Minimax Analysis of monetary policy under model uncertainty

Alexei Onatski,
Economics Department of Harvard University,
July 1999, Last revised January 2000.

I thank Jim Stock for getting me interested in the topic, for constant support, helpful discussions, and suggestions. I am grateful to Peter von zur Muehlen and Vladislav Karguine for interesting discussions and editorial help. All possible mistakes are my own.

Abstract:

Recently there have been several studies that examined monetary policy under model uncertainty. These studies formulated uncertainty in a number of different ways. One of the prominent ways to formulate model uncertainty is to form a non-parametric set of perturbations around some nominal model where the set is structured so that the uncertainty is focused on potentially important weaknesses of the model. Unfortunately, previous efforts were unable to compute exact optimal policy rules under this general formulation of uncertainty. Moreover, for those special cases when the robust rules were computed, the degree of their aggressiveness was often counterintuitive in light of conventional Brainard/Bayesian wisdom that policy under uncertainty should be conservative. This paper,therefore, consists of three different exercises concerning minimax analysis of policy rules under model uncertainty. First, the minimax approach is compared with the Bayesian one in a stylized Brainard (1967) setting. Strong similarities between recommendations of the two approaches are found. Next, a more realistic setting such as in Onatski and Stock (1999) is considered. A characterization of the worst possible models corresponding to the max part of the minimax scheme is given. It is shown that the worst possible models for very aggressive rules, such as the H-infinity rule, have realistic economic structure whereas those for passive rules, such as the actual Fed's policy, are not plausible. Thus, the results of minimax analysis presented in Onatski and Stock (1999) might be biased against the passive rules. Finally, exact optimal minimax policy rules for the case of slowly time-varying uncertainty in the case of the Rudebusch and Svensson's (1998) model are computed. The optimal rule under certainty turns out to be robust to moderate deviations from Rudebusch and Svensson's model.

.1 Introduction

The question of robustness of monetary policy rules to model uncertainty has recently received much attention, both from practitioners and academic researchers. Uncertainty about the workings of the economy in the new European environment, steady decline of the natural unemployment rate in the USA, and the recent Asian crisis have all contributed to this interest. Alan Blinder (1998) outlined one approach to deal with model uncertainty. Speaking from the perspective of a member of the Board of Governors of the Federal Reserve System, he suggested choosing a set of models that might be good approximations of reality, "... simulate a policy on as many of these models as possible, throw out the outlier(s), and average the rest..." The recently proposed minimax approach seems to recommend doing exactly the opposite. Keep outliers and make sure a policy works reasonably well in the worst possible case.

This paper looks closely at the minimax approach and tries to answer several questions. First, how do policy recommendations based on minimax differ from those sounded on the Bayesian approach? Do the minimax recommendations make sense in a simple Brainard (1967) setting for policy analysis? Second, what are the outlier models the minimax approach takes care of? Particularly, do these outliers have any economic meaning for the minimax analysis performed by Onatski and Stock (1999)? Finally, is computation of the optimal minimax policy feasible for general model uncertainty? If yes, what are the minimax policy recommendations?

The paper consists of three exercises corresponding to the above three groups of questions. First, I consider a static policy environment with only one policy target and one policy instrument. As Brainard's Bayesian analysis showed, uncertainty about model parameters leads to more conservative policy rules in this simple setting. I found that the optimal minimax policy also becomes less aggressive as the uncertainty rises. However, dissimilar to Bayesian case, there exist two regimes of the optimal rule's reaction to the amount of uncertainty. When the uncertainty is small the rule stays equal to the certainty equivalence rule, after certain threshold the minimax policy becomes more and more passive.[1]

The second exercise considers the framework for minimax analysis of model uncertainty proposed by Onatski and Stock (1999). Rudebusch and Svensson's (1998) two-equation macroeconometric model of the US economy is used as a core around which a non-parametric set of plausible models is built. Attention is restricted to the sets of models that could be obtained from the core by changing its parameters and/or adding arbitrarily many more lags of inflation, output gap, and real interest rate to the dynamic Phillips curve and aggregate demand equations that constitute the core model. The cumulative changes made are required to be small. Only 2-parameter policy rules of the Taylor type are considered. According to these rules, the nominal interest rate is set to be a linear combination of inflation and the output gap.

I analyzed particularly bad deviations from the Rudebusch-Svensson model that lead to instability of one of the policy maker's variables of interest: inflation, output gap, or changes in the nominal interest rate. I found that the worst possible lag structure of the deviations is exponentially decaying and showed how to calculate the rate of decay. My

[1] As is well known, for more general settings than that of Brainard, optimal Bayesian rules as well as optimal minimax rules need not to be less aggressive than the optimal certainty rule. See, for example, Chaw (1975), Sargent (1998), Stock (1998).

calculations show that the decay is implausibly slow for relatively passive policy rules and reasonably fast for relatively aggressive policy rules. The economy's worst possible impulse responses for the aggressive (passive) rules look as periodic fluctuations with ever increasing amplitude and period of about 5 (40) year. These findings indicate that the results of minimax analysis presented in Onatski and Stock (1999) might be biased against the passive rules.

Finally, the third exercise calculates exact optimal minimax rules for the case of slowly time varying uncertainty about Rudebusch and Svensson's model. The rule suggests more aggressive response to the output gap and less aggressive response to inflation than the certainty rule does. The conventional linear quadratic optimal rule reported by Rudebusch and Svensson demonstrates a moderate degree of robustness to model uncertainty.

Summing up, the analysis allows me to conclude that, first, the minimax approach, though very different methodologically from the Bayesian approach, could lead to similar policy recommendations. Second, the worst possible cases that minimax takes care of often make good economic sense. However, in Onatski and Stock (1999) setting, the worst models for passive policy rules have unrealistically long memory. Third, calculation of exact optimal minimax policy rules is feasible for a slowly time varying uncertainty specification. The optimal certainty rule for Rudebusch and Svensson's model is moderately robust to this form of uncertainty. Aggressive policy rules, though robust to uncertainty about the structure of the noise process and about parametric model uncertainty,[2] are not robust against lag structure uncertainty and might lead to frequent and increasing business cycles.

---

[2] See Sargent (1998), Stock (1998)

The remainder of the paper is organized as follows. Section 2 analyzes optimal minimax policy rules in a simple Brainard (1967) setting. Section 3 describes the worst possible deviations from the Rudebusch-Svensson model. In section 4, I compute exact minimax optimal rules for lag structure uncertainty about the core model. Section 5 concludes.

.2 Comparison of minimax and Bayesian policy implications

The minimax approach to policy analysis under uncertainty is methodologically very different from the Bayesian approach. The latter treats uncertainty as stochastic indeterminance summarized in a prior distribution. It recommends a policy rule that minimizes expected posterior risk. The minimax approach treats uncertainty as ignorance. It assumes that a policy maker faces a whole set of possible alternatives and minimizes a loss, assuming the worst possible realization from the set. In the Bayesian case, the size of uncertainty is associated with such characteristics of the prior distribution as its standard deviation. In the minimax case, it is associated with the diameter of the set of possible alternatives.

The difference between the two approaches suggests that their policy recommendations may differ substantially. This section compares the recommendations in a simple Brainard setting. Somewhat surprisingly, it finds a similarity between optimal minimax and Bayesian policy rules.

I consider here the case of one target and one instrument, to make the analysis as simple and transparent as possible. Assume that a variable of the policy maker's interest, Y, is equal to $aP + u$, where P is a policy variable, u is an exogenous variable, and a is a multiplier determining the degree to which policy affects the target variable. Let the policy maker's utility function be $-(Y-Y^*)^2$, where Y^* is some desired level of Y. In a world of certainty, the optimal policy would be

$$P^* = (Y^* - u)/a. \tag{.1}$$

The policy maker may, however, face uncertainty about the model. This uncertainty may come from two sources. On the one hand, the policy maker may be uncertain about effect of the exogenous variable, u, on Y. On the other hand she might be uncertain about the value of the coefficient a. This uncertainty may be modeled by assuming that a and/or u are random variables. If only u is uncertain, then the policy maker is in a situation of certainty equivalence. To formulate the optimal policy, she could consider u as being equal to its expectation and proceed as in the certainty case. The optimal policy will be:

$$P^* = (Y^* - \bar{u})/a, \tag{.2}$$

where $\bar{u}$ is the mean value of u.

As Brainard showed in his paper, in the case when a is uncertain, a less aggressive policy is optimal. More precisely, if a and u are random variables with variances σ_a^2, σ_u^2, means $\bar{a}, \bar{u}$, and coefficient of correlation, ρ, then the following rule is optimal:

$$P^* = \frac{\bar{a}(Y^* - \bar{u}) - \rho\sigma_a\sigma_u}{\bar{a}^2 + \sigma_a^2}. \tag{.3}$$

In the simplest case when $\rho = 0$ the formula becomes:

$$P^* = \frac{\bar{a}(Y^* - \bar{u})}{\bar{a}^2 + \sigma_a^2}. \tag{.4}$$

One can see that the optimal policy under uncertainty (.4) is, indeed, less aggressive than the optimal policy under certainty equivalence (.2). The reason behind this "cautiousness" is simple. The expected squared error, $(Y - Y^*)^2$, that one would like to minimize, consists of two parts. One part is the squared deviation of the average of Y from the target Y^*, and the other part is the variance of Y. The policy that makes the former part minimal is the optimal policy under certainty equivalence. The latter part of the expected error is made minimal by a zero policy. The optimal policy is in between, with the precise position depending on the relative importance of the two parts of the expected squared error captured by the coefficient of variation, $\sigma_a/\bar{a}$, of a.

