap is subject to substantial later revisions. The standard deviation of the revisions is comparable to the standard deviation of the gap series itself. The major part of the revisions is associated with the revisions of potential output series.

Typically, the real-time output gap uncertainty was modeled as an additive measurement error entering policy equation. This measurement error was found to be very persistent and was modeled as AR(1) process with large autoregressive root. It was often assumed that the error is uncorrelated with the true (or the final revision of) output gap. This implies (see discussion in Rudebusch (2000a)) that the real-time estimates of the gap are just noisy estimates of the true gap as opposed to efficient estimates based on the incomplete information available in the real time.

Here I will model the potential output uncertainty in a different way. I will assume that the real-time output gap estimate, y_t^{rt}, is related to the final estimate, y_t, through the following equation:

$$y_t^{rt} = y_t + E_t \Delta(L) y_t + \zeta_t, \tag{4.6}$$

where Δ is a mixed linear time invariant operator. This relationship implies that the error in the real-time potential output estimate is correlated to the true output gap so that the error may represent "news" as opposed to the "noise" as described above.

There are several reasons to model uncertainty in this way. The first reason is purely technical: additive shocks to the model that do not feedback on endogenous variables would not change stability properties of the model. Hence, the index of robustness would underestimate importance of the real-time data uncertainty. Making the real-time noise feeding back on y creates a possibility for changing stability properties and hence would, in some sense, address importance of the real-time data uncertainty.

The second reason is that even though empirical real-time data uncertainty studies (see Rudebusch (2000a)) find only a weak correlation between $y_t^{rt} - y_t$ and y_t, some theoretical models of the real-time data uncertainty imply that a substantial correlation must exist. For example, if the real-time estimates of the gap are obtained using some efficient procedures such as Kalman filtering described in Kuttner (1994) and the final estimates are obtained using the Kalman smoother, then a significant correlation exists. I run a regression of $y_t^{rt} - y_t$ on y_t where y_t^{rt} corresponds to Kuttner's estimates of the gap obtained using the filter and y_t corresponds to estimates of the gap obtained using the smoother. I estimated the regression coefficient to be -0.37 with t-statistics being -12.58.

Finally, the changes in the potential output are poorly understood. Stationarity of the potential output growth is questionable. Policy makers may, therefore, fear unprecedented behavior of the potential output that makes policy inadequate. If, for example, the potential output is overestimated in the time of recessions and underestimated in the time of booms then policy makers risk to overreact to the available information which may stimulate economic instability.

For the potential output uncertainty, the index of robustness measures the smallest (in L_∞ sense) operator Δ that brings instability or indeterminacy in the model. We may roughly interpret the $E_t\Delta(L)y_t$ part of (4.6) as representing "news" in the error and ζ_t as representing noise in the error. Given that the variance of $y_t^{rt} - y_t$ and y_t is about the same (as is suggested by empirical studies) the size of Δ provides an upper bound on the portion of the variance in the error due to the news. For example, if the index of robustness for a rule is equal to, say, 0.5 then no more than[20] $100\%*(0.5)^2 = 25\%$ of variance in the destabilizing real-time gap error is associated

[20]Precisely how much of the variance corresponds to the "news" depends on the spectral characteristics of Δ and y_t.

with the "news" component of the error and the rest 75% or more correspond to the "noise" component. The destabilization (or indeterminacy) of the economy under the real-time gap error is caused by a particularly unfortunate correlation structure of the news part of the error.

In the numerical section below I combine the data uncertainty with the simple specification and the econometric error uncertainty. For such a combined uncertainty, interpretation of the index of robustness depends on the weights given to the different sources of uncertainty. Below, I weight the different uncertainties equally. Loosely speaking, this means that the policy maker is equally afraid of "bad" deviations of the parameters inside $100\%{*}F_{\chi^2(12)}(12) = 55\%$ confidence ellipsoid and "bad" serial correlation structure in the news about the output gap when the "news" constitute 100% of the real-time error in the output gap.

