WORKING PAPER

DIPARTIMENTO DI ECONOMIA PUBBLICA

Working Paper n. 98

Francesco Giuli

Robust Control in a Sticky Information Economy

Roma, Aprile 2007

UNIVERSITA' DEGLI STUDI DI ROMA
"LA SAPIENZA"

Robust Control in a Sticky Information Economy

Francesco Giuli*

University of Rome I
francesco.giuli@uniroma1.it

ABSTRACT *This paper analyzes the behavior of a central bank under strong ("Knightian") uncertainty when the short run trade-off between output and inflation is represented by the Sticky Information Phillips Curve recently proposed by Mankiw and Reis (2002). By solving the robust control problem analytically, this paper elucidates the economic mechanisms at play in a sticky information economy and shows how and why the robust monetary policy in this economy differs from the optimal one identified by Ball, Mankiw and Reis (2005).*

1 Introduction

Increasing research activity has recently focused on the effects of uncertainty in the conduct of monetary policy. Whereas the traditional literature dealt with uncertainty by adding exogenous disturbances to a linear-quadratic economic framework, which guarantees that the certainty equivalence holds true, a more interesting kind of uncertainty faced by central bankers, known as model uncertainty, is progressively attracting the interest of scholars.

In Brainard's (1967) classic analysis, model uncertainty, understood in the sense that the marginal effect of a policy instrument on a macroeconomic outcome is described by a parameter distribution, leads to cautious policy. This result has been recently challenged by a body of research investigating how monetary policy should be conducted when the central bank knows the structural relations of the economy (the reference model), but she faces uncertainty about the value of the parameters or about the stochastic structure of the model. This problem has been tackled using robust control techniques: a min-maximizing central bank, aiming to avoid poor performances associated to "unfortunate" parameter configurations, derives robust monetary policy rules under the worst possible specification (the worst case model, or scenario). In Dynamic Stochastic General Equilibrium (DSGE) models, with monopolistic competitive producers experiencing frictions in price setting which cause monetary non-neutrality (i.e., the New Keynesian approach), this behavior leads, in general to an anti-attenuation result (i.e., the policy-maker reacts more vigorously to shocks), but this conclusion not always applies.

The anti-attenuation result typically emerges in closed economies where uncertainty is modelled through a set of additive perturbations controlled by a fictitious "evil agent" who tries to maximize the central bank's loss by setting a specification error of bounded size (i.e., the *unstructured model uncertainty* proposed by Hansen and Sargent, HS, 2004), whereas the opposite may occur in open economies (Leitemo and Söderström, 2004a; 2004b). In any case, the optimal trade-off between targets is not affected by the preference for robustness (i.e., Walsh's (2004) equivalence result). Yet, even if the desired trade-off is the same, target variables turn out to be more volatile. The same non univocal conclusions are reached in forward looking New Keynesian models under *parameter uncertainty* where, differently from HS, the evil agent controls one or more coefficients of the structural equations (Giannoni, 2002). In such a context, the robust policy may require to react more or less strongly than in the certainty-equivalence environment to fluctuations in inflation and in the output gap according to the assumptions which are made on the central bank's objective function or on the parameter which is subject to uncertainty (Kara, 2002). Min-max policy prescriptions are hence fragile with respect to changes in the description of the economy, starting from the very characterization of

uncertainty.

This conclusion provides a strong motivation to analyze the behavior of a central bank under strong ("Knightian") uncertainty when the short run trade-off between output and inflation is represented by the Sticky Information Phillips Curve (SIPC) recently proposed by Mankiw and Reis (MR, 2002). By solving the robust control problem analytically, this paper shows how and why the robust monetary policy in this economy differs from the optimal one identified by Ball, Mankiw and Reis (BMR, 2005).

The paper is organized as follows. In section 2 we summarize the BMR's (2005) model, highlighting the optimality of a price targeting rule.

In section 3 we compute the robust policy under the unstructured approach. In line with most of the literature[1], if the central bank and the evil agent act simultaneously, in the worst case model targets variability increase with the preference for robustness. When the policy maker employs this robust instrument rule but the actual misspecification is zero (the *approximated model* solution), an anti-attenuation result obtains: the policy maker varies his instrument more vigorously than he would in the certainly equivalent case. This over-reaction produces a greater or smaller stabilization of the price level depending on the value of the desired trade-off in the rational expectation model. When in this model the policy maker reacts to the cost shock by injecting (mopping up) money, in the approximated model he reacts by injecting (mopping up) even more money. Thus, whereas in sticky price economies, where it is optimal to target inflation, the central bank never reacts to a cost shock by producing a price deflation, this may well happen in a sticky information economy under a price targeting regime. If the central bank acts as a Stackelberg leader with respect to the evil agent, the Walsh equivalence result is no longer valid: the policy maker is less aggressive in the stabilization of the price level. The macro-economic consequences of this behavior in the approximating model solutions are driven by the interaction of two forces which may operate in opposite directions: the first one is the anti-attenuation effect in the use of the nominal policy instrument (the same at play under the Nash timing); the second one is due to the new desired trade-off that favours output stabilization. We conclude the section by tackling the issue of the optimal design of institutions, i.e., the central bank's degree of conservativeness that maximizes social welfare.

Section 4 derives the robust policies according to Giannoni's parametric approach and shows that, also in this case, the central bank seeks to stabilize the price level less than in the model without parameter uncertainty.

Section 5 concludes.

2 The Ball Mankiw Reis model

MR (2002) propose a relation between output and inflation based on the sticky information hypotheses. The SIPC is based on the idea that symmetric monopolistically competitive firms choose their optimal price in each period, but

---

[1] See Walsh (2004), and Leitemo and Söderström (2004a)

the information set they use when solving the profit maximization problem may not be the current one. Information is sticky in the sense that firms update it sporadically. The process that guides information gathering is similar to that formulated by Calvo (1983) for price adjustment: In the latter model, in each period, only a fraction $(1 - \omega)$ of firms can adjust their prices; in the former model in each period only a fraction $(1 - \omega)$ of firm obtains a new information set, while the remaining ω continue to fix prices on the basis of the old one. The log-linearization of the first order conditions of the consumers' and the producers' optimization problem yields the Phillips curve and the price level equations[2]:

$$\pi_t = \frac{(1 - \omega)}{\omega} \left[\alpha x_t + u_t \right] + (1 - \omega) \sum_{k=0}^{\infty} \omega^k E_{t-1-k} \left(\pi_t + \alpha \Delta x_t + \Delta u_t \right) \quad (1)$$

$$p_t = (1 - \omega) \sum_{k=0}^{\infty} \omega^k E_{t-k} \left(p_t + \alpha x_t + u_t \right) \quad (2)$$

Inflation (π_t) depends on the current output gap $(\widehat{x}_t)$ and the cost push shock (u_t), the past expectations on current inflation, the growth of the output gap and the shock. α is a combination of fundamental parameters which represents the sensitivity of firm's optimal price to the expected deviation of output. The price level depends both on current and past expectations on the output gap and the cost push shock[3].

As MR (2002) and BMR (2005) assume, the demand side of the economy is represented by a quantity theory equation:

$$x_t = m_t - p_t \quad (3)$$

where m_t is the log of the money supply. The policy instrument for period t is chosen at time $t - 1$, before the realization of u_t; the policy decisions hence affect the economy with a lag.

As for the welfare function, BMR define the current period welfare as the average level of utility across all households. A second-order approximation of households' loss is found to be a function of the variability of output around its flexible price level and of the cross-sectional variability of output across different firms (Woodford, 2003):

$$w_t = w \left(Var \left(x_t \right), Var_j \left(y_{jt} - y_t \right) \right)$$

As the natural output differs from efficient output only by a constant independent of policy, the variability of output around the natural level measures the variability around efficient output. The cross-sectional variability of output across different firms enters the loss function since variability at the firm level is

_____

[2] For a complete derivation of the SIPC see, among others, Khan and Zhu (2002) and Tranbandt (2003).

[3] BMR consider also productivity and demand shocks, but we omitt these for simplicity.

inefficient because it creates variability in the labor supply around the efficient level. Following Woodford (2003), the period loss can be formally written as:

$$w_t = Var(x_t) + bE[Var_j(p_{jt} - p_t)] + K + \| \varnothing \| \qquad (4)$$

where b is a combination of fundamental parameters and K collects the terms independent of policy (for the sake of simplicity, we set $K = 0$). In equation 4, we have used:

$$Var_j(y_{jt} - y_t) = \zeta^2 Var_j(p_{jt} - p_t)$$

to replace the cross-sectional variability of output across different firms with the price variability, ζ is the elasticity of demand.

Since $Var_j(p_{jt} - p_t) = \sum_{j=1}^{\infty} f(\omega, j)(p_{t+k} - E_{t+k-j}p_{t+k})^2$ [4],the central bank thus attempts to minimize:

$$W_t = \sum_{k=0}^{\infty} \beta^k w_{t+k} = \sum_{k=0}^{\infty} \beta^k x_{t+k}^2 + b \sum_{j=1}^{\infty} f(\omega, j)(p_{t+k} - E_{t+k-j}p_{t+k})^2 \qquad (5)$$

where $f(\omega, j) = \frac{(1-\omega)\omega^j}{(1-\omega^j)(1-\omega^{j+1})}$.

Equation (5) shows that the terms in the loss function depend on aggregate variables. Under sticky information, social welfare departs from the common micro-founded representation derived under sticky prices, where the cross-sectional variability of prices is determined by the current and lagged values of squared inflation (see, e.g., Woodford, 2003).

2.1 Optimal monetary policy in BMR

The optimal policy can be found by minimizing the loss function (5) with respect to the policy instrument subject to the demand side and the supply side of the model. BMR show that by combining equations (2) and (3) it is possible to determine a linear relation between p_t and the money supply target m_t. Since m_t is determined one period in advance and the shocks are unforecastable, there exists a linear relation between the money supply target and the expectation at time $t - 1$ (when the policy instrument is set) of the price level at time t ($E_{t-1}p_t$). Hence we can interpret $E_{t-1}p_t$ as the policy instrument.

The general moving average, MA(∞), representation of the price level and the output equilibrium paths are, respectively, $p_t = \sum_{j=0}^{\infty} \phi_j \epsilon_{t-j}$ and $\hat{x}_t = \sum_{j=0}^{\infty} \varphi_j \epsilon_{t-j}$, where ϵ_t are the innovations of the model; ϕ_j and φ_j are unknown coefficients. Because of the lag in the policy transmission mechanism the robust policy process is given by $E_{t-1}p_t = \sum_{j=1}^{\infty} \phi_j \epsilon_{t-j}$ and $E_{t-1}x_t = \sum_{j=1}^{\infty} \varphi_j \epsilon_{t-j}$.

