An aggregate welfare maximizing interest rate rule under heterogeneous expectations

Tim Hagenhoff

tim.hagenhoff@uni-bamberg.de

University of Bamberg, Feldkirchenstraße 21, 96052 Bamberg

March 10, 2019

Abstract

I implement fully optimal monetary policy under heterogeneous expectations as in Di Bartolomeo et al. (2016) by deriving an explicit interest rate rule under commitment. Implementation requires the derivation of agent's consumption decisions that incorporate the higher-order beliefs assumption of Branch and McGough (2009). As a result, "rational" agents are not sophisticated enough to have model-consistent individual consumption expectations, as assumed in Di Bartolomeo et al. (2016), even though they forecast aggregate variables correctly on average. Further, I show that the optimal interest rate rule yields substantial welfare gains compared to a rule that is derived from a conventional inflation-targeting objective as in Gasteiger (2014). The implementation of the non-optimal inflation-targeting rule already requires an increase of 14.5 percent of steady-state consumption to compensate for the higher welfare losses relative to the optimal interest rate rule when only ten percent of the population form (naive) backward-looking expectations. Welfare gains become substantially high when the underlying economy features a high degree of bounded rationality. Finally, I illustrate that consumption dispersion increases with the central bank's aggressiveness towards inflation.

Keywords: Optimal monetary policy, policy implementation, heterogeneous expectations

JEL classifications: E52, D84.

1 Introduction

The New Keynesian model emphasizes the ability of monetary policy to stabilize the macroeconomy by taking into account agent's expectations. However, optimal monetary policy is usually studied within a framework that assumes agents to form their expectations rationally (Clarida et al., 1999; Woodford, 1999; McCallum, 1999). Yet, econometric studies based on inflation expectation surveys show that the data favors heterogeneous expectations with a certain degree of bounded rationality (Branch, 2004, 2007; Pfajfar and Santoro, 2010). Starting from this observation, several New Keynesian models were designed that include heterogeneous expectations (Branch and McGough, 2009, 2010; De Grauwe, 2011; Massaro, 2013; Hommes et al., 2015).

A natural follow-up question is: How should central banks set interest rates optimally given its knowledge about the heterogeneity in expectations? To answer this question I derive an optimal interest rate rule under commitment based on the Branch and McGough (2009) framework and a model-consistent welfare criterion following Di Bartolomeo et al. (2016). To describe the micro level explicitly, I derive consumption Euler equations that adequately account for the assumption of Branch and McGough (2009) about higher-order beliefs. More specifically, the underlying model of this paper incorporates two types of agents. The more sophisticated agents, that I call "rational forecasters", are able to forecast *aggregate* variables consistent with the model predictions but are not smart enough to understand the micro level fully. In contrast, bounded rational forecasters use a simple backward-looking heuristic instead for forecasting. Such backward-looking heuristics are consistent with evidence from laboratory experiments (Assenza et al., 2014; Pfajfar and Žakelj, 2016).

While optimal monetary policy under homogeneous rational expectations is well known and extensively studied, the strand of literature dealing with monetary policy under heterogeneous expectations is rather new. Recent advances in the literature are made by Gasteiger (2014), Gasteiger (2018), Di Bartolomeo et al. (2016) and Beqiraj et al. (2017). Beqiraj et al. (2017) investigate fully optimal monetary policy under heterogeneous expectations based on the framework developed in Massaro (2013) that includes agents that forecast over all future periods up to infinity. On the other hand, Gasteiger (2014), Gasteiger (2018) and Di Bartolomeo et al. (2016) follow the Euler-equation-learning approach of Branch and McGough (2009). Gasteiger (2014) and Gasteiger (2018) explore interest rate rules derived from an ad-hoc inflation-targeting objective while Di Bartolomeo et al. (2016) provide an extension based on a model-consistent central bank objective.

Although Di Bartolomeo et al. (2016) implicitly assume expectations-based reaction functions, they

do not derive them. Thus, the literature has so far not provided a fully optimal interest rate rule under heterogeneous expectations based on the Branch and McGough (2009) model. Deriving an interest rate rule is important as well-grounded policy advice does require an actual rule for policymakers to apply. Further, such policy rules provide an intuitive way of identifying which variables are important in determining the interest rate and how their influence differs especially with the degree of heterogeneity. I will focus on the commitment case as it is typically superior to discretion.

Additionally, I explore the role of the higher-order beliefs assumption of Branch and McGough (2009) for agent's individual consumption decisions. Consumption decisions have to be made explicit as the central bank's objective function introduced by Di Bartolomeo et al. (2016) depends on consumption dispersion. This approach allows me to clarify the properties of the expectations operator, E_t^R, of "rational" forecasters in Branch and McGough (2009). It is assumed that rational forecasters, by using E_t^R, predict aggregate variables consistent with the model predictions which can, however, not be the case for expectations about the distribution of individual consumption. This is a straightforward consequence of the higher-order beliefs assumption which puts a particular (non-rational) structure on the agents believe about other agents' individual consumption expectations. In particular, all agents believe that all other agents form the same expectations about their individual consumption as they do. Hence, even "rational forecasters" are not smart enough to sophisticatedly forecast individual consumption. The final consumption equations only depend on aggregate variables and can, therefore, be used to substitute for individual consumption in the optimal interest rate rule.

However, if an Euler equation with model-consistent individual consumption expectations as in Di Bartolomeo et al. (2016) is applied, the first-order conditions of the central bank problem under commitment can only be reduced to a second-order difference equation in one of the Lagrange-multipliers to which the solution is fairly complicated and exponentially depends on time. Consequently, a meaningful interest rate rule under commitment in this case cannot be derived. Further, it would not be possible to substitute for individual consumption so that practical implementation would require individual consumption to be observable. However, applying the consumption equation that appropriately accounts for the higher-order beliefs assumption makes the derivation of a meaningful interest rate rule under commitment possible.

Moreover, I compare the optimal interest rate rule to a micro-founded version of the policy rule in Gasteiger (2014). As already indicated, the author derives an interest rate rule under commitment that recognizes the heterogeneity in expectations in the private sector equations but is optimized under a conventional ad-hoc inflation-targeting objective. The resulting interest rate rule is sub-optimal but also

much simpler than the rule derived in this paper. While it is straight-forward that the fully optimal rule must yield lower welfare losses than the non-optimal rule derived from the inflation-targeting objective, it is not clear *by how much*.

The welfare analysis shows that the optimal interest rate rule generates substantial welfare gains. The implementation of the non-optimal inflation-targeting rule already requires an increase of 14.5 percent of steady-state consumption to compensate for the higher welfare losses relative to the optimal interest rate rule when only ten percent of the population form (naive) backward-looking expectations. The welfare gains of the optimal interest rate rule crucially depend on the relative fraction of agent types. The optimal interest rate rule performs relatively better the higher the fraction of bounded rational forecasters. However, by numerically optimizing the weights on inflation and the shock with respect to the model-consistent welfare criterion in the less complicated inflation-targeting rule gives a relatively good approximation of the fully optimal policy already.

Finally, I find that consumption dispersion is not necessarily lower under the optimal rule compared the non-optimal inflation-targeting rule, even though the former explicitly incorporates consumption heterogeneity as opposed to the latter. This is because consumption dispersion increases with the central bank's aggressiveness towards inflation, as rational and bounded rational forecasters' consumption decisions become more unequal for larger increases in the policy rate. As a consequence, there is also a (local) trade-off between minimizing welfare losses and reducing consumption dispersion.

The remainder of the paper is organized as follows. The underlying model including the modified consumption rules are presented in Section 2. The optimal interest rate rule under heterogeneous expectations is derived in the subsequent section. Section 4 shows the impulse responses under optimal monetary policy with an emphasis on the micro-behavior followed by the welfare analysis in Section 5. Finally, the conclusion is given in Section 6.

2   Model

In this section, I introduce the underlying model and derive consumption decision that incorporate the higher-order beliefs assumption of Branch and McGough (2009). It is assumed that the economy is populated by an exogenous fraction α of rational forecasters (R) which have rational (model-consistent) expectations with respect to *aggregate* variables and a fraction $1 - \alpha$ of boundedly rational forecasters (B) that employ a simple backward-looking heuristic. The general forecasting rule of bounded rational forecasters takes the form of $E_t^B x_{t+1} = \theta^2 x_{t-1}$ while rational forecasters simply use the expected value,

i.e. $E_t^R x_{t+1} = E_t x_{t+1}$, for *forecasting the output gap and inflation*. Backward-looking expectations for $\theta < 1$ are called steady-state-reverting, for $\theta = 1$ naive and for $\theta > 1$ trend-setting. Steady-state-reverting expectations constitute a stabilizing force while trend-setting expectations imply a further amplification of macroeconomic variables.

Assuming perfect consumption insurance *within* each of the two agent groups the model can be expressed in terms of two representative agents (RA). Both RA's maximize their individual expected discounted lifetime utility $E_t^\tau \sum_{t=0}^\infty \beta^t U_t$ given their subjective expectations E_t^τ with $\tau \in \{R, B\}$. However, as in Branch and McGough (2009) agents follow Euler equation learning, i.e. they disregard their intertemporal budget constraint as an optimality condition and solely base their consumption decision on the variational intuition of the Euler equation. The period utility function is of CES-form and is given by

$$U_t = \frac{(C_t^\tau)^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}} - \frac{(Y_t^\tau)^{1+\eta}}{1+\eta} \tag{1}$$

with C_t^τ being consumption of type τ, Y_t^τ the output that each RA τ produces, $\frac{1}{\sigma}$ the coefficient of relative risk aversion and η the elasticity of marginal disutility of producing output. Agents must satisfy their real budget constraint

$$C_t^\tau + B_t^\tau = \frac{1 + i_{t-1}}{\Pi_t} B_{t-1}^\tau + \Psi_t^\tau \tag{2}$$

with B_t^τ being real bond holdings, i_{t-1} the nominal interest rate in $t-1$, Π_t gross inflation and Ψ_t^τ real income of agent τ.

I will now turn to the derivation of agent's consumption decisions. All agents in this economy are assumed to believe that all other agents will form the same expectations as they do. This is the higher-order beliefs assumption A6 in Branch and McGough (2009). The authors explicitly emphasize that this assumption is *necessary* for aggregation. This assumption, however, also implies that "rational" forecaster are not fully rational in the conventional sense of the rational expectations hypothesis and therefore cannot have rational individual consumption expectations, as I will explicitly show below. Assuming rational forecaster to possess rational individual consumption expectation *implies too much rationality* to be consistent with the *necessary* higher-order beliefs assumption, as is the case in Di Bartolomeo et al. (2016). Incorporating this assumptions into the consumption decisions of agents is crucial as a meaningful interest rate rule under commitment cannot be derived otherwise (see Appendix C).