4.2.4. Shock uncertainty

Finally, I study uncertainty about serial correlation of the shocks to the reference model. I assume that ε_t and η_t may be arbitrarily serially correlated, but have finite variance. Then for each policy rule I compute the expected loss under the worst possible serial correlation of the shocks[21]. The rule that minimizes such a worst possible loss is called H_∞ control. It is a limit of minimum entropy control rules

[21] I assume that policy maker's loss is equal to

$$L_t = (1-\beta)\sum_{i=0}^{\infty}\beta^i\left(\pi_t^2 + y_t^2 + 0.5(i_t - i_{t-1})^2\right)$$

as in Rudebusch (2000). I consider the case when $\beta = 0.99$.

when a parameter regulating degree of uncertainty aversion tends to the breakdown value (see Hansen and Sargent (2000)).

4.3. Numerical Results

Table 1 reports the index of robustness for simple policy rules that are optimal under no uncertainty about the reference model. Several specifications for the policy rules are considered. Panel A of the table corresponds to the Taylor-type rules reacting to the current data and to the data lagged by one quarter. For the rules reacting to the current data, both interest rate smoothing and no smoothing ($g_i = 0$) case are considered. Results for the Taylor rule ($g_\pi = 1.5$, $g_y = 0.5$) are reported as a benchmark.

Panel B of the table corresponds to the nominal income growth rules. Panel C describes results for three different specifications of the forecast based rules. I consider reaction to 1 year ahead forecast of inflation, 2 years ahead forecast of inflation and reaction to 1 year ahead forecasts of both inflation and the output gap. Panel D corresponds to the optimal H_∞ Taylor-type rule.

The last three columns of the table represent different uncertainty specifications. The first of these columns reports indices of robustness and the confidence levels of the corresponding confidence ellipsoids (not parallelepipeds) for econometric error uncertainty described in section 4.2.1. The next column corresponds to the uncertainty associated with few additional lags added to the reference model combined with the econometric error uncertainty as explained in section 4.2.2. The last column adds potential output uncertainty as described in section 4.2.3.

The indices of robustness for the policy rules optimal under no model uncertainty are surprisingly low (except for the forecast based rules). Particularly striking results are those for the combined simple specification uncertainty and the econometric error uncertainty. The optimal Taylor-type rule reacting to the current data leads to the dynamic instability for deviations of the parameters from the point estimates that belong to as small as 62% confidence ellipsoid. Neither interest rate smoothing, nor reaction to the lagged data improves the robustness. The benchmark Taylor rule results in instability for parameters inside 29% confidence ellipsoid. Orphanides' nominal income growth rule destabilizes the economy for as small as 5% confidence deviations of the parameters from the point estimates!

Introducing just a few additional lags to the model may change the robustness characteristics of the policy rules dramatically. The confidence level of the smallest "bad" deviations of the parameters from the point estimates reported in the "simple specification error" column are typically less than half the size of those reported in the "econometric error" column. Adding real-time uncertainty about potential output makes policy rules even less robust. For example, for the optimal Taylor-type rule reacting to the current data, instability is brought by deviation of the parameters from the point estimate inside $100\%{*}F_{\chi^2(12)}(0.52^2 * 12) = 0.6\%$ confidence ellipsoid (vs. 62% for the combined specification error and econometric error uncertainty!) given that about $100\%{*}0.52^2 = 27\%$ variation of the real time output gap error corresponds to the news arbitrarily correlated to the final revision of the output gap.

It is instructive to compare the rows of the table corresponding to the optimal

Taylor rule and the benchmark Taylor rule. We see that the robustness of the benchmark rule (believed to better correspond to actual historical policy than the optimal rule) is much less sensitive to the different choices of the uncertainty formulation. The benchmark rule looks much less robust than the optimal rule for the econometric error uncertainty specification. But it is in fact more robust than the optimal rule under the most general type of uncertainty considered. As we will see below, in general, relative robustness of different rules considerably varies with different uncertainty specifications. The rules with relatively sluggish response to inflation and the output gap tend to become relatively more robust when relatively more encompassing formulation of uncertainty is chosen.