In the minimax setting, uncertainty about a and u is modeled differently. The policy maker considers a and u as non-random but unknown quantities that belong to a given set Ω.[3] The size and the shape of this set has the same function as the variances σ_a^2 and σ_u^2 and the correlation ρ, in the sense that they define a degree of "dispersion" and "interdependence" of possible values of a and u. The position of the set gives information about a and u similar to what one gets from the mean values $\bar{a}$ and $\bar{u}$.[4]

For example, on the one hand, we can model uncertainty about a and u by saying that a and u are independent uniformly distributed random variables with means $\bar{a}, \bar{u}$, and variances σ_a^2, σ_u^2; on the other hand, we can similarly say that a and u are unknown numbers from the support of the above distributions, that is from intervals

$$\Omega_a = [\bar{a} - \delta_a, \bar{a} + \delta_a] \qquad and \qquad \Omega_u = [\bar{u} - \delta_u, \bar{u} + \delta_u], \tag{.5}$$

---

[3] Such uncertainty is called Knightian uncertainty.

[4] In principle, if u plays the role of noise in the model and one has good reasons to believe that it should be modeled as a random variable, then it is possible to stay with this assumption in the minimax design and consider only a as being a non-random unknown quantity.

where $\delta_a = \sqrt{3}\sigma_a, \delta_u = \sqrt{3}\sigma_u$. The minimax problem that the policy maker faces is to find a policy P that minimizes maximum loss for a and u from Ω_a and Ω_u,that is:

$$\min_P \max_{a\in\Omega_a, u\in\Omega_u} (aP + u - Y^*)^2. \tag{.6}$$

In the following, I first find the minimax optimal policy for this particular specification of uncertainty. After that, I will formulate a result for a more general specification of Ω.

Let us, first, consider the case when only u is unknown. The squared error subject to minimax analysis is a quadratic convex function of u; therefore it attains its maximum in one of the extreme points of Ω_u. We can rewrite the minimax problem as follows:

$$\min_P \max\{(aP + \bar{u} + \delta_u - Y^*)^2, (aP + \bar{u} - \delta_u - Y^*)^2\},$$

or, equivalently,

$$\min_P \left\{(aP + \bar{u} - Y^*)^2 + \delta_u^2 + |2\delta_u(aP + \bar{u} - Y^*)|\right\}.$$

A policy, $P = (Y^* - \bar{u})/a$, minimizes both the first and the third terms of the above expression, therefore, it is optimal. Note that this optimal policy has the same form as the optimal policy (.2) in the situation of certainty equivalence in the previous section: minimax and certainty equivalent strategies are therefore the same in the case of additive uncertainty.

Consider now the case when both a and u are uncertain. As before, note that the squared error, subject to minimax, is a quadratic convex function of u and of a. Therefore the minimum must be attained in one of the four points: $(\bar{a} \pm \delta_a, \bar{u} \pm \delta_u)$. Consider the situation when $Y^* < \bar{u}$ and $\delta_a < |\bar{a}|$.

Figure 1 shows the space of variables a and u. The line $aP + u - Y^* = 0$ represents the set of points, a and u, for which the squared error is equal to zero. For any point, a_1, u_1, the corresponding square error is equal to the square of the vertical distance from the point to the line. When P changes, the line rotates around point $(0, Y^*)$. To solve the minimax problem, we need to find a P such that the maximum of vertical distances from the points $(\bar{a} \pm \delta_a, \bar{u} \pm \delta_u)$ to the line is minimal.

For the situation under consideration when P is positive and large, the worst that could happen is a considerable overshooting of the target Y^*. Accordingly, the maximally "distant" point is $(\bar{a} + \delta_a, \bar{u} + \delta_u)$. When P decreases, the maximal distance start to decrease. When P crosses zero, the "maximally distant" attribute switches to the point $(\bar{a} - \delta_a, \bar{u} + \delta_u)$. Indeed, as policy becomes negative, considerable overshooting will happen, if the policy multiplier is small and u is big.

The maximal distance is minimum for the line depicted in the figure. This happens when the vertical distance from the point $(\bar{a} - \delta_a, \bar{u} + \delta_u)$ becomes equal to the vertical distance from the point $(\bar{a} + \delta_a, \bar{u} - \delta_u)$. The overshooting and undershooting situations are balanced so that $(\bar{a} - \delta_a)P + \bar{u} + \delta_u - Y^* = -(\bar{a} + \delta_a)P - \bar{u} + \delta_u + Y^*$, that is the optimal P is equal to $(Y^* - \bar{u})/\bar{a}$. One can see that the optimal minimax policy rule is equal to the certainty rule for the uncertainty specification under consideration and when $Y^* < \bar{u}$ and $\delta_a < |\bar{a}|$.

More generally, let Ω consists of all points (a, u) such that

$$\sigma_u^2(a - \bar{a})^2 - 2\rho\sigma_u\sigma_a(a - \bar{a})(u - \bar{u}) + \sigma_a^2(u - \bar{u})^2 \leq r\sigma_u^2\sigma_a^2(1 - \rho^2). \qquad (.7)$$

This set is an ellipse with the center in $(\bar{a}, \bar{u})$, such that the "dispersions" of possible a and u are proportional to σ_a^2 and σ_u^2, respectively, and the "interdependence" of possible a and u is regulated by a coefficient ρ. Of course, Ω would have been just a confidence ellipse had a and u been normally distributed with variances, σ_a^2 and σ_u^2, and correlation, ρ. This choice of Ω is natural and intuitive and might well summarize scarce knowledge about a and u that a policy maker might possess. The following proposition is true.

Proposition 1. *If the size of uncertainty, r, is less than $(1+k)\bar{a}^2/\sigma_a^2$, where $k = (1-\rho^2)/\left(\frac{\sigma_a}{\sigma_u}\frac{(\bar{u}-Y^*)}{\bar{a}} - \rho\right)^2$, then the optimal minimax policy is:*

$$P = \frac{(Y^* - \bar{u})}{\bar{a}}. \tag{.8}$$

Otherwise, if ρ is less (greater) than $\frac{\sigma_a}{\sigma_u}\frac{(\bar{u}-Y^*)}{\bar{a}}$, then the optimal minimax policy is:

$$P = -\frac{\sigma_u}{\sigma_a}\left(\rho \pm \bar{a}\sqrt{\frac{1-\rho^2}{\sigma_a^2 r - \bar{a}^2}}\right). \tag{.9}$$

The proof of the proposition is along the lines described above and is omitted.

Figure 2 plots optimal Bayesian (dashed line) and optimal minimax rules for different sizes of uncertainty, r. To calculate Bayesian rules, I assume that a and u are distributed uniformly inside ellipse (.7).[5]

Note that (.7) implies that the sign of the policy effect, a, is unambiguous if and only if the size of uncertainty, r, is less than $\bar{a}^2/\sigma_a^2$. The parameter, k, in the proposition is always positive. Therefore, if there is no uncertainty about the sign of the policy effect, then the minimax approach recommends using the certainty equivalent rule. This recommendation corresponds to the flat portion of the minimax policy lines at figure 2. If uncertainty becomes larger than certain threshold, P becomes less and less aggressive as r rises.

The intuition behind proposition 1 is as follows. The optimal minimax policy tries to balance between two equally bad cases. One of the bad scenarios is overshooting the target

[5] Three cases are considered: $\rho = 0.5, \rho = 0$, and $\rho = -0.5$. The rest of the parameters for all of the cases are: $\frac{Y^*-\bar{u}}{\bar{a}} = \frac{\sigma_a}{\sigma_u} = \frac{\sigma_a}{\bar{a}} = 1$.

Y^*, the other is undershooting Y^*. When uncertainty is relatively small, a deviation from the certainty equivalence policy improves one of the bad cases but worsens the other, so one should stay at the certainty equivalence. However, when the sign of the policy multiplier becomes uncertain there is room for improvement in both cases.

Indeed, consider the situation when the gap, $Y^* - \bar{u}$, is positive and $\rho = 0$. The two worst possible cases are: very large a and u on the one hand and very small a and u on the other hand. When the sign of a becomes uncertain, it might become negative in the latter bad case. Therefore, reducing P will improve not only the overshooting case but also the undershooting one.