Further, the central bank's welfare criterion can be re-written using market clearing in a way that it only depends on the consumption decision of rational forecasters. Hence, for now I focus on the

consumption Euler equation for $\tau = R$ which is given by

$$(C_t^R)^{-\frac{1}{\sigma}} = \beta E_t^R \left[(C_{t+1}^R)^{-\frac{1}{\sigma}} \frac{1+i_t}{\Pi_{t+1}} \right]. \tag{3}$$

Log-linearizing (3) gives

$$c_t^R = E_t^R c_{t+1}^R - \sigma(i_t - E_t^R \pi_{t+1}) \tag{4}$$

where lower case letters indicate log-deviations from steady state. Forward iteration yields

$$c_t^R = E_t^R c_\infty^R - \sigma \sum_{k=0}^{\infty} (i_{t+k} - E_t^R \pi_{t+k+1}). \tag{5}$$

It is assumed that rational forecasters know that market clearing $y_t = \alpha c_t^R + (1-\alpha)c_t^B$ holds and that bounded rational forecasters will also satisfy their consumption Euler equation. Writing market clearing one period forward and inserting equation (5), and equivalently the forward-iterated consumption Euler equation for bounded rational forecasters, gives

$$E_t^R y_{t+1} = E_t^R \Big[\alpha (E_t^R c_\infty^R - \sigma E_{t+1}^R \sum_{k=1}^{\infty} (i_{t+k} - \pi_{t+k+1})) $$
$$+ (1-\alpha)(E_t^B c_\infty^B - \sigma E_{t+1}^B \sum_{k=1}^{\infty} (i_{t+k} - \pi_{t+k+1})) \Big]. \tag{6}$$

Note that (6) contains higher-order beliefs, i.e. beliefs of rational forecasters E_t^R about the beliefs of bounded rational forecasters E_t^B and E_{t+1}^B. In order to arrive at the IS curve that has the same functional form as in the model under homogeneous rational expectations, Branch and McGough (2009) need to impose a specific (non-rational) structure on higher-order beliefs on consumption. This assumption states that agent's believe that all other agents will forecast their individual consumption in the same way they do. Mathematically and in the context of rational forecasters: $E_t^R E_{t+k}^B c_{t+l} = E_t^R c_{t+l}$ with $l > k$. Hence, bounded rational expectations just drop out under this assumption. Further, note that making an alternative assumption, e.g. allowing rational forecasters to be fully rational, would result in a different system of aggregate equations (see Hagenhoff and Lustenhouwer, 2019).

Using the higher-order beliefs assumption and the law of iterated expectations at the individual level yields

$$E_t^R y_{t+1} = E_t^R y_\infty - \sigma \sum_{k=1}^{\infty} (i_{t+k} - E_t^R \pi_{t+k+1}). \tag{7}$$

It becomes obvious that (7) cannot hold under conventional rational expectations, i.e. when $E_t^R = E_t$

6

would hold, as bounded rational expectations, E_t^B and E_{t+1}^B, would *not drop out* and thus show up in (7). Mathematically written: $E_t E_t^B = E_t^B$ and $E_t E_{t+1}^B = E_{t+1}^B$, which would contradict the higher-order beliefs assumption of Branch and McGough (2009). Equation (7) would only hold under conventional rational expectations when boundedly rational forecasters were absent, i.e. under *homogeneous rational expectations*. In this case (6) would collapse to (7) without any further assumption.

It follows that rational forecasters in this model are not smart enough to have model-consistent expectations with respect to the distribution of individual consumption *when there is heterogeneity*. Thus, assuming model-consistent individual consumption expectations in the Euler equation of rational forecaster as in Di Bartolomeo et al. (2016) is inconsistent with the underlying framework.

Using (7) to replace the infinite sum in (5), one obtains

$$c_t^R = E_t^R y_{t+1} + E_t^R(c_\infty^R - y_\infty) - \sigma(i_t - E_t^R \pi_{t+1}) \tag{8}$$

which is the true consumption decision of rational forecasters satisfying the higher-order beliefs assumption from above.

Equation (8) could have been derived for the general case of agent τ as the assumption on higher-order beliefs holds for both agent types. In the general case (8) reads

$$c_t^\tau = E_t^\tau y_{t+1} + E_t^\tau(c_\infty^\tau - y_\infty) - \sigma(i_t - E_t^\tau \pi_{t+1}). \tag{9}$$

Still, it has to be defined what $E_t^\tau(c_\infty^\tau - y_\infty)$ is. In Branch and McGough (2009) these terms drop out when aggregating (9) and when the assumption that agents agree on expected differences in expected limiting consumption is used (A7 in Branch and McGough (2009)). An assumption consistent with A7 in Branch and McGough (2009) is to assume that agents believe to be back in steady state in the long-run. In this case $E_t^\tau(c_\infty^\tau - y_\infty) = 0$ holds and thus (9) becomes

$$c_t^\tau = E_t^\tau y_{t+1} - \sigma(i_t - E_t^\tau \pi_{t+1}). \tag{10}$$

From (10) one can infer that agents only forecast aggregate variables when making consumption decisions. Note that, as rational forecasters have rational expectations with respect to aggregate variables, E_t^R can be replaced by E_t in the consumption decision of rational forecasters.

Using goods market clearing and (10) the IS curve is given by

$$y_t = \alpha E_t y_{t+1} + (1-\alpha)\theta^2 y_{t-1} - \sigma(i_t - \alpha E_t \pi_{t+1} - (1-\alpha)\theta^2 \pi_{t-1}). \tag{11}$$

Further, all agents produce output under monopolistic competition. Calvo pricing is assumed where a fixed fraction ξ_p of agents cannot reset their prices in a given period (Calvo, 1983). Price dispersion arises because, first, optimal prices are different *between* expectation types since they depend on *expected* future marginal costs and, second, they differ *within* each type due to the fact that only a fraction of firms can reset prices. Thus, the Phillips-curve can be derived as

$$\pi_t = \alpha\beta E_t \pi_{t+1} + (1-\alpha)\beta\theta^2 \pi_{t-1} + \kappa y_t + e_t \tag{12}$$

with $\kappa = \frac{(1-\xi_p)(1-\beta\xi_p)(\eta+\sigma^{-1})}{\xi_p(1+\epsilon\eta)}$ where ϵ is the price elasticity of demand for a differentiated good. As in Di Bartolomeo et al. (2016) the Phillips curve is augmented with a random cost-push shock e_t.[1]

Note that inflation and output exhibit some degree of *persistence* due to the presence of backward-looking expectations. The degree of persistence depends on the fraction of bounded rational forecasters $1-\alpha$ and their forecasting coefficient θ. The higher the two parameters the higher the degree of persistence. Further, as rational forecasters use the aggregate equations (11) and (12) to forecast output and inflation, they are aware of this persistence. Hence, even if a *transitory* cost-push shock hits the economy, rational forecasters will expect inflation to be above the steady state in the next period. This non-zero inflation expectation then feeds back into current inflation via (12) and thus causes an *amplification* of inflation (and via the central bank in output). This amplification mechanism is strongest for intermediate values of α as already investigated by Gasteiger (2018). The reason is that for large values of α only a minority of agents is backward-looking and thus persistence becomes less pronounced while for low values of α there are not enough rational forecasters through which the amplification works. Hence, this model associates an amplification of macroeconomic variables with the presence of bounded rational agents.

The model is calibrated as in Di Bartolomeo et al. (2016) for the US economy following Rotemberg and Woodford (1997) with the time unit being one quarter.

$\alpha = 0.7$	$\theta = 1$	$\beta = 0.99$	$\sigma = 6.25$	$\epsilon = 7.84$	$\eta = 0.47$	$\xi_p = 0.66$

Table 1: Baseline calibration

---

[1]This shock can, for instance, be micro-founded by assuming an exogenous time-varying wage mark-up as in Galí (2015).

3 An optimal interest rate rule

In this section I derive, first, an optimal interest rate rule from a model-consistent welfare criterion and, second, a non-optimal rule under a conventional inflation-targeting objective as in Gasteiger (2014).

3.1 Loss functions

As in Gasteiger (2014), Gasteiger (2018) and Di Bartolomeo et al. (2016) a paternalistic central bank is assumed, i.e. the central bank's aim is to maximize social welfare. I follow the approach of Di Bartolomeo et al. (2016) where the central bank exploits its detailed knowledge about the heterogeneity in expectations and minimizes a social welfare loss that is a second-order approximation of household utility (1). The intertemporal second-order approximated aggregate welfare loss can be derived as

$$\mathcal{W} = -\frac{\bar{C}U_C}{2} \sum_{t=0}^{\infty} \beta^t L_t + t.i.p. \tag{13}$$

with

$$L_t = \left(\eta + \frac{1}{\sigma}\right) y_t^2 + (\epsilon^2 \eta) var_i(p_t(i)) + \frac{1}{\sigma} var_i(c_t(i)). \tag{14}$$

and

$$var_i(p_t(i)) = \delta \pi_t^2 + \frac{\delta \xi_p (1-\alpha)}{\alpha} \left[\pi_t - \beta \theta^2 \pi_{t-1} - \kappa \frac{c_t^B + \eta \sigma y_t}{1 + \eta \sigma} \right]^2 \tag{15}$$

$$var_i(c_t(i)) = \alpha(1-\alpha)(c_t^R - c_t^B)^2. \tag{16}$$

where $\delta = \frac{\xi_p}{(1-\beta\xi_p)(1-\xi_p)}$ is a measure of price stickiness.

Using market clearing to eliminate c_t^B, (15) and (16), (14) can be rewritten as

$$L_t = \Gamma_1 y_t^2 + \Gamma_2 \pi_t^2 + \Gamma_3 \pi_{t-1}^2 + \Gamma_4 (c_t^R)^2 \tag{17}$$

$$+ \Gamma_5 y_t c_t^R + \Gamma_6 \pi_t c_t^R + \Gamma_7 \pi_{t-1} c_t^R + \Gamma_8 \pi_t \pi_{t-1} + \Gamma_9 \pi_t y_t + \Gamma_{10} \pi_{t-1} y_t$$

where the Γ_x-coefficients are given in the Appendix B.1.

Under homogeneous rational expectations, i.e for $\alpha = 1$, price dispersion reduces to $var_i(p_t(i)) = \delta \pi_t^2$ and $var_i(c_t(i))$ to zero. Hence, in this case (14) reduces to

$$L_t^{\alpha=1} = \left(\eta + \frac{1}{\sigma}\right) y_t^2 + \epsilon^2 \eta \delta \pi_t^2. \tag{18}$$

Equation (14) is called the model-consistent loss function and (18) the conventional inflation-targeting loss function.

As agents want to smooth their consumption over time that is due ot the concave nature of their utility function, they dislike volatility in general. However, the weight that is placed on inflation in second-order approximated utility functions in the canonical New Keynesian model such as (18) is usually considerably higher compared to the weight on output (see Woodford, 2003 or Galí, 2015). This reflects that price dispersion, due to inflation and sticky prices, is *the* source of inefficiency in the baseline model, quickly resulting in relatively large welfare losses. In case of the conventional inflation-targeting loss (18), the weight on inflation is indeed very high and roughly 160 times the weight on output under baseline calibration.[2]

However, when producers have heterogeneous expectations, price dispersion arises not only due to sticky prices but also because they have different expectations regarding future inflation and marginal costs, as reflected by (15). The weight on contemporaneous inflation in the model-consistent loss function (17) is roughly 230 times of the weight on contemporaneous output under baseline calibration with a 70 percent of rational forecasters, and increases to 270 times of the weight on contemporaneous output for 50 percent of rational forecasters. Additionally, the weights on lagged inflation and on the interaction between contemporaneous and lagged inflation are non-negligible. Hence, inflation results in even higher welfare losses through the price dispersion channel under heterogeneous expectations. However, even though inflation explains most of the results in the the welfare analysis in Section 5, there is still a trade-off between inflation and output (and consumption dispersion) under a cost-push shock that can be important in some cases.

Further, an interesting novelty of (14) is the appearance of consumption dispersion, i.e. the cross-sectional variance in consumption $var_i(c_t(i))$. It should be noted, however, that the weight on consumption dispersion in (14) is even smaller compared to the weight on output. This indicates that agents might accept a certain degree of heterogeneity when the economy is relatively stable. This finding is also in line with Debortoli and Galí (2017) who derive a model-consistent loss function in a Two Agent New Keynesian (TANK) model and show that it depends on a measure of heterogeneity but where the corresponding weight relative to inflation and output is also very low.

---

[2]The fact that the weight on inflation is high relative to the weight on output is general and robust with respect to the calibration.