Given the equation for the price level:

$$p_t = (1-\omega) \sum_{k=0}^{\infty} \omega^k E_{t-k}(p_t + ax_t + u_t)$$

---

[4] see BMR (2005) for technical details.

where $u_t = \sum_{j=0}^{\infty} \rho_j \epsilon_{t-j}$ is an arbitrary stationary process, the price level can be rewritten as:

$$\sum_{j=1}^{\infty} \phi_j \epsilon_{t-j} = \sum_{j=1}^{\infty} \Lambda^j \left(\phi_j \epsilon_{t-j} + a\varphi_j \epsilon_{t-j} + \rho_j \epsilon_{t-j} \right)$$

where $\Lambda^j = (1 - \omega) \sum_{k=0}^{j} \omega^k$.

Since this expression must hold for all possible realizations of ϵ_{t-j}, it follows that:

$$\varphi_j = \frac{1}{a} \left(\frac{1 - \Lambda^j}{\Lambda^j} \phi_j - \rho_j \right) \tag{6}$$

Since $p_t - E_{t-i} p_t = \sum_{j=0}^{i-1} \phi_j \epsilon_{t-j}$, we can write the objective function as:

$$L = \left[\sum_{j=1}^{\infty} \varphi_j^2 + b \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=1}^{i-1} \phi_j^2 \right] \sigma_\epsilon^2. \tag{7}$$

By minimizing (7) with respect to ϕ_j subject to (6), we obtain the optimal coefficients:

$$\phi_j^* = \frac{1}{a^2 b + \frac{1 - \Lambda^j}{\Lambda^j}} \rho_j \quad \text{for } j > 0 \tag{8}$$

and thus:

$$\varphi_j^* = - \frac{ab}{a^2 b + \frac{1 - \Lambda^j}{\Lambda^j}} \rho_j \quad \text{for } j > 0 \tag{9}$$

given the lag in the policy transmission mechanism the central bank cannot reply contemporaneously to the shock, hence in period $t = 0$ the price and the output coefficients are given by:

$$\phi_0 = \frac{1 - \omega}{1 - (1 - \omega)(1 - a)} \quad \text{and} \quad \varphi_0 = -\phi_0$$

The relation between the optimal coefficients is given by:

$$\varphi_j^* = -ab\phi_j^* \tag{10}$$

which implies the following targeting rule:

$$E_{t-1} p_t = k_t - \frac{1}{ab} E_{t-1} x_t \tag{11}$$

where k_t can be any deterministic path. As stressed by Hall (1984) and BMR (2005), optimal policy can be described as an *elastic price standard*, i.e. the central bank allows the price level to deviate from its target when output deviates from its natural rate. To understand this result, consider a cost push shock at time $t = 0$. Since policy responds with a lag, the informed firms immediately rise their prices and the aggregate price level increases. Consider an auto-correlated shock. In the following period the uninformed firms continue to

set prices at zero, but for the informed firms policy now produces its effect and induces them to set prices according to the optimal trade-off (11). As long as the shock persists the central bank allows output and the price level to deviate from their targets. When the shock vanishes all firms set the same zero price, whereas, under inflation targeting, in order to avoid disinflation, the central bank must induce the informed firms to set positive prices. This would cause output variability, since informed and uninformed firms are still setting different prices.

The desired targets trade-off is implemented through the following instrument rule in MA(∞)representation:

$$m_t^* = \sum_{j=1}^{\infty} \eta_j^* \epsilon_{t-j}$$

where the coefficients η_j^{rn} are given by:

$$\eta_j^* = \phi_j^* + \varphi_j^* = \frac{1-ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \tag{12}$$

For the purpose of our analysis it is important to highlight that the optimal policy predicts a trade-off between the target variables that can be implemented though either an increase or a reduction in the nominal policy instrument depending on the value of αb. If $\alpha b < 1$, in front of a cost shock the policy maker reacts by increasing the nominal money supply because the desired targets trade-off implies that an increase in the price level requires a less than proportional decrease in the output gap. By contrast if $\alpha b > 1$, the optimal trade-off predicts a more than proportional decrease in the output gap which, in its turn, implies a negative response of the nominal money supply. Under an inflation targeting regime such a policy behavior is sub-optimal, as inflation stabilization never requires a decrease in the price level.

3 The *Unstructured* model uncertainty approach

3.1 The Nash timing

In this section we apply the HS (2004) robust control techniques to the New-Keynesian model with sticky information summarized in the previous section. The *unstructured* model uncertainty approach conceives the specification errors as a serially correlated shock process stemming from omitted variables in the structural model of the economy. These shocks are assumed to be of bounded size, which can be viewed as a measure of the central bank's preference for robustness. The lack of a prior distribution for the shock process and the preference for robustness induce the policy maker to adopt a min-max strategy which can be represented through a "mind" game played by the central bank and a fictitious "evil agent", who represents the policy maker's fear concerning specification errors. The "evil agent" chooses the amount of misspecification so

as to maximize the central bank loss. In this subsection we assume that the two players act simultaneously while in the following subsection we analyze the Stackelberg equilibrium.

In order to model parameter uncertainty, we introduce in the price equation (2) a second source of disturbance denoted by z_t, which is added to the standard cost push shock u_t in the central bank's reference model.

Thus, the misspecificated price equation becomes:

$$p_t = (1 - \omega) \sum_{k=0}^{\infty} \omega^k E_{t-k} \left(p_t + \alpha x_t + u_t + z_t \right) \tag{13}$$

Similarly to the other variables, the general MA(∞) representation for the "optimal misspecification" chosen by the evil agent can be written as $z_t = \sum_{j=1}^{\infty} \gamma_j \epsilon_{t-j}$, with γ_j unknown coefficients. By substituting it in the price level equation we get:

$$\sum_{j=1}^{\infty} \phi_j \epsilon_{t-j} = \sum_{j=1}^{\infty} \Lambda^j \left[\phi_j \epsilon_{t-j} + a\varphi_j \epsilon_{t-j} - \left(\rho_j \epsilon_{t-j} + \gamma_j \epsilon_{t-j} \right) \right] \tag{14}$$

Since this expression must hold for all possible realizations of ϵ_{t-j}, it follows that:

$$\varphi_j = \frac{1}{a} \left(\frac{1 - \Lambda^j}{\Lambda^j} \phi_j - \rho_j - \gamma_j \right) \tag{15}$$

Since $p_t - E_{t-i} p_t$ is still given by $\sum_{j=0}^{i-1} \phi_j \epsilon_{t-j}$, we can rewrite the cost function as:

$$L_r = \left[\sum_{j=1}^{\infty} \varphi_j^2 + b \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=0}^{i-1} \phi_j^2 \right] \sigma_\epsilon^2 - \sum_{j=1}^{\infty} \theta \gamma_j^2 \sigma_\epsilon^2.$$

where θ represents the preference for robustness, or equivalently the set of models available to the evil agent against which the policy maker wants to be robust. When the degree of misspecification goes to zero, θ goes to infinity and we turn to the standard rational expectations model. An increase in the preference for robustness means a decrease in θ[5].

By minimizing the new objective function subject to equation (15) we obtain:

$$\phi_j = \frac{1}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \left(\rho_j + \gamma_j \right) \tag{16}$$

and thus:

$$\varphi_j = -\frac{ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \left(\rho_j + \gamma_j \right) \tag{17}$$

---

[5] See Hansen and Sargent (2004), Giordani and Soderlind (2004), and Leitmo and Soderstrom (2004)

8

Even if γ_j is still undetermined, at this stage we can note that the relation between the output and the price level coefficients is given by:

$$\varphi_j = -ab\phi_j \tag{18}$$

Hence, the optimal price-output trade-off is not affected by the preference for robustness, that is, the robust optimal targeting rule is the same as the optimal rule without fear of specification errors. This is the equivalence result derived by Walsh (2004) in sticky price New-Keynesian models.

The optimal amount of misspecification is obtained by maximizing the loss function with respect to γ_j, subject to equation (15):

$$\gamma_j = -\frac{\frac{1-\Lambda^j}{\Lambda^j}}{a^2\theta - 1}\phi_j + \frac{1}{a^2\theta - 1}\rho_j$$

By substituting equation (16) into the one above, we get:

$$\gamma_j^{rn} = \frac{b}{\theta\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right) - b}\rho_j \tag{19}$$

the amount of misspecification is increasing in the central bank's preference for robustness (θ), and it is negatively related to a, which captures the sensitivity of firms' optimal price to the expected deviation of output (when a decreases, larger movements in output are needed to affect inflation). As for parameter b (i.e., the weight associated to price variability in the loss function) we have $\frac{\partial \gamma_j^{rn}}{\partial b} > 0$; this means that when offsetting relative price variability becomes more costly the amount of misspecification increases.

By substituting equation (19) in equations (16) and (17) we obtain the price and the output coefficients in the worst case scenario:

$$\phi_j^{rn} = \frac{1}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}}\rho_j \tag{20}$$

$$\varphi_j^{rn} = -\frac{ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}}\rho_j \tag{21}$$

When $\theta \to \infty$, ϕ_j^{rn} and φ_j^{rn} collapse to the optimal coefficients in the absence of model misspecification (see equations (8) and (9)). When the fear of misspecification is very low (i.e., θ close to infinity) $\phi_j^{rn} > 0$ and $\varphi_j^{rn} < 0$: a cost push shock increases inflation and decreases output. When θ is so small that ϕ_j^{rn} and φ_j^{rn} have opposite signs, the policy maker faces a quite unrealistic situation, as a cost push shock would cause deflationary pressure. This is why, following most of the literature, we focus on small degrees of misspecification[6].

---

[6] See Leitemo and Söderström (2004), and Giordani and Soderlind (2004).

$\theta = a^{-2}$ is the cut-off value below which the model predicts unreasonable impulse responses. The relations between the misspecification, and the output and price coefficients are respectively:

$$\frac{\gamma_j^{rn}}{\varphi_j^{rn}} = -\frac{1}{a\theta} \tag{22}$$

$$\frac{\gamma_j^{rn}}{\phi_j^{rn}} = \frac{b}{\theta} \tag{23}$$

Moreover we can note that:

$$\frac{\partial \left|\varphi_j^{rn}\right|}{\partial \theta} < 0$$

$$\frac{\partial \phi_j^{rn}}{\partial \theta} < 0$$

thus, in line with standard results, both the price level and the output gap are more volatile in the worst case scenario, when the preference for robustness increases.

Remember that the analytical solutions in the worst case model are derived by assuming that the central bank's worst fear about parameter configuration results to be justified ex-post. This implies that we cannot say whether the greater volatility of macroeconomic variables (in comparison with the rational expectation solution) represents a welfare cost ascribable to the central bank's fear about misspecification or is exactly due to the realization of the specification errors. In other words, in the worst case model we cannot isolate the effect of the pure fear for misspecification on the policy behavior from the effect of the actual realization of a misspecificated Phillips curve[7]. Hence, any comparison between the rational expectations and the worst case model have to be taken *cum grano salis*, because they are slightly different models. This is why, following Hansen and Sargent (2004) and Leitemo and Söderström (2004a, 2004b), in order to isolate the consequences for the economy of the fear of misspecification we need to focus on the *approximating model solution.* This is found by assuming that the central bank sets its instrument with the aim to be robust against model misspecification but that the actual misspecification is zero.