In case when $\rho \neq 0$ the implications of formula (.9) are parallel to those of formula (.3) for the optimal Bayesian rule. For example, similar to Brainard's (1967) findings, if the gap $Y^* - \bar{u}$ is positive and ρ becomes positive, then it pays for a policy maker to use a less active policy P. If ρ is sufficiently positive, the policy maker might even "go in the wrong direction" making policy P negative, which is the same result as in Brainard (1967). Such a situation is illustrated by the upper graph of figure 2. In the limit, when uncertainty about a and u goes to infinity, the optimal minimax policy becomes equal to the optimal Bayesian policy.

To summarize, the minimax policy recommendations turn out to be similar to the Bayesian recommendations in the simple setting considered above. The important difference between the minimax optimal policy and the Bayesian one is that the former stays equal to the certainty equivalent rule if there is no uncertainty about the sign of the policy multiplier, a. Only after the uncertainty becomes large enough, the minimax policy rules become less active and equal to Bayesian rules in the limit when uncertainty goes to infinity.

.3 The worst possible cases

The rest of the paper is concerned with minimax analysis of policy rules in the framework described by Onatski and Stock (1999). It is assumed that a policy-maker views a model, M_0, as a core or nominal model of the economy. She believes, however, that the model is only an approximation to a truer model. She considers a non-parametric set of models, $\{M_\Delta\}$, as a set of possible alternatives to M_0. Models of this set are indexed by the operator, $\Delta \in D$, corresponding to the difference in dynamics described by M_Δ and M_0. The goal of the policy-maker is to choose a rule, P, from a set, $\{P\}$, that minimizes expected loss (risk), R, assuming that the worst possible model, M_Δ, is the true model:

$$\min_{\{P\}} \sup_{\Delta \in D} R(P, M_\Delta). \tag{.10}$$

In this paper, I consider the Rudebusch-Svensson (1998) model to be the nominal model of the economy. It consists of two equations, estimated econometrically with U.S. data:

$$\begin{aligned}
\pi_{t+1} &= \underset{(.08)}{.70}\ \pi_t - \underset{(.10)}{.10}\ \pi_{t-1} + \underset{(.10)}{.28}\ \pi_{t-2} + \underset{(.08)}{.12}\ \pi_{t-3} + \underset{(.03)}{.14}\ y_t + \varepsilon_{t+1} \\
y_{t+1} &= \underset{(.08)}{1.16}\ y_t - \underset{(.08)}{.25}\ y_{t-1} - \underset{(.03)}{.10}\ (\overline{i_t} - \overline{\pi}_t) + \eta_{t+1}
\end{aligned}$$

where 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. Variables $\overline{\pi}$ and $\bar{i}$ stand for four-quarter moving averages of inflation and federal funds rate, respectively. The coefficients on lagged inflation in the first equation sum to one, so the long run Phillips curve is assumed to be vertical.

It is assumed that a policy-maker can control the federal funds rate using a simple Taylor-type rule:

$$i_t = g_\pi \overline{\pi}_t + g_y y_t.$$

The policy maker's loss is as in Rudebusch and Svensson:

$$L_t = \pi_t^2 + y_t^2 + 1/2(i_t - i_{t-1})^2,$$

so the risk R is equal to $Var(\pi_t) + Var(y_t) + 1/2Var(i_t - i_{t-1})$ where expectation is taken over the shocks ε and η uncertainty.

Let L denote the lag operator and $A(L)$ denote a lag polynomial. Then the nominal model could be rewritten in the following form:

$$\begin{aligned} \pi_{t+1} &= A_{\pi\pi}(L)\pi_t + A_{\pi y}(L)y_t + \varepsilon_{t+1}, \\ y_{t+1} &= A_{yy}(L)y_t - A_{yr}(L)(\overline{i_t} - \overline{\pi}_t) + \eta_{t+1}. \end{aligned}$$

It is reasonable to believe that the dynamic links between past inflation and the output gap on the one hand and present inflation and output gap on the other are undermodeled. This might be the case because some important variables, such as, for example, the exchange rate, are omitted, or not all relevant lags are included, to mention only the most obvious reasons. One could, therefore, believe that a truer model has a form,

$$\pi_{t+1} = (A_{\pi\pi}(L) + \Delta_{\pi\pi})\pi_t + (A_{\pi y}(L) + \Delta_{\pi y})y_t + \varepsilon_{t+1}, \tag{.11}$$

$$y_{t+1} \quad = \quad (A_{yy}(L) + \Delta_{yy})y_t - (A_{yr}(L) + \Delta_{yr})(\overline{i_t} - \overline{\pi}_t) + \eta_{t+1}, \qquad (.12)$$

where Δ_{ij} are some causal dynamic operators.

Equations (.11, .12) constitute a perturbed model, M_Δ, indexed by a perturbation, Δ, from the set

$$D_r = \{\Delta : \Delta = diag\{\Delta_{ij}/\rho_{ij}\}, \|\Delta\| \leq r\},$$

so that the perturbations of each particular channel, Δ_{ij}, have norms, less than $r\rho_{ij}$. If the policy-maker believes that the nominal model is a good approximation to the truth, then the size, r, of the perturbation Δ could be chosen to be small. The scaling factors, ρ_{ij}, regulate relative importance of the individual channel perturbations, Δ_{ij}, and will be ignored in the following for simplicity of notations.

The operators from the set, D_r, can be quite broadly defined, ranging from constant operators (multiplication by a constant diagonal matrix) to linear time invariant operators (that could be represented by a diagonal matrix of infinite L polynomials) or even to non-linear time-varying operators. In this section of the paper, I restrict attention to linear time invariant perturbations, $\Delta(L)$. The norm $\|\Delta\|$ is taken to be the H_∞ norm, that is $\|\Delta\| = \sup_{|z|\leq 1} \sigma\left(\Delta(z)\right)$, where σ is the largest singular value of the matrix $\Delta(z)$. In the case under consideration, $\Delta(z)$ is a diagonal matrix so the largest singular value is equal to $\max_{ij} \sup_{|z|\leq 1} \left|\Delta_{ij}(z)\right|$.

Computation of the exact minimax optimal rules, P, satisfying (.10) will be considered in the next section. Here I will concentrate on the analysis of particularly bad deviations, Δ, that destabilize the economy for a given policy rule, P.

Suppose that a policy rule, P, stabilizes all models, M_Δ, such that $\Delta \in D_{r_0}$, but for any $r > r_0$, there exists a $\Delta_r \in D_r$, such that the model, M_{Δ_r}, is unstable under P. Let us call the number, r_0, the radius of affordable perturbations for the rule P. Stability of M_Δ is equivalent to the condition that the characteristic roots of the model's MA representation, $M_\Delta(L)$, lie inside the unit circle. If the core model is stable then all roots of $M_0(L)$ lie inside the circle. When $\|\Delta\|$ starts to rize the roots might come closer and closer to the boundary until one of them hits the unit circle at some point $e^{i\omega_0}$ when $\Delta = \Delta_{r_0}$. In fact, because $M_\Delta(L)$ must have real coefficients, a pair of complex conjugated roots hit the circle simultaneously at $e^{i\omega_0}$ and $e^{-i\omega_0}$, where $\omega_0 \in [0, \pi)$.

A method of finding radii of affordable perturbations was outlined in Onatski and Stock (1999) where such radii were found for Taylor-type rules. Those interested in details can find them in the above mentioned paper and references therein. A byproduct of the method is the destabilizing frequency, ω_0, and the value of $\Delta_{r_0}(z)$ at $z = e^{i\omega_0}$, $\Delta_{r_0}(e^{i\omega_0}) = diag\{r_{ij}e^{i\omega_{ij}}\}$.[6] Note that $r_{ij} \leq r_0$, because the H_∞ norm of Δ_{r_0} must be equal to r_0.

Figure 3 summarizes Onatski and Stock's findings. One can see that the highest radii of affordable perturbations are associated with policy rules with small reaction to inflation and medium reaction to output gap. The star at the picture denotes a policy rule with $g_\pi = 1.5$ and $g_y = 0.5$ that was proposed by Taylor as a rule approximating the Fed's rule well. The square denotes a policy rule that is optimal when there is no uncertainty about the Rudebusch-Svensson model.

[6] Strictly speaking, the method only allows to find upper and lower bounds on the radius that turn out to be very close each to other for the problem at hand. The destabilizing frequency corresponds to the upper bound for the radius.

In the rest of this section of the paper I would like to utilize the knowledge of the destabilizing frequency and the “value” of the perturbation at this frequency to find the destabilizing perturbation itself. It is especially interesting to know if the corresponding perturbed model, which is the worst possible model for a given policy rule, has any economic sense.

Formally, I am looking for causal time invariant operators, Δ_{ij}, having a representation in the form of infinite L polynomials with real coefficients and such that $\|\Delta_{ij}\| \leq r_0$, and $\Delta_{ij}(e^{i\omega_0}) = r_{ij}e^{i\omega_{ij}}$.

