ON ROBUST MONETARY POLICY

Filippo Altissimo*, Stefano Siviero** and Daniele Terlizzese**

* European Central Bank, Directorate General Economics
** Bank of Italy, Economic Research Department.*

January 13, 2005

*The opinions expressed here are those of the authors and do not necessarily reflect views of the European Central bank or of the Banca d’Italia. Any remaining errors are of course the sole responsibility of the authors.
1 Introduction

The notion of “robust monetary policy”—a monetary policy whose stabilisation properties remain relatively good irrespective of the true model of the economy—has recently attracted considerable attention. The main reason for the interest in the subject is well exemplified in a recent paper by Levin, Wieland and Williams (2003, LWW). Considering five different and widely used models of the U.S. economy, LWW show that the monetary policy rule that would be optimal (within a fairly general class of rules) for each single model could be a real disaster if one of the other models were to be the true representation of the economy. Levin and Williams (2001, LW) conduct a similar experiment using three different models of the U.S. economy. This kind of results is not typical of the U.S.. Coenen (2003) and Adalid, Coenen, McAdam and Siviero (2004) find a similar lack of robustness considering two (respectively, four) alternative models of the euro area economy.

The various models considered differ along several dimensions. Yet, when testing the robustness of monetary policy rules an important role seems to be played by differences in the degree of inertia (of both inflation and output). This makes these findings particularly worrying, at least from a monetary policy perspective, since the degree of nominal inertia of the economy is still subject to theoretical as well as empirical controversies. Mankiw (2001) recently noted the tension between the forward-looking characterization that emerges in modern, optimization-based models and the inertial characterization that seems to be required to fit the data (as he wrote, “... the assumption of adaptive expectations is, in essence, what the data are crying out for.”).

Therefore, identifying a monetary policy rule that performs reasonably well across economies differing for their degree of inflation inertia seems to a be a particularly pressing task for both monetary policy theorists and for monetary policymakers.

What constitutes a robust policy is however not unanimously agreed either. One possible approach, taken for example by LWW and by Coenen, is to search for the monetary policy rule that minimises the loss function on average across different models. This corresponds to a Bayesian uniform prior
over the different models. A different approach, taken by LW (who adopt also the Bayesian approach) and by Giannoni (2002), is to search for the monetary policy rule whose largest loss is minimised. Onatski and Williams (2003) also compute both Bayesian and minimax monetary policy rules.

Whether a Bayesian or a minimax approach is taken often makes a substantive difference as to the resulting ”robust” monetary policy rule. Considering for example the results presented in LW (excluding the extreme case in which no weight at all is placed on output stabilisation), the Bayesian monetary policy generates a loss that, across the three different models considered, is never more than 35% larger than the minimal loss achievable in each particular model. Conversely, the minimax monetary rule can easily lead to losses that are 80-100% larger than the optimized loss. Therefore, the minimax approach seems rather poor in identifying a truly “robust” policy (at least as soon as the set of models across which robustness is sought is large enough, a point already made by Sims, 2001). Conversely, the Bayesian approach, although much more successful in generating a robust policy, requires the use of an explicit prior, which might be problematic, particularly when the uncertainty concerning the alternative models of the economy is deep, as in the case we are considering.

In this paper we explore a third approach to select a robust monetary policy, which does not require the use of a prior and leads to a policy whose losses remain even closer to the minimal losses across alternative models of the economy than those obtained with the Bayesian approach. The approach we consider is borrowed from the theoretical computer science literature, where it is termed competitive ratio. Brafman and Tennenholtz (1999) provide an axiomatic foundation for such an approach, which relies on a

---

1In fact, a further relevant difference in the literature concerns the type of model uncertainty against which robustness is desired. On the one hand, LWW, LW, Coenen (2003) and Adalid, Coenen, McAdam and Siviero (2004) address the issue of policy robustness across a set of non-nested models, which embed different paradigms about controversial issues such as expectation formation and inflation inertia. On the other, Giannoni (2002) and Onatski and Williams (2003), in line with the approach to robustness put forward by Hansen and Sargent (2002), explore a much narrower set of models, essentially derived from small perturbations around a given model of the economy.
generalised version of Savage’s “sure thing principle”. In practice, we will compute for each policy a relative loss, by scaling its absolute loss in each model (or, more generally, in each state of the world) by the minimal loss which is achievable in that model, and we will select the policy whose maximal relative loss is smallest. It is worth noting that this approach is very close to the interpretation that Savage (1954) gave of Wald’s (1950) original formulation of the minimax criterion. Savage thought that what Wald meant was the minimisation of the maximal difference from the highest achievable utility in each state of the world, a difference that he termed “regret” (thus interpreted, the minimax decision criterion is also known as Savage’s minimax regret). Here, instead of the difference with the performance of the best (state contingent) action, we consider the ratio. It is worth pointing out that the axiomatisation proposed in Brafman and Tennenholtz (1999) highlights the close connection among minimax, minimax regret and competitive ratio. Given the loaded meaning of the word “competitive” in economics, we will refer to the latter decision criterion as “relative minimax”.

An interesting and important point to note is that while we are, to the best of our knowledge, the first to explore the implications of adopting the relative minimax as a decision criterion for monetary policy (and more generally for economic decisions), the relative loss is commonly used as a diagnostic criterion in the recent literature assessing the robustness of various rules (selectively quoted above). Taking for example LWW, the performance of a given monetary policy rule in the various models considered is assessed by showing its loss relative to the optimized loss for each particular model, i.e. in terms of its relative loss. The novelty of our paper is that we use the relative loss also to design a robust rule.

