UNIVERSITÄT BIELEFELD

Faculty of Business Administration and Economics

Working Papers in Economics and Management

No. 06-2023
Sept. 2023

Macroeconomic stability of price level targeting in a model of heterogeneous expectations

Tomasz Makarewicz

Bielefeld University
P.O. Box 10 01 31
33501 Bielefeld – Germany

ISSN 2196-2723

→ www.wiwi.uni-bielefeld.de

Macroeconomic stability of price level targeting in a model of heterogeneous expectations

Tomasz Makarewicz[1]

[1]Universität Bielefeld

September 20, 2023

Abstract

This paper studies the efficacy of Price Level Targeting (PLT) under the Heuristic Switching Model (HSM). PLT has been considered as an alternative to the traditional inflation targeting in the aftermath of the 2008 Great Recession. However, this policy was not thoroughly tested in real life, whereas the experimental evidence remains contradictory. This study contributes to the discussion by extending the HSM by Hommes and Lustenhouwer (2019) with the PLT Taylor rule.

The properties of the PLT rule under the behavioral expectations are mixed. On the one hand, the fundamental steady state becomes the unique fixed point of the system in this setup. However, the stability of this fixed point is highly sensitive to the calibration of the underlying New Keynesian economy, and the exact specification of the HSM. As a result, policy makers can use this policy only if they understand well the way in which the economic agents form their expectations.

JEL codes: E32, E52, C62.

Keywords: monetary policy, price level targeting, behavioral expectations, learning, heuristic switching model.

1 Introduction

This paper studies the efficacy of the Price Level Targeting (henceforth PLT) monetary policy in a Brock-Hommes model of heterogeneous expectations (Brock and Hommes, 1998). PLT was proposed as an alternative to the inflation targeting in the wake of the Great Recession of 2008, when the conventional policies seemed to be failing: the world economy went through a surprisingly deep and long downturn period, even though major central banks were aggressively cutting their interest rates down to the so called Zero-Lower Bound. To complicate the matter, the standard macroeconomic framework was unable to convincingly explain the financial mechanism behind the recession, nor to predict its scale (Lindé et al., 2016; Trichet, 2018). These events spurred two controversial and related debates: how to improve macroeconomic models and whether to abandon the traditional policy in favor of some unconventional measures, such as the aforementioned PLT (Evans, 2012).

Under PLT and in contrast with the conventional inflation targeting, central bank reacts to a one-time positive (negative) inflationary shock with a spell of high (low) interest rates, until the excess (deficient) price level rolls back to the intended trajectory.[1] This serves as a powerful anchor under Rational Expectations (Covas and Zhang, 2010) and adaptive learning (Honkapohja and Mitra, 2020). As a result, PLT can, under correct circumstances, outperform inflation targeting in terms of realized volatility of the business cycle (Ambler, 2009; Svensson, 1996; Vestin, 2006). However, PLT has only once been put into practice, in Sweden in 1930ties (Berg and Jonung, 1999). Its introduction would be, as Evans (2012) framed it, "a leap of faith" based on untested theory – which itself became a subject of critique by the policy makers and economic literature.

One obvious and immediate reaction by macroeconomists to the 2008 crisis was to put a higher emphasis the role of the financial frictions and banking sector in the DSGE framework (among a plethora of examples, see Brázdik and Marsal, 2011; Brzoza-Brzezina and Kolasa, 2013; Gambacorta and Signoretti, 2014; Gerali et al., 2010; Merola, 2015). However, some critique cut deeper into the conventional macroeconomic theory (Evans and Honkapohja, 2012; Levine, 2020; Woodford, 2012). Before the Great Recession, the standard macroeconomic

[1] See Jääskelä (2005) for an introduction into PLT and Bernanke and Woodford (2005) for a thorough discussion of the inflation targeting.

modeling was based on the New Keynesian (henceforth NK) framework with Rational Expectations: economic actors process local information and events in a timely fashion through some "behind the curtains" learning process, which leads them to coordinate on "as-if" model-consistent equilibria (Woodford, 2013). After the Great Recession, this assumption came into scrutiny for two important reasons: it faces severe empirical challenges and goes against the consensus within the literature on learning models (Kirman, 2014; Poledna et al., 2023).