It is easy to build such operators. Assume first that $\omega_{ij} \in [0, \pi)$, and consider operators having the so called Blashke form

$$\Delta_{ij}(L) = r_{ij}\frac{L - x_{ij}}{1 - x_{ij}L}, \tag{.13}$$

where $x_{ij} \in (-1, 1)$. Clearly, $\Delta_{ij}(z)$ is an analytic function inside the unit circle with absolute value r_{ij} on the circle. Thus, $\|\Delta_{ij}\| = r_{ij} \leq r_0$. . When x_{ij} goes from -1 to 1, the value of $\Delta_{ij}(z)$ at $e^{i\omega_0}$ goes from r_{ij} to $-r_{ij}$ counterclockwise along the circle with radius r_{ij}. Define x_{ij} as

$$x_{ij} = \frac{e^{i\omega_0} - e^{i\omega_{ij}}}{1 - e^{i\omega_0}e^{i\omega_{ij}}}.$$

It is not difficult to see that x_{ij} are indeed between -1 and 1, and that $\Delta_{ij}(e^{i\omega_0}) = r_{ij}e^{i\omega_{ij}}$.

If $\omega_{ij} \in (-\pi, 0)$ then the sign of the right hand side in (.13) must be changed to ”–” and x_{ij} must be defined as

$$x_{ij} = \frac{e^{i\omega_0} + e^{i\omega_{ij}}}{1 + e^{i\omega_0}e^{i\omega_{ij}}}.$$

In both cases, the worst possible perturbations, Δ, have ARMA(1,1) form and, therefore, exponentially decaying lag structure. An immediate implication of this result is that unreasonable perturbations that emphasize the importance of, say, the one thousandth lag in the model equations are not the worst. The worst possible operator (.13) can be expressed in the form of an infinite L polynomial as follows:

$$\Delta_{ij}(L) = r_{ij}\left(-x_{ij} + (1-x_{ij}^2)L + (1-x_{ij}^2)x_{ij}L^2 + (1-x_{ij}^2)x_{ij}^2L^3 + ...\right). \qquad (.14)$$

One can see that the rate of decay in the lag structure of Δ_{ij} is determined by x_{ij}. Small by absolute value x_{ij} implies short memory whereas $|x_{ij}|$ close to one indicates long memory. The size of the lag coefficients in (.14) is small for both extremes, $|x_{ij}| \to 1, x_{ij} \to 0.$

I computed the worst possible Δ_{ij} for the policy rules with $g_\pi \in [1.25, 7]$ (grid of 0.25), and $g_y \in [0.125, 4.5]$ (grid of 0.125). The rates of decay, x_{ij}, and position of ω_{ij} vary substantially in the chosen range of policies. However, it is possible to distinguish between two groups of policy rules.

First, for relatively "passive" rules, very roughly, the rules with $g_\pi < 3, g_y < 2,$ the values of all individual channel perturbations Δ_{ij} at destabilizing frequency, $\Delta_{ij}(e^{i\omega_0}) = r_{ij}e^{i\omega_{ij}}$, lie in the upper half of the complex plane. That is $\omega_{ij} \in [0, \pi)$ and, therefore, all Δ_{ij} are given by (.13). Parameters x_{ij} turn out to be positive numbers close to 1. Informally, this means that the memory of the economy in the worst case is very long for these "passive" rules. Formula (.14) together with equations (.11,.12) imply that the perturbed economy's reaction to key economic variables relative to the core economy is higher in total, but more "spread out" through time so that the immediate reaction of the economy is smaller.

Second, for policy rules with very active response to the output gap, $g_y > 2,$ the values of all Δ_{ij}, except Δ_{yr}, at the destabilizing frequencies lie in the lower part of the complex plane. Therefore, $\Delta_{\pi\pi}, \Delta_{\pi y},$ and Δ_{yy} are given by (.13) with sign of the right hand side changed to "–". Values of x_{ij} are positive and quite low (lower than 0.7) for those rules with relatively low response to inflation, but, as response to inflation rises, these values become larger. For the rules with very active response to inflation, x_{ij} become close to 1. One can see from (.11,.12) and (.14) (with sign of the right hand side changed to "–") that the immediate reaction of the perturbed economy to changes in inflation and the output gap is greater than in the nominal model, whereas the total reaction of the economy is lower: the economy overshoots its long run response. Such a situation is dangerous for active rules because under these rules, the economy starts to swing from expansion to recession and back to expansion.

In what follows I restrict attention to two particularly interesting rules representing the groups of policy rules described above. The "passive" rules are represented by Taylor rule, $g_\pi = 1.5, g_y = 0.5$. The "active" rules are represented by the optimal H_∞ rule, $g_\pi = 6.42, g_y = 2.75$. The optimal H_∞ rule is an interesting rule to analyze because it is an extreme case among very aggressive rules shown to be robust to different kinds of uncertainty. Sargent (1998) showed that aggressive rules are robust to uncertainty about the nature of shock process. Stock (1998) showed that such rules are robust to uncertainty about slope of the Phillips curve and the output reaction to real interest rate in the Rudebusch-Svensson model.

The parameters x_{ij} of the worst possible perturbations, Δ_{ij}, for Taylor rule and the optimal H_∞ rule are given in the table below.

	$x_{\pi\pi}$	$x_{\pi y}$	x_{yr}	x_{yy}
Taylor rule	0.91	0.93	0.97	0.71
Optimal H_∞ rule	0.96	0.97	0.71	0.92

One can see that both worst possible perturbations are characterized by very long memory. Thus, if one excludes long memory perturbations from the set of plausible ones, the radii of affordable perturbations for the two rules might become much higher. In such a case the rules might start to look much more robust. This conjecture turns out to be false for the optimal H_∞ rule but might be true for Taylor rule.

To see this, let truncated worst possible perturbations, $\bar{\Delta}_{ij}$, corresponding to a rule P be obtained from Δ_{ij} by keeping only first three terms in (.14). The truncated worst possible perturbations are no longer destabilize the economy. However, they still destabilize the economy when scaled by a number greater than 1. If the norm of the so scaled $\bar{\Delta}_{ij}$ is only marginally higher than the radius of affordable perturbations for P then the exclusion of the long memory perturbations from the set of plausible ones does not affect the radius much.

I computed the norms of the truncated worst possible perturbations for Taylor rule and the optimal H_∞ rule.[7] For Taylor rule, the norm is 1.45 that is substantially higher than the corresponding radius of affordable deviations, 1.01. On the contrary, for the optimal H_∞ rule, the norm is 0.95 that is only marginally higher than the radius, 0.91.

Figures 4 and 5 show the impulse responses of inflation and the output gap for Taylor rule and the optimal H_∞ rule respectively. The solid line represents the impulse responses in the case of nominal model. The dashed line represents the impulse responses for the "truncated" perturbed model. The fluctuations caused by impulse shocks in the "truncated"

[7] I did not truncate Δ_{yy} for Taylor rule and Δ_{yr} for H_∞ rule because for $x_{ij} = 0.71$ half decay of the corresponding polynomial's coefficients happens after three quarters, which is rather fast.

worst possible economy regulated by Taylor rule have very long period, about 40 years. The corresponding fluctuations for the optimal H_∞ rule have period of only 5 years.

To summarize, the analysis of the worst possible cases shows that the corresponding worst possible perturbations have exponentially decaying lag structure. Furthermore, the analysis suggests that the policy rules could be roughly divided into two groups, "passive" and "active" rules. The worst possible economies regulated by the "active" rules could be characterized by larger than estimated immediate response to inflation and the output gap which is reversed later, so that the economy swings into frequent ever increasing business cycles. On the contrary, the worst possible perturbations for "passive" rules are characterized by less than estimated immediate response of the economy to the key variables. However, this response is substantially "spread out" through time, so that the perturbed economy has unrealistically long memory resulting in very long destabilizing cycles.

.4 Optimal minimax rules

In this section I return to the minimax problem (.10). For each policy, P, and each size of uncertainty, r, I would like to find a maximum risk

$$\max_{\Delta \in D_r} R(P, M_\Delta).$$

A rule that provides the minimum of this maximum is the optimal minimax rule for the uncertainty represented by D_r.

There exist technical difficulties that do not permit a solution to (.10) in its exact form. The situation is, roughly, that the stochastic definition of shocks does not fit well in the minimax methodology. It is, however, possible to reformulate (.10) in a way convenient for minimax analysis. This requires modeling shock sequences not as realizations of a white noise process but as points in some set that provides tight characterization of such realizations. Below, I first describe such a characterization, recently proposed by Paganini (1996). After that, I consider a modified minimax problem.

A researcher decides that a sequence might be a realization of white noise process if she cannot reject the corresponding statistical hypothesis. It is, therefore, quite natural to assume that a sequence is close to the realization of a white noise process if it belongs to the set complementary to the rejection region of a test.

One of the tests for the white noise structure of a sequence is the Bartlett cumulative periodogram test. It consists of accumulating the periodogram of the sequence and comparing the result to a linear function. Let $v(t)$ be a sequence of length N. Denote by $V(k)$ a discrete Fourier transform of $v(t)$

$$V(k)=\sum_{t=0}^{N-1}v(t)e^{-\frac{i2\pi}{N}kt}.$$

Then the periodogram of $v(t)$ is defined by $s_v(k)=|V(k)|^2$, $k=0,...,N-1$. I will call a sequence $v(t)$ white noise with inaccuracy θ if it belongs to the following set:

$$W_{N,\theta}=\left\{v\in R^N:\left|\frac{1}{N\|v\|^2}\sum_{k=0}^{m-1}s_v(k)-\frac{m}{N}\right|\leq\theta,1\leq m\leq N\right\}.$$

The actual realizations of the white noise process belong to the set with asymptotic probability 1. More precisely, the following proposition is true.[8]

[8] See Paganini (1996), Theorem 4.8, p.58.