3.2 An optimal interest rate rule under commitment

The central bank is assumed to set its interest rate so as to minimize the model-consistent loss function (17) or, respectively, the conventional inflation-targeting objective (18) subject to the private sector equations

$$y_t = \alpha E_t y_{t+1} + (1-\alpha)\theta^2 y_{t-1} - \sigma[i_t - \alpha E_t \pi_{t+1} - (1-\alpha)\theta^2 \pi_{t-1}] \tag{19}$$

$$\pi_t = \alpha\beta E_t \pi_{t+1} + (1-\alpha)\beta\theta^2 \pi_{t-1} + \kappa y_t + e_t \tag{20}$$

$$c_t^R = E_t y_{t+1} - \sigma(i_t - E_t \pi_{t+1}). \tag{21}$$

Minimizing the conventional inflation-targeting objective (18) subject to the Phillips curve (20) under timeless commitment[3] gives the *sub-optimal* inflation-targeting interest rate rule

$$i_t = \gamma_1 y_{t-1} + \gamma_2 E_t y_{t+1} + \gamma_3 \pi_{t-1} + \gamma_4 E_t \pi_{t+1} + \gamma_5 e_t \tag{22}$$

with

$$\gamma_1 = \frac{(1-\alpha)\theta - \alpha}{\sigma} + \alpha\frac{\delta\epsilon^2\eta\kappa^2}{1 + \eta\sigma + \delta\epsilon^2\eta\kappa^2\sigma} \tag{23}$$

$$\gamma_2 = \alpha\frac{1}{\sigma} - \frac{(1-\alpha)}{\sigma}\left[\frac{\beta^2\theta^2(1+\eta\sigma)}{1 + \eta\sigma + \delta\epsilon^2\eta\kappa^2\sigma}\right] \tag{24}$$

$$\gamma_3 = (1-\alpha)\left[\frac{\theta^2(1+\eta(\sigma+\delta\epsilon^2\kappa(\beta+\kappa\sigma)))}{1 + \eta\sigma + \delta\epsilon^2\eta\kappa^2\sigma}\right] \tag{25}$$

$$\gamma_4 = \alpha\left[1 + \frac{\beta\delta\epsilon^2\eta\kappa}{1 + \eta\sigma + \delta\epsilon^2\eta\kappa^2\sigma}\right] \tag{26}$$

$$\gamma_5 = \frac{\delta\epsilon^2\eta\kappa}{1 + \eta\sigma + \delta\epsilon^2\eta\kappa^2\sigma}. \tag{27}$$

Equation (22) is similar to the rule derived by Gasteiger (2018) and Gasteiger (2014).[4] Note that timeless commitment introduces persistence even in the absence of bounded rational forecaster, i.e. γ_1 reduces to $\frac{\delta\epsilon^2\eta\kappa^2}{1+\eta\sigma+\delta\epsilon^2\eta\kappa^2\sigma} - \frac{1}{\sigma}$ and does not vanish for for $\alpha = 1$.

The *optimal* commitment interest rate rule can be obtained by minimizing (17) subject to the private

---

[3]The timeless commitment approach of Woodford (2003) assumes that the optimal commitment policy was implemented in the distant past so as to omit the first period's optimality condition. The problem of the latter is that it renders the policy time-inconsistent. Hence, the drop of the initial period's optimality condition solves this problem.

[4]The author uses a non-micro-founded version of (18), i.e. $L_t = \frac{1}{2}(\pi_t^2 + \omega y_t^2)$. Setting $\omega = \frac{\sigma\eta+1}{\epsilon^2\eta\delta\sigma}$ and calculating through the optimization problem yields the interest rate rule (22).

sector equation (19), (20) and (21) under timeless commitment as

$$i_t = \Omega_1 E_t \pi_{t+1} + \Omega_2 E_t \pi_{t+2} + \Omega_3 \pi_{t-3} + \Omega_4 \pi_{t-2} + \Omega_5 \pi_{t-1} + \Omega_6 E_t y_{t+1} + \Omega_7 E_t y_{t+2} \qquad (28)$$

$$+ \Omega_8 y_{t-2} + \Omega_9 y_{t-1} + \Omega_{10} E_t c^R_{t+1} + \Omega_{11} E_t c^R_{t+2} + \Omega_{12} c^R_{t-2} + \Omega_{13} c^R_{t-1} + \Omega_{14} e_t$$

where the reaction coefficients Ω_x and derivations are given in the Appendix B.2. A first inspection of (28) shows that the central bank reacts to output and inflation as usual but also to individual consumption of rational forecasters due to the consumption inequality dimension. However, a more striking observation is that the central bank finds it optimal to react to lags and leads of all variables ranging from $t-2$ to $t+2$ (and additionally $t-3$ for inflation). This will be clarified further below. Note that under homogeneous rational expectations, $\alpha = 1$, all coefficients associated with heterogeneous expectations vanish as well as the additional coefficients due to commitment, except for y_{t-1} which can be seen in table (5) in the Appendix B.4.

As the central bank is assumed to be rational in the conventional sense, it fully understands the functioning of the economy, including the feedback of its policy on private sector rational expectations. To gain more intuition for the fully optimal interest rate rule (28), I, for now, shut down this channel, i.e. the central bank observes and reacts to heterogeneous expectations but does not incorporate the feedback on private sector rational expectations.[5] In this case, the interest rate rule is given by

$$i_t = \Omega^*_1 y_{t-1} + \Omega^*_2 E_t y_{t+1} + \Omega^*_3 E_t y_{t+2} + \Omega^*_4 \pi_{t-1} + \Omega^*_5 E_t \pi_{t+1} + \Omega^*_6 E_t \pi_{t+2}$$

$$+ \Omega^*_7 E_t c^R_{t+1} + \Omega^*_8 E_t c^R_{t+2} + \Omega^*_9 e_t \qquad (29)$$

where the reaction coefficients are given in the Appendix B.2.

By comparing (29) to (28) it becomes clear that all variables with timing $t-2$ and $t-3$ in (28) are because the central bank includes the feedback between its policy and rational expectations when calculating its optimal policy.

Moreover, to understand the appearance of the $t+2$ terms consider figure (1) which displays the reaction of the inflation expectations of both agent types following a one standard deviation i.i.d. cost-push shock under the policy rule (28). Since all subsequent shock realizations are zero and rational forecasters know the true *aggregate* equations, they have *de facto* perfect foresight. Thus, rational forecasters' expectations in $t = 1$ about inflation in $t+1$ will be correct, i.e. $E_t \pi_{t+1} = \pi_{t+1}$. However, the

---

[5]Such a behavior would, for instance, be consistent with a bounded rational central bank that operates under some sort of "limited" commitment. However, I use this only for the sake of exposition.
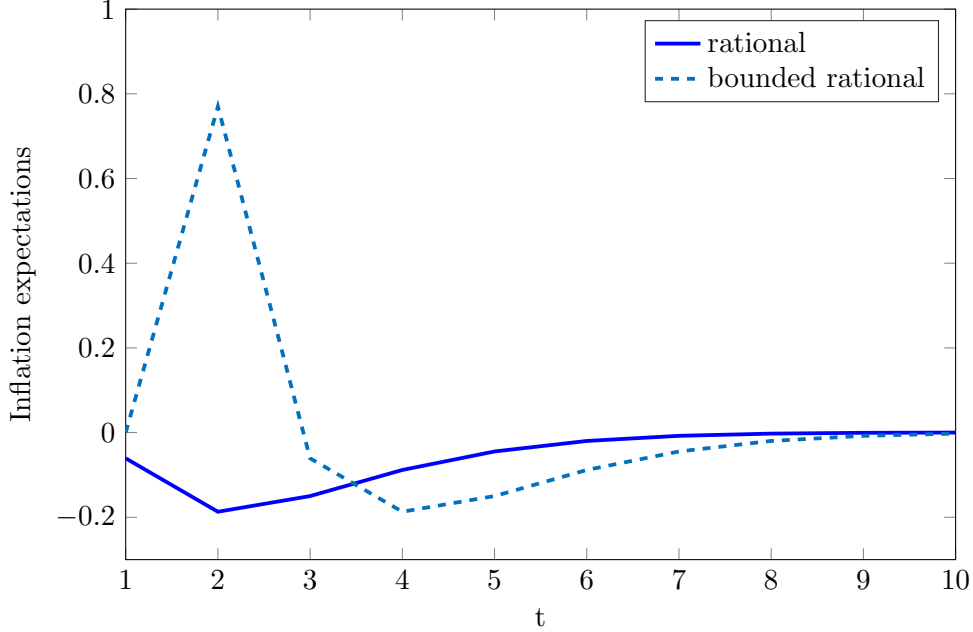
Figure 1: Individual inflation expectations in percentage deviation from steady state following a single, non-autocorrelated cost-push shock of one percent with monetary policy given by (28).

backward-looking expectations of bounded rational forecasters in t about inflation in $t+1$ will be zero, i.e. $E_t^B \pi_{t+1} = \pi_{t-1} = 0$ (where $\theta = 1$ for simplicity). In period $t+1$ bounded rational forecasters will expect inflation to *increase* drastically in $t+2$ as their expectations are based on the period where the shock hits, i.e. $E_{t+1}^B \pi_{t+2} = \pi_t$. On the contrary, rational forecasters correctly expect inflation to *decrease* further as the central bank increases the nominal interest rate a second time (see impulse responses in Figure 2 in section 4). Thus, the different expectations in $t+1$ about $t+2$ diverge transitorily. Therefore, the central bank should set interest rates so as to align the two expectation types in order to minimize the adverse effects of different expectations on price and consumption dispersion.

Moreover, it seems, at first glance, that for practical implementation the optimal interest rate rule (28) requires to observe individual consumption of rational forecaster which is, however, not observable in reality. As already indicated, an advantage of the consumption decision (10) is that it only depends on aggregate variables as a result of the explicit incorporation of the higher-order beliefs assumption. Therefore, it is possible to substitute for individual consumption so that the optimal interest rate rule is merely a function of several leads and lags of aggregate variables.

4   Impulse Responses

This section briefly describes the simulation outcomes under baseline calibration. Determinacy issues are not discussed as the model is determinate for all considered parameter constellations. There are two

reasons for this. First, an expectations-based interest rate rule is derived, i.e. it properly accounts for private sector expectations which are known to perform exceptionally well as opposed to fundamentals-based reaction functions (Evans and Honkapohja, 2006). Second, the interest rate rule is derived from the fully model-consistent loss function. Hence, the proposed interest rate rule is a good proxy for the fully optimal (non-linear) policy.

The impulse responses of a one percent i.i.d cost-push shock with monetary policy given by (28) are depicted in figure 2.