A direct test of the RE hypothesis on market data is difficult, since the researcher may be unable to observe many relevant variables that concern the market structure, information flows and attitude of the agents. Individual beliefs, in particular, are notoriously difficult to measure and thus to interpret (see Armantier et al., 2013, for an example about inflation surveys). The learning literature has thus focused on laboratory experiments to study empirical expectations and test the RE hypothesis. In such a laboratory experiment it is relatively simple to set up the experimental economy in a manner that offers a straightforward theoretical prediction, and the researcher can also control the information flow and observe any relevant beliefs and decisions of the subjects (Colasante et al., 2017; Plott and Sunder, 1982; Weizsäcker, 2010). Experimental studies on forecasting cast serious doubts about the validity of the RE framework (Nuzzo and Morone, 2017). The most well known strand of this literature consists of the so called Learning-to-Forecast experiments (henceforth LtF experiments; see Hommes, 2011, for a comprehensive introduction), in which the value of the relevant economic variable depends on subjects' forecasts through some predetermined economic model like a DSGE economy (Assenza et al., 2014). LtF experiments demonstrate that subjects may converge to the rational (model-consistent) forecast, but this is by no means a necessary outcome. Instead, one can often find a large degree of heterogeneity, backward-looking (myopic) expectations and non-fundamental dynamics (Anufriev et al., 2022).

The experimental dynamics depend on the feedback between forecasts and the realized variables, with two important factors highlighted by the literature (Bao et al., 2021). Firstly, the more complicated the underlying experimental economy is, the more difficult it is for the subjects to form accurate forecasts or coordinate on the RE equilibria (see Kopányi et al., 2019, for a recent example). Secondly, the sign of the feedback matters (Bao and Hommes, 2019; Heemeijer et al., 2009). If the effect of the average forecast on the realized variable is positive (ie. a positive partial derivative in the underlying model), subjects tend to steer away from the RE equilibria and coordinate on trend chasing dynamics, where waves of optimism (pessimism) result in an increase (decrease) of the forecasted variable, feeding back into the initial optimism (pessimism). Interestingly, in such a scenario subjects tend to be quite well coordinated, and if the feedback itself is not overly complicated, their expectations could be surprisingly accurate (Hommes, 2011).

The positive feedback is important in the context of monetary policy, for example inflation and output gap in standard DSGE models depend positively on their respective forecasts. Experiments on monetary policy often result in off-equilibrium business cycles that the monetary authorities cannot easily control (among growing number of examples, see Ahrens et al., 2017; Arifovic and Petersen, 2017; Assenza et al., 2021; Galí et al., 2021; Hommes et al., 2019; Kryvtsov and Petersen, 2013, 2021; Lustenhouwer and Salle, 2022; Mauersberger, 2021). Moreover, Hommes and Makarewicz (2021) and Kostyshyna et al. (2022) have directly tested PLT in two parallel experiments, with unclear results. Both studies highlight a confluence of the aforementioned results: oscillatory dynamics remain prevalent, the central bank may find it difficult to establish its credibility, while the subjects are often unable to effectively use all the relevant information (including reported price level deviations). This suggests that, in contrast with the traditional inflation targeting, PLT can anchor subject expectations in the RE solution only if the PLT rule is surprisingly strict. It should be finally noted that those experiments suggest that even if the economic agents eventually learn the rational response to PLT, it may take them a substantial spell of time, which may be a practical limitation for policy makers. These insights were further confirmed by Salle (2021).

How relevant are these results for policy makers? A growing empirical literature demonstrates that the LtF experiments, with their non-rational outcomes, have at least some degree of external validity (Bao et al., 2021). For example, Cornand and Hubert (2020) show that actual market forecasters and experimental subjects have similar forecasting strategies and errors. This suggests that it may be prudent to study policy choices under behavioral models of expectations, even if for the sake of safety and robustness. This is particularly important if we take the “as-if” approach on its face value: even if the economic agents eventually do converge to the RE solution, larger shocks or changes to the monetary policy may cause transitory non-RE dynamics, which are significant enough that they need to be accounted for by central banks. Such a model uncertainty is an obvious issue for the PLT policy, since neither central banks nor the markets have any substantial experience with dealing with or communicating it (*cf.* Kryvtsov and Petersen, 2021).