Given the simple demand side of the model we can immediately find the MA(∞) for the instrument rule under the worst case scenario:

$$m_t^{rn} = \sum_{j=1}^{\infty} \eta_j^{rn} \epsilon_{t-j}$$

where the coefficients η_j^{rn} are given by:

$$\eta_j^{rn} = \phi_j^{rn} + \varphi_j^{rn} = \frac{1 - ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \tag{24}$$

---

[7] Equivalently we cannot say if the implementation of the optimal tarde-off (which is unaffected by the preference for robustness) is indeed reached through a more or less aggressive attitude in responding to the cost push shocks.

the approximating model solution (indicated by the superscript a) is then obtained by substituting equation (24) in the price level equation (14) setting $\gamma_j = 0$:

$$\phi_j^{an} = \frac{a + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)} \rho_j$$

and

$$\varphi_j^{an} = -\frac{a^2 b + ab\frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)} \rho_j$$

By comparing the coefficients of the rational expectation and of the approximating model we get:

$$\phi_j^{an} - \phi_j^* = \frac{\frac{b}{\theta}(1-ab)}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1}{a}\frac{1-\Lambda^j}{\Lambda^j} + 1\right)\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right)} \rho_j \gtrless 0 \quad \text{if} \quad b \lessgtr a^{-1}$$

(25)

$$\left|\varphi_j^{an}\right| - \left|\varphi_j^*\right| = -\frac{\frac{1-\Lambda^j}{\Lambda^j}\frac{b}{\theta}(1-ab)}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right)} \rho_j \lessgtr 0 \quad \text{if} \quad b \lessgtr a^{-1}$$

(26)

and

$$\eta^{rn} - \eta^* = \frac{\frac{b}{\theta}(1-ab)}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right)} \rho_j \gtrless 0 \quad \text{if} \quad b \lessgtr a^{-1} \qquad (27)$$

Thus, the consequence for the economy of an instrument rule sett with the aim to be robust against model misspecification when the actual misspecification is zero depends upon the central bank degree of conservatism. When $ab < 1$, the central bank in the original rational expectation model reacts to a cost push shock by expanding the nominal money supply (see equation (12)). In fact the desired trade-off (see equation (10)) implies that an increase in the price level is counterbalanced by a less then proportional contraction in output, that is, in the real money supply . On the contrary, when $ab > 1$, in front of a cost push shock the central bank will destroy money in order to induce a more than proportional output contraction. In the approximating model the central bank tries to implement the optimal targets' trade-off (see equation (18)) but, due to the concerns for model misspecification, the implied instrument rule is not designed for the approximating model but for the worst case one. This fear of misspecification produces an anti-attenuation result in the use of the policy instrument: when in the original model the policy maker reacts to a cost shock by injecting money, in the approximating model it overreacts by injecting even more money. If the optimal policy is to tight the nominal policy instrument, in the approximating model the nominal money reduction will be even greater (see equation (27)). Hence, the consequences for the economy of this policy behavior are to allow the price level (output gap) to deviate from the steady state more

11

(less) than under the rational expectation model when $b < a^{-1}$ and the opposite is true when $b > a^{-1}$ (see equations (25) and (26)). There is also a third case that emerges when $ab = 1$. In such a situation both models produce the same results, in fact the anti-attenaution principle cannot emerge when there is no policy reaction in the original model, that is $\eta^* = 0$. This means that in this particular parameter configuration the endogenous model trade-off between the price-level and the output gap reflects exactly the desired trade-off of the central bank.

To summarize, the effect of the fear of misspecification on the target variables is not univocal. Due to the price targeting regime, the anti-attenuation principle in the use of the policy instrument can result in a lower, in a greater or in the same price level stabilization depending on the targets trade-off in the central bank objective function. This, in its turn, implies that an increases in the preference for robustness causes the output gap and the reaction of the policy instrument (price level) to be less (more) volatile when $b < a^{-1}$ and *viceversa when* $b > a^{-1}$ as showed by the signs of the following derivatives:

$$\frac{\partial \phi_j^{an}}{\partial \theta} = -\frac{\frac{b}{\theta^2}\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right) a\,(1 - ab)}{\left[\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)\right]^2} \rho_j \lessgtr 0 \quad \text{if } b \lessgtr a^{-1}$$

$$\frac{\partial \left|\varphi_j^{an}\right|}{\partial \theta} = \frac{\frac{b}{\theta^2}\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)(1 - ab)\frac{1-\Lambda^j}{\Lambda^j}}{\left[\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)\right]^2} \rho_j \gtrless 0 \quad \text{if } b \lessgtr a^{-1}$$

$$\frac{\partial \eta^{rn}}{\partial \theta} = \frac{-\frac{b}{\theta^2}(1 - ab)}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)^2} \lessgtr 0 \qquad\qquad \text{if } b \lessgtr a^{-1}$$

All the above results are illustrated in figure 1 which shows the impulse responses of the money supply, the price level and the output gap under the approximating, the rational expectation and the worst case model. The graphs in the left column are drawn under our benchmark parametrization[8] and thus $b < a^{-1}$, the right hand column considers the same parametrization but with $b = 10 > a^{-1}$.

3.2   *The Stackelberg Solution*

Following Leitemo and Söderström (2004a) in this section we analyze a different timing of the game and assume that the central bank, by acting as a Stackelberg leader, designs policies taking into account the evil agent's optimal decision for

---

[8] As a baseline in the whole paper, we assume a calibration that is commonly used in the sticky information literature: $a = 0.16$ for the sensitivity of price to expected output gap; $(1 - \omega) = 0.25$ for the sticky information parameter; $b = 0.8$ for the weight assigned to the relative price variability in the central bank's objective function; $\rho = 0.9$ is the coefficient of the AR(1) process for the cost push shock. $\theta = 100$ is the preference for robustness.

misspecification[9]. From the previous section we know that the evil agent's reaction function is:

$$\gamma_j = -\frac{\frac{1-\Lambda^j}{\Lambda^j}}{a^2\theta - 1}\phi_j + \frac{1}{a^2\theta - 1}\rho_j$$

by substituting it into the price level equation constraint (15) and rearranging we get:

$$\varphi_j = \frac{1}{a}\frac{a^2\theta}{1 - a^2\theta}\left(-\frac{1-\Lambda^j}{\Lambda^j}\phi_j + \rho_j\right) \tag{28}$$

by minimizing the central bank loss function (7) with respect to ϕ_j subject to equation (28) we obtain the robust price coefficients under the Stackelberg timing in the worst case model, denoted by the superscript "rs":

$$\phi_j^{rs} = \frac{1}{\left(\frac{a^2\theta - 1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}}\rho_j \tag{29}$$

using the above equation in the constraint (28) we obtain the robust output coefficients:

$$\varphi_j^{rs} = -\frac{\left(\frac{a^2\theta - 1}{a^2\theta}\right)ab}{\left(\frac{a^2\theta - 1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}}\rho_j \tag{30}$$

Note firstly that when $\theta \to \infty$, both ϕ_j^{rs} ,φ_j^{rs} collapse to the optimal ones (see equations (8), (9)); secondly, as $\left(\frac{a^2\theta - 1}{a^2\theta}\right) \leq 1$, the robust price coefficients are always greater that the optimal ones, that is $\left|\phi_j^{rs}\right| > \left|\phi_j^*\right|$, but the difference between the robust and the optimal output coefficients is non-linear:

$$\left|\varphi_j^{rs}\right| \lessgtr \left|\varphi_j^*\right| \text{ if } \frac{1-\Lambda^j}{\Lambda^j} \gtrless b\frac{(a^2\theta - 1)}{\theta}$$

given that $\frac{1-\Lambda^j}{\Lambda^j}$ is a combination of probability that decreases over time, the robust output coefficients can be initially smaller than the optimal ones (as it happens with our benchmark parametrization) and the opposite is true after some periods. By comparing the Stackelberg robust coefficients with the ones obtained under the Nash timing (equations (20) and (21)) the following relations hold:

$$\begin{aligned}\left|\phi_j^{rs}\right| &> \left|\phi_j^{rn}\right| \\ \left|\varphi_j^{rs}\right| &< \left|\varphi_j^{rn}\right|\end{aligned}$$

Thus, irrespectively of the central bank's degree of conservatism, output gap stabilization under the Stackelberg timing is greater than under the Nash timing and the opposite is true for price level deviations. We can clarify the economic

---

[9]This alternative timing was originally suggested by Carl Walsh (see Leitemo and Söderström, (2004a), footnote 10).

mechanism that leads to this unusual result after showing the specification error and the robust targeting rule. Using the price robust coefficient in the evil agent's reaction function, the specification error under the Stackelberg timing is:

$$\gamma_j^{rs} = \frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b}{(a^2\theta-1)\left(\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right)}\rho_j$$

as it happens under the Nash timing, γ_j^{rs} is positively related with the preference for robustness, i.e., $\frac{\partial \gamma_j^{rs}}{\partial \theta} < 0$, but after some manipulations it is easy to verify that the first mover advantage of the central bank leads to a lower specification error:

$$\left|\gamma_j^{rs}\right| < \left|\gamma_j^r\right|$$

This result is obtained by employing the following targeting rule, i.e., the ratio between equations (30) and (29):

$$\varphi_j^{rs} = -\left(\frac{a^2\theta-1}{a^2\theta}\right)ab\phi_j^{rs} \tag{31}$$

The relations between the misspecification, and the output and the price coefficients are:

$$\frac{\gamma_j^{rs}}{\varphi_j^{rs}} = -\frac{1}{a\theta} \tag{32}$$

$$\frac{\gamma_j^{rs}}{\phi_j^{rs}} = \frac{a^2\theta-1}{a^2\theta}\frac{b}{\theta} \tag{33}$$

Differently from what occurs under the Nash timing, the robust trade-off between the target variables (31) is now affected by the preference for robustness. The equivalence result does not hold because the central bank anticipates the evil agent's behavior. The relation between the specification error and the output coefficient is the same as under the Nash timing (see equations (32) and (22)), while the ratio between the specification error and the price coefficient is now lower (see equations (33) and (23)) meaning that the central bank lets the price level deviate from the target, due to the specification shock errors, more than it does under the Nash timing. The same occurs for the cost push shock, as showed by the robust targeting rule (see equations (18) and (31)).