Proposition 2. *Let* $v(0), v(1), ..., v(N-1), ...$ *be i.i.d., zero mean, Gaussian random variables. If* $\theta_N \sqrt{N} \to \infty$ *when* $N \to \infty$, *then*

$$\Pr\left((v(0), ..., v(N-1)) \in W_{N\theta}\right) \to 1.$$

Now let $v(t)$ be an infinite square summable sequence of n-dimensional vectors. Let $s_v(\omega)$ be the spectral density of $v(t)$. Denote by $\|v\|_2^2$ the sum of squared components of $v(t)$. That is $\|v\|_2^2 = \sum_{t=-\infty}^{\infty} \|v(t)\|^2$, where the norm in the right hand side is simply a length of an n-dimensional vector. Denote by I_n an n-dimensional identity matrix. Finally, for any matrix A, denote $\|A\|_\infty = \max_{i,j} |A_{ij}|$.

As in the finite-horizon, one-dimensional case, I call the infinite sequence $v(t)$ white noise with inaccuracy θ, if it belongs to the following set:

$$W_\theta^n = \left\{ v \in l_2(R^n) : \sup_{s \in [0, 2\pi]} \left\| \int_0^s s_v(\omega) d\omega - s \|v\|_2^2 \frac{I_n}{n} \right\|_\infty < 2\pi\theta \right\}.$$

Let us view a model, $M_\Delta(P)$, controlled by a policy rule, P, as an operator transforming the 2-dimensional sequence of shocks, $v = \{(\varepsilon_t, \eta_t)'\}$[9], to the sequence of 3-dimensional vectors of target variables, $z = \left\{ \left(\pi_t, y_t, (i_t - i_{t-1})/\sqrt{2}\right)' \right\}$. Formally, we can represent this transformation by equation $z = M_\Delta(P)v$.

Let us define a "norm" of the model, $M_\Delta(P)$, presented as an operator from the set, W_θ^2, to the space $l_2(R^3)$ as

$$\|M_\Delta(P)\|_{W_\theta^2} = \left\{ \sup \|M_\Delta(P)v\|_2 : v \in W_\theta^2, \frac{1}{2} \|v\|_2^2 \le 1 \right\}.$$

The following proposition is true.[10]

---

[9] In the following, I assume that the shocks are scaled so that their variance-covariance matrix is unity.

[10] See Paganini (1996), Lemma 4.11, p 68.

Proposition 3. $\|M_\Delta(P)\|^2_{W^2_\theta} \overset{\theta\to 0}{\to} R(P, M_\Delta)$.

Note that $\|M_\Delta\|_{W^2_\theta}$ is non-decreasing with respect to θ. Therefore, proposition 3 implies that if $\sup_{\Delta\in D}\|M_\Delta(P)\|^2_{W^2_\theta} < \gamma$, then the risk, $R(P, M_\Delta)$, is less than γ for any $\Delta \in D$.

Let us denote

$$\gamma_r(P) = \inf_{\theta>0} \sup_{\Delta\in D_r} \|M_\Delta(P)\|^2_{W^2_\theta}.$$

The last observation together with proposition 3 suggests that a good policy rule would minimize function $\gamma_r(P)$. Therefore, the minimax problem (.10) could be reformulated as follows.[11]

$$\min_{P\in\{P\}} \gamma_r(P). \tag{.15}$$

There is one more modification to the minimax problem that I need to introduce. This concerns the nature of the set of perturbations D. Let us consider a set of linear slowly time varying perturbations to the nominal model M_0 :

$$D_{r\nu} = \{\Delta = diag\{\Delta_{ij}/\delta_{ij}\} : \Delta \text{ is } LTV, \|L\Delta - \Delta L\| \le \nu, \|\Delta\| < r\},$$

where the norm on the right hand side is the induced norm of a linear time varying (LTV) operator from the space, l_2, to itself. If $\nu = 0$, then the set becomes a familiar set of linear time invariant perturbations. If $\nu > 0$, then the set is wider, allowing lag specifications of, M_Δ, to slowly vary in time with the rate of variation measured by ν.

Consider the following problem:

$$\min_{P\in\{P\}} \inf_{\nu>0} \gamma_r(P). \tag{.16}$$

[11] Whether the solution to (.15) differs from that to (.10) depends on the uniformity properties of convergence in proposition 3. This question is left for future research.

The solution to this problem will yield a policy rule that is most robust against arbitrarily slowly time varying uncertainty. More precisely, let P^* be a solution to (.16). Then for any other policy rule, P, there exist a $\nu > 0$ and a $\theta > 0$, such that the maximum over $\Delta \in D_{r\nu}$ of the norm of $M_\Delta(P)$ is larger than that of the norm of $M_\Delta(P^*)$. In other words, the maximum risk over arbitrarily close to white noise sequences and arbitrarily slowly time varying perturbations of size less than r is higher for rule, P, than for rule, P^*.

I found solution to problem (.16) by computing $\gamma_r(P)$ as outlined in Paganini (1996), Proposition 6.2. Below I analyze the solution obtained. Figures 6 and 7 show $\inf_{\nu>0} \gamma_r(P)$ for policy rules of the Taylor-type for $r = 0.5$ and 1 respectively. The expression was calculated for all rules with $g_\pi \in [1.25, 7.25]$ (grid of 0.25), and $g_y \in [0.125, 4.5]$ (grid of 0.125). The star, square, and circle points in the pictures correspond to the Taylor rule, the optimal certainty rule, and the optimal H_∞ rule respectively. The isolines are marked by the corresponding levels of the worst possible risk.

For $r = 0.5$, the optimal minimax rule is $g_\pi = 2.8$, $g_y = 2.1$. The corresponding worst possible risk is just below 21. The optimal rule is not far from the certainty rule, $g_\pi = 2.7, g_y = 1.6$. The worst possible risk for the certainty rule is somewhere between 21 and 25. For comparison, the nominal model risk for the rule is just above 11. The Taylor rule and the optimal H_∞ rule have approximately equal worst possible risks, about 50. If there were no uncertainty about Rudebusch and Svensson's model, the risk for the Taylor rule would be about 17; that for the optimal H_∞ rule would be about 19.

When r doubles, (see figure 7) the optimal minimax rule becomes more responsive to the output gap, $g_y = 2.3$, and less responsive to inflation, $g_\pi = 2.3$. The worst possible risk more than doubles for the optimal rule, becoming equal to 50. The Taylor rule and especially the H_∞ rule become absolutely unacceptable, whereas the worst possible risk for the certainty rule is somewhere about 100.

The worst possible risk for the optimal minimax rule rises quickly with the size of perturbations, r. For example, for $r = 1.25$, the optimal rule is $g_\pi = 2, g_y = 2.3$. The associated worst possible risk is 102.

Many questions are left for future research. One of them is how well solution to problem (.16) approximates that to original problem (.10). Another one is how to find the worst possible perturbations. The analysis of the previous section does not apply to the present situation because here we analyze robust performance and not simply robust stability problem. It would also be interesting to analyze robustness with respect to perturbations other than linear slowly time varying ones.

.5 Conclusion

This paper describes three different exercises concerning the minimax analysis of policy rules under model uncertainty. First, the minimax approach is compared with the Bayesian one in simple Brainard's (1967) setting. Strong similarities between recommendations of the two approaches are found. Similar to the Bayesian rules, the optimal minimax rules become less aggressive when uncertainty about the policy multiplier rises. However, in contrast to the Bayesian case, there exist two regimes of the optimal minimax rule's reaction to the amount of uncertainty. When the uncertainty is small the rule stays equal to the certainty equivalence rule. The minimax policy becomes more and more passive only after the sign of the policy effect on the target becomes uncertain.

In the second exercise I analyze the worst possible deviations from Rudebusch and Svensson's model. The perturbed models differ from the nominal one in that arbitrarily more lags of inflation, the output gap, and the real interest rate are added to the Phillips curve and the aggregate demand equations, and coefficients on the existent lags may be different. I find the smallest perturbations from the described class that destabilize the economy for different Taylor-type policy rules. I show that these perturbations have an exponentially decaying lag structure, characterize the rates of decay and compute impulse-response functions for the worst possible economies for two representative rules.

The worst possible cases for aggressive rules manifest themselves in frequent and ever increasing business cycles. These worst possible cases could be characterized by a relatively high contemporaneous sensitivity of the economy to inflation and the output gap and the reaction of the output gap to the real interest rate that is more spread-out through time. The aggressive rules were previously shown to be robust against the structure of the noise process and parametric uncertainty. The analysis performed in this paper suggests that these rules are not robust to the model's lag structure uncertainty.