The most significant challenge to the non-rational approach is the famous "wilderness of bounded rationality" (Sargent, 2008; Sims, 1980): there are many models of individual learning, which have significantly different dynamic properties and policy implications. A solution to this issue is offered by the LtF literature, which identifies as the most promising alternative to the RE framework the so called Heuristic Switching Model (henceforth HSM; see Brock and Hommes, 1998, 2001, for an introduction). The idea behind HSM is that the agents rely on simple rules of thumb, heuristics like naive or trend following expectations. The agents are intelligent in *how* they chose those heuristics, namely they tend to switch to rules with a better forecasting performance in the past. For instance, if the economy starts to stabilize after a spell of oscillations, they could switch from a chartist to a fundamental rule, and vice versa if the economy becomes unstable again (Anufriev and Hommes, 2012b). HSM has been successfully applied to a plethora of experimental data (see Anufriev and Hommes, 2012a, for a representative example), as well as a number of empirical data sets, ranging from financial markets (see Boswijk et al. (2007); Franke and Westerhoff (2012); Kukacka and Barunik (2017) for typical examples and Lux and Zwinkels (2018) for a thorough literature review), through housing market (eg. Bolt et al., 2019; Kouwenberg and Zwinkels, 2014), to macroeconomic applications (among many examples, see Cornea-Madeira et al., 2019; Grazzini et al., 2017; Jang and Sacht, 2021; Kukacka et al., 2018). Despite the relative novelty of this literature, it shows that policy makers should at least consider a possibility that the economic agents have myopic and heterogeneous expectations (Deak et al., 2020; Jump and Levine, 2019).

This paper adds to the macroeconomic literature on HSM (Massaro, 2013), by investigating the efficacy of the PLT rule under behavioral expectations. I investigate a simple DSGE economy in the spirit of Hommes and Lustenhouwer (2019), with a PLT rule, and compare its dynamic properties under four models of expectations: i) RE, ii) naive, iii) two-type HSM as in Hommes and Lustenhouwer (2019) (where agents can switch between the naive and fundamental forecast) and iv) four-type HSM (with fundamental, adaptive and two types of chartist rules). The first variant of the HSM is popular in the behavioral macroeconomic literature, while the second is taken from Anufriev and Hommes (2012b) and has been shown as a solid representation of the subjects' behavior in LtF experiments.

This study yields three important results. Firstly, under the PLT rule and all the considered expectation models, the fundamental equilibrium becomes the unique steady state of the economy. This contrast the inflation targeting rule, under which behavioral expectations can yield non-fundamental fixed points.

Secondly, stability conditions for the PLT rule are sensitive to the calibration of the economy and, to complicate the issue, are different between the behavioral and rational models of expectations. Furthermore, local stability does not guarantee the global one, and hence a poor choice of the policy rule parameters may lead to a false sense of stability until a larger external shock hits the economy.

Thirdly, the study confirms that the strength of the trend following rules has an overall negative impact on the stability properties of the model under the behavioral expectations. This suggest that if we allowed the agents to actually learn the degree of the trend following (see Anufriev et al., 2019), the dynamic properties of the PLT rule may be even more obfuscated.

How should policy makers interpret these results? This study suggests that PLT is associated with a high degree of model uncertainty. On the other hand, the traditional inflation targeting has much more robust properties between different models of expectations and in laboratory experiments. As a result, policy makers may find it easier to use the inflation targeting as a standard policy, and search for alternative crisis measures elsewhere. Finally, these results explain the mixed findings of the experimental studies on PLT – they were to be expected, as the rule is sensitive to calibration and thus likely to the experimental design as well.

This paper is organized in the following fashion. The baseline NK economy and the PLT rule are introduced in Section 2. Section 3 presents the four models of expectation formation (RE, naive and two variants of the HSM). The existence and uniqueness of the fundamental steady state is discussed in Section 4, whereas Section 5 deals with the stability properties of this fixed point. Section 6 discusses the role of the intensity of choice in the model. The last Section summarizes the paper. Finally, Appendices A through C contain technical derivations of the model and Appendix D displays additional results for contemporaneous and backward-looking PLT rules.