In order to investigate the performance of the relative minimax approach, we consider, for much of this paper, a very simple class of economies, in fact a single Neo-Keynesian hybrid Phillips curve, with the degree of inflation inertia varying between zero and one. We assume that the policy maker can make a credible commitment to follow a given policy rule, chosen within a class that nests the optimal (model-dependent) policy, and that its objective is to minimise the expected value of a discounted weighted average of infla-
tion and output variability. For each of the possible values of the degree of inflation inertia we compute the optimal monetary policy rule, and the corresponding value of the loss (the optimal loss corresponding to that particular degree of inertia). We then compute the loss that each of these rules would generate in an economy with a different degree of inflation inertia, and we find, much as in LWW, that in general the optimal rules are not very robust. Rules that are optimal when the economy is very inertial, when implemented in economies with little or no inertia yield losses that are twice as large as the optimal loss. Conversely, rules that are optimal for economies with little or no inertia, when implemented in very inertial economies yield losses that are up to seven times as large as the optimal loss. A modicum of robustness can be obtained only for rules that are optimal in economies with a degree of inertia close to 0.5, yielding losses that at most are about 40% higher than the optimal loss. This result is clearly suggestive of the possibility to obtain a relatively robust monetary policy rule with a (flat prior) Bayesian approach, as stressed in LWW.

We then proceed in computing three alternative “robust” policies, corresponding to three different concepts of robustness: standard minimax, (flat-prior) Bayesian and relative minimax. The basic result of the paper is summarised in Figure 1 (see section 4). The relative minimax rule generates losses, as the degree of inertia of the economy varies, that depart from the optimal loss by at most 12%; the losses generated by the Bayesian rule can be up to 50% larger than the optimal ones; the losses generated by the minimax rule can be twice as large as the optimal ones in some state of the world. Moreover, while the stabilisation performance of the last two rules, relative to the performance of the inertia-contingent optimal rule, deteriorates the more the economy becomes forward looking, that of the relative minimax rule is much more stable, with no discernible trend.

Finally we check the sensitivity of those results to a few generalisations. First we confirm them with different parametrisation of the model and of the preferences. Second, we show that very similar results also emerge when we consider economies that differ simultaneously across two dimensions: the degree of inflation inertia and the responsiveness of inflation to output gap.
changes. Finally, we check the performance of the three notions of robustness across four different models of the euro area economy. These models differ widely along several dimensions, and we no longer have the explicit solution for the optimal (model contingent) monetary policy rule. We therefore restrict the analysis to a class of simple interest rate rules. Again, the results broadly confirm the good performance of relative minimax.

The paper is organized as follows. Section 2 highlights potential problems with the minimax concept and motivates the notion of relative minimax as an alternative decision criterion. Sections 3 and 4 assess the performance of the three alternative notions of robust choices considering a very simple model of the economy, with no strong claim of empirical realism. Section 5 addresses the same question considering four different, estimated models of the euro area. The final section concludes.

2 Minimax revisited

To illustrate the rationale for relative minimax let us consider a simple example in which there are only two possible actions and two states of the world: action 1, yielding a loss of 1 in state 1 and a loss of 100 in state 2, and action 2 yielding a loss of 4 in state 1 and a loss of 99 in state 2 (see Table 1).

<table>
<thead>
<tr>
<th></th>
<th>state 1</th>
<th>state 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>action 1</td>
<td>1</td>
<td>100</td>
</tr>
<tr>
<td>action 2</td>
<td>4</td>
<td>99</td>
</tr>
</tbody>
</table>

A Bayesian decision maker would trade-off the losses associated with each action in the various states by weighting them with the probability of the states. Indeed, the loss associated to a given action in a given state can be interpreted (up to an affine transformation) as the probability of obtaining,
with that action when that state occurs, the worst of all possible consequences (and with the complementary probability the best of all possible consequences; see for example Pratt, Raiffa and Schlaifer, 1964).\(^2\) When we multiply this loss by the probability of the state, and we sum across states, we obtain the unconditional probability of obtaining the worst consequence when the given action is taken, and these unconditional probabilities (or any affine transformation of them) can be safely compared across actions. The Bayesian decision maker would prefer action 1 as long as he attributes a likelihood of at least \(\frac{1}{4}\) to state 1. Therefore, a “flat-prior” Bayesian (attributing probability \(\frac{1}{2}\) to each state) would select action 1.

In case the decision maker were not able (or not willing) to attribute a likelihood to the different states, one approach to reach a decision would be to select the rule that minimises the maximum loss (minimax). In the example of Table 1, the minimax solution would pick action 2. Most people would however object to such a choice. Indeed, action 2 is in state 1 four times worse than action 1, and it is only slightly better than action 2 in state 2. The reason why we might feel uncomfortable with the choice of action 2 can perhaps be clarified as follows. The minimax criterion asks us to select, for each action, the largest loss. To achieve this selection, we compare the losses on an equal footing, disregarding any reference to the state in which they would emerge. It therefore makes no difference whether a given large loss is obtained in one state of nature in which all losses tend to be large, or rather in one state where the losses are in general small.

In the example, action 1 would be discarded by the minimax criterion based on its performance in one state (state 2) where the best that can be achieved (99) is not much better than what action 1 can deliver (100). It would therefore seem sensible not to penalise overly action 1 for this small departure from the best. Instead, action 2 yields a relatively large loss (4) in

\(^2\)In fact, Pratt, Raiffa and Schlaifer (1964) show that the utility of a given consequence (rather than the loss associated with a given consequence) can be constructed by comparing the consequence to a “canonical” lottery (a roulette lottery, in the terminology of Anscombe and Aumann, 1963) with prizes given by the best and the worst of all possible consequences. In the text we loosely adapt this interpretation to the case of losses.
one state in which a much better outcome can be achieved. This is totally neglected by the minimax criterion, as the different order of magnitude of the losses in the two states makes what happen in state 1 irrelevant. To put it slightly differently, while the minimax criterion rests on the implicit assumption that there is no way to trade-off losses across states (as the likelihood of the various states cannot be assessed, and therefore each state has to be considered in isolation), it paradoxically ends up trading-off the losses one for one, since all the losses are directly compared, with no allowance being made for the state in which they would arise.