Finally, the paper finds optimal minimax policy rules for arbitrarily slowly time varying uncertainty. The optimal rules turn out to be less responsive to inflation and more responsive to the output gap than the optimal certainty rule.

.6 Literature

Brainard, W. (1967), "Uncertainty about the Effectiveness of Policy", *American Economic Review*, 57, 411-425.

Blinder, A. (1998), Central Banking in Theory and Practice, The MIT press.

Chow, G.C. (1975), Analysis and Control of Dynamic Economic Systems, John Wiley & Sons.

Onatski, A. and Stock J.H. (1999), "Robust Monetary Policy under Model Uncertainty in a Small Model of the US Economy", *Proceedings of Federal Reserve Bank of San Francisco*, 1, May.

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

Rudebusch, G.D. and L.E.O. Svensson (1998), "Policy Rules for Inflation Targeting", forthcoming in J. Taylor (ed.), Policy Rules for Inflation Targeting (Chicago: University of Chicago Press for the NBER).

Sargent, T. (1998), "Comment on Ball", forthcoming in J. Taylor (ed.), Policy Rules for Inflation Targeting (Chicago: University of Chicago Press for the NBER).

Stock, J.H.(1998), "Comment on Rudebusch and Svensson", forthcoming in J. Taylor (ed.), Policy Rules for Inflation Targeting (Chicago: University of Chicago Press for the NBER).

Figure 1

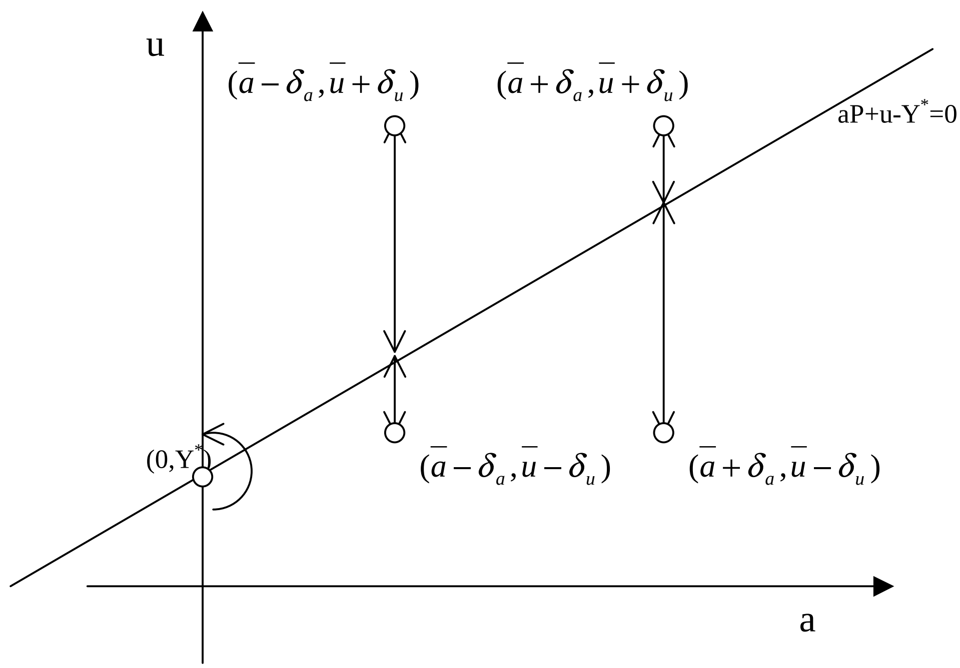

u
$(\bar{a}-\delta_a, \bar{u}+\delta_u)$
$(\bar{a}+\delta_a, \bar{u}+\delta_u)$
aP+u-Y*=0
(0,Y*)
$(\bar{a}-\delta_a, \bar{u}-\delta_u)$
$(\bar{a}+\delta_a, \bar{u}-\delta_u)$
a

Figure 2

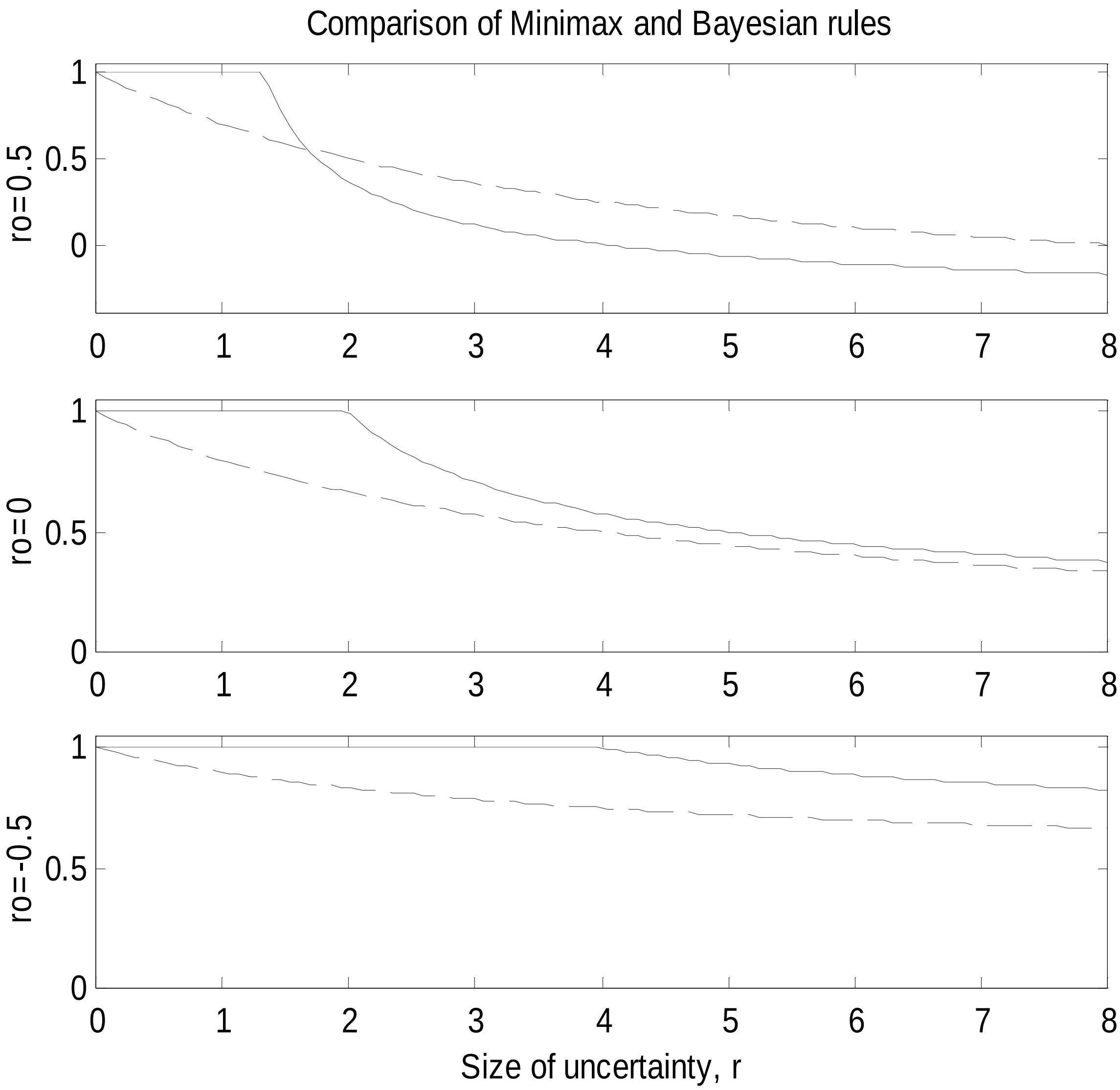

Comparison of Minimax and Bayesian rules
ro=0.5
ro=0
ro=-0.5
1
0.5
0
0
1
2
3
4
5
6
7
8
Size of uncertainty, r