2 New Keynesian economy

Consider standard 3-equation New Keynesian economy with Euler equation and Philips Curve, which will be henceforth denotes as the NK model. We will focus on its deterministic skeleton, assuming no exogenous shocks, which implies no dynamics under the RE solution. Under log-linearization, the model is given by:

$$y_t = y^e_{t+1} - \sigma^{-1}\left(r_t - \pi^e_{t+1}\right) \tag{1}$$

$$\pi_t = \beta\pi^e_{t+1} + \kappa y_t = \kappa y^e_{t+1} - \mu r_t + \nu\pi^e_{t+1} \tag{2}$$

where π, y and r denote deviation of inflation, output gap and interest rate from the target π^* and the fundamental (full employment steady state) levels Y^* and R^* respectively, β represents consumers' discount factor, σ denotes consumers' elasticity of substitution, and κ, $\mu = \kappa/\sigma$ and $\nu = \beta + \kappa/\sigma$ are composite coefficients based on the parameters of the model. Finally, operator x^e_{t+1} denotes expectations of variable x_{t+1} at period t, which may be rational (with $x^e_{t+1} = E_t\{x_{t+1}\}$) or otherwise behavioral.

The realized output gap depends on the interest rate of the central bank. In this paper, agents have to learn to form expectations, which may lead them to some non-rational (model inconsistent) beliefs. The central bank realizes that and hence focuses on the so called expectations management, ie. it reacts to market expectations for the next period.[2] The traditional focus on inflation targeting implies the following Taylor rule:

$$R_t = 1 + (R^* - 1)\left(\frac{\Pi^e_{t+1}}{\Pi^*}\right)^{\frac{\phi_p R^*}{R^*-1}}\left(\frac{Y^e_{t+1}}{Y^*}\right)^{\frac{\phi_y R^*}{R^*-1}},$$

where $\Pi_t = P_t/P_{t-1}$ denotes inflation index and $\Pi^* = 1 + \pi^*$ is the gross inflation target. This

[2] Alternatively, the central bank could focus on the contemporaneous or lagged variables, see Appendix A and following for details. These two model variants have largely similar qualitative results, see Appendix E.

rule can be log-linearized into much simpler and well known form

$$r_t = \phi_p \pi^e_{t+1} + \phi_y y^e_{t+1}. \tag{3}$$

Under Price Level Targeting (henceforth PLT), the target is based on deviation from what the monetary authority defined as its intended price path, given by

$$P^*_t = \Pi^* P^*_{t-1}$$

for initial price level P^*_0.[3] In other words, under the PLT the central banks commits to a whole *trajectory* of prices, instead of a period-to-period inflation rate. Denote price level deviation as

$$D_t = \frac{P_t}{P^*_t}. = \frac{\Pi_t D_{t-1}}{\Pi^*}$$

Hence the monetary policy rule becomes

$$R_t - 1 = (R^* - 1)\left(D^e_{t+1}\right)^{\frac{\phi_p R^*}{R^*-1}} \left(\frac{Y^e_{t+1}}{Y^*}\right)^{\frac{\phi_y R^*}{R^*-1}}, \tag{4}$$

which can be log-linearized into

$$r_t = \phi_p d^e_{t+1} + \phi_y y^e_{t+1}, \tag{5}$$

where the price deviation under log-linearization becomes

$$d_t = d_{t-1} + \pi_t, \tag{6}$$

$$d^e_{t+1} = d_t + \pi^e_{t+1} = d_{t-1} + \pi_t + \pi^e_{t+1}. \tag{7}$$

Substituting the price deviation into equations (1–2), we obtain the final version of the NK

[3] In the simulations we normalized $P^*_0 = 1$. The monetary authority could in principle have time-varying intended slope with $\pi^*_t \neq \pi^*_s$ for some $t \neq s$. We will leave this issue for future research.