This less aggressive attitude in stabilizing the price level irrespectively of the degree of conservatism stems from the central bank awareness that a more stable output gap would dampen the persistence of the inflationary process and of the shock error. Due to the Phillips curve a more stable output gap can be achieved only by allowing the price level to absorb an higher part of the cost push shock. The central bank foresees that the specification error is not randomly chosen but it is "optimally" designed by the evil agent, whose objective is to maximize the targets variability. By internalizing the evil agent's

reaction function the policy maker optimally trades-off the marginal benefit, in terms of a lower specification error due to the output gap stabilization, and the marginal cost of an higher price variability. These results are illustrated in figure 2 that shows the impulse responses of the price level, the output gap and the specification error under both the Nash and the Stackelberg timing for a given preference for robustness ($\theta = 100$). In the first quarters after the shock, when deviations from targets account more in the loss function, under the Stackelberg timing there is a consistent gain in terms of output gap stabilization. This, in its turn, causes a lower specification error while the price level pattern is quite similar to the price level response under the Nash timing. After some periods the price level exhibits higher variability, but the output gap deviations are still smaller than under the Nash timing. Hence, the gain in terms of targets stabilization in the first periods after the shock more than offset the higher price variability in following periods.

How can the central bank reach this result in the first periods after the shock? The reason lies in the dynamic properties of the SIPC. Coibion (2006) has emphasized two features that play a key role in determining the inertial behavior of the inflation process in the sticky information model, that is, the frequency of information updating and the degree of real rigidities. Both serve to obtain small price adjustments in the first periods after the shock since few firms know the shock and their optimal price is largely unaffected. A high degree of real rigidities means in fact a low elasticity of the firm's optimal price to expected deviation in the output gap or, in other words, that "firms care relatively more about the overall price level...than about.... the aggregate demand"[10]. In such a situation a central bank that foresees the evil agent behavior will put more weight on the output gap stabilization in order to immediately dampen the specification error without provoking a sudden increases in price level variability.

In order to fucus on the consequences for the economy of a central bank policy action designed to be robust against model misspecification when the actual misspecification is zero, we now solve the approximating model under the Stackelberg timing.

The robust instrument rule in MA(∞) representation is:

$$m_t^{rs} = \sum_{j=1}^{\infty} \eta_j^{rs} \epsilon_{t-j}$$

with

$$\eta_j^{rs} = \phi_j^{rs} + \varphi_j^{rs} = \frac{1 - \left(\frac{a^2\theta-1}{a^2\theta}\right) ab}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \tag{34}$$

and thus $\eta_j^{rs} \gtrless 0$ if $b \lessgtr \frac{1}{a} \left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}$. By substituting the coefficients (29), (30)

---

[10] See Coibion (2006, p. 4)

and (34) in the price level equation without specification errors, we obtain:

$$\phi_j^{as} = \frac{a + \frac{1-\Lambda^j}{\Lambda^j} - \left(\frac{a^2\theta-1}{a^2\theta}\right)a^2b\left(1 - \frac{a^2\theta-1}{a^2\theta}\right)}{\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2b + \frac{1-\Lambda^j}{\Lambda^j}\right]}\rho_j$$

and

$$\varphi_j^{as} = -\frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)ab\left[\frac{1-\Lambda^j}{\Lambda^j} + \left(\frac{a^2\theta-1}{a^2\theta}\right)a\right]}{\left(\frac{1-\Lambda^j}{\Lambda^j} + a\right)\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2b + \frac{1-\Lambda^j}{\Lambda^j}\right]}$$

By comparing the coefficients of the rational expectation and of the approximating model, after a tedious algebra, we obtain the following conditions:

$$\phi_j^{as} - \phi_j^* > 0 \ \ \text{if} \ \ 0 < b < b_j^{as}$$

$$\phi_j^{as} - \phi_j^* < 0 \ \text{if} \ \ b > b_j^{as}$$

and

$$\left|\varphi_j^{as}\right| - \left|\varphi_j^*\right| > 0 \ \ \text{if} \ \ b > b_j^{as}$$

$$\left|\varphi_j^{as}\right| - \left|\varphi_j^*\right| < 0 \ \ \text{if} \ \ 0 < b < b_j^{as}$$

where

$$b_j^{as} = \frac{1}{a}\left[1 + \left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}\left(1 + \frac{1}{a}\frac{1-\Lambda^j}{\Lambda^j}\right)\right]$$

given that $1 + \left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}\left(1 + \frac{1}{a}\frac{1-\Lambda^j}{\Lambda^j}\right) > 1$, when the central bank acts as a Stackelberg leader, the range value of the degree of conservatism that allows the emergence of an higher (lower) price level (output gap) variability is greater with respect to the Nash one (i.e., $b < \frac{1}{a}$). The reason is that under the Stackelberg timing the policy maker implements a trade-off between the targets that stabilize the output gap more than under the Nash timing because this policy behavior minimizes the specification error. Hence, the consequences for the economy of such a policy behavior are given by the interaction of two forces: On the one hand the anti-attenuation result in the use of the nominal policy instrument induces the central bank to react to the cost shocks in the same direction as in the rational expectation model but with greater intensity (see the previous section); on the other hand, and differently from the Nash timing, the central bank now seeks to implement a target trade-off which depends on its preference for robustness and it always stabilizes the output gap more than under the Nash timing, we name this result the *trade-off effect*. To show the consequences of this interaction, we can divide the central bank degree of conservatism in four regions:

1. $0 < b \leqslant \frac{1}{a}$. When $b < \frac{1}{a}$ we know that in the rational expectation model the central bank reacts to a cost push shock by increasing the nominal money supply (equation (12)). Hence, under the Stackelberg timing in the approximating model both the anti-attenuation result and the desired targets

16

trade-off act in the direction of increasing the policy instrument response, with the consequence to reach a greater stabilization of the output gap, that is, $\phi_j^{as} > \phi_j^*$, $\left|\varphi_j^{as}\right| < \left|\varphi_j^*\right|$ and $\eta_j^{rs} > \eta_j^* > 0$ (figure 3a). When $b = \frac{1}{a}$ the anti-attenuation result does not operate because in the rational expectation model the central bank does not react to the cost shock, but the desired trade-off induces the central bank to react to the cost shock by increasing the nominal money supply and thus we obtain again: $\phi_j^{as} > \phi_j^*$, $\left|\varphi_j^{as}\right| < \left|\varphi_j^*\right|$ and $\eta_j^{rs} > \eta_j^*$.

2. $\frac{1}{a} < b \leqslant \frac{1}{a}\left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}$. In this region the anti-attenuation result and the trade-off effect act in opposite directions. The first effect decreases the nominal policy instrument, as in the rational expectation model we have $\eta_j^* < 0$, whereas the trade-off effect leads to the opposite reaction (equation (31)). The latter effect dominates over the former, leading to higher output stabilization, i.e., $\phi_j^{as} > \phi_j^*$, $\left|\varphi_j^{as}\right| < \left|\varphi_j^*\right|$ and $\eta_j^{rs} \gtrless 0 > \eta_j^*$. When $b = \frac{1}{a}\left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}$ we get $\eta_j^{rs} = 0$, meaning that the central bank does not react to the cost shock because the two forces (for what concerns the instrument reaction) are in exact balance. Anyway given that $\eta_j^* < 0$ we still have $0 = \eta_j^{rs} > \eta_j^*$ and thus $\phi_j^{as} > \phi_j^*$ and $\left|\varphi_j^{as}\right| < \left|\varphi_j^*\right|$ (figure 3a).

3. $\frac{1}{a}\left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1} < b < b_j^{as}$. The anti-attenuation result and the trade-off effect act in opposite directions. The latter continues to dominate over the former, that is, $\phi_j^{as} > \phi_j^*$ and $\left|\varphi_j^{as}\right| < \left|\varphi_j^*\right|$, but differently from the previous case, the anti-attenuation result is great enough to induce a contraction in the policy instrument, i.e., $\eta_j^{rs} < 0$. In any case, the nominal money contraction is lower than under the rational expectation model, that is, $\left|\eta_j^{rs}\right| < \left|\eta_j^*\right|$ (figure 3b).

4. $b > b_j^{as}$. Opposite to regions 2 and 3, now the anti-attenuation result dominates over the trade-off effect. Hence the policy instrument reaction ($\eta_j^{rs} < 0$) is greater than under the rational expectation model, leading to an higher price level stabilization, that is, $\phi_j^{as} < \phi_j^*$, $\left|\varphi_j^{as}\right| > \left|\varphi_j^*\right|$ and $\left|\eta_j^{rs}\right| > \left|\eta_j^*\right|$. When $b = b_j^{as}$ the two models produce the same results, i.e., $\phi_j^{as} = \phi_j^*$ and $\left|\varphi_j^{as}\right| = \left|\varphi_j^*\right|$ meaning that the consequences for the economy (and not for the policy instrument as it happens when $b = \frac{1}{a}\left(\frac{a^2\theta-1}{a^2\theta}\right)^{-1}$) of the interaction of the two forces are exactly countervailed, i.e., $\eta_j^{rs} = \eta_j^* < 0$. It is worth to notice that b_j^{as} is not fixed once and for all, but it depends from $\frac{1-\Lambda^j}{\Lambda^j}$ that is a combination of probability that decreases over time. It is hence possible that a given degree of conservatism falls in the the 3th region in the first quarters after the shock and in the 4th later on. In such a situation the trade-off effect dominates over the anti-attenuation result as long as $b < b_j^{as}$ and the opposite is true when $b > b_j^{as}$[11]. This is the case showed in figure 3b.

---

[11] Due to the stickyness in the diffusion of information the peak of the policy instrument response occurs after some periods. Hence, in the first quarters after the shock the small (absolute) value of the policy instrument in the rational expectation model generates a mild over-reaction (region 3) which increases over time (region 4).

By comparing the Stackelberg robust coefficients under the approximating model with the ones obtained under the Nash timing it is easy to check that $\phi_j^{as} > \phi_j^{an}$ and $\left|\varphi_j^{as}\right| < \left|\varphi_j^{an}\right|$; in both schemes we have the same anti-attenuation result, whose sign depends upon the desired trade-off in the rational expectation model, i.e., $ab \gtrless 1$. When the central bank has the first mover advantage it stabilizes the output gap more than under the Nash timing and this force acts always in the same direction, leading to $\eta_j^{rs} > \eta_j^{rn}$.

3.3   Welfare analysis

This section analyzes the rationale for delegating monetary policy to an independent central banker whose preferences does not coincide with the social ones. Since Rogoff's (1985) contribution it is widely acknowledged that in the presence of an inflation bias, produced by the policy maker's desire to push output above its natural level, the monetary authorities should be more conservative than society in order to maximize a micro-founded social welfare function. The sticky price New-Keynesian models reach the same conclusion, but the rationale is different. Even though in this kind of models usually there is no inflation bias, the forward looking nature of inflation and the presence of a time inconsistency problem induce a "stabilization bias" that justifies the appointment of a conservative central banker[12].

