This paper is preliminary and incomplete.

Minimax analysis of model uncertainty: comparison to Bayesian approach, worst possible economies, and optimal robust monetary policies.

Alexei Onatski,
Department of Economics, Harvard University,
July 29, 1999

I thank Jim Stock for helpful suggestions, discussions, and support. All possible mistakes are my own.

Abstract

This paper looks closely at the minimax approach to analysis of monetary policy rules under model uncertainty. It recommends the approach as a sensible and powerful alternative to the standard Bayesian analysis. Despite the apparent differences in the two strategies, the resulting optimal policy rules might be not far apart. This is shown in a simple Brainard's (1967) setting. The minimax analysis of monetary policy rules is performed under linear slowly time varying model uncertainty built around Rudebusch and Svensson's (1998) model. The exact minimax optimal policy rules are found. The rule optimal under certainty turns out to be very robust to quite large deviations from Rudebusch and Svensson's model. However, the minimax optimal rules tend to respond less to inflation and more to output gap. To check the plausibility of the structure of the model uncertainty I constructed the worst possible models that the minimax approach takes care of. It is shown that for many policy rules these models correspond to reasonable economic situations. This is the case for very aggressive optimal H_∞ rule. Though robust to uncertainty about structure of the noise process and to parametric model uncertainty the rule is shown to perform poorly if there exists a lag structure uncertainty about the core 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 workings of the economy in the new European environment, steady decline of the natural unemployment rate in the USA, and 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 to choose 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..." 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 of the minimax differ from those of the Bayesian approach? Do the minimax recommendations make sense? Second, what are the minimax policy recommendations after all? So far, the optimal minimax policy rules were obtained only for very special cases of model and shock uncertainty. These rules turned out to be very aggressive. Does this aggressiveness result hold for the general model uncertainty case? Last, but not least, do the worst possible cases, the minimax takes care of, have any economic meaning? What are these worst possible cases?

To answer the first group of questions I consider a simple Brainard's (1967) setting for policy analysis. I assume that there is 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 his simple setting. The more uncertain a policy maker is the more conservative optimal policy rules are.

I found that in Brainard's setting the optimal minimax policy rules have strong similarities with the optimal Bayesian ones. The optimal minimax policy does become less aggressive as uncertainty about the parameters rises. However, unlike the Bayesian optimal rules, the minimax rules react to changes in the amount of the uncertainty in a very discrete way. The sign of the policy effect on the target must become uncertain before the optimal minimax rule starts to become more passive than the optimal certainty rule.[1]

Concerning the second group of the questions this paper concentrates on the framework for the minimax analysis of model uncertainty proposed by Onatski and Stock (1999). The Rudebusch and Svensson's (1998) two equations 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. Only policy rules of Taylor type are considered. According to these rules, nominal interest rate is set to be a linear combination of inflation and output gap.

This paper finds exact optimal minimax policy rules for the general case of model uncertainty described above. It shows that the optimal minimax rules tend to respond more

[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).

aggressively to the output gap and less aggressively to the inflation than the certainty rule does. The conventional linear quadratic optimal rule reported by Rudebusch and Svensson demonstrates a high degree of robustness to model uncertainty. Deviations of minimax rules from this optimal rule are very small for quite large degree of the uncertainty.

To answer the third group of the questions I analyze the structure of particularly bad deviations from the Rudebusch-Svensson that lead to instability of one of the policy maker's variables of interest: inflation, output gap, or changes in the nominal interest rate. Sensibility of the minimax recomendations depends crucially on the nature of these bad deviations. If, for example, the bad scenario assumes importance of, say, the thousandth lag of inflation then the minimax recommendations might be of little value.

The paper shows that the worst possible lag structure of the deviations is exponentially decaying. The rate of decay varies very much for different policy rules. The worst possible deviations for very aggressive rules lead to frequent and increasing business cycles. It is shown that a highly aggressive H_∞ rule could lead to instability of the economy for a deviation from the core model which adds exponentially decaying lags in the effect of the real interest rate on the output gap. Half of the decay happens after three quarters, so the deviation seems plausible.

The performed 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, for the case of the Rudebusch and Svensson's model the optimal certainty rule is very robust to model uncertainty. Aggressive policy rules, though robust to uncertainty about structure of the noise process and parametric model uncertainty,[2] are not robust against lag structure uncertainty and might lead to frequent and increasing business cycles. Third, the worst possible cases that the minimax takes care of often have good economic sense. However, for some policy rules these cases have too slow rate of lags' decay. The results of the paper suggest that for some of these rules a "truncation" of the worst possible cases might lead to only marginally better ones.

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

.2 Comparison of the minimax and Bayesian policy implications

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

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

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

I consider here the case of one target and one instrument to make analysis as simple and transparent as possible. Assume that a variable of policy maker's interest Y is equal to $aP+u$, where P is policy variable, u is exogenous variable and a is a multiplier determining a 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 can, however, face uncertainty about the model. This uncertainty may come from two sources. On the one hand the policy maker might be uncertain about effect of exogenous variable u on Y. On the other hand she might be uncertain about the value of coefficient a. This uncertainty might be modeled by assuming that a or/and u are random variables. If only u is uncertain than the policy maker is in the situation of the 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 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 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 squared deviation of 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 zero policy. Optimal policy is in between, with precise position depending on the relative importance of the two parts of the expected squared error that is captured by coefficient of variation $\sigma_a/\bar{a}$ of variable a.

In the minimax setting uncertainty about a and u is modeled differently. 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 would carry the same function as variances σ_a^2 and σ_u^2 and 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 give information about a and u similar to that one can get from mean values $\bar{a}$ and $\bar{u}$.[4]

---

[3] Such an uncertainty is called Knightian uncertainty.

[4] In principle, if u plays a 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 non-random unknown quantity.