A way to remain more consistent to the “no trade-off” idea that motivates the minimax would be to introduce a state-by-state benchmark, obtained by collecting the smallest losses that can be achieved in each state, and assessing the performance of the various actions relative to this ideal benchmark. By “normalising” for the state-dependent smallest loss we implicitly take into account the fact that losses in different states are different, and therefore we feel less uncomfortable in comparing “normalised” losses across states. This idea is common to the “minimax regret” approach and to what we call, in this paper, the “relative minimax” approach. In Table 2 and Table 3, we transform the original problem presented in Table 1 according to the “absolute normalisation” of the minimax regret approach (obtained by subtracting from the entries in each column of Table 1 the minimum of the column) and the “relative normalisation” of the relative minimax approach (obtained by dividing the entries in each column of Table 1 by the minimum of the column), respectively.

<table>
<thead>
<tr>
<th></th>
<th>state 1</th>
<th>state 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>action 1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>action 2</td>
<td>3</td>
<td>0</td>
</tr>
</tbody>
</table>
If we now pick the action whose largest “normalised loss” is minimal (i.e. if we apply the minimax criterion to the transformed problems in Table 2 or 3) we end up selecting action 1 in both cases. Although both approaches can be given a well specified axiomatic foundation (as shown by Brafman and Tennenholtz, 1999), we believe that the relative normalisation is more appropriate, as it is “unit free”, while the regret approach remains more dependent on large differences in the losses across states. To illustrate this point, consider for example Table 4, where the original decision problem presented in Table 1 has been only slightly modified.

<table>
<thead>
<tr>
<th>Table 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>state 1</td>
</tr>
<tr>
<td>action 1</td>
</tr>
<tr>
<td>action 2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>state 1</td>
</tr>
<tr>
<td>action 1</td>
</tr>
<tr>
<td>action 2</td>
</tr>
</tbody>
</table>

It is immediate to check that the minimax regret choice would become action 2 (which is still the minimax solution), while the relative regret choice would still be, as before, action 1. Clearly, the reason why the minimax regret criterion is affected by the change in the losses presented in Table 4 is because the “absolute normalisation” does not completely get rid of the differences in scale of the losses. In turn, this makes the comparison across states of the resulting “normalised losses” somewhat awkward, since it is still affected by the different nature of the losses in different states. Since we believe that the intuitive justification for preferring action 1 holds even when confronted with Table 4 (action 1 is only slightly worse than the best that can be done in state 2, and avoid a four-fold worsening in state 1), we find the relative minimax a better guide to action. In the following sections,
we will explore in a number of cases the performance of the relative minimax criterion, and compare it with a (flat prior) Bayesian approach and simple minimax.

2.1 Minimax revisited: notation

Before moving to that task, we introduce here for the sake of clarity some notation. Let $X$ be the set of actions available to a decision maker (with typical element $x$), $\Gamma$ be the set of (unknown) states of the world (with typical element $\gamma$). Also, let there be a mapping $X \times \Gamma \rightarrow C$ that defines the consequence $c \in C$ that obtains when action $x$ is taken and state $\gamma$ prevails. Whether or not the consequences are specified in such a way that the decision maker can express his preferences over $C$ irrespective of the prevailing state is the subject of some controversies (for example, the Savage approach requires the consequences to be defined independently from the state and the action; other approaches relax one or the other of these requirements, allowing for the preferences to be defined directly over $X \times \Gamma$). For our purposes, that are merely expository, we will adopt the more general formulation, assuming that the preferences of the decision maker be defined over $X \times \Gamma$ and be represented by a loss function $L(x, \gamma)$. Depending now on which additional assumptions we make on the preferences of the decision maker (and hence on his behaviour), we obtain different decision criteria.

First, consider a Bayesian decision maker (that obeys Savage’s axioms). He can construct a subjective probability distribution over the states $\Gamma$, $f(\gamma)$. Given this probability distribution a Bayesian decision maker acts in order to minimize its expected loss, i.e.:

$$\min_{x \in X} E(L(x, \gamma)) = \min_{x \in X} \sum_{\gamma \in \Gamma} L(x, \gamma) f(\gamma).$$

The Bayesian solution hinges on the ability to assign a meaningful probability to the states of the world. One possibility, often interpreted as reflecting an high degree of uncertainty, is to consider all the states as equally likely ($f(\gamma) = \frac{1}{n} \forall \gamma$, where $n$ is the cardinality of $\Gamma$). This is what we will refer to, in the following, as a flat-prior Bayesian criterion (when there is no
ambiguity we will neglect the “flat-prior” qualification). Considering all the states as equally likely leads, loosely speaking, to seek protection equally against all alternatives. For this reason the flat-prior Bayesian criterion is often interpreted as a robust decision criterion.

However, it is certainly possible to imagine that the degree of understanding of the problem faced by the decision maker (possibly because of a total lack of similarity with problems previously faced) is such that he feels uncomfortable in assigning any probability at all. Even the statement that all possibilities are equally likely could be perceived to be too precise. Usually in the literature this situation has been indicated as Knightian uncertainty (uncertainty where numerical probabilities cannot be assigned).

In these circumstances, a widely advocated decision criterion is to take the action whose worst-case loss is minimal. This decision criterion is referred to in the literature as minimax and it embodies an extremely risk-averse attitude to decision making. In mathematical term, the minimax solution is the action \( x \) which solves the minimization part of:

\[
\min_{x \in X} \max_{\gamma \in \Gamma} L(x, \gamma).
\]

There are links between the Bayesian decision maker and the minimax one. As pointed out by Hansen and Sargent (2002), if the Bayesian decision attributes all probability weight to the worst state associated with the minimax decision, then the Bayesian and the minimax decisions are identical. Moreover, the minimax solution can be shown to be the limit to a Bayesian decision problem as the degree of risk-aversion goes to infinity (see for example Adam, 2003). However, it is clear that the minimax criterion is less demanding on the decision maker, and it protects against the worst-case. Moreover, note that it is a purely ordinal criterion: any other loss \( L'(x, \gamma) \) that does not modify the ordering induced by \( L(x, \gamma) \) leads to the same choice. For these reasons the minimax criterion is often credited for being a “robust” decision criterion.