Welfare analysis in the robust control literature is in its infancy. The existing contributions on this topic[13] consider a social planner with the same preference for robustness as the policy maker. This assumption stems from the awareness that the max-min policy problem formulation implies the hypothesis that the policy maker and the private sector share the same reference model and the same degree of preference for robustness. Otherwise there would be a discrenpancy between the policy maker's and private sector's expectations about future inflation and output. As stressed by Walsh "An interesting area for future research would be to allow private agents and the policy maker to have different worst-case models"[14]. Even though this would enrich the analysis of the optimal institutional design in the presence of strong uncertainty, in this preliminary welfare analysis in a sticky information model we limit ourselves to provide an analytical relation linking the social and the central bank degree of conservatism with the commonly shared degree of preference for robustness.

In BMR (2005), there is neither an inflation bias, since the central bank does not have an overly ambitious output target in the objective function, nor a stabilization bias, since the lag in the policy transmission mechanism and the backward looking nature of the SIPC assure that the discretionary and commitment solution coincide. In this context there is no rationale for appointing a Rogoff-conservative central banker, i.e., the social preference has to be translated into the central banker's objective function. We now wish to determine

---

[12] See Clarida et al. (1999) and Walsh (2003).

[13] See Kilponen (2003),Tillmann (2006) and Gaspar and Vestin (2004).

[14] See Walsh (2004).

whether and how the model misspecification affects this result. A micro-founded loss function can be written in the following form:

$$L_s = \left[\sum_{j=1}^{\infty} \varphi_j^2 + b^s \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=1}^{i-1} \phi_j^2 \right] \sigma_\epsilon^2 \qquad (35)$$

where b^s is the weight society assigns to relative price variability.

Starting from the Nash timing we substitute equations (20) and (21) into (35), and obtain the equilibrium value:

$$L_s = \left[\sum_{j=1}^{\infty} \left(\frac{ab}{a^2b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \right)^2 + b^S \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=0}^{i-1} \left(\frac{1}{a^2b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \right)^2 \right] \sigma_\epsilon^2$$

The optimal trade-off between the targets in the central bank objective function is obtained by minimizing the above expression with respect to degree of conservativeness, i.e., the parameter b. After some algebra, we obtain the following relation linking the socially desired trade-off (b^S) with the optimal one (b^{rn}) (see appendix A for details):

$$b^{rn} = \left(1 - \frac{1}{a^2 \theta} \right) b^S \qquad (36)$$

According to equation (36), the minimization of social welfare requires the central bank to be more populist than society; the optimal degree of conservatism decreases when the preference for robustness increases. Of course under the rationale expectation model $(\theta \rightarrow \infty)$ we have $b^{rn} = b^S$.

How can we interpret this unconventional result? We know that the specification error is negatively related with the output gap; we also know that when the central bank can exploit the first mover advantage it chooses to stabilize the output gap more then it would do under the Nash timing, i.e., when the policy maker is not aware of the evil agent's behavior. We have interpreted this result through the central bank's desire to trade-off more inflation variability in favor of smaller output gap deviations, thus dampening the specification error.

The same rationale is at the basis of our delegation result. The social planner in charge of the design of the central bank objective function, who foresees both the behavior of the evil agent and the central banker's robust policies, in order to minimize the specification error optimally assigns a greater relative weight on output gap stabilization. In this sense, the social planner and the central banker under the Stackelberg timing act in an equivalent way but through different instruments (respectively b and the couple $(p_t, x_t))$[15]. Not surprisingly, under

---

[15] This result might inspire novel interpretations for a number of stylized facts such as the so-called Great Moderation (Bernake 2004), that is the joint reduction of output and price volatility that occured in the 1980's. This fact contrasts with the output-price volatility trade-off which is implicit in most monetary policy models. Some authors suggest that this volatility puzzle can be explained through a shift in the focus of monetary policy (Clarida et al., 2000). According to this interpretation, during the 1970's the central bankers' instrument

the Stackelberg timing there are no information advantages the social planner could exploit and that are not considered by the central bank. By substituting the coefficients (29) and (30) in the social loss function and maximizing with respect to b it is easy to check that the preference of the central bank must be the same as society's (see Appendix B for details).

4   The Parametric approach

In this section we use the *parametric* procedure developed by Giannoni (2002)[16] where the source of uncertainty is identified in one or more parameters of the economic model. As in the *unstructured* approach the policy maker has multiple priors about the distribution of the true model and he is uncertainty-adverse. Once again, the policy maker adopts a min-max strategy which can be represented by a "mind" game played by the central bank and the fictitious "evil agent", where the latter now controls the numerical values of the uncertain parameters over a given domain.

In what follows we determine the robust policy rule when the central bank is uncertain about the key parameter of the sticky information Phillips curve, that is a, which represents the sensitivity of firm's optimal price to the expected deviation of output, and which dramatically affects the dynamic properties of the model[17]. We assume that the policy maker is able to determine the maximum and the minimum value for a (i.e., $a \in (a_L, a_H)$ with $a_L < a_H$), but that he has multiple priors about the probability distribution.

The solution procedure for this min-max problem is based on 3 steps:

Step 1. Determine the optimal coefficients for known parameters, i.e., the policy problem tackled in section 2, and obtain (8) and (9). For clarify, it is useful to rewrite the optimal price and the optimal output gap coefficients:

$$\phi_j^* = \frac{1}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j$$

$$\varphi_j^* = -\frac{ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j$$

---

rules did not satisfy the Taylor principle, thus leading to sunspot equilibria which exhibit high targets volatility. By contrast, during the 1980's, more appropriate Taylor rules allowed to pin down the volatility of both inflation and output. Our finding suggests another interpretation which does not rely on a wrong conduct of monetary policy. In our context, the preference for robustness affects both the targets volatility and the optimal trade-off between targets in the central bank's loss function. A reduction in the preference for robustness decreases the volatility of both the output gap and the price level (see equations (20) and (21)), but at the same time increases the degree of the central banker's optimal conservatism. Hence, both effects lower the price volatility while acting in opposite direction on the output volatility. If the reduction in the preference for robustness dominates the increase in the degree of conservatism we obtain the Great Moderation result.

[16] See also Kara (2002) for an application to Clarida et al.(1999).

[17] See Keen (2004) and Coibion (2006).

The optimal trade-off is then:

$$\phi_j^* = -\frac{1}{ab}\varphi_j^* \tag{37}$$

Step 2. Determine the candidate worst parameter by maximizing the loss function with respect to a subject to the optimal coefficients.

From section 2, we know that the loss function can be written as:

$$L\left(\varphi,\phi,a\right) = \left[\sum_{j=1}^{\infty}\varphi_j^2 + b\sum_{i=1}^{\infty}\left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right)\sum_{j=0}^{i-1}\phi_j^2\right]\sigma_\epsilon^2 \tag{38}$$

The maximization problem is:

$$\underset{a}{Max} \quad L\left(\varphi,\phi,a\right)$$

$$s.t. \quad \phi_j^* = \frac{1}{a^2b + \frac{1-\Lambda^j}{\Lambda^j}}\rho_j$$

$$\varphi_j^* = -\frac{ab}{a^2b + \frac{1-\Lambda^j}{\Lambda^j}}\rho_j$$

By substituting the constraints into the objective function and by taking the derivative with respect to a, we obtain:

$$\frac{\partial L}{\partial a}\left(\varphi,\phi,a\right) = \frac{\partial \varphi_1^*}{\partial a}\sigma_\epsilon^2 + \frac{\partial \varphi_2^*}{\partial a}\sigma_\epsilon^2 + ... + b\sigma_\epsilon^2\left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1}\right)\frac{\partial \phi_0^*}{\partial a} + b\sigma_\epsilon^2\left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2}\right)\left[\frac{\partial \phi_0^*}{\partial a} + \frac{\partial \phi_1^*}{\partial a}\right] +$$

$$+b\sigma_\epsilon^2\left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3}\right)\left[\frac{\partial \phi_0^*}{\partial a} + \frac{\partial \phi_1^*}{\partial a} + \frac{\partial \phi_2^*}{\partial a}\right] + b\sigma_\epsilon^2\left(\frac{1}{\Lambda^3} - \frac{1}{\Lambda^4}\right)\left[\frac{\partial \phi_0^*}{\partial b} + \frac{\partial \phi_1^*}{\partial b} + \frac{\partial \phi_2^*}{\partial b} + \frac{\partial \phi_3^*}{\partial b}\right] + ..$$

After some manipulations we get the FOC:

$$\frac{\partial L\left(\varphi,\phi,a\right)}{\partial a} = \sum_{j=1}^{\infty}\frac{\partial \varphi_j^*}{\partial a}\sigma_\epsilon^2 + \sigma_\epsilon^2 b\sum_{j=0}^{\infty}\frac{1-\Lambda^j}{\Lambda^j}\frac{\partial \phi_j^*}{\partial a} = 0$$

By substituting the partial derivatives this becomes:

$$\frac{\partial L\left(\varphi^*,\phi^*,a\right)}{\partial a} = \sigma_\epsilon^2\frac{\frac{1-\Lambda}{\Lambda}ab^2 - a^3b^3}{\left(a^2b + \frac{1-\Lambda}{\Lambda}\right)^3}\rho_1^2 + \sigma_\epsilon^2\frac{\frac{1-\Lambda^2}{\Lambda^2}ab^2 - a^3b^3}{\left(a^2b + \frac{1-\Lambda^2}{\Lambda^2}\right)^3}\rho_2^2 + ...$$

$$-\sigma_\epsilon^2 b\frac{\frac{1-\Lambda}{\Lambda}2ab}{\left(a^2b + \frac{1-\Lambda}{\Lambda}\right)^3}\rho_1^2 - \sigma_\epsilon^2 b\frac{\frac{1-\Lambda^2}{\Lambda^2}2ab}{\left(a^2b + \frac{1-\Lambda^2}{\Lambda^2}\right)^3}\rho_2^2 - .. =$$

$$= \sigma_\epsilon^2\sum_{j=1}^{\infty}\frac{-a^3b^3 - \frac{1-\Lambda^j}{\Lambda^j}ab^2}{\left(a^2b + \frac{1-\Lambda^j}{\Lambda^j}\right)^3}\rho_j^2$$

In order to determine the worst parameter specification we need to study the sign of this derivative which is composed by an infinite sum of negatives terms (remember that $\frac{1-\Lambda^j}{\Lambda^j}$ is a combination of probabilities and all the other parameters are positives). Hence we get:

$$\frac{\partial L\left(\varphi^*, \phi^*, a\right)}{\partial a} < 0$$

The central banks' fear of misspecification is represented by $a = a_L$, and the robustly coefficients are:

$$\phi_j^r = \frac{1}{a_L^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \tag{39}$$

$$\varphi_j^r = -\frac{a_L b}{a_L^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \tag{40}$$

The robustly optimal trade-off is thus:

$$\phi_j^r = -\frac{1}{a_L b} \varphi_j^r \tag{41}$$

Step 3. Check that the vector $\left(\phi_j^r, \varphi_j^r, a_L\right)$ is also a *global* Nash equilibrium (see appendix C).