The aggregate behavior of the model is straight-forward: taking into account subjective expectations, the real interest rates of both agent types, $r_t^\tau = i_t - E_t^\tau \pi_{t+1}$, increase due to an increase of the nominal rate by the central bank. Hence, both agent types cut their individual consumption which leads to a quite severe recession which counteracts the cost-push shock to some extent. Consequently, inflation increases by less than one percent. Thus, the central bank finds it optimal to be extraordinarily hawkish with respect to inflation which comes at the cost of a significant recession.
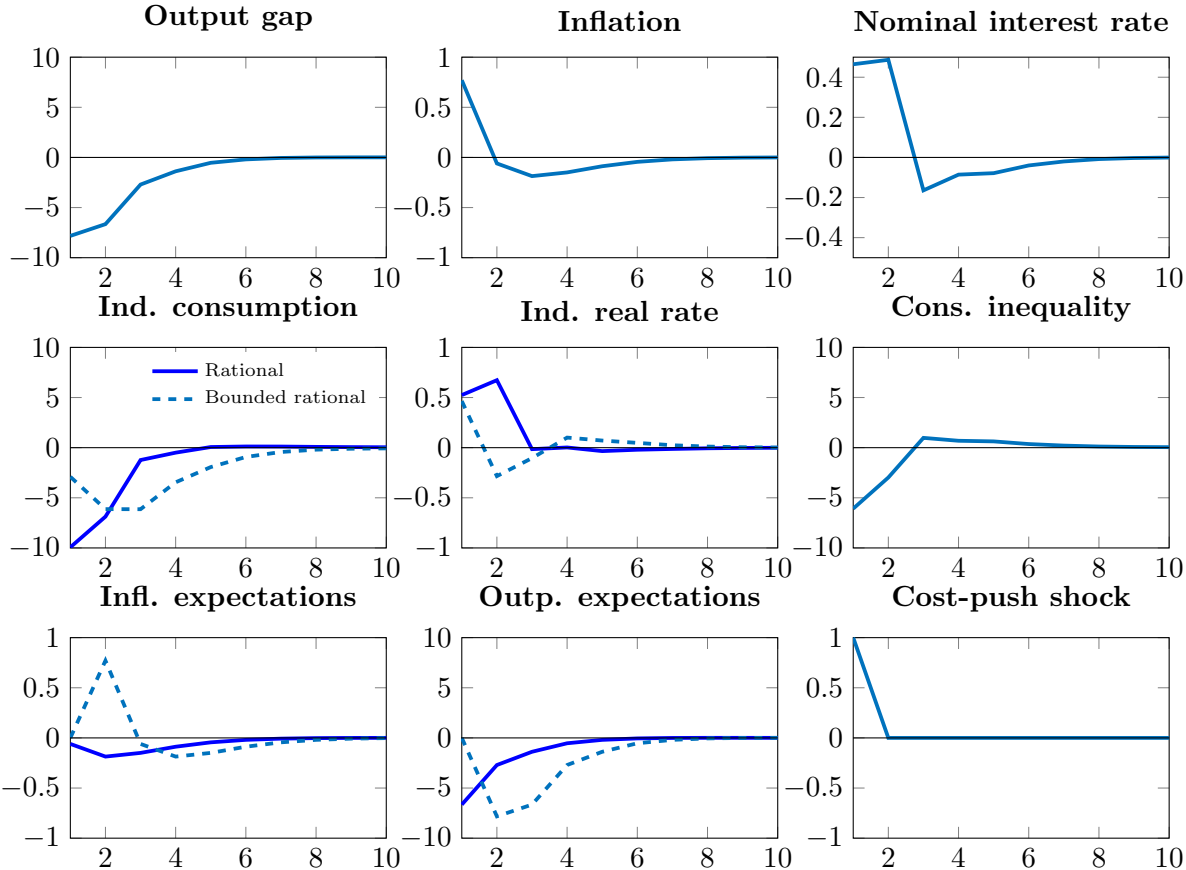

Figure 2: Impulse responses in %-deviations from steady state following a single, non-autocorrelated cost-push shock.

On the individual level, the disparity between the consumption adjustment paths of both agent types

becomes obvious. While bounded rational forecasters cut their consumption by approximately three percent on impact, rational forecasters decrease consumption by almost ten percent. This is, first, because of substantially negative rational output gap expectations and, second, due to a slightly higher subjective real interest rate. On the other hand, as bounded rational forecasters are backward-looking, their output gap expectations are zero on impact and, therefore, cut their consumption because of the increase in the subjective real interest rate only. This results in a consumption cut that is far smaller compared to rational forecasters and thereby in significant inequality in individual consumption on impact.[6]

Thus, bounded rational forecaster seem to be better off than rational ones at first. However, bounded rational forecasters make less smart decisions than rational forecasters *by definition.* This becomes clear when looking at the following periods where bounded rational forecasters have to pay for their initially higher consumption by giving up a lot of future consumption. Specifically, one can observe that bounded rational output gap expectations in the second period $(t+1)$ *drop* drastically to $E_{t+1}^B y_{t+2} = y_t$ which is approximately minus 8 percent, resulting in a further cut of consumption. This is the case even though the subjective real interest rate of bounded rational forecasters becomes negative which is due to the jump of their inflation expectations to $E_{t+1}^B \pi_{t+2} = \pi_t$. At the same time, output gap expectations of rational forecasters *increase* as the output gap recovers. From this period onwards rational forecasters are able to consume more than bounded rational ones for a prolonged time.

5 Welfare evaluation

This section provides a comparison between the optimal interest rate rule (28) and the non-optimal inflation-targeting rule (22) in terms of welfare and a brief discussion on consumption dispersion. In particular, I analyze the welfare consequences of these rules following a one-percent i.i.d. cost-push shock as before. It should be noted that shocks to inflation directly (and hence to price dispersion) induce high welfare losses. The reason is, first, that price dispersion leads to dispersion in imperfectly substitutable individual production and, therefore, to losses in the final consumption bundle and, second, that output itself needs to be contracted in order to bring down inflation. Further, I will restrict the analysis in this Section to the case of naive expectations, i.e. $\theta = 1$, of bounded rational forecasters for simplicity.

---

[6]I define consumption inequality here as $ci_t = \alpha c_t^R - (1-\alpha)c_t^B$ which should not be confused with the cross-sectional variance of consumption.

5.1 Optimal vs. inflation-targeting rule

In the following, I show *to what extent* the optimal interest rate rule (28) yields lower welfare losses compared to rule (22) depending on the fraction of rational forecasters. To that end, I compute consumption equivalent welfare losses following Ravenna and Walsh (2011). Let

$$\mathcal{W}^O = -\frac{\bar{C}U_c}{2}E_t\sum_{t=0}^{\infty}\beta^t L_t^O + t.i.p. = -\frac{\bar{C}U_c}{2(1-\beta)}L^O + t.i.p. \tag{30}$$

be the welfare loss under the optimal commitment policy (28), and

$$\mathcal{W}^{IT} = -\frac{\bar{C}U_c}{2}E_t\sum_{t=0}^{\infty}\beta^t L_t^{IT} + t.i.p. = -\frac{\bar{C}U_c}{2(1-\beta)}L^{IT} + t.i.p. \tag{31}$$

be the welfare loss under the non-optimal inflation-targeting objective (22), where instantaneous losses are measured by (17) in both cases. The welfare loss of implementing policy (22) instead of (28) can be measured as the percentage increase of steady state consumption, E, satisfying

$$\frac{U((1+E)\times\bar{C})}{1-\beta} + \mathcal{W}^{IT} = \frac{U(\bar{C})}{1-\beta} + \mathcal{W}^O. \tag{32}$$

Inserting consumption utility $U(\bar{C}) = \frac{\bar{C}^{1-\frac{1}{\sigma}}}{1-\frac{1}{\sigma}}$, (30), (31) and $U_c = \bar{C}^{-\frac{1}{\sigma}}$, and solving for E gives[7]

$$E = \left(1 - \frac{\sigma-1}{2\sigma}(L^O - L^{IT})\right)^{\frac{\sigma}{\sigma-1}} - 1. \tag{33}$$

Table 2 shows absolute losses L^O and L^{IT} (second and third column) as well as the consumption equivalent welfare costs, E, (fifth column) for different fractions of rational agents, α. The other columns of Table 2 are discussed further below.

A first, more trivial observation is that absolute welfare losses increase with the fraction of bounded rational forecasters for both interest rate rules. The higher the fraction of naive forecasters, the more persistent the deviations of variables from steady state and, therefore, the higher the long-run variances. Further, welfare losses are naturally lowest under the optimal interest rate rule (28). Deriving the interest rate rule from the conventional inflation-targeting (IT) objective (18) is costly in terms of consumption equivalents, E. The implementation of the non-optimal inflation-targeting rule (22) already requires an increase of 14.5 percent of steady-state consumption to compensate for the higher welfare losses relative to the optimal interest rate rule (28) when only ten percent of the population form (naive) backward-

---

[7]Note that the terms independent of policy (*t.i.p.*) are the same for both $\mathcal{W}^{IT}$ and $\mathcal{W}^O$ and, therefore, cancel each other.

α	L^{T*}	L^O	L^{IT}	L^{IT*}	E	E^*
0.95	161.65	127.077	127.138	127.09	0.030	0.006
0.9	183.82	144.988	145.274	144.99	0.145	0.034
0.7	277.56	243.455	246.632	244.02	1.744	0.290
0.5	400.36	394.547	402.428	395.45	4.693	0.47

Table 2: Column 2-4 show absolute welfare losses for *numerically optimized* simple Taylor rule (T^*), the optimal (O) interest rate rule (28), the inflation-targeting (IT) rule (22) and the inflation-targeting rule where the coefficients on inflation and the cost-push shock are *numerically optimized* (IT^*). Column 5-6 depict consumption equivalent welfare costs of implementing the non-optimal inflation-targeting (IT) interest rate rule (22) relative to the optimal (O) commitment policy (28) in percent, E, and also for the *numerically optimized* inflation-targeting (IT^*) rule relative to the optimal rule, E^*. All shown values are calculated for different fractions of rational agents, α.

looking expectations. Welfare costs become substantially higher for higher fractions of bounded rational forecasters (lower α). Hence, the optimal interest rate rule (28) yields considerable welfare gains when the underlying economy features high degree of bounded rationality.

The difference between the two rules can be explained by their relative ability to stabilize inflation. The first two rows in Table 3 depict the variances of inflation and output for different values of rational forecasters, α, for both interest rate rules. In general, the optimal interest rate rule (first row) yields lower inflation but higher output volatility across all fractions of rational forecasters, α. This is a straightforward implication of the relatively higher weights on inflation in the model-consistent loss function (17) compared to the conventional inflation-targeting objective (18), as discussed in Section 3.1. While the differences in output and inflation volatility are relatively low for $\alpha = 0.95$, they become substantial for higher fractions of bounded rational forecasters, resulting in quite extreme consumption equivalents.

Further, as discussed in Section 3.2, the optimal interest rate rule (28) depends on much more leads and lags of all variables compared to the non-optimal inflation-targeting rule (22). This raises the question whether it is the absence of these leads and lags that makes (22) sub-optimal or if it is rather an issue of weighting the different variables in the interest rate rule, or both. To answer this question, I *numerically optimize* the weights on inflation and the cost-push shock with respect to welfare (17) while "fixing" the coefficients on output to the analytically derived ones, (23) and (24), for simplicity. The corresponding absolute welfare losses, L^{IT*}, and consumption equivalent welfare costs relative to the optimal commitment policy, E^*, can be found in columns four and six in Table 2, respectively.

Consumption-equivalent welfare costs, E^*, *substantially* decrease relative to the consumption-equivalent welfare costs, E. This indicates that the inflation-targeting interest rate rule (22) is sub-optimal rather because of the sub-optimal weighting, especially for inflation. Hence, the optimal interest rate rule can be approximated relatively well by using the functional form of the inflation-targeting rule (22) and "adjusting" the weights accordingly.

	$Var(y)$				$Var(\pi)$			
rule	$\alpha = 0.95$	$\alpha = 0.9$	$\alpha = 0.7$	$\alpha = 0.5$	$\alpha = 0.95$	$\alpha = 0.9$	$\alpha = 0.7$	$\alpha = 0.5$
Optimal	41.23	50.66	115.17	231.69	0.58	0.60	0.66	0.67
IT	38.62	44.80	88.98	179.89	0.59	0.62	0.75	0.83
IT^*	41.37	50.93	116.82	232.55	0.58	0.60	0.66	0.66
T^*	42.00	59.98	149.45	240.37	0.78	0.78	0.70	0.65

Table 3: Theoretical variances of inflation and output under the optimal (O) rule (28), the inflation-targeting (IT) objective (22), the *numerically optimized* inflation-targeting (IT^*) rule and the *numerically optimized* simple Taylor rule (T^*) for different values of rational forecaster α.