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

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

where $\delta_a = \sqrt{3}\sigma_a, \delta_u = \sqrt{3}\sigma_u$. The minimax problem that policy maker faces is to find such 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 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

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

Policy P equal to $(Y^* - \bar{u})/a$ minimizes both the first and the third terms of the above expression, therefore, this policy is optimal. Note that the optimal policy has the same form as the optimal policy (.2) in the situation of certainty equivalence in the previous section.

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 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 such P 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 large positive 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 the considerable overshooting will happen if the policy multiplier is small and u is big.

The maximal distance is minimum for the line depicted on 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 considered uncertainty specification 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 ellips with the center in $(\bar{a}, \bar{u})$ and such that the "dispersions" of possible a and u are proportional to σ_a^2 and σ_u^2 respectively and "interdependence" of possible a and u is regulated by a coefficient ρ. Of course, Ω would have been just a confidence ellips 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, intuitive, and might well summarize scarce knowledge about a and u that a policy maker might posess. 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:

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

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

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

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 assumed that a and u are distributed uniformly inside ellips (.7).[5]

[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: $\frac{Y^*-\bar{u}}{\bar{a}} = \frac{\sigma_a}{\sigma_u} = \frac{\sigma_a}{\bar{a}} = 1$.

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 the uncertainty becomes larger than the threshold $(1+k)\bar{a}^2/\sigma_a^2$, and in the simplest case of $\rho = 0$, the policy P becomes less active than the certainty equivalence rule. Indeed, at the threshold expressions (.8) and (.9) are equal. Absolute value of expression (.9) is decreasing in r. So when r rises P becomes less aggressive. See figure 2, middle graph.

Intuition behind proposition 1 is simple. The optimal minimax policy tries to balance between two worst possible cases. One of the bad scenarios is overshooting the target Y^*, the other is its undershooting. 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 a 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: too large a and u on the one hand and too 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 the Brainard's 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 way" making policy P negative, which is the same result as in Brainard. 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 go 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 economy. She, however, believes that the model in 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 the set are indexed by 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 $\{P\}$ so that to minimize expected loss (risk)

R assuming 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 Rudebusch-Svensson (1998) model as the nominal model of economy. The model consists of two equations, estimated econometrically using the 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 variable y stands for the gap between output and potential output, π is inflation and i is 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 $\overline{i}$ stand for four-quarter inflation and federal funds rate respectively. The coefficients on the 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 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 a 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 output gap on 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, 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

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

where Δ_{ij} are some dynamic operators. If the policy-maker believes that the nominal model is good approximation to the truth then the operators Δ_{ij} might be chosen to have small norms, less than, say, $r\delta_{ij}$. The last two equations correspond to a model M_Δ where $\Delta = diag\{\Delta_{ij}/\delta_{ij}\}$. The possible deviations constitute a set $D_r = \{\Delta : \|\Delta_{ij}/\delta_{ij}\| < r\}$

The operators from the set D_r could have very broad nature from constant operators (multiplication by a constant) and linear time invariant operators (that could be represented by an infinite L polynomial) to nonlinear time-varying operators. In this paper I restrict attention to linear time invariant operators. This means that models M_Δ differ from Rudebusch and Svensson's model by their lag specification. They could have arbitrarily more lags of inflation, output gap, and real interest rate in the Phillips curve and the aggregate demand equations. Of course, the coefficients on new lags and the difference between the coefficients on the old ones must be small for deviations Δ to have small norm.

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. Such deviations would destabilize 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 exist a $\Delta_r \in D_r$ such that the model M_{Δ_r} is unstable under P. Let us call the number r_0 a radius of affordable perturbations for the rule P. Onatski and Stock found the radia of affordable perturbations for the Taylor type rules.[6] Figure 3 summarizes their results. The authors took δ_{ij} to be equal to standard deviations of the estimates of $A_{ij}(0)$. One can see that the highest radia 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 the optimal when there is no uncertainty about Rudebusch-Svensson model.

A radius of affordable perturbations for a given rule P carries only partial information about quality of the rule. To assess this quality one probably needs to know not only the radius but the perturbations that bring P at the verge of instability themselves. Are these perturbations plausible from economic point of view?

---

[6] The exact calculation of the radius of affordable perturbations is proved to be an NP hard problem. However, it is easy to find lower and upper bounds, r_l and r_u, on the radius that turns out to be very close each to other for the problem at hand.

Recall that any perturbation Δ has a diagonal structure $diag\{\Delta_{ij}/\delta_{ij}\}$. I would ignore the scaling by δ_{ij} in the following for simplicity of notations. Denote $\Delta_{ij}(L)$ the representation of the diagonal components of the worst possible perturbation in the form of infinite $L-$ polynomial. The program used for computing the radius of affordable perturbations r supplies the following information about operators Δ_{ij}. First, $\|\Delta_{ij}\| \leq r$. Second, the set of equalities

$$\Delta_{ij}(e^{i\omega_0}) = r_{ij}e^{i\omega_{ij}} \tag{.11}$$

hold for some $\omega_0 \in [0, \pi)$ and $r_{ij} \leq r$. Any LTI (linear time invariant) operator with the diagonal structure and the diagonal components satisfying the above restrictions qualifies for the worst possible perturbation.

It is easy to build such an operator. Indeed, consider operators of the diagonal structure with the diagonal components of the so called Blashke form $\Delta_{ij} = \pm r_{ij}\left(L - x_{ij}\right)/\left(1 - x_{ij}L\right)$. It is not difficult to see that there exist a $x_{ij} \in (-1; 1)$ such that equalities (.11) hold and $\|\Delta_{ij}\| = r_{ij} \leq r$.[7] If $\omega_{ij} \in [0, \pi)$ then the sign in the formula for Δ_{ij} is "+", if $\omega_{ij} \in (-\pi, 0)$ then the sign is "˘". In both cases $x_{ij} = \left(e^{i\omega_0} - e^{i\omega_{ij}}\right)/\left(1 - e^{i(\omega_0+\omega_{ij})}\right)$.