Results: Differently from the HS approach under the Nash timing, the robust trade off (41) is now steeper than that without specification errors (37). Hence, it is optimal for the policy maker to let the price level deviate from the target (when output deviates from the natural rate) more than in the case with known parameter[18]: to interpret this, note that a low a means a worsened price level-output trade-off, since the gain in terms of price stabilization of a unit output loss is now lower. This less aggressive attitude leads to more price variability ($\frac{\partial \phi_j^r}{\partial a} < 0$). As for output stabilization, the effect of a reduction in a is not univocal: When $\frac{1-\Lambda^j}{\Lambda^j} \gtrless a^2 b$ then $\frac{\partial \left|\varphi_j^r\right|}{\partial a} \gtrless 0$; as $\frac{1-\Lambda^j}{\Lambda^j}$ decreases over time, in the first quarters after the shock we may have $\left|\varphi_j^r\right| < \left|\varphi_j^*\right|$ but the inequality is the opposite after some periods.

This result is not surprising. As emphasized by Coibion (2006) and Keen (2004), a is the key parameter in determining the inertial behavior of the target variables. In particular, when a decreases, the largest impact of the shock occurs with a greater delay, due to the strategic complementarity in the firms' pricing decisions. Hence, in the first periods after the shock the robust target rule induces a softer adjustment in the output gap.

A simple intuition for this attenuation result is the following: a positive cost push shock leads the central bank to offset it trough an output contraction. When robustness concerns on the slope of the Phillips curve are introduced the central bank is aware that part of its effort will be frustrated by the evil agent's best response (i.e., $a = a_L$). This reduction in the policy action effectiveness

---

[18] See Kara (2002) for a similar result.

induces the policy maker to place more weight on the target variable where the evil agent's action is less harmful, that is, output stabilization.

In the absence of uncertainty, the robust optimal trade-off (41) can be replicated by assigning the weight $\frac{a_L}{a}b$, instead of b, to the variability of relative prices in the central bank's objective function (7). In this case, a less conservative central banker (in the sense of Rogoff, 1985) is behaviorally equivalent to a min-maximizer central banker (that is, the optimal trade-off (37) is equal to the robust optimal trade-off (41)).

5   Concluding Remarks

The analysis carried out in this paper confirms most of the results obtained in sticky price models, including the well known fragility with respect to different approaches to model misspecification, but also elucidates the economic mechanisms at play in a sticky information economy under strong uncertainty and provides some new results which can be summarized as follows.

1 Under the HS approach, if the central bank and the evil agent act simultaneously, in the worst case model targets variability increases with the preference for robustness, whereas the desired trade-off between targets is the same as in the rational expectation model: Walsh's equivalence result holds. In line with the standard interpretation, this result stems from the central bank's overestimation of the inflationary consequences of cost shocks due to its fear of misspecification. This, in its turn, causes an higher output contraction. When the central bank wants to be robust against specification errors but the actual misspecification is zero (the approximating model), the robust instrument rule overreacts to the cost shocks in the following way. When in the rational expectation model the policy maker responds to the cost shock by injecting money (the desired trade-off implies that an increases in the price level is counterbalanced by a less then proportional contraction in output), in the approximating model it overreacts by injecting more money, thus leading to a greater output stabilization. By contrast, when in the rational expectations model the price targeting regime requires to respond to the cost shock by decreasing money (the desired trade-off implies that an increases in the price level is counterbalanced by a more then proportional contraction in output), in the approximating model the robust instrument rule overreacts by decreasing even more the money supply, thus producing a greater price level stabilization.

2 If the central bank has the first mover advantage, the optimal trade-off between targets is affected by the preference for robustness because the central bank internalizes the evil agent's behavior. In order to minimize the specification error, the central banker, who can now exploit the trade-offs between the specification error and the target variables, seeks to stabilize the output gap at the cost of higher price variability (the trade-off effect).

The consequences for the economy in the approximating model solution depend on the interaction between the trade-off effect, which always acts in the same direction, and the anti-attenuation effect, which acts as in the Nash timing. Given this interaction, there exists a range of degrees of conservatism for which the price level stabilization is initially lower with respect to the rational expectation model, while the opposite is true after some periods. The explanation of the trade-off effect relies on the dynamic properties of the sticky information model, while the possibility for the anti-attenuation effect to produce either tighter or looser policies is due to the price targeting regime adopted by the central bank.

3 Our welfare analysis suggests that, under the Nash timing, the maximization of a micro-founded social welfare function requires to appoint a central banker who is less conservative than society. The optimal degree of conservatism decreases when the preference for robustness increases. This finding relies neither on the classical "inflation bias" nor on the "stabilization bias" motivations. Knowing the relation between the output gap and the specification error, a welfare maximizer social planner assigns a weight to the price level stabilization that is lower than that of society. Under the Stackelberg timing, there are no information advantages the social planner can exploit, since they are already fully considered by the central bank; hence the policy-maker's preference must be the same as that of society.

4 Under the parametric approach (as under the Stackelberg timing in the HS approach), the central banker is more cautious in stabilizing the price level, since the gain in terms of price stabilization of a unit output loss is lower than the rational expectations model. This approach leads to higher persistency in the behaviour of the target variables since the uncertain parameter (a), which captures the sensitivity of a firm's optimal price to the expected deviation of the output gap, plays a key role in determining the inertial behavior of the economy.

References

[1] Ball, L., N.G, Mankiw, and R. Reis, (2005), "Monetary Policy for Inattentive Economies," *Journal of Monetary Economics*, 52(4), 703-725.

[2] Bernake, B.S., (2004), "The Great Moderation," Remarks by Governor Ben S. Bernake at the meetings of the EEA, Washington.

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

[4] Calvo, G.A., (1983), "Staggered Prices in a Utility Maximizing Framework," *Journal of Monetary Economics*, 12, 383-398.

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

[6] Clarida, R., Gali, J. and M. Gertler, (2000), "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory," *Quarterly Journal of Economics,* 115, 147-180.

[7] Coibion, O. (2006), "Inflation inertia in Sticky Information Models," *Contributions to Macroeconomics,* 6(1). pp 1-29.

[8] Gaspar, V. and D. Vestin (2004): "Imperfect Knowledge, Learning and Conservatism",unpublished, European Central Bank.

[9] Giannoni, M.P. (2002), "Does Model Uncertainty Justify Caution? Robust Optimal Monetary Policy in a Forward Looking Model," *Macroeconomic Dynamics*, 6, 111-114.

[10] Giordani, P. and P. Soderlind, (2004), "Solution of Macromodels with Hansen-Sargent Robust Policies: Some Extensions," *Journal of Economic Dynamics and Control*, 28(12): 2367-2397.

[11] Hansen, L.P. and T.J. Sargent, (2004), "Misspecification in Recursive Macroeconomic Theory," University of Chicago.

[12] Leitmo, K. and U. Söderström, (2004a), "Robust Monetary Policy in the New-Keynesian Framework," forthcomong, *Macroeconomic Dynamics*.

[13] Leitmo, K. and U. Söderström, (2004b), "Robust monetary policy in a small open economy," Working Paper No. 30, CEPR, .

[14] Kara A.H. (2002), "Robust Targeting Rules for Monetary Policy," Central Bank of Turkey Working Paper No. 60.

[15] Keen, B.D., (2004), "Sticky Price and Sticky Information Price Setting Models: What is the Difference?," mimeo, Texas Tech University.

[16] Kilponen, J. (2003), "A Positive Theory of Robust Monetary Policy and Robust Control," Discussion Paper, No. 18, Bank of Finland.

[17] Mankiw, N.G. and R. Reis, (2002), "Sticky information Versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve, " *The Quarterly Journal of Economics,* vol. 117(4), 1295-1328.

[18] Onatski, A. and J. Stock (2002), "Robust monetary policy under model uncertainty in a small model of the U.S. economy," *Macroeconomic Dynamics*, 6: 85-110.

[19] Onatski, A. and N. Williams (2003), "Modelling Model Uncertainty," *Journal of the European Economic Association,* 1(5), 1087-1122.

[20] Rogoff, K. (1985), "The optimal degree of commitment to an intermediate monetary target," *Quarterly Journal of Economics*, 100: 1169-1189.

[21] Tillmann, P. (2006), "Does Model Uncertainty Justify Conservatism? Robustness and the Delegation of Monetary Policy," manuscript, University of Bonn.

[22] Walsh, C.E., (2003), "Monetary Theory and Policy," MIT Press, Cambridge.

[23] Walsh, C.E., (2004), "Robust Optimal Interest Rule and Robust Control: An Equivalence Result," *Journal of Money, Credit, and Banking*, 36: 1105-1114

[24] Woodford, M. (2003), "Interest&prices: Foundations of a theory of monetary policy," Princeton: Princeton University Press.

Appendix A

In this appendix we demonstrate the delegation result under the Nash timing obtained in section 3. By substituting the robust coefficients ((20) and (21)) into the society's loss function we obtain:

$$L_s = \sum_{j=1}^{\infty} \left(\frac{ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \right)^2 \sigma_\epsilon^2 +$$

$$+ b^s \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=0}^{i-1} \left(\frac{1}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \right)^2 \sigma_\epsilon^2$$

The optimal degree of central bank conservatism is obtained by minimizing the above expression with respect to the parameter b.

$$\frac{\partial L_s^r}{\partial b} = \frac{\partial \varphi_1^{r2}}{\partial b} \sigma_\epsilon^2 + \frac{\partial \varphi_2^{r2}}{\partial b} \sigma_\epsilon^2 + ... + \left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1} \right) \frac{\partial \phi_0^{r2}}{\partial b} + b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} \right) \left[\frac{\partial \phi_0^{r2}}{\partial b} + \frac{\partial \phi_1^{r2}}{\partial b} \right] +$$

$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} \right) \left[\frac{\partial \phi_0^{r2}}{\partial b} + \frac{\partial \phi_1^{r2}}{\partial b} + \frac{\partial \phi_2^{r2}}{\partial b} \right] +$$

$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^3} - \frac{1}{\Lambda^4} \right) \left[\frac{\partial \phi_0^{r2}}{\partial b} + \frac{\partial \phi_1^{r2}}{\partial b} + \frac{\partial \phi_2^{r2}}{\partial b} + \frac{\partial \phi_3^{r2}}{\partial b} \right] + ... = 0$$

After some manipulations we get:

$$\frac{\partial \varphi_1^{r2}}{\partial b} \sigma_\epsilon^2 + \frac{\partial \varphi_2^{r2}}{\partial b} \sigma_\epsilon^2 + ... + b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1} + \frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} + \frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + ... + \frac{1}{\Lambda^\infty} \right) +$$

$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} + \frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + ... + \frac{1}{\Lambda^\infty} \right) \frac{\partial \phi_1^{r2}}{\partial b} +$$