Finally, for the sake of comparison, I add a numerically optimized simple Taylor rule with contemporaneous inflation and output to the analysis. Absolute welfare losses, L^{T*}, are shown in the first column of Table 2 and the corresponding variances of inflation and output can be found in the last row of Table 3. For a fraction of 70 percent of rational forecasters, or higher, the numerically optimized simple Taylor rule is the worst performing among all alternatives. However, interestingly, for 50 percent of rational forecasters it trumps the analytically derived inflation-targeting rule (22). In this case, it even yields lower inflation volatility than the optimal commitment policy. This is, however, not welfare maximizing which indicates that, although inflation is the most important driver of welfare, there is still a welfare-relevant trade-off between inflation and output as output volatility is highest among all interest rate rules for $\alpha = 0.5$.

5.2 Inflation and welfare vs. consumption dispersion

In this section, I briefly discuss the issue of distribution, measured by the cross-sectional variance in consumption (16), and monetary policy. It should be noted, however, that *welfare losses* due to consumption dispersion are in general very low under the model-consistent loss function (17), where individual utilities are weighted equally and simply summed up. Of course, different conclusions may be reached when a

social planer would attach higher weights to "inequality".

Table 4 shows consumption dispersion (CD) and the variance of inflation for both the optimal interest rate rule (O) and inflation-targeting rule (IT) for different fractions of rational forecasters, α. Interestingly, consumption dispersion is *higher* under the optimal interest rate rule compared to the (analytically derived) inflation-targeting rule for all α. This is the case even though the former explicitly incorporates consumption heterogeneity as opposed to the latter. At the same time, the optimal interest rate rule yields lower inflation volatility, as discussed in the previous section. This indicates that the model additionally implies a trade-off between stabilizing inflation and consumption heterogeneity.

α	CD^O	CD^{IT}	$var(\pi)^O$	$var(\pi)^{IT}$
0.95	1.904	1.104	0.580	0.590
0.9	4.071	2.567	0.597	0.618
0.7	17.805	12.700	0.664	0.750
0.5	39.071	25.812	0.670	0.833

Table 4: Consumption dispersion (CD) $var_i(c_t(i))$ under the optimal (O) interest rate rule (28) and the non-optimal inflation-targeting (IT) rule (22) for different fractions of rational forecaster α.

The trade-off becomes more evident when considering Figure 3. Figure 3 depicts consumption dispersion, the inflation variance (right ordinate) and absolute welfare losses (left ordinate) in case of a simple Taylor rule against different values of the coefficient on inflation.[8] The higher the coefficient on inflation, the lower inflation volatility and the higher consumption dispersion. Therefore, it is not surprising that this also implies a local trade-off between minimizing welfare losses and reducing consumption dispersion.[9] Minimizing welfare losses requires the central bank to get a tight grip on inflation causing a substantial drop of individual consumption over time. As both agents react quite differently to the increase in the policy rate, as discussed in Section 4, consumption dispersion increases with the central bank's aggressiveness towards inflation.

Another observation worth mentioning is that consumption dispersion is minimized when the coefficient on inflation in the Taylor rule is one, i.e. when $i_t = \pi_t$ holds. This was analytically shown by Hagenhoff and Lustenhouwer (2019) in a model with fully rational agents and bounded rational agents

---

[8]The coefficient on output is zero which is optimal under the model-consistent loss function (17).

[9]On the left side of the minimum, welfare losses decrease because of decreasing inflation. At the same time, consumption dispersion increases. At some point, however, welfare losses increase again as output volatility (not shown in Figure 3) becomes substantially higher as the central bank needs to contract output further to achieve further reductions in inflation. Therefore, the trade-off between welfare and consumption dispersion arises only locally, i.e. on the left side of the minimum where welfare losses and inflation decrease simultaneously.
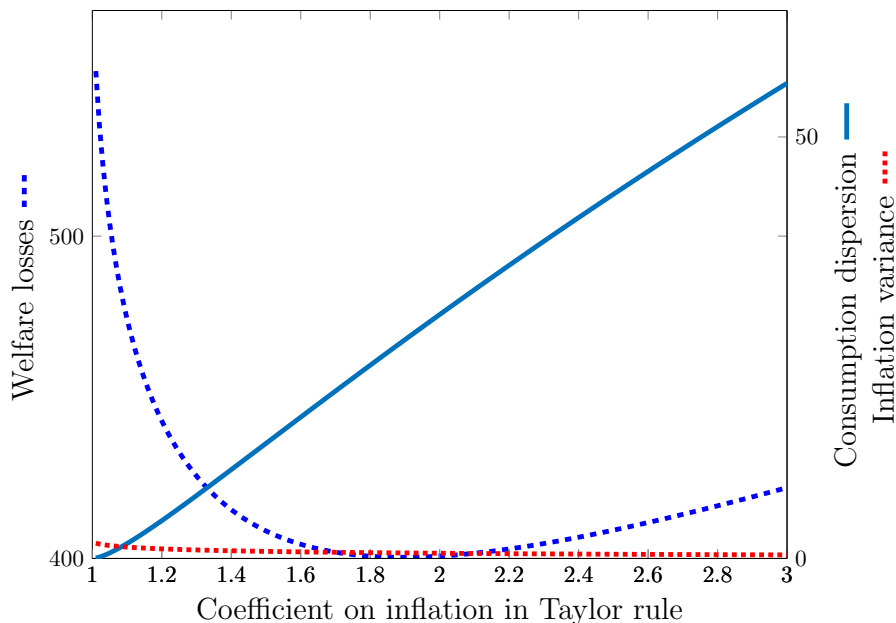
Figure 3: Trade-off between minimizing welfare, inflation and consumption dispersion under a simple Taylor rule and 70% of rational agents.

similar to this paper. Thus, the appearance of the same finding in this model serves as a robustness check for Hagenhoff and Lustenhouwer (2019).

6   Conclusion

In this paper, I propose a fully optimal interest rate rule under heterogeneous expectations where the central bank commits to its policy from a timeless perspective. This rule incorporates the more complex nature of price dispersion and consumption dispersion under heterogeneous expectations as identified by Di Bartolomeo et al. (2016). Further, this rules performs considerably better than a micro-founded version of the interest rate rule as in Gasteiger (2014). The implementation of the non-optimal inflation-targeting rule already requires an increase of 14.5 percent of steady-state consumption to compensate for the higher welfare losses relative to the optimal interest rate rule when only ten percent of the population form (naive) backward-looking expectations.

I additionally explore the properties of the expectations operator of "rational" agents in the Branch and McGough (2009) framework and find that the consumption Euler equation that includes model-consistent individual consumption expectations as in Di Bartolomeo et al. (2016) is inconsistent with the higher-order beliefs assumption of Branch and McGough (2009). This assumption puts a specific (non-rational) structure on higher-order beliefs which implies that not even "rational forecasters" understand the micro level fully. Therefore, I derive consumption decisions that account for this particular assumption

which makes it the implementation of the optimal commitment policy by an interest rate rule possible in the first place.

Finally, I illustrate that the model implies a local trade-off between maximizing welfare and reducing consumption dispersion. The reason is that consumption dispersion increases with the central bank's aggressiveness towards inflation, as rational and bounded rational forecasters consumption decisions become more unequal with more aggressive inflation-targeting. Because inflation is the most important determinant of welfare, the central bank has to allow for a certain heterogeneity in consumption to maximize welfare.

Acknowledgements

References

Assenza, T., Heemeijer, P., Hommes, C., Massaro, D., et al. (2014). Managing self-organization of expectations through monetary policy: a macro experiment. *CeN-DEF Working Paper*, pages 14–07.

Beqiraj, E., Di Bartolomeo, G., and Serpieri, C. (2017). Rational vs. long-run forecasters: Optimal monetary policy and the role of inequality. *Macroeconomic Dynamics*, page 1–16.

Branch, W. A. (2004). The theory of rationally heterogeneous expectations: Evidence from survey data on inflation expectations*. *The Economic Journal*, 114(497):592–621.

Branch, W. A. (2007). Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control*, 31(1):245–276.

Branch, W. A. and McGough, B. (2009). A new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 33(5):1036 – 1051.

Branch, W. A. and McGough, B. (2010). Dynamic predictor selection in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 34(8):1492 – 1508.

Calvo, G. A. (1983). Staggered prices in a utility-maximizing framework. *Journal of monetary Economics*, 12(3):383–398.

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

De Grauwe, P. (2011). Animal spirits and monetary policy. *Economic theory*, 47(2-3):423–457.

Debortoli, D. and Galí, J. (2017). Monetary policy with heterogeneous agents: Insights from tank models. *Manuscript, September.*

Di Bartolomeo, G., Di Pietro, M., and Giannini, B. (2016). Optimal monetary policy in a new keynesian model with heterogeneous expectations. *Journal of Economic Dynamics and Control*, 73:373 – 387.

Evans, G. W. and Honkapohja, S. (2006). Monetary policy, expectations and commitment. *The Scandinavian Journal of Economics*, 108(1):15–38.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications.* Princeton University Press.

Gasteiger, E. (2014). Heterogeneous expectations, optimal monetary policy, and the merit of policy inertia. *Journal of Money, Credit and Banking*, 46(7):1535–1554.

Gasteiger, E. (2018). Optimal constrained interest-rate rules under heterogeneous expectations. *Working Paper.*

Hagenhoff, T. and Lustenhouwer, J. (2019). The rationality bias. *BERG Working Paper Series.*

Hommes, C., Lustenhouwer, J., et al. (2015). Inflation targeting and liquidity traps under endogenous credibility. *CeNDEF working paper*, (15-03).

Massaro, D. (2013). Heterogeneous expectations in monetary DSGE models. *Journal of Economic Dynamics and Control*, 37(3):680–692.

McCallum, B. T. (1999). Issues in the design of monetary policy rules. *Handbook of macroeconomics*, 1:1483–1530.

Pfajfar, D. and Santoro, E. (2010). Heterogeneity, learning and information stickiness in inflation expectations. *Journal of Economic Behavior & Organization*, 75(3):426–444.

Pfajfar, D. and Žakelj, B. (2016). Inflation expectations and monetary policy design: Evidence from the laboratory. *Macroeconomic Dynamics*, 22(4):1035–1075.

Ravenna, F. and Walsh, C. E. (2011). Welfare-based optimal monetary policy with unemployment and sticky prices: a linear-quadratic framework. *American Economic Journal: Macroeconomics*, 3(2):130–62.

Rotemberg, J. J. and Woodford, M. (1997). An optimization-based econometric framework for the evaluation of monetary policy. *NBER macroeconomics annual*, 12:297–346.

Woodford, M. (1999). Optimal monetary policy inertia. *The Manchester School*, 67:1–35.

Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton University Press.

A   Implementation under the conventional inflation-targeting objective

The policy problem under commitment and the conventional inflation-targeting objective is given by

$$\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \frac{1}{2} \left[\left(\eta + \frac{1}{\sigma} \right) y_{t+s}^2 + \epsilon^2 \eta \delta \pi_{t+s}^2 \right] \tag{34}$$

$$+ \lambda_{t+s} [\pi_{t+s} - \alpha \beta E_t \pi_{t+s+1} - (1-\alpha)\beta \theta^2 \pi_{t+s-1} - \kappa y_{t+s} - e_{t+s}]$$

$$\frac{\partial \mathcal{L}}{\partial \pi_{t+s}} : \ E_t \left\{ \epsilon^2 \eta \delta \pi_{t+s} + \frac{\lambda_{t+s}}{2} - (1-\alpha)\beta^2 \theta^2 \frac{\lambda_{t+s+1}}{2} - \alpha \frac{\lambda_{t+s-1}}{2} \right\} \overset{!}{=} 0 \tag{35}$$

$$\frac{\partial \mathcal{L}}{\partial y_{t+s}} : \ E_t \left\{ \beta^s \left[\left(\eta + \frac{1}{\sigma} \right) y_{t+s} - \frac{\kappa}{2} \lambda_{t+s} \right] \right\} \overset{!}{=} 0. \tag{36}$$

Combining and solving for inflation gives

$$\pi_t = -\frac{1+\eta\sigma}{\sigma \epsilon^2 \eta \delta \kappa} [y_t - \alpha y_{t-1} - (1-\alpha)\beta^2 \theta^2 E_t y_{t+1}] \tag{37}$$

where the index s was dropped as the central bank employs timeless commitment. Combining with the Phillips and IS curve yields (22).