Non-robustness of nominal income growth rules is much more impressive than that of the Taylor-type rules. The rules look non-robust even for the least general econometric error uncertainty. The index of robustness for nominal income growth rules deteriorates when potential output uncertainty is added even though the rules do not depend on the estimates of the level of the potential output (only on the growth of potential output). It is because I do not put special restrictions on the uncertainty about the growth of the potential output. If we believe that the growth must be much less uncertain than the level then we may want to compare robustness of the Taylor-type rules and the nominal income growth rules using columns "potential output error" and "simple specification error" respectively. Such a comparison reveals that in the most favorable situation for the nominal income growth rules, their robustness may be comparable to that of the Taylor-type rules.

The rules based on the forecasts of inflation and the output gap are significantly more robust than the rest of the rules studied in the table. Even for the most encompassing uncertainty specification the index of robustness for these rules is very large. Indeterminacy of equilibrium is not a real threat for the forecast based rules under the model uncertainty and the reference model studied in this paper.

The final panel of Table 1 reports indices of robustness for the Taylor-type rule specifically designed to be robust against shock uncertainty. We observe a striking non-robustness of this rule to the model uncertainty. At first, this observation seems to be extremely puzzling given that any uncertainty may be represented in the "shock form" as described in section 2.2 of the paper. However, some thought reveals the nature of this apparent inconsistency. The matter is that the shock uncertainty (and, more generally, the minimum entropy uncertainty) is ill-suited for description of dynamic uncertainty, that is the one that can feed back on the endogenous variables in the model. The shock uncertainty assumes that the uncertainty size does not depend on the particular policy rule studied. However, the effects of uncertainty in dynamic channels of the reference model may be amplified (or dampened) through the policy feedback on the endogenous variables of the model. Hence, the size of uncertainty in general depends on the particular policy rule chosen, and therefore such a dynamic uncertainty cannot be consistently represented in the shock form that does not account for this dependence.

Figures 1 through 4 present results in more detail. They show contour plots of the upper bounds for the index for different rules in the whole range of the parameters studied. Parameters of the rules are chosen in the following domain: $g_\pi, g_{n2}, g_y \in [0, 6]$ (grid of 1), $g_{n1} \in [0, 8]$ (grid of 1/3) and $g_i \in [-1, 1]$ (grid of 0.5). I experimented with the size of the grid and chose the reported one because it represents my solutions well. Figure 1 corresponds to the Taylor-type rule based on current observations of inflation and the output gap (J, K=0). The four subdivisions of the figure present indices of robustness under econometric error uncertainty, combined econometric and simple specification error uncertainty, general specification error uncertainty (as described in section 4.2.2), and the combined econometric, simple specification error, and potential output error uncertainty. We see that relatively more aggressive rules become non-robust when uncertainty about quality of the real-time data is introduced. Without this uncertainty, rules with aggressive reaction to the output gap and moderate reaction to inflation are relatively more robust.

The index of robustness is discontinuous with respect to g_π. It drops to zero if g_π becomes smaller than 1. This fact can be explained as follows. When g_π is only marginally larger than 1, dynamic uncertainty in the system hardly matters for "destabilization strength" of the aggregate supply shocks ε_t. It is because the supply shocks are almost fully accommodated by the policy so that only a tiny portion of the shock can be amplified through the endogenous dynamics of the economy.[22]

As was mentioned before, there is no simple statistical interpretation of the index of robustness for the general specification error uncertainty. However, qualitatively, the pattern of robustness for different policy rules under this type of uncertainty is

[22] Of course, policy rules with the inflation response close to unity are associated with high conventional loss (even though they do not easily result in dynamic instability).

similar to that for the other types of uncertainty analyzed. Rules with relatively more sluggish reaction to inflation are relatively more robust. The degree of the "most robust" response to the output gap under the general specification error uncertainty varies with different weightings of the different dynamic channels of uncertainty.

Figure 2 and 3 present results for the nominal income growth rules. The index of robustness for the nominal income growth rules is generally much smaller than that for the Taylor-type rules. However, in the most favorable situation when we assume that the target for the nominal income growth does not depend on the potential output estimates, the most robust nominal income rules are more robust than the most robust Taylor-type rules.