So we see that the worst possible perturbations could be always thought of as LTI operators with exponentially decaying lag structure. Therefore, unreasonable deviations that emphasize importance of, say, one thousandth lag in the model equations are not the worst. An operator $\Delta = \pm r\left(L - x\right)/\left(1 - xL\right)$ could be expressed in the form of the infinite $L-$polynomial as

$$\Delta(L) = \pm r\left(-x + (1-x^2)L + (1-x^2)xL^2 + (1-x^2)x^2L^3 + (1-x^2)x^3L^4 + ...\right)$$

[7] Here the norm of operators is H_∞ norm, that is $\|\Delta\| = \sup_{|z|>1}|\Delta(z)|$

If x turns out to be close by absolute value to 1 then almost all "effect" of the operator is a multiplication by a constant $\mp rx$.[8] If x turns out to be close to 0 then almost all "effect" is an application of $\pm rL$ operator. For x in between the two extremes $\Delta(L)$ is an infinite $L-$polynomial with fast decaying lags. For example, for $x = 0.7$ the half decay of the polynomial's coefficients happens after three quarters.

I computed the x's and the signs of the worst possible operators 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 degree of x's and the signs vary for different policy rules. However, it is possible to distinguish between several groups of policy rules.

First, for relatively "passive" rules, very roughly, the rules with $g_\pi < 3, g_y < 2$, the worst possible deviations Δ_{ij} all "have positive sign". This means that in the worst possible case the economy's reaction to key economic variables is higher in total, but more "spread out" through the time so that the immediate reaction of the economy is smaller.

The worst cases' memory of the economy is very long for these "passive" rules. It is not unusual to have the rates of decay, x_{ij}, greater than 0.90. For example, for the rule proposed by Taylor, $g_\pi = 1.5, g_y = 0.5$, the rates of decay $x_{\pi\pi}, x_{\pi y}, x_{yr}$, and x_{yy} are equal to 0.91,0.93,0.97 and 0.71 respectively.

The rules have relatively large radia of affordable perturbations. If one excludes long memory models from the set of plausible possibilities then the radius of affordable perturbations for the "passive" rules will be even higher. For the Taylor rule I considered a truncated

---

[8] In such a case the extreme percistence of the tiny tail of the polynomial might be important for destabilizing of the economy under the rule P. Whether this is the case should be checked by trying to "destabilize economy" by adding only "first order" disturbance $\mp rx$ to the nominal model.

version of the worst possible case with $\Delta_{ij}(L) = r\left(-x_{ij} + (1 - x_{ij}^2)L + (1 - x_{ij}^2)x_{ij}L^2\right)$ except Δ_{yy} that I left untruncated because 0.71 is rather fast degree of decay. The truncated worst possible cases start to destabilize the economy when r becomes 1.45. It is not much higher than the "untruncated radius" that is equal to 1.01.

Figure 4 shows the impulse responses of the inflation and output gap for the Taylor rule. The solid line represent the impulse response for the nominal model. The dashed line represents the impulse responses in the worst possible case. The fluctuations caused by impulse shocks in the truncated worst possible case have very long period, about 40 years. The short run structure of the responses is similar to that in the nominal case.

The figure suggests that the worst possible cases for "passive" rules manifest themselves in a very prolonged fluctuations of the economy that never settle down. In such a context the assumption that the policy-maker could not change a policy rule becomes important. It is, however, likely that a policy-maker's credibility would survive slowly time varying rules. In such a situation the worst possible cases described here might be not too menacing.

The second group of policy rules that has distinct structure of the worst possible perturbations consists roughly of the rules with very active responce to the output gap, $g_y > 2$. For these rules the sign of all delta's except Δ_{yr} is "$\smile$". This means that in the worst possible case the immediate reaction of the economy to changes in inflation and output gap is increased whereas total reaction of the economy is decreased. The economy overshoots its long run responce. 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.

The memory of the economy in the worst possible cases is rather low for the rules with little response to inflation. It, however, rises with g_π. For the rules with very large inflation response the $x_{\pi\pi}, x_{\pi y}, x_{yy}$ become close to 1 whereas x_{yr} is about 0.70. For example, for the H_∞ optimal rule, $g_\pi = 6.42, g_y = 2.75$, the rates of decay $x_{\pi\pi}, x_{\pi y}, x_{yr}$, and x_{yy} are equal to 0.96,0.97,0.71 and 0.92 respectively.

The H_∞ optimal rule is interesting rule to analyze because it is an extreme case of the very aggressive rules shown to be robust to different kinds of uncertainty. Sargent (1998) showed that the aggressive rules are robust to uncertainty about the nature of shock process. Stock (1998) showed that the such rules are robust to uncertainty about slope of the Phillips curve and the output reaction to real interest rate in the context of the Rudebusch and Svensson model.

I considered a truncated version of the worst possible case with $\Delta_{\pi\pi} = \Delta_{\pi y} = \Delta_{yy} = r$. I found that under these truncated perturbations the radius of affordable perturbations for the optimal H_∞ rule is 0.95, that is only slightly larger than 0.91, the radius of affordable perturbations in the unrestricted case. Figure 5 shows the impulse responses of the economy controled by the optimal H_∞ rule for the case of the nominal model and the case of the modified (as described above) worst possible perturbation to the nominal model. One can see that in the worst possible case the economy starts to swing without settling down in response to inflation or output gap shock. The period of fluctuations is about 5 years that is quite short.

The analysis suggests that the aggressive rules are not robust to uncertainty in the lag specification of the model. If one assumes that the reaction of the economy to real interest rate is more “spread out ” through the time and the economy is more easily excited in that the contemporaneous reaction to inflation and output gap is increased then the aggressive rules destabilize the economy quite easily.

.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

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

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

There exist technical difficulties that do not allow to solve (.10) in its exact form. The matter is, roughly, that the stochastic definition of shocks does not fit well in the minimax methodology. The calculation of risk uses averaging, not minimizing, the maximum loss. It is, however, possible to reformulate (.10) in a “completely minimax” way. This requires modeling the shock sequences not as realizations of the white noise process but as points in some set that provide tight characterization of such realizations. Below I, first, describe such a characterization that was 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 corresponding statistical hypothesis. It is, therefore, quite natural to assume that a sequence is close to a realization of white noise process if it belongs to the set complementary to the rejection region of a test.