B Optimal monetary policy

B.1 Rewriting the model-consistent loss function

The period loss function L_t is given by

$$
\begin{aligned}
L_t = {} & \frac{\sigma\eta + 1}{\sigma} y_t^2 + \frac{\alpha(y_t - c_t^R)^2}{(1-\alpha)\sigma} \\
& + \epsilon^2 \eta \delta \left\{ \pi_t^2 + \frac{\xi_p(1-\alpha)}{\alpha} \left[\pi_t - \beta\theta^2 \pi_{t-1} - \kappa y_t - \frac{\alpha\kappa(y_t - c_t^R)}{(1+\eta\sigma)(1-\alpha)} \right]^2 \right\}.
\end{aligned}
\tag{38}
$$

which can be rewritten using market clearing to eliminate c_t^B as

$$
\begin{aligned}
L_t = {} & \frac{\sigma\eta + 1}{\sigma} y_t^2 + \frac{\alpha(y_t - c_t^R)^2}{(1-\alpha)\sigma} \\
& + \epsilon^2 \eta \delta \left\{ \pi_t^2 + \frac{\xi_p(1-\alpha)}{\alpha} \left[\pi_t - \beta\theta^2 \pi_{t-1} - \kappa y_t - \frac{\alpha\kappa(y_t - c_t^R)}{(1+\eta\sigma)(1-\alpha)} \right]^2 \right\}.
\end{aligned}
\tag{39}
$$

By multiplying out, we get

$$
\begin{aligned}
L_t = {} & \Gamma_1 y_t^2 + \Gamma_2 \pi_t^2 + \Gamma_3 \pi_{t-1}^2 + \Gamma_4 (c_t^R)^2 \\
& + \Gamma_5 y_t c_t^R + \Gamma_6 \pi_t c_t^R + \Gamma_7 \pi_{t-1} c_t^R + \Gamma_8 \pi_t \pi_{t-1} + \Gamma_9 \pi_t y_t + \Gamma_{10} \pi_{t-1} y_t
\end{aligned}
\tag{40}
$$

with

$$\Gamma_1 = \frac{((\alpha-1)\eta\sigma - 1)(\alpha(\eta^2\sigma^2(\delta\epsilon^2\kappa^2\xi_p - 1) - 1 - 2\eta\sigma) - \delta\epsilon^2\eta\kappa^2\xi_p\sigma(1 + \eta\sigma))}{(1-\alpha)\alpha\sigma(1+\eta\sigma)^2} \tag{41}$$

$$\Gamma_2 = \frac{\delta\epsilon^2\eta(\alpha + \xi_p - \alpha\xi_p)}{\alpha} \tag{42}$$

$$\Gamma_3 = \frac{(1-\alpha)\beta^2\delta\epsilon^2\eta\theta^4\xi_p}{\alpha} \tag{43}$$

$$\Gamma_4 = \frac{\alpha(1 + \eta\sigma(2 + \delta\epsilon^2\kappa^2\xi_p) + \eta^2\sigma^2)}{(1-\alpha)\sigma(1+\eta\sigma)^2} \tag{44}$$

$$\Gamma_5 = \frac{2(\alpha + 2\alpha\eta\sigma + \alpha\eta^2\sigma^2(1 - \delta\epsilon^2\kappa^2\xi_p) + \delta\epsilon^2\eta\kappa^2\xi_p\sigma(1+\eta\sigma))}{(\alpha-1)\sigma(1+\eta\sigma)^2} \tag{45}$$

$$\Gamma_6 = \frac{2\delta\epsilon^2\eta\kappa\xi_p}{1+\eta\sigma} \tag{46}$$

$$\Gamma_7 = -\frac{2\beta\delta\epsilon^2\eta\theta^2\kappa\xi_p}{1+\eta\sigma} \tag{47}$$

$$\Gamma_8 = \frac{2(\alpha-1)\beta\delta\epsilon^2\eta\theta^2\xi_p}{\alpha} \tag{48}$$

$$\Gamma_9 = \frac{2\delta\epsilon^2\eta\kappa\xi_p((\alpha-1)\eta\sigma - 1)}{\alpha + \alpha\eta\sigma} \tag{49}$$

$$\Gamma_{10} = \frac{2\beta\delta\epsilon^2\eta\theta^2\kappa\xi_p(1 + \eta\sigma(1-\alpha))}{\alpha + \alpha\eta\sigma}. \tag{50}$$

B.2   Optimal interest rate rule

The policy problem under full commitment takes the following form:

$$\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \Bigg[\Gamma_1 y_{t+s}^2 + \Gamma_2\pi_{t+s}^2 + \Gamma_3\pi_{t+s-1}^2 + \Gamma_4(c_{t+s}^R)^2$$

$$+\Gamma_5 y_{t+s}c_{t+s}^R + \Gamma_6\pi_{t+s}c_{t+s}^R + \Gamma_7\pi_{t+s-1}c_{t+s}^R + \Gamma_8\pi_{t+s}\pi_{t+s-1} + \Gamma_9\pi_{t+s}y_{t+s} + \Gamma_{10}\pi_{t+s-1}y_{t+s}$$

$$+ \lambda_{1,t+s}[y_{t+s} - \alpha E_t y_{t+s+1} - (1-\alpha)\theta^2 y_{t+s-1} + \sigma[i_{t+s} - \alpha E_t\pi_{t+s+1} - (1-\alpha)\theta^2\pi_{t+s-1}]]$$

$$+ \lambda_{2,t+s}[\pi_{t+s} - \alpha\beta E_t\pi_{t+s+1} - (1-\alpha)\beta\theta^2\pi_{t+s-1} - \kappa y_{t+s} - e_{t+s}]$$

$$+\lambda_{3,t+s}[c_{t+s}^R - E_t y_{t+s+1} - \phi b_{t+s-1}^R + \sigma(i_{t+s} - E_t\pi_{t+s+1})] \Bigg]. \tag{51}$$

The first order conditions are

$$
\frac{\partial \mathcal{L}}{\partial y_{t+s}} : E_t \Big\{ \beta^s [2\Gamma_1 y_{t+s} + \Gamma_5 c_{t+s}^R + \Gamma_9 \pi_{t+s} + \Gamma_{10} \pi_{t+s-1} + \lambda_{1,t+s} - \kappa \lambda_{2,t+s}] \tag{52}
$$
$$
- \beta^{s+1}(1-\alpha)\theta^2 \lambda_{1,t+s+1} - \beta^{s-1}[\alpha \lambda_{1,t+s-1} + \lambda_{3,t+s-1}] \Big\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial \pi_{t+s}} : E_t \Big\{ \beta^s [2\Gamma_2 \pi_{t+s} + \Gamma_6 c_{t+s}^R + \Gamma_8 \pi_{t+s-1} + \Gamma_9 y_{t+s} + \lambda_{2,t+s}]
$$
$$
+ \beta^{s+1}[2\Gamma_3 \pi_{t+s} + \Gamma_7 c_{t+s+1} + \Gamma_8 \pi_{t+s+1} + \Gamma_{10} y_{t+s+1} - (1-\alpha)\theta^2 \sigma \lambda_{1,t+s+1} \tag{53}
$$
$$
- (1-\alpha)\beta\theta^2 \lambda_{2,t+s+1}] - \beta^{s-1}[\alpha\sigma \lambda_{1,t+s-1} + \alpha\beta \lambda_{2,t+s-1} + \sigma\ \lambda_{3,t+s-1}] \Big\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial c_{t+s}^R} : E_t \Big\{ \beta^s [2\Gamma_4 c_{t+s}^R + \Gamma_5 y_{t+s} + \Gamma_6 \pi_{t+s} + \Gamma_7 \pi_{t+s-1} \tag{54}
$$
$$
+ \lambda_{3,t+s}] \Big\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial i_{t+s}} : E_t \Big\{ \beta^s \sigma \lambda_{1,t+s} + \beta^s \sigma \lambda_{3,t+s} \Big\} \overset{!}{=} 0. \tag{55}
$$

Again, the index s can be dropped assuming commitment from a timeless perspective. Using $\lambda_{3,t} = -\lambda_{1,t}$ the FOCs can equivalently be written as

$$
2\Gamma_1 y_t + \Gamma_5 c_t^R + \Gamma_9 \pi_t + \Gamma_{10} \pi_{t-1} + \lambda_{1,t} - \kappa \lambda_{2,t} - (1-\alpha)\beta\theta^2 \lambda_{1,t+1} -
$$
$$
\beta^{-1}(\alpha - 1)\lambda_{1,t-1} \overset{!}{=} 0 \tag{56}
$$

$$
2\Gamma_2 \pi_t + \Gamma_6 c_t^R + \Gamma_8 \pi_{t-1} + \Gamma_9 y_t + \lambda_{2,t}
$$
$$
+ 2\Gamma_3 \beta \pi_t + \Gamma_7 \beta c_{t+1} + \Gamma_8 \beta \pi_{t+1} + \Gamma_{10} \beta y_{t+1} - (1-\alpha)\beta\theta^2 \sigma \lambda_{1,t+1} \tag{57}
$$
$$
- (1-\alpha)\beta^2 \theta^2 \lambda_{2,t+1} + \beta^{-1}(1-\alpha)\sigma \lambda_{1,t-1} - \alpha \lambda_{2,t-1} \overset{!}{=} 0
$$

$$
2\Gamma_4 c_t^R + \Gamma_5 y_t + \Gamma_6 \pi_t + \Gamma_7 \pi_{t-1} - \lambda_{1,t} \overset{!}{=} 0. \tag{58}
$$

Eliminating the Lagrange multipliers yields the reduced-form FOC

$$
\Delta_1^c \pi_t + \Delta_2^c \pi_{t+1} + \Delta_3^c \pi_{t+2} + \Delta_4^c \pi_{t-3} + \Delta_5^c \pi_{t-2} + \Delta_6^c \pi_{t-1} + \Delta_7 y_t + \Delta_8^c y_{t+1} \tag{59}
$$
$$
+ \Delta_9^c y_{t+2} + \Delta_{10}^c y_{t-2} + \Delta_{11}^c y_{t-1} + \Delta_{12} c_t^R + \Delta_{13}^c c_{t+1}^R + \Delta_{14}^c c_{t+2}^R + \Delta_{15}^c c_{t-2}^R + \Delta_{16}^c c_{t-1}^R \overset{!}{=} 0.
$$

with

$$\Delta_1 = \Gamma_6 + \Gamma_9 + (1-\alpha)(-1+2\alpha)\beta\Gamma_6\theta^2 + 2\Gamma_2\kappa + 2\beta\Gamma_3\kappa \tag{60}$$

$$+ (\alpha-1)\beta\theta^2(\Gamma_7 + \beta(\Gamma_{10}+\Gamma_7) + \Gamma_7\kappa\sigma) \tag{61}$$

$$\Delta_2 = \beta(\Gamma_8\kappa + (\alpha-1)\theta^2(\beta(\Gamma_9 + (\alpha-1)\beta\Gamma_7\theta^2) + \Gamma_6(1+\beta+\kappa\sigma))) \tag{62}$$

$$\Delta_3 = (\alpha-1)^2\beta^3\Gamma_6\theta^4 \tag{63}$$

$$\Delta_4 = \frac{(\alpha-1)\alpha\Gamma_7}{\beta} \tag{64}$$

$$\Delta_5 = \frac{\alpha^2\Gamma_6 + \Gamma_7 + \Gamma_7\kappa\sigma - \alpha(\Gamma_6 + \Gamma_7 + \beta(\Gamma_{10}+\Gamma_7) + \Gamma_7\kappa\sigma)}{\beta} \tag{65}$$