Another alternative, which we will consider in this paper, is the decision criterion known in the theoretical computer science literature as competitive ratio (as mentioned above, we will call it relative minimax). First, define the
relative loss of action $x$ in state $\gamma$ as the original loss $L(x, \gamma)$ divided by the minimum achievable loss in that state, i.e. define $R(x, \gamma) = \frac{L(x, \gamma)}{\min_{x' \in X} L(x', \gamma)}$.

Second, the relative minimax decision is the one that solves the minimization part of:

$$\min_{x \in X} \max_{\gamma \in \Gamma} R(x, \gamma).$$

Unlike the Bayesian criterion, and similar to the minimax one, this criterion does not require a quantitative measure of the probability of the states. However, it is not purely ordinal as is the minimax. It can therefore be considered as a sort of compromise between the other two criteria.

3 The robustness of optimal rules

Consider a basic New-Keynesian model with staggered price setting behaviour and some form of indexation (see for example: Woodford, 2003), which is described by the aggregate supply relation:

$$\pi_t = (1 - \gamma)\beta E_t(\pi_{t+1}) + \gamma \beta \pi_{t-1} + \lambda x_t + e_t,$$

where $x$ denotes the output gap, $\pi$ is inflation, $e$ is an exogenous shock that for simplicity will be assumed to be an independent white noise process. $\gamma$ is the parameter indexing the degree of inertia of the economy. With $\gamma = 0$ the economy has no inflation inertia (it is completely forward-looking), a case that has been studied, among others, by Clarida, Galí and Gertler (1999). At the other extreme, with $\gamma = 1$ the economy is completely inertial (backward-looking), a case that has been explored by Rudebush and Svensson (1999). Model (1) nests all intermediate cases. We will assume, for simplicity, that the output gap is directly under the control of the monetary authority, as for example in Clarida, Galí and Gertler (1999) or Woodford (2003). This allows us to neglect the aggregate demand equation in solving for the (optimal or robust) $x$ (as long as no weight is given in the loss function to the
values assumed by the interest rate). Moreover, we sidestep in this way the thorny issue of the implementability of the optimal rule, as it is well known that the interest rate rule that implements the output gap “rule” might be indeterminate (see the discussion in Clarida, Galí and Gerlter, 1999, and in Woodford, 2003).

The policymaker is assumed to have the following loss function:

\[ L_t = (1 - \beta) E_t \sum_{\tau=0}^{\infty} \beta^\tau [(\pi_{t+\tau})^2 + \alpha x_{t+\tau}^2], \]

(2)

where \( \alpha \) is a parameter that reflect the weights attached by the policymaker to the variability of the output gap relative to deviations of inflation from the central bank’s target, assumed to be zero for convenience. For \( \beta \to 1 \) the loss converges to a weighted sum (with weight \( \alpha \)) of the unconditional variances of inflation and output gap.

We will assume that the central bank can credibly commit itself to follow a given policy rule, and that it observes the realisation of the shock before choosing its instrument. Under these assumptions, it can easily be shown that the optimal monetary policy, for given values of the model and preference parameters \([\alpha, \beta, \gamma, \lambda]\), takes the form:

\[ x_t^* (\gamma) = ax_{t-1} + b\pi_{t-1} + ce_t, \]

(3)

where the coefficients \( a, b, c \) are appropriate functions of \([\alpha, \beta, \gamma, \lambda]\), although for ease of notation we highlight in the optimal rule only the dependence on \( \gamma \). In particular, when \( \gamma = 0 \), equation (3) specializes to:

\[ x_t^* (0) = \frac{\alpha}{\lambda^2 p_f + \alpha \beta} x_{t-1} - \frac{\lambda}{\lambda^2 p_f + \alpha \beta} e_t, \]

(4)

with \( p_f = \frac{\lambda^2 - \alpha (\beta - 1) + \sqrt{(\lambda^2 - \alpha (\beta - 1))^2 + 4 \lambda^2 \alpha \beta}}{2 \lambda^2} \), whereas when \( \gamma = 1 \) equation (3) specializes to:

\[ x_t^* (1) = - \frac{\lambda (1 + \beta p_b)}{\alpha + \lambda^2 (1 + \beta p_b)} (\beta \pi_{t-1} + e_t), \]

(5)

with \( p_b = \frac{-\lambda^2 + \alpha (\beta^2 + 1) \sqrt{(\lambda^2 + \alpha (\beta^2 + 1))^2 - 4 \alpha^2 \beta}}{2 \lambda^2 \beta} \). For intermediate values of \( \gamma \) the expressions for \( a, b, c \) are much more complicated. With the notation
introduced in the previous Section, we have that under (4) and assuming that $\gamma = 0$ the value of (2), i.e. the optimized loss, is given by:

$$L(x^*(0), 0) = \frac{\alpha \beta}{\lambda^2 p_f + \alpha \beta},$$  \hspace{1cm} (6)

while under (5) and assuming that $\gamma = 1$ the optimized loss is given by:

$$L(x^*(1), 1) = \frac{p_b}{\beta}. \hspace{1cm} (7)$$

In the following we will assume that the central bank is uncertain about the parameters of model (1). For the sake of simplicity, in this Section we will restrict the only uncertainty the central bank faces concerning the state of the economy to be the degree of inertia $\gamma$, and we will consider a number of values for $\gamma$ ranging from 0 to 1. In Section 4 we will extend the analysis to consider the case in which the central bank is uncertain about both $\gamma$ and $\lambda$.