One of the tests for 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 $V(k)$ 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 acual realizations of the white noise process belong to the set with asymptotic probability 1. More precisely the following proposition is true.[9]

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 $\|v\|_2^2$ 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 n-dimensional vector. Denote I_n an n-dimensional identity matrix. Finally, for any matrix A denote $\|A\|_\infty = \max_{i,j} |A_{ij}|$.

9 See Paganini (1996), Theorem 4.8, p.58

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

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

Lut us view a model $M_\Delta(P)$ controled by a policy rule P as an operator transforming 2-dimensional sequence of shocks $v = \{(\varepsilon_t, \eta_t)'\}$ [10] to 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 $M_\Delta(P)$ as operating 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.[11]

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

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

Let us denote

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

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

[10] In the following I assume that the shocks are scaled so that there variance-covariance matrix is unity.

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

follows.[12]

$$\min_{P \in \{P\}} \gamma_r(P) \tag{.12}$$

There is one more qualification to the minimax problem that is due here. This concerns the nature of the set of perturbations D. Let us consider a set of slowly linear time varying (LTV) perturbations to the nominal model M_0 :

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

where the norm in the right hand side is induced norm of 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. It allows lag specifications of M_Δ to slowly vary in time with the rate of variation measured by ν.

Consider the following problem:

$$\min_{P \in \{P\}} \inf_{\nu > 0} \gamma_r(P) \tag{.13}$$

Solution to this problem will give a policy rule that is most robust against, so to speak, arbitrarily slowly time varying uncertainty. More precisely, let P^* be a solution to (.13). Then for any other policy rule P there exist a $\nu > 0$ and $\theta > 0$ 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^*.

Figures 6 and 7 show $\inf_{\nu>0} \gamma_r(P)$ for policy rules of 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

[12] Whether solution to (.12) differs from that to (.10) depends on the uniformity properties of convergence in proposition 3.

of 0.25), and $g_y \in [0.125, 4.5]$ (grid of 0.125). The star, square, and circle points at 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 risk for the rule under certainty 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 the optimal minimax rule becomes more responsive to 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. It becomes equal to 50. The Taylor rule and especially H_∞ become absolutely unacceptable, whereas the worst possible risk for the certainty rule is somwhere 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.

.5 Conclusion

This paper discribes three different excersises concerning the minimax analysis of policy rules under model uncertainty. First, the minimax approach is compared with the Bayesian one in a simple Brainard's setting. Strong similarities between recommendations of the two approaches are found. Similar to the Bayesian rules the optimal minimax rules do become less aggressive when uncertainty about policy multiplier rises. However, unlike the optimal Bayesian rules, the optimal minimax rules react to changes in uncertainty about the multiplier in a very discrete way. The sign of the effect of the policy on the target variable must become uncertain before the optimal minimax rules start to be more passive than the certainty equivalence rule.

The second excersise concerns with analysis of the worst possible deviations from Rudebusch and Svensson's model. The class of linear time invariant deviations is considered. The perturbed models differ from the Rudebusch-Svensson in that arbitrarily more lags of inflation, output gap, and real interest rate are added into the Phillips curve and the aggregate demand equations and coefficients on the existent lags might be different. The paper finds the smallest perturbations from the described class that destabilize the economy for different policy rules of Taylor-type. It is shown that these perturbations have exponentially decaying lag structure.

The worst possible cases for aggressive rules manifest themselves in a frequent and ever increasing business cycles. These worst possible cases could be characterized by relatively high contemporaneous sensitivity of the economy to the inflation and output gap and more spread-out through the time reaction of the output gap to the real interest rate. 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 the 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 output gap than the optimal certainty rule. The latter is robust to quite large degree of uncertainty.

.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.

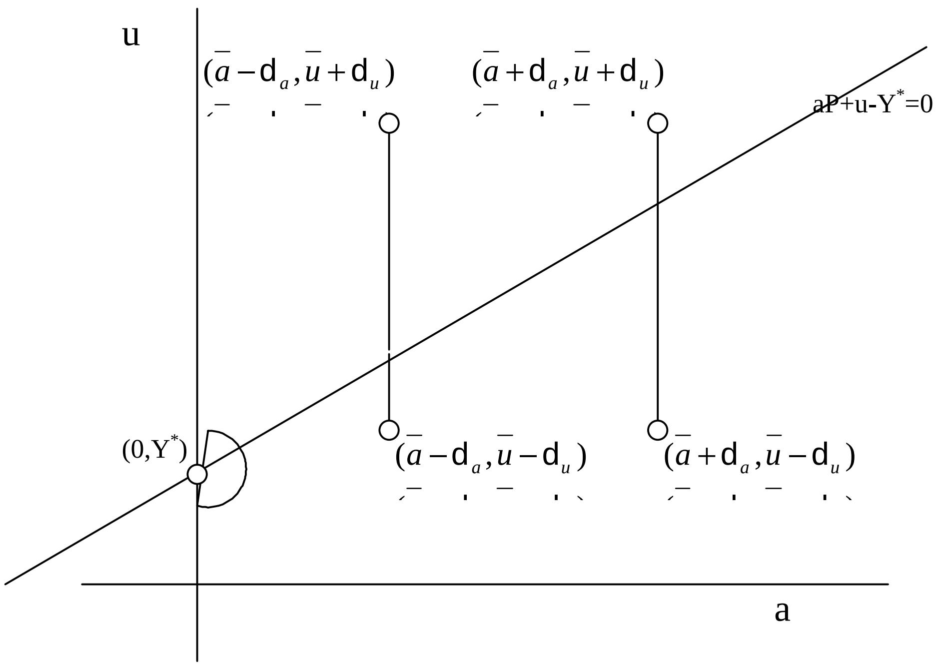
Paganini, F. (1996), Sets and Constraints in the Analysis of Uncertain Systems, California Institute of Technology, Ph.D. thesis.