$$\Delta_6 = \Gamma_{10} + \Gamma_7 - \alpha(\Gamma_6+\Gamma_9) + (-1+(3-2\alpha)\alpha)\beta\Gamma_7\theta^2 + \Gamma_8\kappa \tag{66}$$

$$- \frac{(\alpha-1)\Gamma_6(1+\kappa\sigma)}{\beta} \tag{67}$$

$$\Delta_7 = 2\Gamma_1 + \Gamma_5 + (1-\alpha)(2\alpha-1)\beta\Gamma_5\theta^2 + \Gamma_9\kappa \tag{68}$$

$$\Delta_8 = \beta(\Gamma_{10}\kappa + (\alpha-1)\theta^2(\Gamma_5 + \beta(2\Gamma_1+\Gamma_5) + \Gamma_5\kappa\sigma) \tag{69}$$

$$\Delta_9 = (\alpha-1)^2\beta^3\Gamma_5\theta^4 \tag{70}$$

$$\Delta_{10} = \frac{(\alpha-1)\alpha\Gamma_5}{\beta} \tag{71}$$

$$\Delta_{11} = \frac{\Gamma_5 + \Gamma_5\kappa\sigma - \alpha(\Gamma_5 + \beta(2\Gamma_1+\Gamma_5) + \Gamma_5\kappa\sigma)}{\beta} \tag{72}$$

$$\Delta_{12} = 2\Gamma_4 + \Gamma_5 - 2(\alpha-1)(2\alpha-1)\beta\Gamma_4\theta^2 + \Gamma_6\kappa) \tag{73}$$

$$\Delta_{13} = \beta(\Gamma_7\kappa + (\alpha-1)\theta^2(\beta\Gamma_5 + 2\Gamma_4(1+\beta+\kappa\sigma)) \tag{74}$$

$$\Delta_{14} = 2(\alpha-1)^2\beta^3\Gamma_4\theta^4 \tag{75}$$

$$\Delta_{15} = \frac{2(\alpha-1)\alpha\Gamma_4}{\beta} \tag{76}$$

$$\Delta_{16} = -\alpha\Gamma_5 - \frac{2\Gamma_4(-1+\alpha+\alpha\beta+(\alpha-1)\kappa\sigma)}{\beta}. \tag{77}$$

Solving (59) for π_t and setting it equal to the NK Phillips curve yields

$$y_t = -\frac{1}{\Delta_7 + \Delta_1\kappa}((\alpha\beta\Delta_1 + \Delta_2)\pi_{t+1} + \Delta_3\pi_{t+2} + \Delta_4\pi_{t-3} + \Delta_5\pi_{t-2}$$

$$+ (\Delta_6 + (1-\alpha)\beta\theta^2\Delta_1)\pi_{t-1} + \Delta_8 y_{t+1} + \Delta_9 y_{t+2} + \Delta_{10}y_{t-2} + \Delta_{11}y_{t-1} \tag{78}$$

$$+ \Delta_{13}c_{t+1}^R + \Delta_{14}c_{t+2}^R + \Delta_{15}c_{t-2}^R + \Delta_{16}c_{t-1}^R)$$

Setting (78) equal to the New IS curve, substituting c_t^R for consumption demand and solving for i_t gives the central bank's reaction function under commitment (28):

$$i_t = \Omega_1 E_t \pi_{t+1} + \Omega_2 E_t \pi_{t+2} + \Omega_3 \pi_{t-3} + \Omega_4 \pi_{t-2} + \Omega_5 \pi_{t-1} + \Omega_6 E_t y_{t+1} + \Omega_7 E_t y_{t+2} \tag{79}$$

$$+ \Omega_8 y_{t-2} + \Omega_9 y_{t-1} + \Omega_{10} E_t c_{t+1}^R + \Omega_{11} E_t c_{t+2}^R + \Omega_{12} c_{t-2}^R + \Omega_{13} c_{t-1}^R + \Omega_{14} e_t$$

with

$$\Omega_1 = \frac{\alpha\beta\Delta_1 + \Delta_2 + \sigma\Delta_{12} + \alpha\sigma(\Delta_7 + \Delta_1\kappa)}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{80}$$

$$\Omega_2 = \frac{\Delta_3}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{81}$$

$$\Omega_3 = \frac{\Delta_4}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{82}$$

$$\Omega_4 = \frac{\Delta_5}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{83}$$

$$\Omega_5 = \frac{\Delta_6 + (1-\alpha)\theta^2(\beta\Delta_1 + \sigma(\Delta_7 + \Delta_1\kappa))}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{84}$$

$$\Omega_6 = \frac{\Delta_{12} + \alpha\Delta_7 + \Delta_8 + \alpha\kappa\Delta_1}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{85}$$

$$\Omega_7 = \frac{\Delta_9}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{86}$$

$$\Omega_8 = \frac{\Delta_{10}}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{87}$$

$$\Omega_9 = \frac{\Delta_{11} + (1-\alpha)\theta^2(\Delta_7 + \Delta_1\kappa)}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{88}$$

$$\Omega_{10} = \frac{\Delta_{13}}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{89}$$

$$\Omega_{11} = \frac{\Delta_{14}}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{90}$$

$$\Omega_{12} = \frac{\Delta_{15}}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{91}$$

$$\Omega_{13} = \frac{\Delta_{16}}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)} \tag{92}$$

$$\Omega_{14} = \frac{\Delta_1}{\sigma(\Delta_7 + \Delta_{12} + \Delta_1\kappa)}. \tag{93}$$

B.3 Taking rational expectations as given

Defining $f_{t+s} = E_t y_{t+s+1}$, $g_{t+s} = E_t \pi_{t+s+1}$ and $h_{t+s} = E_t c^R_{t+s+1}$, the policy problem is

$$
\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \left[\Gamma_1 y_{t+s}^2 + \Gamma_2 \pi_{t+s}^2 + \Gamma_3 \pi_{t+s-1}^2 + \Gamma_4 (c^R_{t+s})^2 \right.
$$

$$
+\Gamma_5 y_{t+s} c^R_{t+s} + \Gamma_6 \pi_{t+s} c^R_{t+s} + \Gamma_7 \pi_{t+s-1} c^R_{t+s} + \Gamma_8 \pi_{t+s} \pi_{t+s-1} + \Gamma_9 \pi_{t+s} y_{t+s} + \Gamma_{10} \pi_{t+s-1} y_{t+s}
$$

$$
+ \lambda_{1,t+s} [y_{t+s} - \alpha f_{t+s} - (1-\alpha)\theta^2 y_{t+s-1} + \sigma[i_{t+s} - \alpha g_{t+s} - (1-\alpha)\theta^2 \pi_{t+s-1}]]
$$

$$
+ \lambda_{2,t+s} [\pi_{t+s} - \alpha\beta g_{t+s} - (1-\alpha)\beta\theta^2 \pi_{t+s-1} - \kappa y_{t+s} - e_{t+s}]
$$

$$
\left. +\lambda_{3,t+s} [c^R_{t+s} - h_{t+s} - \phi b^R_{t+s-1} + \sigma(i_{t+s} - g_{t+s})] \right]. \tag{94}
$$

The first order conditions are

$$
\frac{\partial \mathcal{L}}{\partial y_{t+s}} : E_t \left\{ \beta^s [2\Gamma_1 y_{t+s} + \Gamma_5 c^R_{t+s} + \Gamma_9 \pi_{t+s} + \Gamma_{10} \pi_{t+s-1} + \lambda_{1,t+s} - \kappa\lambda_{2,t+s}] \right. \tag{95}
$$

$$
\left. - \beta^{s+1}(1-\alpha)\theta^2 \lambda_{1,t+s+1} \right\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial \pi_{t+s}} : E_t \left\{ \beta^s [2\Gamma_2 \pi_{t+s} + \Gamma_6 c^R_{t+s} + \Gamma_8 \pi_{t+s-1} + \Gamma_9 y_{t+s} + \lambda_{2,t+s}] \right.
$$

$$
+ \beta^{s+1} [2\Gamma_3 \pi_{t+s} + \Gamma_7 c_{t+s+1} + \Gamma_8 \pi_{t+s+1} + \Gamma_{10} y_{t+s+1} - (1-\alpha)\theta^2 \sigma \lambda_{1,t+s+1} \tag{96}
$$

$$
\left. - (1-\alpha)\beta\theta^2 \lambda_{2,t+s+1}] \right\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial c^R_{t+s}} : E_t \left\{ \beta^s [2\Gamma_4 c^R_{t+s} + \Gamma_5 y_{t+s} + \Gamma_6 \pi_{t+s} + \Gamma_7 \pi_{t+s-1} + \lambda_{3,t+s}] \right\} \overset{!}{=} 0
$$

$$
\frac{\partial \mathcal{L}}{\partial i_{t+s}} : E_t \left\{ \beta^s \sigma\lambda_{1,t+s} + \beta^s \sigma\lambda_{3,t+s} \right\} \overset{!}{=} 0. \tag{97}
$$

Since the central bank acts under timeless commitment, the index s can be dropped. Using $\lambda_{3,t} = -\lambda_{1,t}$ the FOCs can equivalently be written as

$$2\Gamma_1 y_t + \Gamma_5 c_t^R + \Gamma_9 \pi_t + \Gamma_{10}\pi_{t-1} + \lambda_{1,t} - \kappa\lambda_{2,t} - (1-\alpha)\beta\theta^2 E_t\lambda_{1,t+1} \stackrel{!}{=} 0 \tag{98}$$

$$2\Gamma_2 \pi_t + \Gamma_6 c_t^R + \Gamma_8 \pi_{t-1} + \Gamma_9 y_t + \lambda_{2,t} + 2\Gamma_3\beta\pi_t + \Gamma_7\beta E_t c_{t+1}^R + \Gamma_8\beta E_t\pi_{t+1} \tag{99}$$
$$+ \Gamma_{10}\beta E_t y_{t+1} - (1-\alpha)\beta\theta^2\sigma E_t\lambda_{1,t+1} - (1-\alpha)\beta^2\theta^2 E_t\lambda_{2,t+1} \stackrel{!}{=} 0$$

$$2\Gamma_4 c_t^R + \Gamma_5 y_t + \Gamma_6 \pi_t + \Gamma_7 \pi_{t-1} - \lambda_{1,t} \stackrel{!}{=} 0. \tag{100}$$