In the baseline calibration of model (1), we take a value of the sensitivity of inflation to output gap changes, $\lambda = 0.1$, which is about midrange among the available estimates; a value of the discount factor that corresponds to a 1% real interest rate, $\beta = 0.99$; we take the case in which inflation and output gap variability have equal weights in the central bank loss function, $\alpha = 1$. We will explore the sensitivity of our results to alternative values.

With the baseline parameter values, the minimal loss when the model is purely forward-looking, $L(x^*(0), 0)$, amounts to 0.9. In the backward-looking case, $L(x^*(1), 1)$, it is equal to 8.23. It is important to notice that these two optimized losses differ by a factor which is almost 10. Larger values for $\lambda$, smaller values for $\beta$ and for $\alpha$, reduce the value of the ratio $L(x^*(1), 1) / L(x^*(0), 0)$, but even under rather extreme assumptions it remains above 2 (for example, with $\lambda = 0.15$, at the top of the range of available estimates, $\beta = 0.95$, corresponding to a real interest rate of more than 5%, and $\alpha = 0.1$, i.e. with the weight of inflation in the loss being 10 times that on the output gap, the ratio $L(x^*(1), 1) / L(x^*(0), 0)$ is almost 2.5). This is precisely the situation that, as argued in the previous Section, makes the minimax approach potentially misleading. Let us now turn to the issue of robustness.

We first assess the robustness of the optimal monetary policy rules. More specifically, we compute for each of a number of possible values of $\gamma$, and for
various values of the preference parameter $\alpha$, the optimal rule $x^*(\gamma)$ (the other parameters are kept at their baseline values; similar results are obtained with alternative values\(^1\)). As a convention, we will denote $\gamma^e$ the degree of inertia present in the economy and $\gamma^p$ the one assumed in designing the policy. We then assess the loss that would be incurred when the rule optimized assuming that the degree of inertia is $\gamma^p$ is implemented in an economy characterised by a degree of inertia $\gamma^e$, $L(x^*(\gamma^p), \gamma^e)$, and we scale it by the optimal loss $L(x^*(\gamma^e), \gamma^e)$.

Table 5, which is similar to Table 5 in LWW, presents the results for five alternative values of $\gamma$, namely 0, 0.25, 0.5, 0.75, 1. The generic cell corresponding to the couple $(\gamma^e, \gamma^p)$ gives the relative loss $L(x^*(\gamma^p), \gamma^e) / L(x^*(\gamma^e), \gamma^e)$. Each column therefore presents five values of the relative loss of the policy rule which would be optimal for the value $\gamma^p$ that appears in the heading of the column, when the true state of the economy is one of the values $\gamma^e$ that appear in the various rows.

\begin{table}
\begin{tabular}{|c|c|c|c|c|c|}
\hline
$\gamma^e \backslash \gamma^p$ & 1.00 & 0.75 & 0.50 & 0.25 & 0.00 \\
\hline
1.00 & 1.000 & 1.001 & 1.045 & 1.470 & 3.415 \\
0.75 & 1.002 & 1.000 & 1.064 & 1.588 & 3.711 \\
0.50 & 1.143 & 1.103 & 1.000 & 1.270 & 1.980 \\
0.25 & 1.659 & 1.574 & 1.170 & 1.000 & 1.035 \\
0.00 & 2.102 & 2.103 & 1.377 & 1.032 & 1.000 \\
\hline
\end{tabular}
\end{table}

\(^1\)In particular, changes in the values of $\lambda$ significantly affect only the last column in each of the panels, corresponding to the relative loss of the rule which is optimal in a purely forward-looking economy (for example, considering the case $\alpha = 1$, the largest relative loss falls to 2.4 for $\lambda = 0.05$, rises to 4.8 for $\lambda = 0.15$). A sizeable reduction in the value of the discount factor, $\beta$, reduces the differences between the relative losses of the various rules and improves the overall performance. The opposite effects are observed for larger values of $\beta$. It should however be noted that a value of $\beta$ significantly smaller than 1 makes the departure from a vertical Phillips curve not negligible, and is therefore an unpalatable assumption. If we distinguish between the discount factor and the slope of the Phillips curve, keeping the latter near-vertical, the results mentioned in the text are not significantly modified even for a discount factor as small as 0.95.
Table 5b: $\alpha = 0.1$

<table>
<thead>
<tr>
<th>$\gamma^e \setminus \gamma^p$</th>
<th>1.00</th>
<th>0.75</th>
<th>0.50</th>
<th>0.25</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.000</td>
<td>1.001</td>
<td>1.054</td>
<td>1.508</td>
<td>7.133</td>
</tr>
<tr>
<td>0.75</td>
<td>1.003</td>
<td>1.000</td>
<td>1.051</td>
<td>1.486</td>
<td>6.633</td>
</tr>
<tr>
<td>0.50</td>
<td>1.097</td>
<td>1.058</td>
<td>1.000</td>
<td>1.167</td>
<td>2.008</td>
</tr>
<tr>
<td>0.25</td>
<td>1.414</td>
<td>1.327</td>
<td>1.111</td>
<td>1.000</td>
<td>1.070</td>
</tr>
<tr>
<td>0.00</td>
<td>1.813</td>
<td>1.686</td>
<td>1.328</td>
<td>1.055</td>
<td>1.000</td>
</tr>
</tbody>
</table>