economy:

$$\pi_t = \frac{\nu - \mu\phi_p}{1+\mu\phi_p}\pi^e_{t+1} + \frac{\kappa - \mu\phi_y}{1+\mu\phi_p}y^e_{t+1} - \frac{\mu\phi_p}{1+\mu\phi_p}d_{t-1},$$
$$y_t = \frac{1-\phi_p(1+\beta)}{\sigma(1+\mu\phi_p)}\pi^e_{t+1} + \frac{\sigma - \phi_y}{\sigma(1+\mu\phi_p)}y^e_{t+1} - \frac{\phi_p}{\sigma(1+\mu\phi_p)}d_{t-1}, \tag{8}$$

together with equations (6) and (7). Remark that unlike under the benchmark inflation targeting, this economy is only partially forward-looking, since the PLT rule introduces additional "stickiness" into the dynamics – central bank is trying to clear the whole accumulated history of price level deviations, instead of just the most recent excess inflation. This serves as a successful anchor the agents' expectations under RE.

3 Models of rational and behavioral expectations

In this Section we will discuss the models of expectations that will be later analyzed: two variants of the behavioral model, as well as two benchmarks, naive and Rational Expectations.

3.1 Rational and naive benchmarks

Rational Expectations are the most sophisticated form of forecasting. The NK model (8) is deterministic, thus RE imply that agents have perfect foresight and – in the absence of any other shocks – immediately converge to the fundamental steady state. Thus, under RE the model has no real dynamics.

Naive expectations, defined as $\left(\pi^e_{t+1}, y^e_{t+1}\right)' = (\pi_{t-1}, y_{t-1})'$, span the other extreme of the level of agents' sophistication in this research. They are often used as a benchmark in experimental literature, since they provide a basic guesstimate on how experimental subjects may react to the experimental economy, before any session is actually run. On the other hand, subjects do exhibit smarter behavior in a typical experimental session than the naive model predicts (Hommes and Makarewicz, 2021).

3.2 Heuristic Switching Model

3.2.1 General setup

The premise behind behavioral models of expectations is that due to some informational, computational or cognitive limitations, agents may be unable to formulate RE, but they are also too sophisticated to use a simplistic rule like the naive benchmark (Anufriev and Hommes, 2012b). Heuristic Switching Model (henceforth HSM) offers a reasonable middle ground between the RE and naive extremes, and also has been validated by the experimental and empirical literature.

The basic idea of HSM is that agents consider a whole set of heuristics, where each is relatively simple, like the naive forecast, trend following expectations or some fundamental rule. On the other hand, agents are sophisticated in how they choose their heuristics: they learn to focus on those with a better past forecasting performance. For instance, if the economy is characterized by a persistent cycle of up- and downturns, the agents may try to extrapolate recent trends, and then will switch to a naive rule instead, if the system stabilizes.

Formally, consider a set $h \in H$ of heuristics

$$\pi^e_{h,t+1} = \pi^e_h\left(\pi^{t-1}, y^{t-1}, \pi^{e,t}, y^{e,t}\right),$$
$$y^e_{h,t+1} = y^e_h\left(\pi^{t-1}, y^{t-1}, \pi^{e,t}, y^{e,t}\right),$$

where the superscript t denotes the whole history of the variable up to period t. Remark that these heuristics have to be backward-looking and agents can use only past data. For instance $\left(\pi^e_{h,t+1}, y^e_{h,t+1}\right)' = (\pi_{t-1} + \tau(\pi_{t-1} - \pi_{t-2}), y_{t-1} + \tau(y_{t-1} - y_{t-2}))'$ is a pair of naive expectations for $\tau = 0$ and trend following rules for $\tau > 0$.

The attraction of heuristic h for variable $x \in \{\pi, y\}$ is typically defined by an AR(1) process of its squared forecasting error

$$U^{x,h}_t = \rho U^{x,h}_{t-1} - \left(x^e_{h,t-1} - x_{t-1}\right)^2 \leqslant 0, \tag{9}$$

where $\rho \in [0, 1)$ is interpreted as the memory parameter. The higher (ie. the closer to zero from below) is the attraction, the lower the perceived past forecasting error of the heuristic is, and thus the higher its role in the average expectations is, where the representative forecast is defined as

$$x_{t+1}^{e} = \sum_{h \in H} n_t^{x,h} x_{h,t+1}^{e}, \tag{10}$$

where

$$n_t^{x,h} = \frac{\exp\left(\gamma U_t^{x,h}\right)}{\sum_{k \in H} \exp\left(\gamma U_t^{x,k}\right)} \tag{11}$$

is the (time-variable) weight of heuristic h and $\gamma \geqslant 0$ denotes intensity of choice. The model can be interpreted in two ways: agents have homogeneous expectations and assign higher weights to more successful heuristics; or each agent picks one heuristic with probabilities defined by (11), which by the Law of Large Number implies that the representative agent tends to switch to more successful heuristics (hence the name of the model).