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).

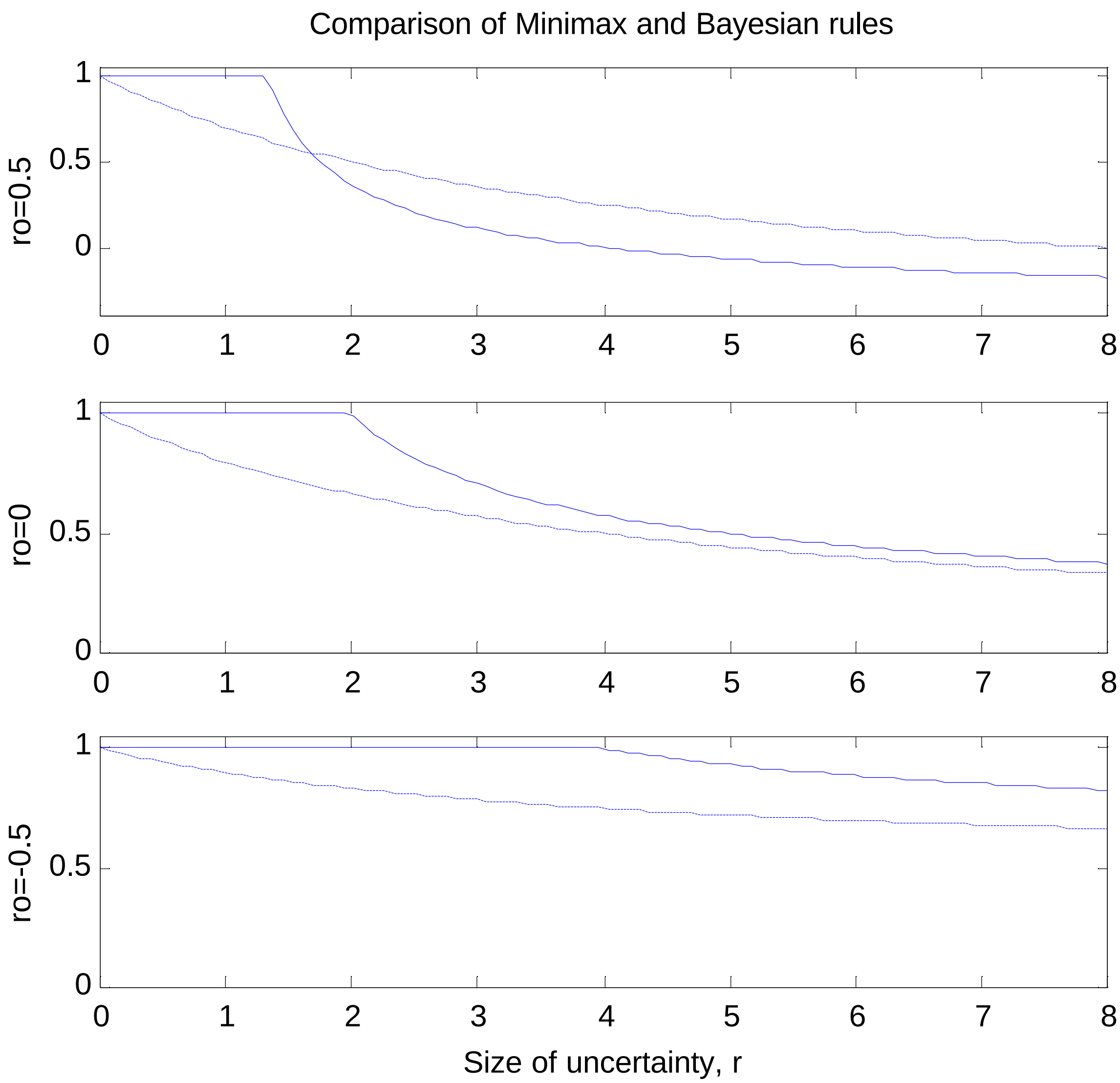
Stock, J.H.(1998), "Comment on Rudebusch and Svensson", forthcoming in J. Taylor (ed.), Policy Rules for Inflation Targeting (Chicago: University of Chicago Press for the NBER).