Table 5c: $\alpha = 2.0$

<table>
<thead>
<tr>
<th>$\gamma^e \setminus \gamma^p$</th>
<th>1.00</th>
<th>0.75</th>
<th>0.50</th>
<th>0.25</th>
<th>0.00</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00</td>
<td>1.000</td>
<td>1.001</td>
<td>1.038</td>
<td>1.402</td>
<td>2.676</td>
</tr>
<tr>
<td>0.75</td>
<td>1.002</td>
<td>1.000</td>
<td>1.059</td>
<td>1.539</td>
<td>2.958</td>
</tr>
<tr>
<td>0.50</td>
<td>1.144</td>
<td>1.110</td>
<td>1.000</td>
<td>1.278</td>
<td>1.839</td>
</tr>
<tr>
<td>0.25</td>
<td>1.664</td>
<td>1.601</td>
<td>1.172</td>
<td>1.000</td>
<td>1.026</td>
</tr>
<tr>
<td>0.00</td>
<td>2.061</td>
<td>1.987</td>
<td>1.354</td>
<td>1.025</td>
<td>1.000</td>
</tr>
</tbody>
</table>

We believe three main lessons can be drawn from this table. First, the monetary policy rule that is optimal for a given model is in general not very robust when implemented in models that differ from the original one for more than just a small perturbation. In particular, the degree of inflation inertia of the economy is one feature that seems to have a large impact on the performance of a monetary policy rule across models. Second, monetary policy rules optimal for economies without inflation inertia show a particularly acute lack of robustness. While the rule that is optimal for a completely inertial economy ($\gamma^p = 1$) would generate a loss which is around 2 times the optimal loss if the true $\gamma^e$ is 0, the rule that is optimal for an economy with no inertia ($\gamma^p = 0$), when implemented in a very inertial economy would generate losses that are 3 to 7 times the optimal loss, depending on the weight attached to output variability. Third, the only rule that seems to have a moderate degree of robustness is that optimized for an economy with an intermediate degree of inertia (in the table, $\gamma^p = 0.5$). For this rule, the loss would depart from the optimal one by at most 30-40%. These lessons are completely consistent with earlier findings in the literature, as presented in LWW, as well as in
LW and Coenen, in spite of the fact that these papers explored richer sets of models and somewhat different classes of monetary policy rules. Therefore, our focus on just the degree of inflation inertia seems not to be overly restrictive, and in particular should not prevent a meaningful discussion of the robustness properties of monetary policy rules.

To this we now turn, computing three alternative "robust" monetary policy rules.

4 Which monetary rule is more robust to alternative degrees of inflation inertia?

We consider a relatively fine grid of values for $\gamma$ in the interval $[0,1]$. In the class of rules of the form (3) we select the three coefficients $a$, $b$, $c$ in such a way as to minimise, across the possible values of $\gamma$, either:

1. the maximum loss a rule would lead to. This is what we term the minimax rule ($x_t^M$); or
2. the (equally weighted) average loss a rule would lead to. This is what we term the Bayesian rule ($x_t^B$); or
3. the maximum relative loss a rule would lead to. This is what we term the relative minimax rule ($x_t^R$).

The rules are computed through numerical optimization. For the baseline calibration, the rules computed are, respectively:

$$x_t^M = 0.039 \times x_{t-1} - 0.791 \times \pi_{t-1} - 0.833 \times e_t,$$  \hspace{1cm} (8)

$$x_t^B = 0.558 \times x_{t-1} - 0.327 \times \pi_{t-1} - 0.446 \times e_t,$$  \hspace{1cm} (9)

$$x_t^R = 0.605 \times x_{t-1} - 0.292 \times \pi_{t-1} - 0.124 \times e_t.$$  \hspace{1cm} (10)

Results for the baseline calibration of the model are presented in Figure 1.
The figure shows the relative loss function of the 3 rules. The minimax rule turns out to be not very robust. Its loss is very close to the optimal loss for values of the inertia parameter above 0.7 (in fact the minimax rule coincides with the rule that is optimal when $\gamma = 0.96$), but it rapidly and significantly departs from the optimal loss for values of $\gamma$ below that threshold, reaching values that are between 50% and 100% larger than the optimal loss for $\gamma$ in the range [0, 0.4]. Therefore, the minimax rule is only really robust for economies with a rather high degree of inertia, and its stabilisation ability deteriorates dramatically in less inertial economies. The Bayesian rule is considerably better. Its losses remain within 5% of the optimal losses for values of $\gamma$ between 0.5 and 1, within 20% for values of $\gamma$ between 0.3 and 0.5, and only for economies with very little inertia depart significantly from the optimal losses, though by much less than with the minimax rule. The stabilisation performance of both rules, however, is clearly better for more inertial economies and increasingly deteriorates as the economy becomes less and less inertial.\footnote{The deterioration is not strictly monotonic for the Bayesian rule, whose loss is very close to the optimal one for values of $\gamma$ close to 0.5} The relative minimax rule is instead different. Its stabilisation performance, relative to the optimal stabilisation implementable in the absence of uncertainty about the degree of inflation inertia, is broadly
constant across the various economies considered, with losses that exceed the optimal ones by at most 12%. The price paid is that the relative minimax rule is never really close to the optimal rule, as its losses are always at least 3% larger than the optimal ones. This is, however, precisely what one would expect from a robust rule.

Figures 2-3 replicate the exercise presented in Figure 1 for different values of the preference parameter $\alpha$ ($\alpha = 0.1$ and $\alpha = 2$; the other parameters are left at their baseline values).

Figure 2: Relative losses of alternative rules ($\alpha = 0.1$)
Figure 3: Relative losses of alternative rules ($\alpha = 2$)

Figures 4-5 do the same for different values of $\beta$ ($\beta = 0.95$ and $\beta = 0.999$; again, the other parameters are left at their baseline values). In the next section the possibility that $\lambda$ (the inflation responsiveness to output gap variations) is uncertain, along with $\gamma$, is explored.