Figure 4 shows the index of robustness for the forecast based rule that responds to 1 year ahead forecast of inflation and the current output gap. The results for other specifications of the forecast based rules considered in this paper ($J = 4, K = 4$; $J = 8, K = 0$) are similar. We see that relative robustness of the forecast based rules depends on the specification of uncertainty chosen. However, in general, the forecast rules are much more robust than the Taylor-type rules responding to the current or past data and the nominal income rules.

Interestingly, the forecast based rules remain to be quite robust to model uncertainty even for purely forward-looking Clarida, Gali and Gertler (1999) reference model[23]. However, in such a case they are relatively less robust than the Taylor-type rules responding to current data. Moreover, indeterminacy become a real issue when

[23] I studied general specification error uncertainty for the CGG model.

policy responds to 1 year ahead forecasts of both inflation and the output gap.

5. Conclusion

The main methodological contribution of this paper is an extension of the robustness analysis proposed by Onatski and Stock (1999) to forward-looking reference models and uncertainty sets. The fact that the most interesting models of the economy include some forward-looking components suggests the importance of such an extension. I propose to characterize the robustness of a given policy rule by the maximal size of the uncertainty set that does not include any unstable models or models having multiple solutions under this particular rule. I modify the ideas of the robust control literature to fit the case of systems with rational expectations.

I apply theoretical results of the paper to analyze simple policy rules under model uncertainty in an empirical New Keynesian models of the US economy discussed in Rudebusch (2000). I address three sets of issues: the degree of policy activism under model uncertainty, the stabilization properties of nominal income rules, and the robustness of forecast-based rules. I find that aggressive policy rules are relatively more robust than cautious rules with respect to uncertainty about point estimates of parameters of the reference model. However, cautious rules look relatively more robust under more broadly specified uncertainty. Nominal income rules are shown to be much less robust than rules responding to inflation and the output gap. The policy rules responding to a forecast of inflation and the current output gap are found to be quite robust even for forecast horizon longer than 1 year.

There are many issues left for future analysis. First, in this paper I assume that the private sector knows the true model of economy whereas a policy maker faces model uncertainty. It would be interesting to put the private agents and policy makers on an equal footing. Second, finding exact minimax rules instead of analyzing stability and uniqueness robustness is a question of practical importance. It is also of interest to try to generalize the technique developed to non-linear and linear time-varying model uncertainty. Finally, a more detailed analysis of the empirical questions studied would be helpful.

6. Appendix

Here I explain how to cast the reference model (4.1, 4.2, 4.3) in form (2.1) and how to represent the confidence parallelepiped uncertainty sets in form (3.4).

Rewrite equations (4.1, 4.2, 4.3) of the reference model in a form:

$$\begin{aligned}
\pi_t &= c_1 E_{t-1}\bar{\pi}_{t+3} + c_2\pi_{t-1} + c_3\pi_{t-2} + c_4\pi_{t-3} + c_5\pi_{t-4} + c_6 y_{t-1} + \varepsilon_t \\
y_t &= c_7 y_{t-1} + c_8 y_{t-2} + c_9(i_{t-1} - E_{t-1}\bar{\pi}_{t+3}) + \eta_t \\
i_t &= g_i i_{t-1} + g_\pi E_t\bar{\pi}_{t+J} + g_y E_t y_{t+K}.
\end{aligned}$$

These equations can be cast in form

$$\sum_{j=0}^{m} E_{t-j} M^j(L) X_t = u_t$$

with the following parameters.

$$
\begin{aligned}
X_t &= (\pi_t, y_t, i_t)', u_t = (\varepsilon_t, \eta_t)', m = 1, \\
M^0(L) &= \begin{pmatrix} 1 - \sum_{i=1}^{4} c_{i+1} L^i & -c_6 L & 0 \\ 0 & 1 - \sum_{i=1}^{2} c_{i+6} L^i & -c_9 L \\ -\frac{g_\pi}{4} \sum_{i=0}^{3} L^{i-J} & -g_y L^{-K} & 1 - g_i L \end{pmatrix}, \\
M^1(L) &= \begin{pmatrix} -\frac{c_1}{4} \sum_{i=0}^{-3} L^i & 0 & 0 \\ \frac{c_9}{4} \sum_{i=0}^{-3} L^i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
$$