Figure 1



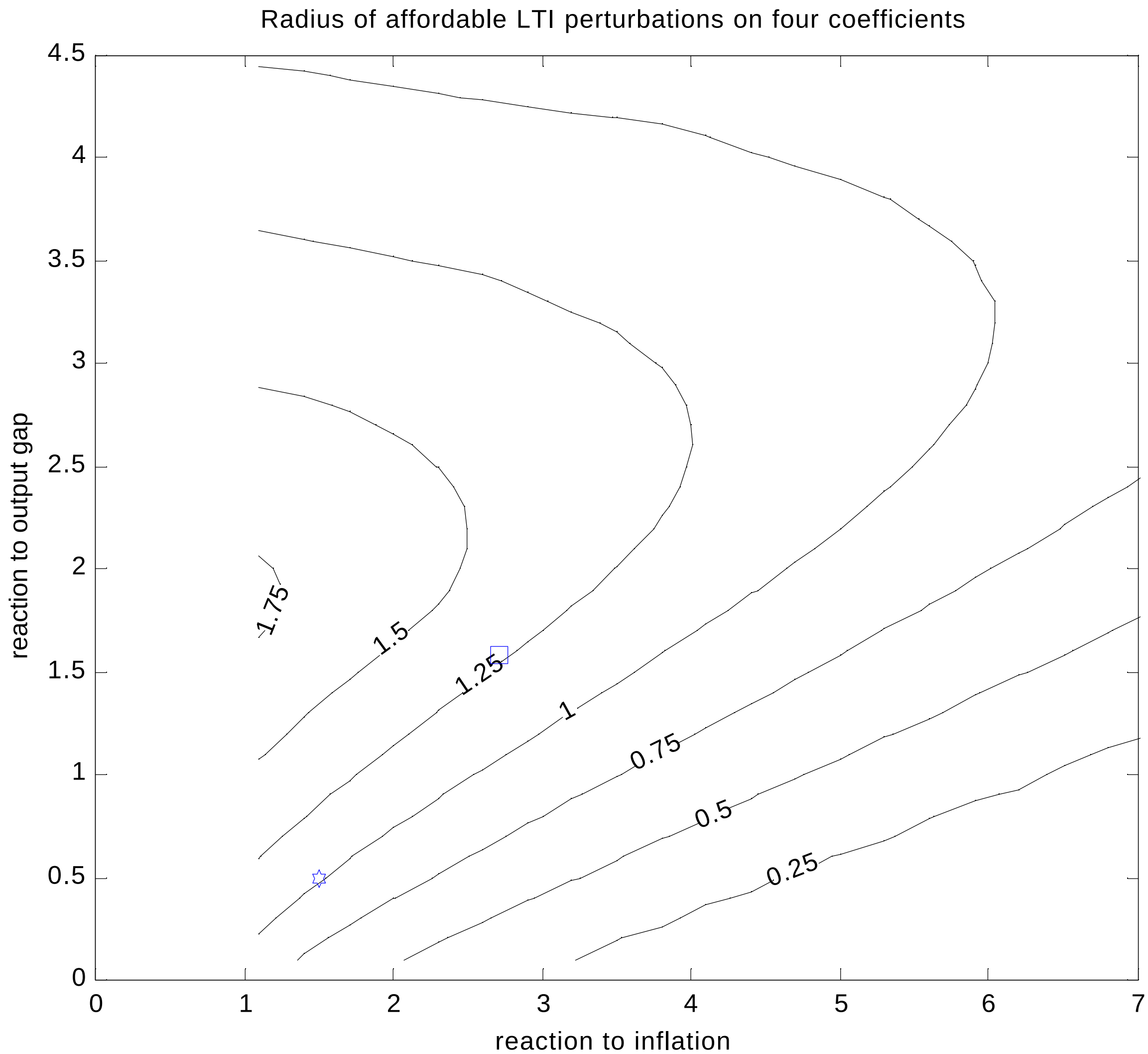
u
$(\bar{a}-d_a,\bar{u}+d_u)$
$(\bar{a}+d_a,\bar{u}+d_u)$
$aP+u-Y^*=0$
$(0,Y^*)$
$(\bar{a}-d_a,\bar{u}-d_u)$
$(\bar{a}+d_a,\bar{u}-d_u)$
a

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
1
2
3
4
5
6
7
1.75
1.5
1.25
1
0.75
0.5
0.25