Figure 4: Relative losses of alternative rules ($\beta = 0.95$)
The results presented in figures 2-5 are essentially the same as those shown in figure 1. In particular, in all cases the relative minimax rule generates losses that, relative to the optimal ones, do not vary much with the degree of inertia, and never exceed the latter by more than 15%. Conversely, the other two approaches to robustness seem to be much more sensitive to the degree of inflation inertia, with the possibility of sizeable (for the Bayesian rule) and big (for the minimax rule) departures from the optimal loss. The only notable difference, compared with the baseline calibration, occurs for small values of the discount factor ($\beta = 0.95$ corresponds to a real interest rate of about 5%). In that case, the relative losses of all three rules are considerably smaller (the largest relative loss, obtained with the minimax rule when the economy is completely forward looking, is of the order of 40%). It should however be pointed out that low values of $\beta$ imply an implausible, significant departure from nominal neutrality (the long run Phillips curve is not vertical). Had we lowered the discount factor while at the same time kept the Phillips curve near-vertical the shrinking of the relative losses would have essentially disappeared.
4.1 Uncertainty on both $\gamma$ and $\lambda$

While extremely simple, model (1) spans, as $\gamma$ varies, a rather wide range of alternative representations of the economy. Indeed, large enough that the optimal monetary policy rule for any of those representations can lead to a very poor stabilisation performance in some of the alternative representations, and large enough to assess the robustness of the policy rules we consider. Yet in this sub-section and in the following section we further enlarge the set of models against which we check the robustness of the various policy rules.

Here we generate a larger set of alternative models of the economy simply by allowing for alternative values of the parameter that measures the inflation responsiveness to output gap variations ($\lambda$), in addition to the range of values for the inflation inertia parameter considered so far. We compute, as before, the minimax, Bayesian and relative minimax rules considering a two dimensional grid for the couples ($\gamma$, $\lambda$). The rules obtained do not differ much from those computed in the previous Section:

$$x^M_t = 0.072 \times x_{t-1} - 0.683 \times \pi_{t-1} - 0.742 \times e_t,$$  \hspace{1cm} (11)

$$x^B_t = 0.561 \times x_{t-1} - 0.317 \times \pi_{t-1} - 0.438 \times e_t,$$  \hspace{1cm} (12)

$$x^R_t = 0.527 \times x_{t-1} - 0.331 \times \pi_{t-1} - 0.129 \times e_t.$$  \hspace{1cm} (13)

We then plot the relative losses of these three rules:
Results, for $\alpha = 1$ and $\beta = 0.99$, are presented in figure 6. Very similar pictures obtain for other values of $\alpha$; as in the previous section, a large drop in $\beta$ would reduce all relative losses and shrink the differences in the performance of the various rules, but would also correspond to a theoretically unpalatable assumption of nominal non-neutrality. From figure 6 we get essentially the message as in the previous section, since the variation in the parameter $\gamma$ seems to be the dominant factor in changing the relative losses of the various rules.

5 Robustness in four estimated models of the Euro area

While the simple model used so far is a standard workhorse in monetary economics and allowed us to span a large set of alternative hypotheses on the
monetary transmission mechanism, it is too simple to have a strong claim of empirical realism. An interesting further check of the performance of the various notions of robustness is obtained when they are faced with alternative, and yet empirically plausible, representations of the economy. This is what we do in this section, in which we apply the competing concepts of robustness illustrated above to four different models of the euro area. All of those models have been developed, estimated and used within the European System of Central Bank (ESCB) in recent years. Those models differ considerably along a number of dimensions: scope; size; degree of aggregation; relevance of forward-looking behavioural mechanisms; adherence to microeconomic foundations. Being so markedly different from one another, they arguably cover a wide range of features that are a priori likely to be of relevance when designing monetary policy rules for the euro area.

The following models are considered:

(1) the DSGE-type model by Smets and Wouters (2004; SWEAR) incorporates a number of recent theoretical advances as to the behaviour of fully optimising forward-looking agents; its degree of adherence to microeconomic foundations is highest within the set of models we consider. Also, the degree of forward-lookingness is relatively high;

(2) the model by Coenen and Wieland (2003; CW) embodies Taylor-style staggered wage contracts; hence, price formation is at least partially forward-looking, as in SWEAR, but the degree of consistency with microeconomic foundations is more relaxed than in the case of the previous model. Both models are meant to capture the aggregate dynamics of a few key macroeconomic variables of the euro area.

It should be stressed that neither of those first two models can be characterized as being totally forward-looking. For example, the inertial component in the Phillips curve appearing in the SWEAR model is of the order of 1/3.

(3) the Area-Wide Model (AWM; see Fagan, Henry and Mestre (2001) for a description of the model, and ECB(2002) for a description of its role in the forecast process of the ESCB), where, by contrast, expectations are essentially backward-looking; the model is once again meant to capture the dynamics of a number of euro area variables; in terms of size, it is relatively
large, as it embodies approximately 70 variables.

(4) the Disaggregate Euro-Area Model (DEAM; see Angelini, Del Gio- 
vane, Siviero and Terlizzese (2002) and Siviero Monteforte (2002)), which 
separately describes consumer price inflation and output gap dynamics in 
the three largest economies of the euro area (Germany, France and Italy), 
and thus differs sharply from the other three models, in that it is the only 
 disaggregate one, allowing for heterogeneity in the structure of those three 
economies; as in the AWM, expectations are backward-looking.

Adalid, Coenen, McAdam and Siviero (2004) use those four models to 
investigate the features of the robust monetary policy for the euro area, if it 
does exist. Consistently with results presented in Section 3, they conclude 
that, while rules derived from forward-looking models perform disastroously 
in backward-looking ones, the reverse does not hold; nevertheless, rules op-
timized for backward-looking models generate a substantial increase in the 
optimal loss when implemented in forward-looking models. Also, they 
find that equal-weighting Bayesian rules tend to perform satisfactorily in the four 
models at hand.