Now, let V be the 8×8 variance-covariance matrix of the point estimates of $c_2, ..., c_8$. Let U be the Choleski factor of V, that is $U'U = V$. Denote the inverse of U as $S = \{s_{ij}\}$. Define Δ^0, Δ^1, $W_i^j(L)$, $j = 0, 1$, $i = 1, 2$ as follows using Matlab notations:

$$
\begin{aligned}
\Delta^0 &= \Delta^1 = \text{diag}(\Delta_1 I_2, ..., \Delta_8 I_2), \text{ where } \Delta_i \text{ are real numbers}, \\
W_1^0 &= W_1^1 = [\text{kron}(\text{ones}(1, 8), [1\ 0]); \text{kron}(\text{ones}(1, 8), [0\ 1]); \text{zeros}(1, 16)],
\end{aligned}
$$

$$
W_2^0(\text{2k-1:2k} \quad ,: \quad) = \begin{pmatrix} -\sum_{i=1}^{4} s_{ki} L^i & -s_{k5} L & 0 \\ 0 & -\sum_{i=1}^{2} s_{k,i+5} L^i & -s_{k8} L \end{pmatrix}, \text{ where } k = 1, ..., 8,
$$

$$
W_2^1(\text{2k-1:2k} \quad ,: \quad) = \begin{pmatrix} \frac{1}{4} \sum_{i=1}^{4} s_{ki} \sum_{j=0}^{-3} L^j & 0 & 0 \\ \frac{1}{4} s_{k8} \sum_{j=0}^{-3} L^j & 0 & 0 \end{pmatrix}, \text{ where } k = 1, ..., 8.
$$

The set of models (2.2) with weighting matrices and perturbation operators defined above and such that $||\Delta^0|| = ||\Delta^1|| < r$ corresponds to the reference model with coefficients $c_1, ..., c_9$ perturbed inside a 100%*$(2\Phi(r) - 1)^8$ confidence parallelepiped as defined in Section 4.

References

[1] Ball, L. (1999) Efficient Rules for Monetary Policy. *International Finance*. Vol. 2 (1). p 63-83.

[2] Bernanke B. S. and M. Woodford (1997) Inflation Forecasts and Monetary Policy, *Journal of Money, Credit, and Banking*, Vol. 29, pp. 653-685.

[3] Brainard, W. (1967) Uncertainty and the Effectiveness of Policy. *American Economic Review* 57, 411-425.

[4] Broze Laurence, C. Gourieroux, and A. Szafarz (1995), Solutions to Multivariate Rational Expectations Models, *Econometric theory,* Vol. 11, pp.229-257.

[5] Christiano L. J. and C. J. Gust (1999) Comment on Levin, Wieland and Williams. In J. Taylor (ed.) *Monetary Policy Rules*, pp. . Chicago: University of Chicago Press.

[6] Clarida, R., Gali, J. and M. Gertler (1999) The Science of Monetary Policy: A New Keynesian Perspective. *Journal of Economic Literature*. Vol. 37 (4). p 1661-1707.

[7] Dahleh, M.A. and I.J. Diaz-Bobillo (1995) *Control of Uncertain Systems: A Linear Programming Approach*. Prentice Hall: Englewood Cliffs, NJ.

[8] Fuhrer J. and Moore G. (1995) Inflation Persistence. *Quarterly Journal of Economics*, Vol. 110, pp.127-59.

[9] Giannoni M. (1999) Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a forward-looking Model. Princeton University. Working paper.

[10] Hansen L. and T. Sargent (2000) *Elements of Robust Control and Filtering for Macroeconomics*, Manuscript.

[11] Kuttner K. (1994) Estimating Potential Output as a Latent Variable, Journal of Business & Economic Statistics, Vol. 12, pp.361-368.

[12] Levin A., V. Wieland and J. Williams (1999) The Performance of Forecast-Based Monetary Policy Rules under Model Uncertainty. Working paper, Board of Governors of Federal Reserve System.

[13] Levin A., V. Wieland and J. Williams (1999a) Robustness of simple monetary policy rules under model uncertainty. In J Taylor (ed.) *Monetary Policy Rules*, pp. . Chicago: University of Chicago Press.