Figure 3

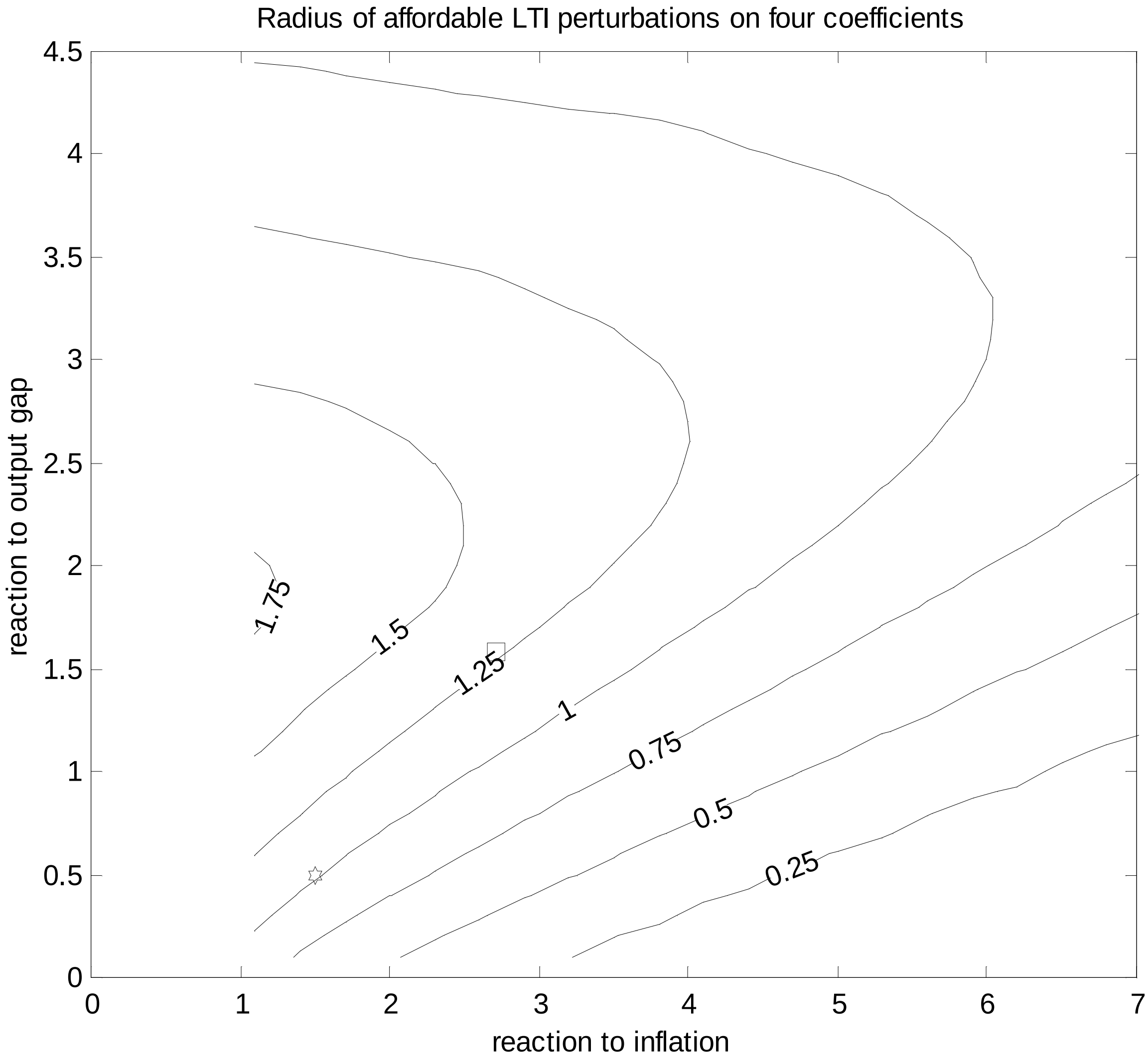
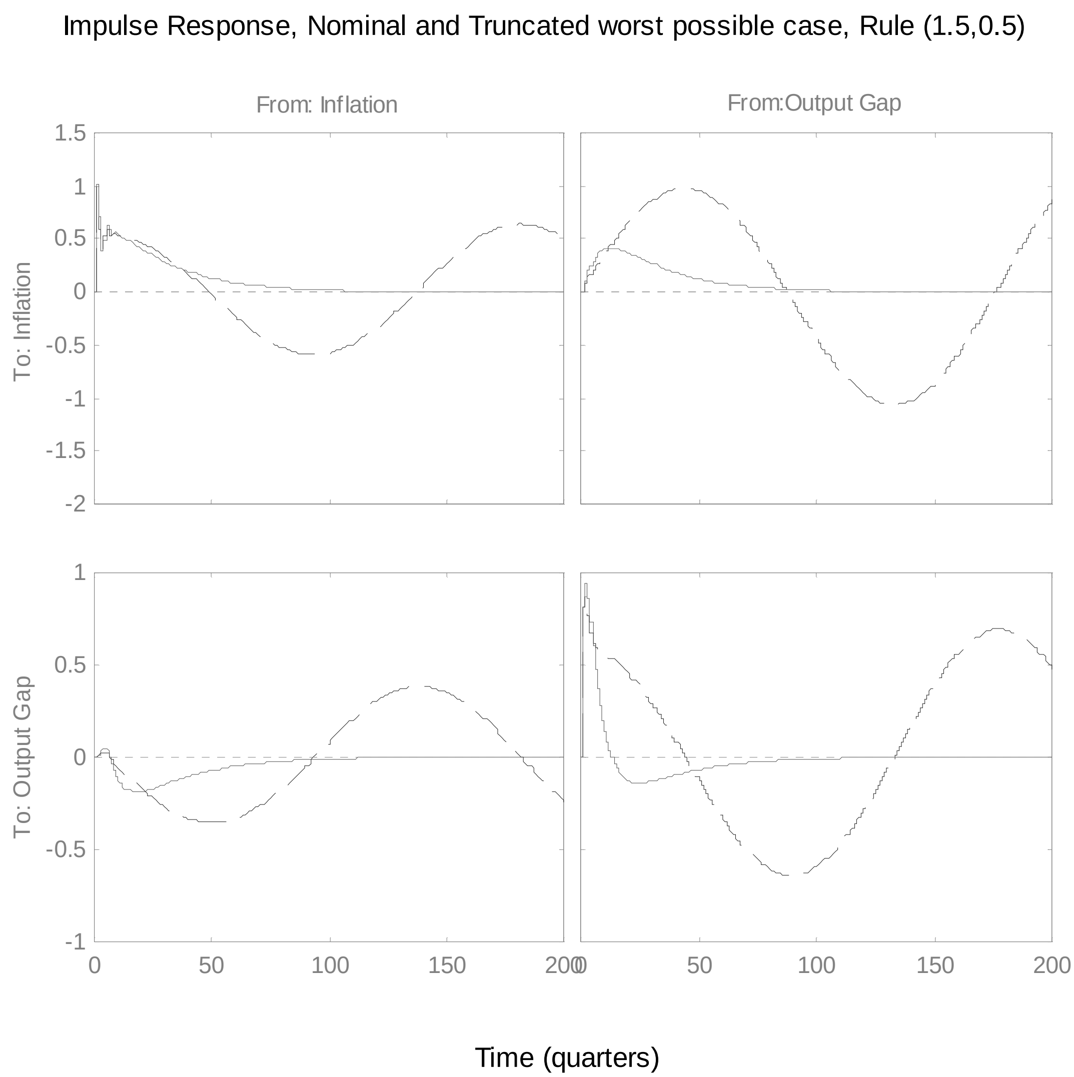

Radius of affordable LTI perturbations on four coefficients
reaction to output gap
reaction to inflation
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
0
1
2
3
4
5
6
7
1.75
1.5
1.25
1
0.75
0.5
0.25

Figure 4



Impulse Response, Nominal and Truncated worst possible case, Rule (1.5,0.5)
From: Inflation
From:Output Gap
To: Inflation
To: Output Gap
1.5
1
0.5
0
-0.5
-1
-1.5
-2
0
50
100
150
200
Time (quarters)