Figure 4

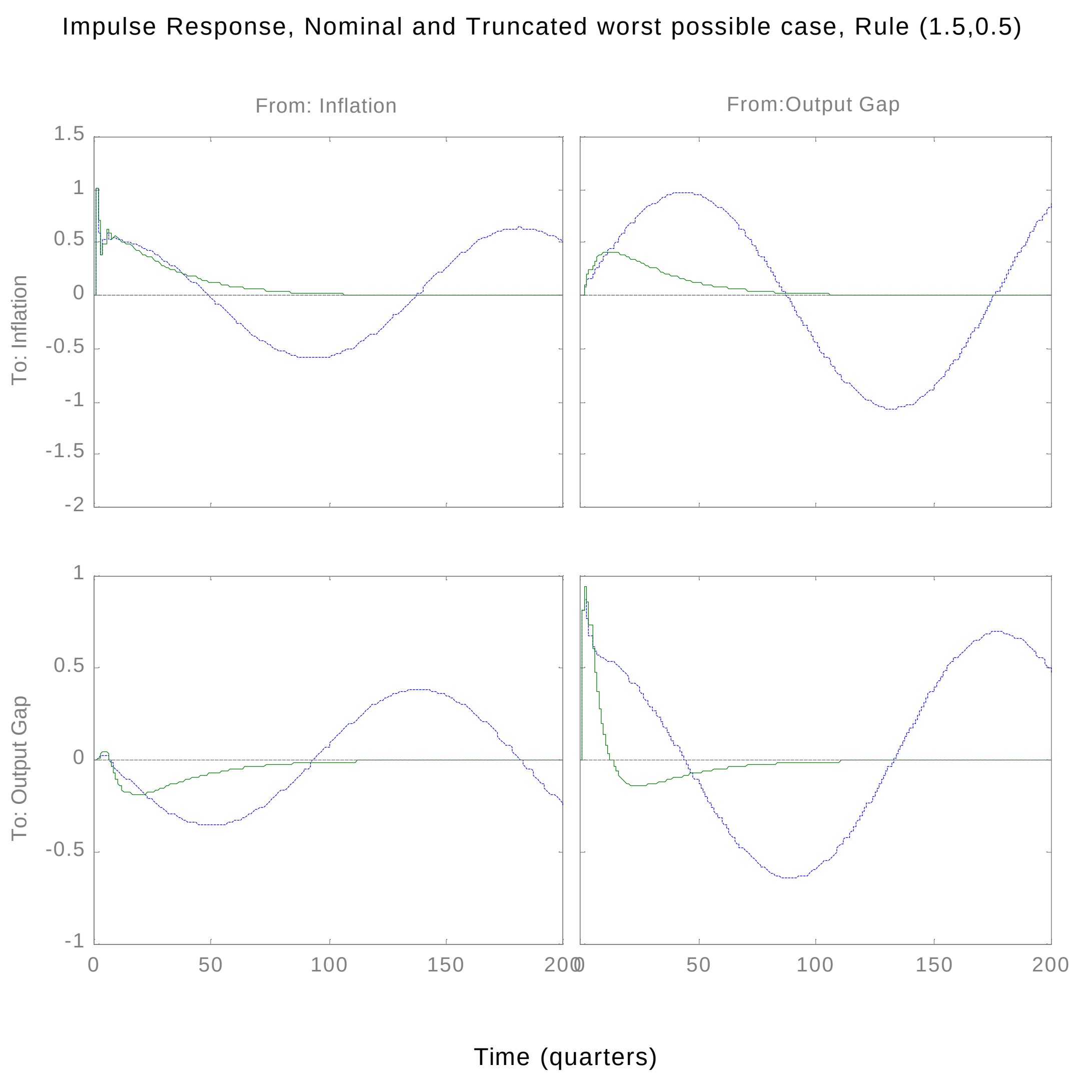

Impulse Response, Nominal and Truncated worst possible case, Rule (1.5,0.5)
From: Inflation
From:Output Gap
To: Inflation
To: Output Gap
1.5
1
0.5
0
-0.5
-1
-1.5
-2
1
0.5
0
-0.5
-1
0
50
100
150
200
0
50
100
150
200
Time (quarters)

Figure 5

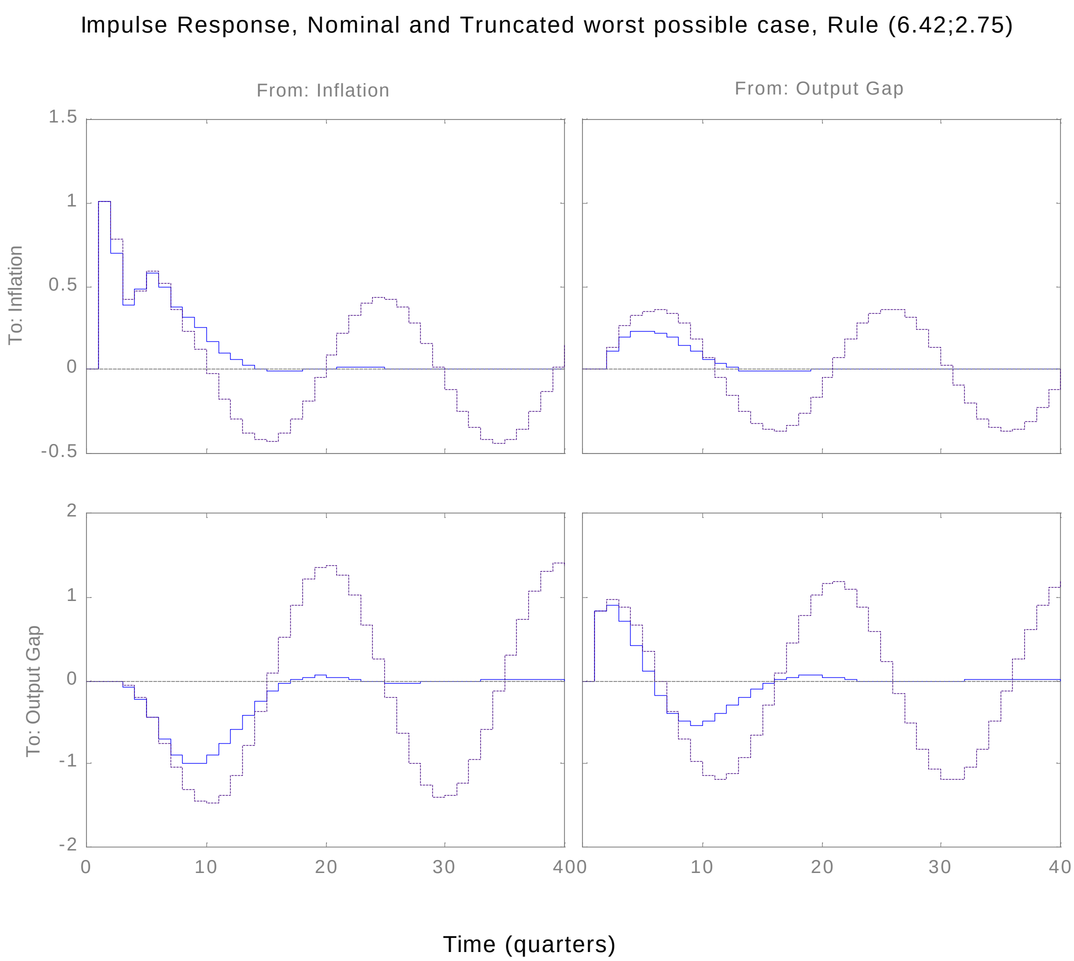

Impulse Response, Nominal and Truncated worst possible case, Rule (6.42;2.75)
From: Inflation
From: Output Gap
To: Inflation
To: Output Gap
1.5
1
0.5
0
-0.5
2
1
0
-1
-2
0
10
20
30
40
0
10
20
30
40
Time (quarters)