The large differences among these four models prevent us from identifying 
a single specification that encompasses the optimal monetary policy rule in 
each of them (this was in contrast possible in the models considered so far). 
We therefore restrict our analysis to simple interest rate rules of the Taylor 
type, where the policymaker is allowed to react only to current inflation and 
output gap, and to the lagged value of the policy instrument (the short term 
interest rate, $i_t$):

$$i_t = ax_t + b\pi_t + ci_{t-1}. \quad (14)$$

We obtain very similar results if the class of rules considered is expanded 
to allow one extra lag for each of those three variables. We also adopt a 
slight generalisation of the loss function (2), in which we also include a term 
penalising the volatility of the policy instrument, $\Delta i_t$, with relative weight 
$\mu$, so that the loss function becomes:
\[ L_t = (1 - \beta)E_t \sum_{\tau = 0}^{\infty} \beta^\tau [ (\pi_{t+\tau})^2 + \alpha x_{t+\tau}^2 + \mu (\Delta i_t)^2]. \]  

(15)

For each model \( m = 1, 2, 3, 4 \), we compute numerically the optimized rule \( i_t = a^m x_t + b^m \pi_t + c^m i_{t-1} \) and the corresponding optimized loss \( L_m^* \). We then compute, as in the previous section, the values of the coefficients \( (a, b, c) \) to minimise:

(i) either the maximal loss across the four models (the minimax rule),
(ii) or the average loss across models (the Bayesian rule),
(iii) or the maximal ratio between the loss and \( L_m^* \) (the relative minimax rule).

Concerning the latter rule, a slight difference with what we did before should be noted. So far, the losses that would obtain when implementing a given rule in a given model were scaled by the loss associated with the rule that achieves the first best in that particular model (what we denoted as the optimal loss). Both the given rule and the optimal (model-dependent) rule belonged to the same class (equation (3)). Now, the scaling factor we use (in each model) is the loss associated with the rule that is optimal, for that model, within a particular restricted class (equation (14)). As before, however, both the rule that we want to assess and the optimized (model-dependent) rule belong to the same class.

Table 6 reports the results, for two different sets of preference parameters. For both sets of preference parameters, the minimax rule is the rule that is also optimal in one of the models (DEAM), and performs rather well in the other backward-looking model in the set (the AWM), while it performs very poorly in the two less inertial models (with relative losses of the order of 50%). The Bayesian rule is much better, with relative losses never exceeding 25%. It shows however a non negligible variability across models; in particular, it tends to be close to the optimal rule for DEAM, and distinctly further from the optimum for the other models. Finally, the relative minimax results in a non-trivial reduction in the largest relative loss. For instance, with \( \alpha = 0.33, \mu = 0.1 \), the largest relative loss falls from 22 to 15 per cent;
similarly, with $\alpha = 1, \mu = 0.1$, the largest relative loss falls from 24 to 17 per cent. Moreover, the relative loss is very stable across models.

These results are qualitatively similar to those obtained before, considering the set of models generated by equation (1) when either $\gamma$ alone or both ($\gamma, \lambda$) were varied. The differences among the three rules are quantitatively less stark.\textsuperscript{6} However, they confirm the good performance of the relative minimax rule.

Table 6: Performance of alternative rules in four models of the euro area.

<table>
<thead>
<tr>
<th></th>
<th>in AWM</th>
<th>in DEAM</th>
<th>in CW</th>
<th>in SWEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimax rule</td>
<td>1.19</td>
<td>1.00</td>
<td>1.50</td>
<td>1.29</td>
</tr>
<tr>
<td>Bayesian rule</td>
<td>1.22</td>
<td>1.05</td>
<td>1.22</td>
<td>1.13</td>
</tr>
<tr>
<td>Relative minimax</td>
<td>1.15</td>
<td>1.15</td>
<td>1.15</td>
<td>1.12</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>in AWM</th>
<th>in DEAM</th>
<th>in CW</th>
<th>in SWEAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Minimax rule</td>
<td>1.05</td>
<td>1.00</td>
<td>1.47</td>
<td>1.45</td>
</tr>
<tr>
<td>Bayesian rule</td>
<td>1.23</td>
<td>1.04</td>
<td>1.24</td>
<td>1.22</td>
</tr>
<tr>
<td>Relative minimax</td>
<td>1.17</td>
<td>1.17</td>
<td>1.17</td>
<td>1.17</td>
</tr>
</tbody>
</table>

Note: for each rule and each model, the table reports the ratio between the loss associated with each rule and the loss delivered by the optimal simple rule computed for each of the four models.

6 Concluding remarks

The quest for robust policy rules in the face of model uncertainty is gaining increasing attention in the recent literature. It is not a new concern, however.\textsuperscript{6}

\textsuperscript{6}We conjecture that this partly results from the more restricted class of rule that is now considered, and partly from the lack, among the four models, of a fully forward-looking one. Indeed, the analysis in the previous section shows that the deterioration of the stabilising performance of minimax and Bayesian rules is particularly acute for very low degrees of inertia.

27
Albert Ando wrote in 1988, in an unpublished manuscript (entitled ”Design of Stabilisation Policies Reconsidered”): “We are, of course, prepared to admit that our estimates may be inaccurate or out of date, or perhaps that the estimated parameters ...may indeed change...in response to policy actions as the Lucas critique would suggest. If such a situation becomes a problem, we should find that the inflation rate is either accelerating or decelerating even when the actual rate of unemployment is equal to [our estimated] NAIRU”.

To cope with this possibility Ando devised stabilisation rules in which the targets (as for example the NAIRU) were supplemented with ”moving” correction terms, as a function of recent observations, inducing changes in the policy instrument that were able to keep the economy close enough to the intended path even when the system turned out to be different from what was originally thought when designing the rule.

In a similar spirit, in this paper we explored the effectiveness of an alternative approach to the design of robust policy rules, borrowed from the theoretical computer science literature. We compared the stabilisation performance of this approach, which we termed in this paper ”relative minimax”, with two alternatives widely used in the economic literature (”pure” minimax and flat-prior Bayesian) and we showed that relative minimax offers better protection against large errors in assessing the true degree of inflation inertia in the economy. We also showed that relative minimax retains its good stabilisation properties across four widely different, estimated models of the euro area economy.
References