Figure 5

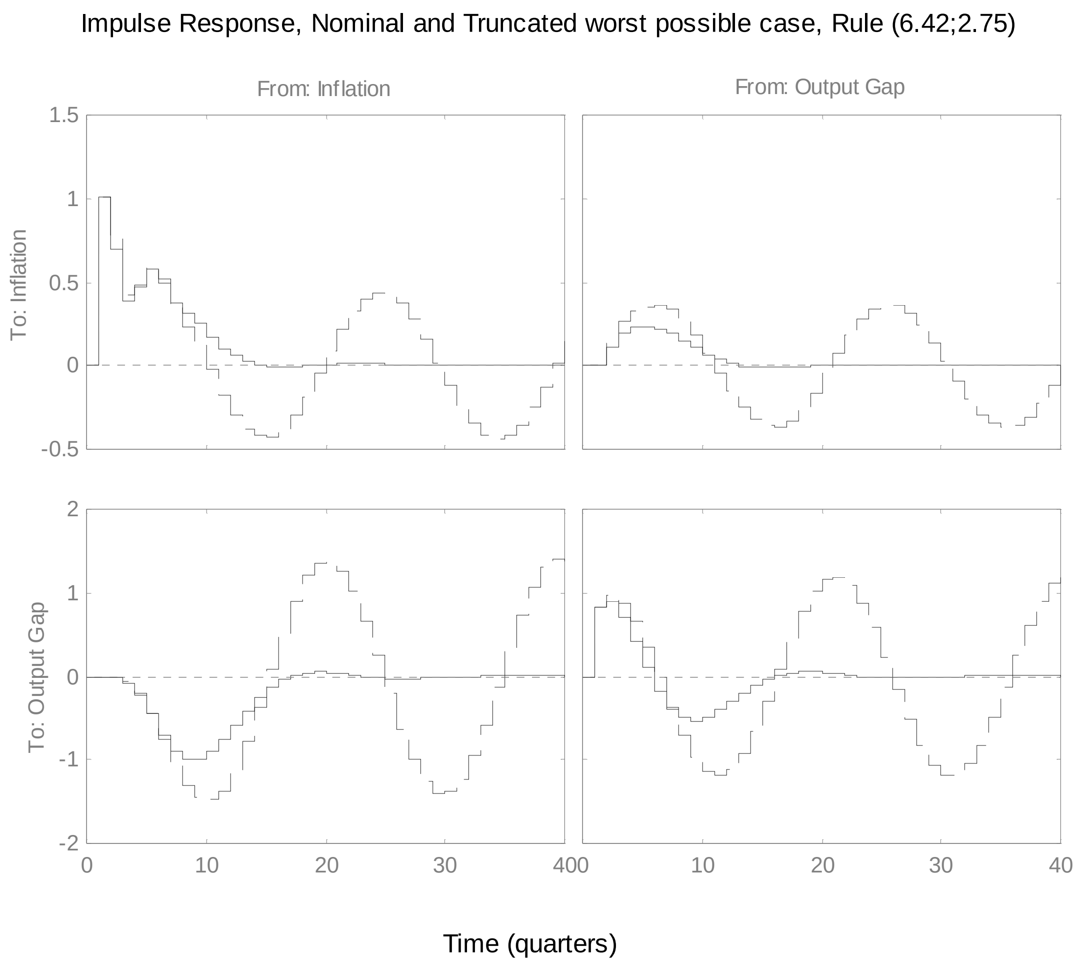
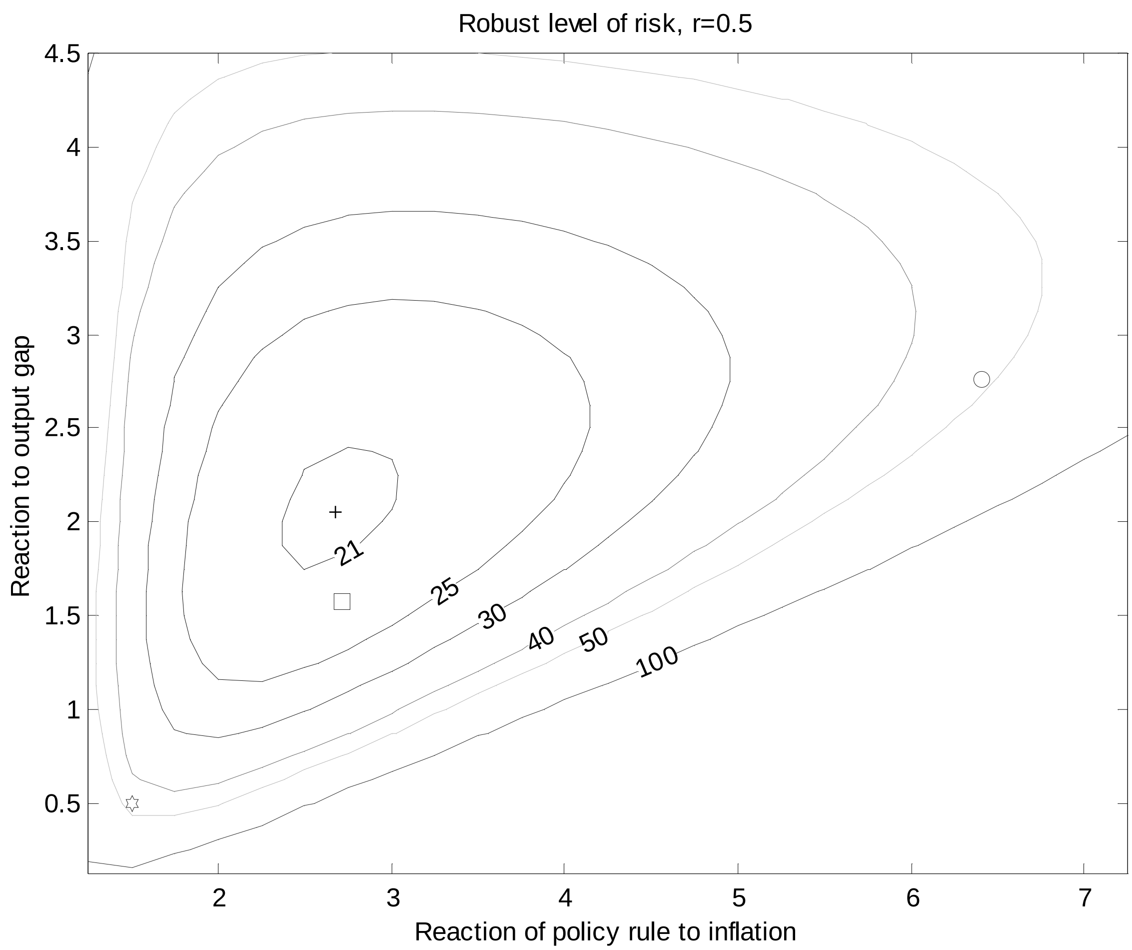
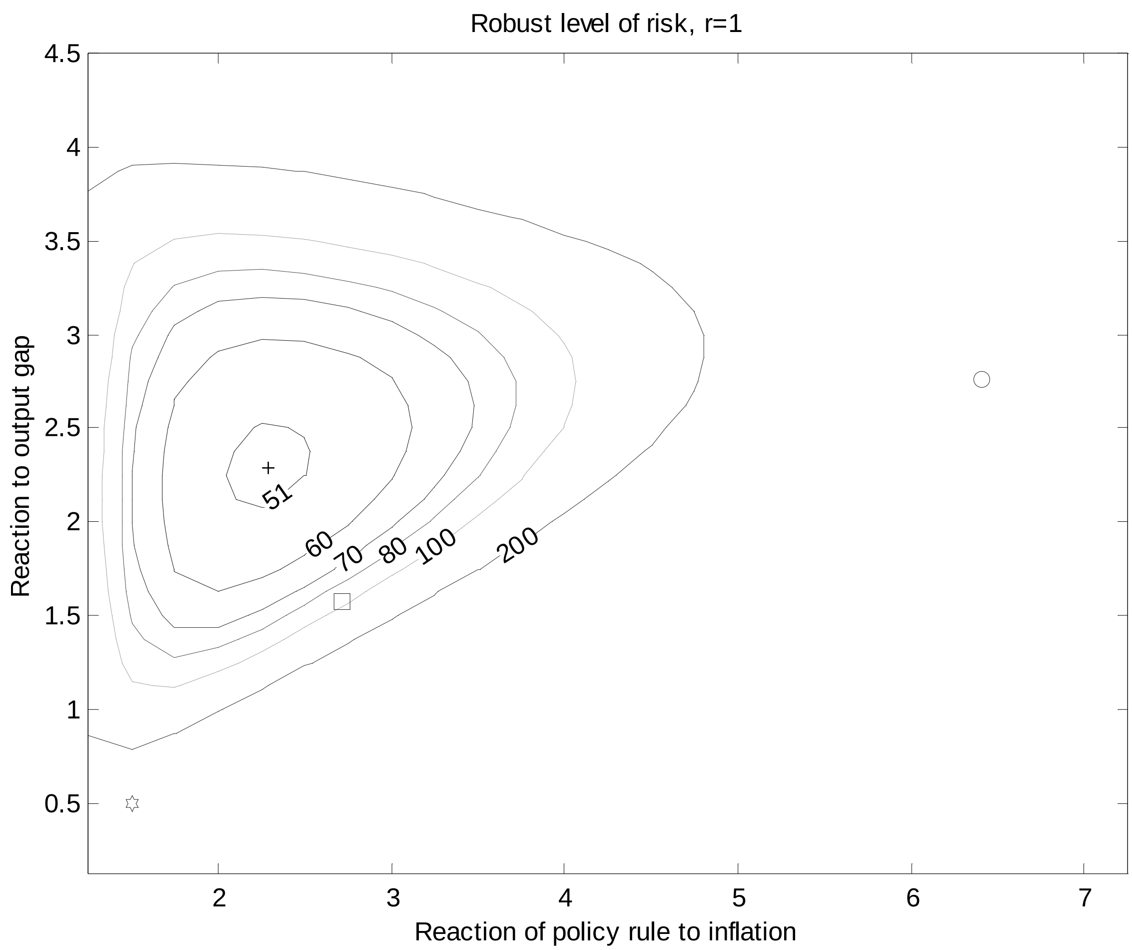

Impulse Response, Nominal and Truncated worst possible case, Rule (6.42;2.75)
From: Inflation
From: Output Gap
To: Inflation
To: Output Gap
1.5
1
0.5
0
-0.5
2
1
0
-1
-2
0
10
20
30
40
0
10
20
30
40
Time (quarters)

Figure 6



Robust level of risk, r=0.5
Reaction to output gap
Reaction of policy rule to inflation
4.5
4
3.5
3
2.5
2
1.5
1
0.5
2
3
4
5
6
7
21
25
30
40
50
100

Figure 7



Robust level of risk, r=1
4.5
4
3.5
3
2.5
2
1.5
1
0.5
2
3
4
5
6
7
Reaction to output gap
Reaction of policy rule to inflation
51
60
70
80
100
200