[14] Mankiw N. G. and R. Reis (2001) "Sticky Information vs. Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve", NBER Working paper 8290.

[15] McCallum, B.T. (1988) Robustness Properties of a Rule for Monetary Policy. *Carnegie-Rochester Conference Series on Public Policy* 29, 175-203.

[16] McCallum, B.T. (1997) The Alleged Instability of Nominal Income Targeting. NBER Working Paper 6291.

[17] McCallum, B.T. (1999) Issues in the Design of Monetary Policy Rules. in John B. Taylor and Michael Woodford (eds.) *Handbook of Macroeconomics*, Amsterdam, North-Holland.

[18] McCallum B. and E. Nelson (1999) Performance of Operational Policy Rules in an Estimated Semiclassical Structural Model. In J. Taylor (ed.) *Monetary Policy Rules*, pp. 15-54, Chicago: Chicago University Press.

[19] Orphanides A. (1999) The Quest for Prosperity without Inflation. Manuscript, Federal Reserve Board.

[20] Orphanides A., V. Wieland (1999) Efficient Monetary Policy Design Near Price Stability, Working paper. Federal Reserve Board.

[21] Onatski A., Stock J. (1999), Robust Monetary Policy under Model Uncertainty in a Small Model of the US Economy. Forthcoming in *Macroeconomic Dynamics.*

[22] Onatski A. (2000) Minimax Monetary Policy: Comparison to Bayesian Approach, Worst Cases, and Exact Minimax Rules. Forthcoming in *Robust Decision Theory in Economics and Finance*, Cambridge University Press.

[23] Onatski A., (2001), A Geometric Criterion for Existence (Uniquness) of Equilibrium in Linear Rational Expectations Models. Manuscript.

[24] Paganini F. (1996) Sets and Constraints in the Analysis of Uncertain Systems. Ph.D. thesis. California Institute of Technology.

[25] Rudebusch, G.D. 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.

[26] Rudebusch, G.D. (2000) Assessing Nominal Income Rules for Monetary Policy with Model and Data Uncertainty. Working paper. Federal Reserve Bank of San Francisco.

[27] Rudebusch, G.D. (2000a) Is the Fed Too Timid? Monetary Policy in an Uncertain World. *Review of Economics and Statistics.*

[28] Sargent, T. (1999) Comment on Ball. In J. Taylor (ed.), *Monetary Policy Rules*, pp. 144-154. Chicago: University of Chicago Press for the NBER.

[29] Stock, J.H. (1999) Comment on Rudebusch and Svensson. In J. Taylor (ed.), *Monetary Policy Rules*, pp. 253-259. Chicago: University of Chicago Press for the NBER.

[30] Svensson, L.E.O. (1997) Inflation Forecast Targeting: Implementing and Monitoring Inflation Targets. *European Economic Review* 41, 1111-1146.

[31] Svensson, L.E.O. (1999) Inflation Targeting: Some Extensions. *Scandinavian Journal of Economics*. Vol. 101 (3). p 337-61.

[32] Taylor, John B. (1993) Discretion versus Policy Rules in Practice. *Carnegie-Rochester Conference Series on Public Poicy* 39, 195-214.

[33] Taylor, John B. (ed.) (1999) *Monetary Policy Rules.* Chicago: University of Chicago Press.

[34] Taylor, John B. (2000) Reassessing Discretionary Fiscal Policy. *Journal of Economic Perspectives*, Vol. 14, No. 3, 21-36.

[35] Tetlow R. and P. von zur Muehlen (2000) Robust Monetary Policy with Misspecified Models: Does Model Uncertainty Always Call for Attenuated Policy? Working paper. Federal Reserve Board.

[36] Whiteman C.H. (1983) *Linear Rational Expectations Models: A User's Guide.* University of Minnesota press.

[37] Woodford, M. (1999) Commentary: How Should Monetary Policy Be Conducted in an Era of Price Stability? Forthcoming, Federal Reserve Bank of Kansas City, New Challenges for Monetary Policy. Proceedings of a symposium held in Jackson Hole, Wyoming, August 1999.