The intensity of choice parameter $\gamma \in [0, \infty^{+}[$ has an important interpretation. If it is equal to zero with $\gamma = 0$, there is no switching and each heuristic gains the same weight of $1/|H|$. Conversely, if it diverges with $\gamma \to \infty$, agents immediately switch to the best performing heuristic. In the intermediary case, the switching is more gradual and one heuristic will dominate the expectations only if its performance is visibly the best. In many application of the HSM, composition of the set of steady states and their stability depend on the value of γ, but – as will become apparent later – this is not the case for the PLT model.

Another important aspect of the HSM is the choice of the heuristic set H. The exact behavior of the HSM can change drastically due to inclusion of one specific heuristic, for instance trend heuristics allow for a strong momentum in forecasts. Experimental and empirical work suggest that among the popular heuristics are naive, trend following, adaptive and fundamental expectations (Hommes, 2011).

In this paper I will consider two important specifications:

2-type HSM with naive and fundamental heuristics

$$\textbf{Naive: } x^N_{t+1} = x_{t-1}, \tag{12}$$

$$\textbf{Fundamental: } x^F_{t+1} = x^f = 0. \tag{13}$$

4-type HSM with adaptive, weak and strong trend following, and fundamental heuristics

$$\textbf{Adaptive: } x^{AD}_{t+1} = \alpha x_{t-1} + (1-\alpha)x^{AD}_t = x^{AD}_t + \alpha\left(x_{t-1} - x^{AD}_t\right), \tag{14}$$

$$\textbf{Weak Trend: } x^{WT}_{t+1} = x_{t-1} + \tau_1\left(x_{t-1} - x_{t-2}\right), \tag{15}$$

$$\textbf{Strong trend: } x^{ST}_{t+1} = x_{t-1} + \tau_2\left(x_{t-1} - x_{t-2}\right), \tag{16}$$

$$\textbf{Fundamental: } x^{FD}_{t+1} = x^f = 0 \tag{17}$$

for some $\alpha \in (0,1)$ and $\tau_2 \geqslant \tau_1 \geqslant 0$ (and typically $\tau_1 > 0$).

3.2.2 2-type model

The 2-type HSM is a common benchmark for policy studies. The fundamental steady state coincides with the central bank's target, hence popularity of the fundamental rule (measured with $n^{x,F}_t$) is often interpreted as the endogenously acquired credibility of the monetary authority. Conversely, if the central bank fails to build up its credibility, agents are more likely to switch to the simple backward-looking naive rule. Hommes and Lustenhouwer (2019) explore this model for the inflation target policy in the same NK model as is presented in this paper.

In the 2-type HSM, there are only two simple heuristics, which allows for some simplification of the notation (cf. Hommes and Lustenhouwer, 2019). Denote the share of fundamental forecasters of inflation and output gap in period t as n^π_t and n^y_t respectively. By definition the shares of naive forecasters are then given by $1 - n^\pi_t$ and $1 - n^y_t$. It is analytically convenient to focus on the index

$$m^x_t = 2n^x_t - 1 \in [-1, 1], \tag{18}$$

for $x \in \{\pi, y\}$, which represent the degree to which the agents find the target of the central bank to be credible: when m^x_t approaches 1 (-1), the agents perfectly trust (distrust) the monetary policy in terms of variable x.