$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + \frac{1}{\Lambda^3} - \frac{1}{\Lambda^4} ... + \frac{1}{\Lambda^\infty} \right) \frac{\partial \phi_j^{r2}}{\partial b} + ... = 0$$

Since $\frac{1}{\Lambda^\infty} = 1$, in general notation we can write:

$$\sum_{j=1}^{\infty} \frac{\partial \varphi_j^{r2}}{\partial b} \sigma_\epsilon^2 + \sigma_\epsilon^2 b^s \sum_{j=0}^{\infty} \frac{1-\Lambda^j}{\Lambda^j} \frac{\partial \phi_j^{r2}}{\partial b} = 0$$

Substituting the partial derivatives:

$$\sigma_\epsilon^2 \sum_{j=1}^{\infty} \frac{ab}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \left(\frac{a \left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right) - \left(a^2 - \frac{1}{\theta} \right) ab}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right)^2} \right) \rho_j +$$

$$+ \sigma_\epsilon^2 b^s \sum_{j=1}^{\infty} \frac{1}{a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}} \rho_j \left(\frac{-\left(a^2 - \frac{1}{\theta} \right)}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right)^2} \right) \rho_j = 0$$

hence:

$$\sigma_\epsilon^2 \sum_{j=1}^{\infty} \left(\frac{a^2 b \left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right) - \left(a^2 - \frac{1}{\theta} \right) a^2 b^2}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right)^3} \right) \rho_j^2 +$$

$$- \sigma_\epsilon^2 b^s \sum_{j=1}^{\infty} \left(\frac{1}{\Lambda^j} - 1 \right) \left(\frac{a^2 - \frac{1}{\theta}}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta} \right)^3} \right) \rho_j^2 = 0$$

$$b^s = \frac{a^2 b \sum_{j=1}^{\infty} \frac{\frac{1-\Lambda^j}{\Lambda^j}}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)^3} \rho_j^2}{\left(a^2 - \frac{1}{\theta}\right) \sum_{j=1}^{\infty} \frac{\frac{1-\Lambda^j}{\Lambda^j}}{\left(a^2 b + \frac{1-\Lambda^j}{\Lambda^j} - \frac{b}{\theta}\right)^3} \rho_j^2}$$

$$\frac{b^s}{b} = \frac{a^2}{\left(a^2 - \frac{1}{\theta}\right)}$$

that is the delegation result in the main test.

Appendix B

In this appendix we demonstrate the delegation result under the Stackelberg timing. The equilibrium value of the loss function is:

$$L_s^{rs} = \tfrac{1}{2} \left[\sum_{j=1}^{\infty} \varphi_j^{rs2} + b^s \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i}\right) \sum_{j=0}^{i-1} \phi_j^{rs2} \right] \sigma_\epsilon^2$$

By minimizing with respect b:

$$\frac{\partial \varphi_1^{rs2}}{\partial b} \sigma_\epsilon^2 + \frac{\partial \varphi_2^{rs2}}{\partial b} \sigma_\epsilon^2 + ... + \left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1}\right) \frac{\partial \phi_0^{rs2}}{\partial b} + b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2}\right) \left[\frac{\partial \phi_0^{rs2}}{\partial b} + \frac{\partial \phi_1^{rs2}}{\partial b}\right] +$$
$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3}\right) \left[\frac{\partial \phi_0^{rs2}}{\partial b} + \frac{\partial \phi_1^{rs2}}{\partial b} + \frac{\partial \phi_2^{rs2}}{\partial b}\right] + ... = 0$$

After some manipulations we get:

$$\frac{\partial \varphi_1^{rs2}}{\partial b} \sigma_\epsilon^2 + \frac{\partial \varphi_2^{rs2}}{\partial b} \sigma_\epsilon^2 + ... + b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^0} - \frac{1}{\Lambda^1} + \frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} + \frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + ... + \frac{1}{\Lambda^\infty}\right) +$$
$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} + \frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + ... + \frac{1}{\Lambda^\infty}\right) \frac{\partial \phi_1^{rs2}}{\partial b} +$$
$$+ b^s \sigma_\epsilon^2 \left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} + \frac{1}{\Lambda^3} - \frac{1}{\Lambda^4} ... + \frac{1}{\Lambda^\infty}\right) \frac{\partial \phi_j^{rs2}}{\partial b} + ... = 0$$

Since $\frac{1}{\Lambda^\infty} = 1$, in general notation we can write:

$$\sum_{j=1}^{\infty} \frac{\partial \varphi_j^{rs2}}{\partial b} \sigma_\epsilon^2 + \sigma_\epsilon^2 b^s \sum_{j=0}^{\infty} \frac{1-\Lambda^j}{\Lambda^j} \frac{\partial \phi_j^{rs2}}{\partial b} = 0$$

and substituting the partial derivatives:

$$\sigma_\epsilon^2 \sum_{j=1}^{\infty} \frac{\left(\frac{a^2\theta-1}{a^2\theta}\right) ab}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \left[\frac{\left(\frac{a^2\theta-1}{a^2\theta}\right) a \left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right] - \left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 \left(\frac{a^2\theta-1}{a^2\theta}\right) ab}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right]^2} \right] \rho_j +$$

$$- \sigma_\epsilon^2 b^s \sum_{j=1}^{\infty} \frac{1-\Lambda^j}{\Lambda^j} \frac{1}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}} \rho_j \frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b + \frac{1-\Lambda^j}{\Lambda^j}\right]^2} \rho_j = 0$$

By expliciting the first two terms of the summations:

$$\frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}}\rho_1 \left[\frac{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}\right]-\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}\right]^2}\rho_1\right]\sigma_\epsilon^2+$$

$$+\frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}}\rho_2 \left[\frac{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}\right]-\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}\right]^2}\rho_2\right]\sigma_\epsilon^2+...$$

$$-\sigma_\epsilon^2 b^s \frac{1-\Lambda}{\Lambda}\frac{1}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}}\rho_1 \frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}\right]^2}\rho_1+$$

$$-\sigma_\epsilon^2 b^s \left(\frac{1-\Lambda^2}{\Lambda^2}\right)\frac{1}{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}}\rho_2 \frac{\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}\right]^2}\rho_2-...=0$$

$$0=\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b\sigma_\epsilon^2 \left(\frac{\frac{1-\Lambda}{\Lambda}}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}\right]^3}\rho_1^2+\frac{\frac{1-\Lambda^2}{\Lambda^2}}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}\right]^3}\rho_2^2+...\right)+$$

$$-\sigma_\epsilon^2 b^s \left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 \left(\frac{\frac{1-\Lambda}{\Lambda}}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda}{\Lambda}\right]^3}\rho_1^2+\frac{\frac{1-\Lambda^2}{\Lambda^2}}{\left[\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b+\frac{1-\Lambda^2}{\Lambda^2}\right]^3}\rho_2^2+...\right)=0$$

$$\left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 b\sigma_\epsilon^2 - \sigma_\epsilon^2 b^s \left(\frac{a^2\theta-1}{a^2\theta}\right)^2 a^2 = 0$$

$$b=b^s$$

that is the delegation result in the main text

Appendix C

In this appendix we demonstrate that the candidate min-max equilibrium $\left(\phi_j^r, \varphi_j^r, a_L\right)$ found in step 3 of section 4 constitutes also a global Nash equilibrium. Using equation (41) we can write the price equation under the robust targeting rule as:

$$\sum_{j=1}^\infty \phi_j^r \epsilon_{t-j}=\sum_{j=1}^\infty \Lambda^j \left(\phi_j^r \epsilon_{t-j}-aa_L b\phi_j^r \epsilon_{t-j}+\rho\epsilon_{t-j}\right)$$

Since this expression must hold for all possible realizations of ϵ_{t-j}, it follows that:

$$\phi_j^r=\frac{1}{1-\Lambda^j\left(1-aa_L\right)}\rho_j \qquad (42)$$

By substituting equations (41) and (42) into the objective function (38) we get:

$$L\left(\phi^r,a\right)=\sum_{j=1}^\infty \left(-\frac{a_L b\rho_j}{1-\Lambda^j(1-aa_L)}\right)^2\sigma_\epsilon^2+$$

$$+b\sum_{i=1}^\infty \left(\frac{1}{\Lambda^{i-1}}-\frac{1}{\Lambda^i}\right)\sum_{j=0}^{i-1}\left(\frac{\rho_j}{1-\Lambda^j(1-aa_L)}\right)^2\sigma_\epsilon^2$$

In order to check if the evil agent still chooses a_L, we need to study the sign of the derivative of the above equation with respect to a:

$$\frac{\partial L(.)}{\partial a} = \sum_{j=1}^{\infty} \left[\frac{-2a_L^3 b^2 \Lambda^j}{(1-\Lambda^j(1-aa_L))^3} \rho_j^2 \right] +$$
$$+ b \sum_{i=1}^{\infty} \left(\frac{1}{\Lambda^{i-1}} - \frac{1}{\Lambda^i} \right) \sum_{j=1}^{i-1} \left[\frac{-2\Lambda^j a_L}{(1-\Lambda^j(1-aa_L))^3} \rho_j^2 \right] \sigma_\epsilon^2$$

Expliciting the first two terms of the summations we get: (remember that $\Lambda^j = (1-\omega) \sum_{k=0}^{j} \omega^k$, $\frac{1-\Lambda^j}{\Lambda^j} = \frac{1-\omega^{j+1}}{\omega^{j+1}}$, $\Lambda^{\infty} = 1$):

$$\frac{\partial L(.)}{\partial a} = \left(\frac{-2a_L^3 b^2 \Lambda}{(1-\Lambda(1-aa_L))^3} \rho_1^2 + \frac{-2a_L^3 b^2 \Lambda^2}{(1-\Lambda^2(1-aa_L))^3} \rho_2^2 + ... \right) \sigma_\epsilon^2 +$$
$$+ b \left(\frac{1}{\Lambda^1} - \frac{1}{\Lambda^2} \right) \frac{-2\Lambda a_L}{(1-\Lambda(1-aa_L))^3} \rho_1^2 \sigma_\epsilon^2 +$$
$$+ b \left(\frac{1}{\Lambda^2} - \frac{1}{\Lambda^3} \right) \left(\frac{-2\Lambda a_L}{(1-\Lambda(1-aa_L))^3} \rho_1^2 + \frac{-2\Lambda^2 a_L}{(1-\Lambda^2(1-aa_L))^3} \rho_2^2 \right) \sigma_\epsilon^2 + ...$$

After some manipulations we obtain:

$$\frac{\partial L(.)}{\partial a} = 2 \frac{-a_L^3 b^2 \Lambda - b\left(\frac{1-\Lambda}{\Lambda}\right) \Lambda a_L}{(1-\Lambda(1-aa_L))^3} \rho^2 \sigma_\epsilon^2 + 2 \frac{-a_L^3 b^2 \Lambda^2 - b\left(\frac{1-\Lambda^2}{\Lambda^2}\right) \Lambda^2 a_L}{(1-\Lambda^2(1-aa_L))^3} \rho^2 \sigma_\epsilon^2 + ...+$$

Since the denominator is positive (remember that $\Lambda^j \leq 1$), the above expression is decreasing in a; hence, in order to maximize the loss function, the evil agent sets $a = a_L$.