Figure 6

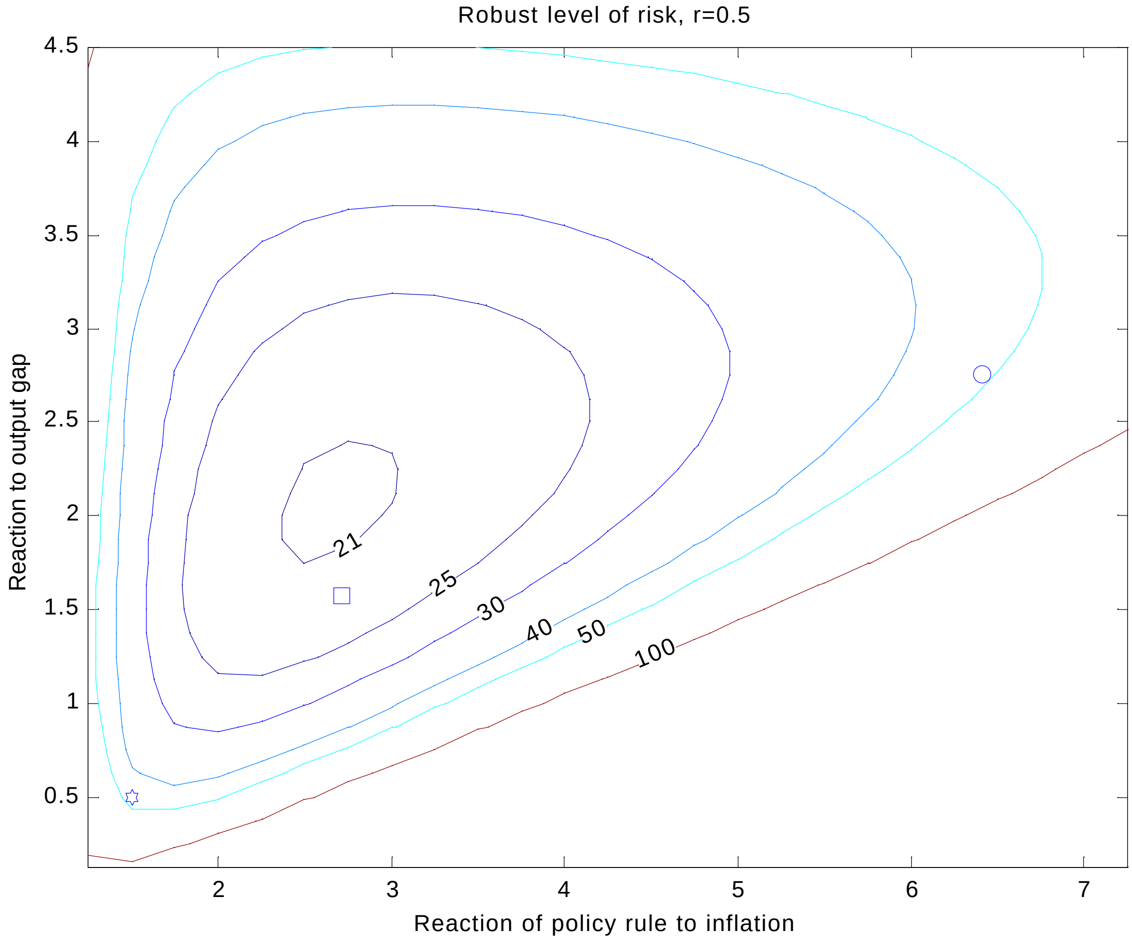
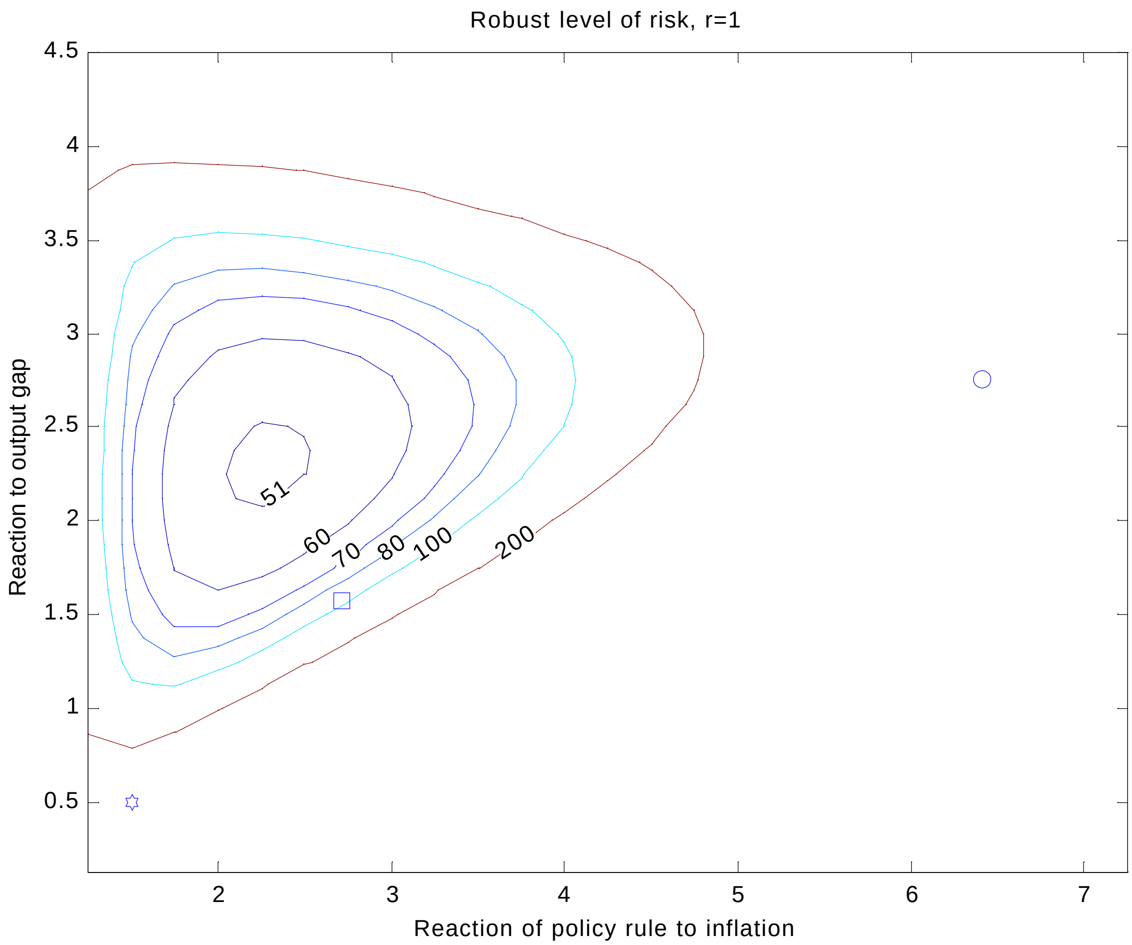

Robust level of risk, r=0.5
Reaction to output gap
Reaction of policy rule to inflation
4.5
4
3.5
3
2.5
2
1.5
1
0.5
2
3
4
5
6
7
21
25
30
40
50
100

Figure 7


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