Eliminating the Lagrange multipliers yields the reduced-form FOC

$$\Delta_1 \pi_t + \Delta_2 E_t\pi_{t+1} + \Delta_3 E_t\pi_{t+2} + \Delta_4 \pi_{t-1} + \Delta_5 y_t + \Delta_6 E_t y_{t+1} + \Delta_7 E_t y_{t+2}$$
$$+ \Delta_8 c_t^R + \Delta_9 E_t c_{t+1}^R + \Delta_{10} E_t c_{t+2}^R \stackrel{!}{=} 0 \tag{101}$$

with

$$\Delta_1 = -\frac{\Gamma_6 + \Gamma_9 + 2\Gamma_2\kappa + 2\beta\Gamma_3\kappa - (1-\alpha)\beta\theta^2(\Gamma_7 + \beta(\Gamma_{10} + \Gamma_7) + \Gamma_7\kappa\sigma)}{\kappa} \tag{102}$$

$$\Delta_2 = \frac{(1-\alpha)\beta\theta^2(\beta(\Gamma_9 - (1-\alpha)\beta\Gamma_7\theta^2) + \Gamma_6(1+\beta+\kappa\sigma))}{\kappa} - \beta\Gamma_8 \tag{103}$$

$$\Delta_3 = -\frac{(\alpha-1)^2\beta^3\theta^4\Gamma_6}{\kappa} \tag{104}$$

$$\Delta_4 = -\frac{\Gamma_{10} + \Gamma_7 + \Gamma_8\kappa}{\kappa} \tag{105}$$

$$\Delta_5 = -\frac{2\Gamma_1 + \Gamma_5 + \Gamma_9\kappa}{\kappa} \tag{106}$$

$$\Delta_6 = -\frac{\beta((\alpha-1)\beta(2\Gamma_1 + \Gamma_5)\theta^2 + \Gamma_{10}\kappa + (\alpha-1)\Gamma_5\theta^2(1+\kappa\sigma))}{\kappa} \tag{107}$$

$$\Delta_7 = -\frac{(\alpha-1)^2\beta^3\theta^4\Gamma_5}{\kappa} \tag{108}$$

$$\Delta_8 = -\frac{2\Gamma_4 + \Gamma_5 + \Gamma_6\kappa}{\kappa} \tag{109}$$

$$\Delta_9 = -\frac{\beta((\alpha-1)\beta\Gamma_5\theta^2 + \Gamma_7\kappa + 2(\alpha-1)\Gamma_4\theta^2(1+\beta+\kappa\sigma))}{\kappa} \tag{110}$$

$$\Delta_{10} = -\frac{2(\alpha-1)^2\beta^3\theta^4\Gamma_4}{\kappa}. \tag{111}$$

Solving (101) for π_t and setting it equal to the NK Phillips curve yields

$$y_t = -\frac{1}{\Delta_5 + \Delta_1\kappa}(\Delta_6 y_{t+1} + \Delta_7 y_{t+2} + (\Delta_2 + \alpha\beta\Delta_1)\pi_{t+1} + \Delta_3\pi_{t+2} + (\Delta_4 + (1-\alpha)\beta\theta^2\Delta_1)\pi_{t-1}$$

$$+ \Delta_8 c_t^R + \Delta_9 c_{t+1}^R + \Delta_{10} c_{t+2}^R + \Delta_1 e_t). \tag{112}$$

Setting (112) equal to the New IS curve, substituting c_t^R for consumption demand and solving for i_t gives the reaction function under the assumption that the central bank takes rational expectations as given

$$i_t = \Omega_1^* y_{t-1} + \Omega_2^* E_t y_{t+1} + \Omega_3^* E_t y_{t+2} + \Omega_4^* \pi_{t-1} + \Omega_5^* E_t \pi_{t+1} + \Omega_6^* E_t \pi_{t+2}$$

$$+ \Omega_7^* E_t c_{t+1}^R + \Omega_8^* E_t c_{t+2}^R + \Omega_9^* e_t \tag{113}$$

with

$$\Omega_1^* = \frac{(1-\alpha)\theta^2(\Delta_5 + \Delta_1\kappa)}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{114}$$

$$\Omega_2^* = \frac{\Delta_6 + \Delta_8 + \alpha(\Delta_5 + \Delta_1\kappa)}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{115}$$

$$\Omega_3^* = \frac{\Delta_7}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{116}$$

$$\Omega_4^* = \frac{\Delta_4 + (1-\alpha)\theta^2(\beta\Delta_1 + \sigma(\Delta_5 + \Delta_1\kappa))}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{117}$$

$$\Omega_5^* = \frac{\alpha\beta\Delta_1 + \Delta_2 + \sigma\Delta_8 + \alpha\sigma(\Delta_5 + \Delta_1\kappa)}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{118}$$

$$\Omega_6^* = \frac{\Delta_3}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{119}$$

$$\Omega_7^* = \frac{\Delta_9}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{120}$$

$$\Omega_8^* = \frac{\Delta_{10}}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)} \tag{121}$$

$$\Omega_9^* = \frac{\Delta_1}{\sigma(\Delta_5 + \Delta_8 + \Delta_1\kappa)}. \tag{122}$$

The Ω-coefficients are expressed in terms of the targeting rule coefficients for simplicity. Writing them in terms of the deep model parameters would yield in part far to big expression.

B.4 Tables

Ω_x	$\alpha = 0.1$	$\alpha = 0.3$	$\alpha = 0.5$	$\alpha = 0.7$	$\alpha = 0.9$	$\alpha \to 1$ (RE)
y_{t-2}	0.001	0.004	0.01	0.018	0.029	0
y_{t-1}	0.115	0.052	-0.016	-0.085	-0.153	-0.139
$E_t y_{t+1}$	-0.08	-0.02	0.045	0.111	0.175	0.160
$E_t y_{t+2}$	-0.005	-0.008	-0.009	-0.007	-0.003	0
π_{t-3}	0.012	0.028	0.034	0.029	0.013	0
π_{t-2}	-0.609	-0.502	-0.38	-0.241	-0.085	0
π_{t-1}	2.747	2.082	1.420	0.792	0.238	0
$E_t \pi_{t+1}$	-0.482	0.115	0.697	1.229	1.639	1.851
$E_t \pi_{t+2}$	0.103	0.064	0.033	0.012	0.001	0
$E_t c^R_{t-2}$	0	-0.003	-0.009	-0.017	-0.029	0
$E_t c^R_{t-1}$	0.004	0.013	0.022	0.031	0.041	0
$E_t c^R_{t+1}$	-0.004	-0.012	-0.021	-0.031	-0.040	0
$E_t c^R_{t+2}$	0.003	0.007	0.008	0.007	0.003	0
e_t	1.275	1.227	1.154	1.056	0.932	0.859

Table 5: Values of reaction coefficients Ω_x^C in the interest rate rule (28) for different values of the share of rational forecasters α.

C   Implementation with model-consistent individual consumption expectations

The policy problem under commitment and the conventional Euler equation with model-consistent individual consumption expectations takes the following form:

$$
\begin{aligned}
\mathcal{L} = E_t \sum_{s=0}^{\infty} \beta^s \Bigg[&\Gamma_1 y_{t+s}^2 + \Gamma_2 \pi_{t+s}^2 + \Gamma_3 \pi_{t+s-1}^2 + \Gamma_4 (c_{t+s}^R)^2 \\
&+ \Gamma_5 y_{t+s} c_{t+s}^R + \Gamma_6 \pi_{t+s} c_{t+s}^R + \Gamma_7 \pi_{t+s-1} c_{t+s}^R + \Gamma_8 \pi_{t+s} \pi_{t+s-1} + \Gamma_9 \pi_{t+s} y_{t+s} + \Gamma_{10} \pi_{t+s-1} y_{t+s} \\
&+ \lambda_{1,t+s}[y_{t+s} - \alpha E_t y_{t+s+1} - (1-\alpha)\theta^2 y_{t+s-1} + \sigma[i_{t+s} - \alpha E_t \pi_{t+s+1} - (1-\alpha)\theta^2 \pi_{t+s-1}]] \\
&+ \lambda_{2,t+s}[\pi_{t+s} - \alpha\beta E_t \pi_{t+s+1} - (1-\alpha)\beta\theta^2 \pi_{t+s-1} - \kappa y_{t+s} - e_{t+s}] \\
&+ \lambda_{3,t+s}[c_{t+s}^R - E_t c_{t+s+1}^R + \sigma(i_{t+s} - E_t \pi_{t+s+1})] \Bigg].
\end{aligned} \tag{123}
$$

The first order conditions are

$$\frac{\partial \mathcal{L}}{\partial y_{t+s}} : E_t \Big\{ \beta^s [2\Gamma_1 y_{t+s} + \Gamma_5 c_{t+s}^R + \Gamma_9 \pi_{t+s} + \Gamma_{10}\pi_{t+s-1} + \lambda_{1,t+s} - \kappa\lambda_{2,t+s}]$$

$$- \beta^{s+1}(1-\alpha)\theta^2\lambda_{1,t+s+1} - \beta^{s-1}\alpha\lambda_{1,t+s-1} \Big\} \overset{!}{=} 0 \tag{124}$$

$$\frac{\partial \mathcal{L}}{\partial \pi_{t+s}} : E_t \Big\{ \beta^s [2\Gamma_2 \pi_{t+s} + \Gamma_6 c_{t+s}^R + \Gamma_8 \pi_{t+s-1} + \Gamma_9 y_{t+s} + \lambda_{2,t+s}]$$

$$+ \beta^{s+1}[2\Gamma_3 \pi_{t+s} + \Gamma_7 c_{t+s+1} + \Gamma_8 \pi_{t+s+1} + \Gamma_{10}y_{t+s+1} - (1-\alpha)\theta^2\sigma\lambda_{1,t+s+1} \tag{125}$$

$$- (1-\alpha)\beta\theta^2\lambda_{2,t+s+1}] - \beta^{s-1}[\alpha\sigma\lambda_{1,t+s-1} + \alpha\beta\lambda_{2,t+s-1} + \sigma\,\lambda_{3,t+s-1}] \Big\} \overset{!}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial c_{t+s}^R} : E_t \Big\{ \beta^s [2\Gamma_4 c_{t+s}^R + \Gamma_5 y_{t+s} + \Gamma_6 \pi_{t+s} + \Gamma_7 \pi_{t+s-1} \tag{126}$$

$$+ \lambda_{3,t+s}] - \beta^{s-1}\lambda_{3,t+s-1} \Big\} \overset{!}{=} 0$$

$$\frac{\partial \mathcal{L}}{\partial i_{t+s}} : E_t \Big\{ \beta^s \sigma \lambda_{1,t+s} + \beta^s \sigma \lambda_{3,t+s} \Big\} \overset{!}{=} 0. \tag{127}$$

Again, the index s can be dropped assuming commitment from a timeless perspective. Using $\lambda_{3,t} = -\lambda_{1,t}$ the FOCs can equivalently be written as

$$2\Gamma_1 y_t + \Gamma_5 c_t^R + \Gamma_9 \pi_t + \Gamma_{10}\pi_{t-1} + \lambda_{1,t} - \kappa\lambda_{2,t} - (1-\alpha)\beta\theta^2\lambda_{1,t+1} -$$

$$\beta^{-1}\alpha\lambda_{1,t-1} \overset{!}{=} 0 \tag{128}$$

$$2\Gamma_2 \pi_t + \Gamma_6 c_t^R + \Gamma_8 \pi_{t-1} + \Gamma_9 y_t + \lambda_{2,t}$$

$$+ 2\Gamma_3 \beta\pi_t + \Gamma_7 \beta c_{t+1} + \Gamma_8 \beta\pi_{t+1} + \Gamma_{10}\beta y_{t+1} - (1-\alpha)\beta\theta^2\sigma\lambda_{1,t+1} \tag{129}$$

$$- (1-\alpha)\beta^2\theta^2\lambda_{2,t+1} + \beta^{-1}(1-\alpha)\sigma\lambda_{1,t-1} - \alpha\lambda_{2,t-1} \overset{!}{=} 0$$

$$2\Gamma_4 c_t^R + \Gamma_5 y_t + \Gamma_6 \pi_t + \Gamma_7 \pi_{t-1} - \lambda_{1,t} + \beta^{-1}\lambda_{1,t-1} \overset{!}{=} 0.. \tag{130}$$

(130) can be used to replace $\lambda_{1,t-1}$ and $\lambda_{1,t+1}$ with $\lambda_{1,t}$ in (128). Then, solving (128) for $\lambda_{1,t}$ and inserting in (129) yields a second-order difference equation in $\lambda_{2,t}$. A solution to this equation can in principle be substituted back into the difference equation, which would give a targeting rule. However, this solution is fairly complicated in which some parameter terms exponentially depend on time. The solution is

available upon request. The resulting targeting rule and, hence, a reaction function would also be of such a complicated form where parameters exponentially depend on time. Consequently, no meaningful interest rate rule under commitment can be derived in this case.