Figure 1: Anti-attenuation result under the Nash timing
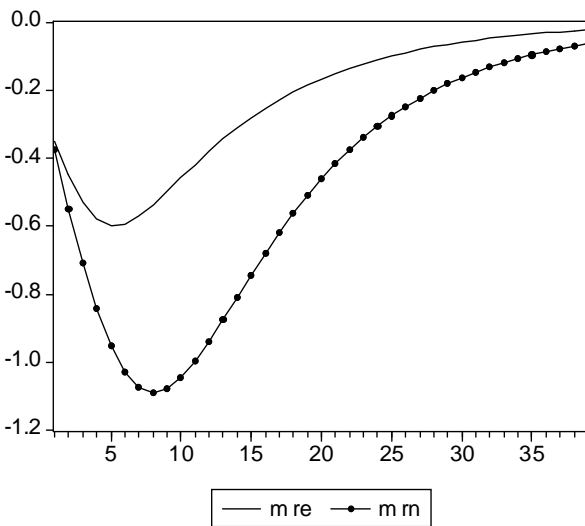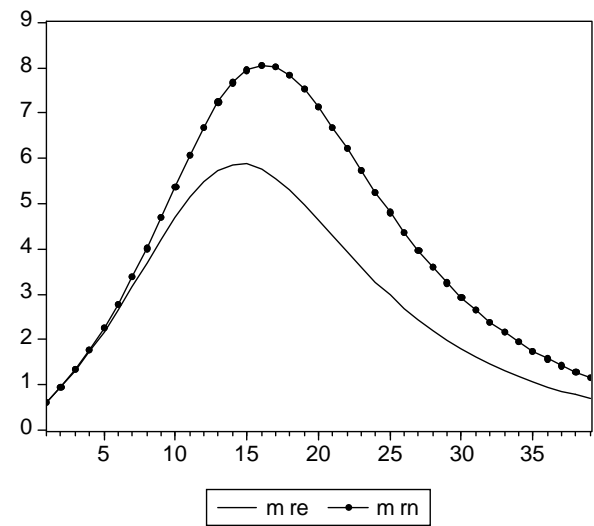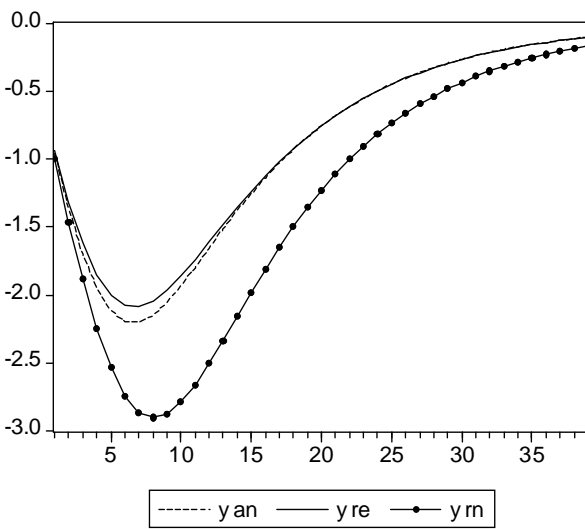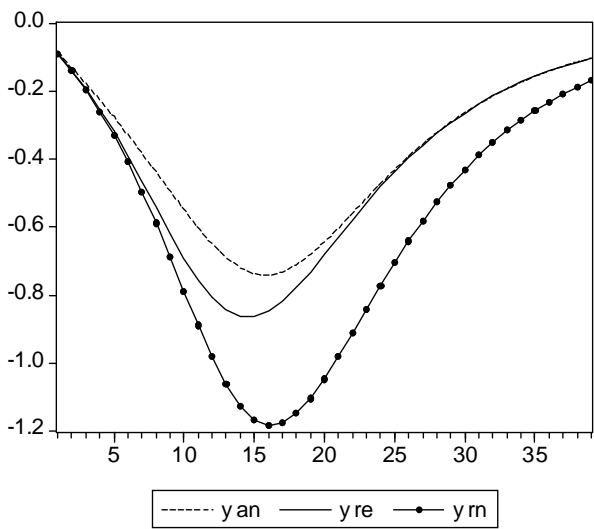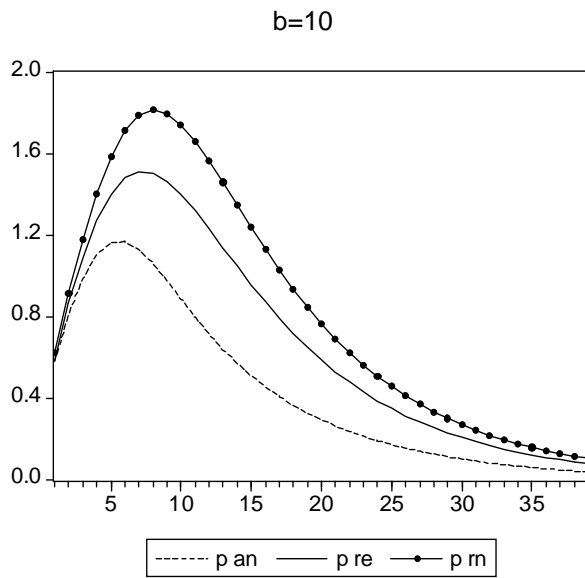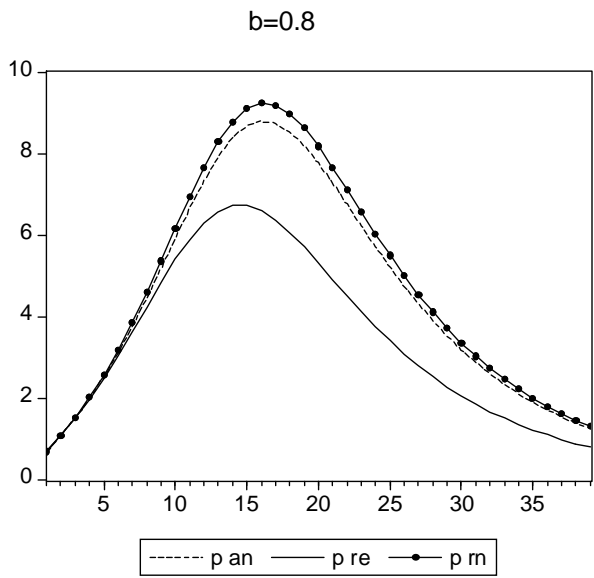
Figure 2: Impulse resonses in the worst case model
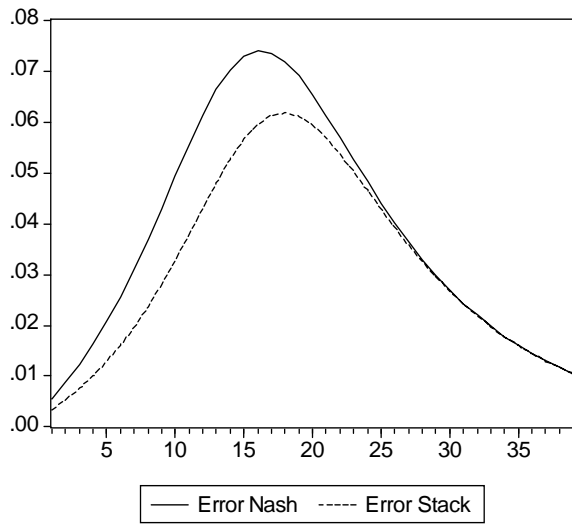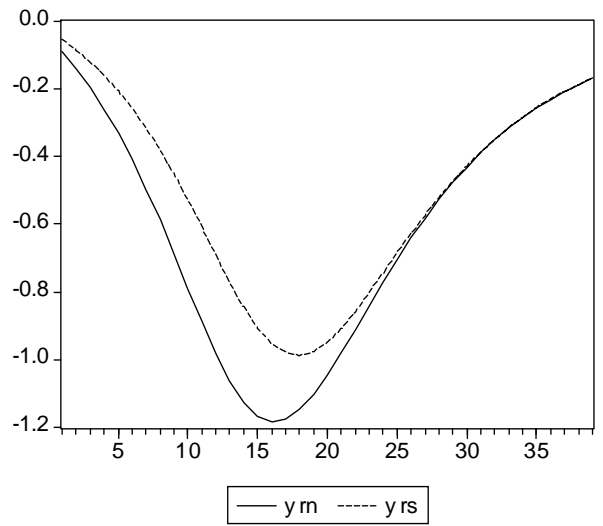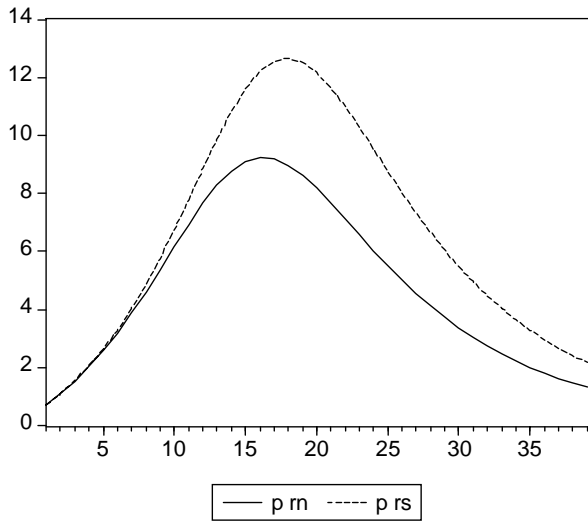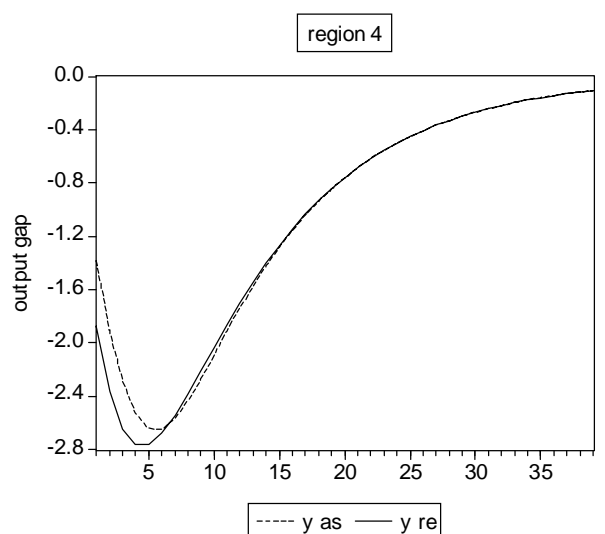
Figure 3a: Approximating model under Stackelberg timing

Figure 3b: Approximating model under Stackelberg timing