For the sake of analytical tractability suppose that the agents judge the two rules based only on their most recent performance, with memory switched off and $\rho = 0$. Thus, the fitness measures $U(\cdot)$ for variable $x \in \{\pi, y\}$ become

$$U_t^{x,F} = -\left(x^* - x_{t-1}\right)^2 = -x_{t-1}^2, \tag{19}$$

$$U_t^{x,N} = -\left(x_{t-3} - x_{t-1}\right)^2 = -x_{t-3}^2 + 2x_{t-3}x_{t-1} - x_{t-1}^2 \tag{20}$$

for the fundamental (denoted with F) and naive (denoted with N) forecasts. Next, the agent switch between the two rules according to logit transformation

$$\begin{aligned} m_t^x =& 2\frac{\exp\left(\gamma U_t^{x,F}\right)}{\exp\left(\gamma U_t^{x,F}\right) + \exp\left(\gamma U_t^{x,N}\right)} - 1 = \tanh\left[0.5\gamma\left(U_t^{x,F} - U_t^{x,N}\right)\right] \\ =& \tanh\left[\gamma\left(0.5x_{t-3}^2 - x_{t-1}x_{t-3}\right)\right]. \end{aligned} \tag{21}$$

Together this yields

$$\pi_{t+1}^e = 0.5\left(1 - m_t^\pi\right)\pi_{t-1} = 0.5\pi_{t-1} + 0.5\tanh\left[\gamma\left(\pi_{t-1}\pi_{t-3} - 0.5\pi_{t-3}^2\right)\right]\pi_{t-1} \tag{22}$$

$$y_{t+1}^e = 0.5\left(1 - m_t^y\right)y_{t-1} = 0.5y_{t-1} + 0.5\tanh\left[\gamma\left(y_{t-1}y_{t-3} - 0.5y_{t-3}^2\right)\right]y_{t-1}. \tag{23}$$

These 2-type expectations can be now substituted into the NK model (8). In total we obtain a 7D system with the state vector $(\pi_t, \pi_{t-1}, \pi_{t-2}, y_t, y_{t-1}, y_{t-2}, d_t)'$, where the extra variable lags are necessary for the $U(\cdot)$ fitness measures.

3.2.3 4-type model

The 4-type HSM is more general and includes the important trend following heuristics (with two possible degrees to which the agents may want to chase the observed trends), as well as adaptive rule than can stabilize the forecast. This model is often found to be a good representation of the laboratory experiments, and may be considered as the most realistic behavioral model of expectations in this study (see Anufriev and Hommes, 2012b, for an experimental application).

In the 4-type model, the average forecast $x^e_{t+1} \in \left\{\pi^e_{t+1}, y^e_{t+1}\right\}$ is given by

$$
\begin{aligned}
x^e_{t+1} =& n^{x,AD}_t \left(\alpha x_{t-1} + (1-\alpha) x^{AD}_t\right) + n^{x,WT}_t \left(x_{t-1} + \tau_1 \Delta x_{t-1}\right) + n^{x,ST}_t \left(x_{t-1} + \tau_2 \Delta x_{t-1}\right) \\
=& \left(\alpha n^{x,AD}_t + (1+\tau_1) n^{x,WT}_t + (1+\tau_2) n^{x,ST}_t\right) x_{t-1} - (\tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t) x_{t-2} \\
&+ (1-\alpha) n^{x,AD}_t x^{AD}_t, \qquad (24)
\end{aligned}
$$

where the weights $n^h_t = \exp(\gamma U^h_t) / \sum_k \exp(\gamma u^k_t)$ for $h, k \in \{AD, WT, ST, FD\}$ are based on

$$U^{x,AD}_t = \rho U^{x,AD}_{t-1} - \left(x^{AD}_t - x_{t-1}\right)^2, \qquad (25)$$

$$U^{x,WT}_t = \rho U^{x,WT}_{t-1} - \left(x_{t-1} - x_{t-3} - \tau_1 \left(x_{t-3} - x_{t-4}\right)\right)^2, \qquad (26)$$

$$U^{x,ST}_t = \rho U^{x,ST}_{t-1} - \left(x_{t-1} - x_{t-3} - \tau_2 \left(x_{t-3} - x_{t-4}\right)\right)^2, \qquad (27)$$

$$U^{x,FD}_t = \rho U^{x,FD}_{t-1} - x^2_{t-1}. \qquad (28)$$

Remark that unlike for the case of the 2-type HSM, we will now consider switching with potentially persistent memory $\rho \geqslant 0$. The term $\tau^x_t \equiv \tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t$ can be interpreted as a realized chartist sentiment for variable $x \in \{\pi, y\}$.

In comparison with the 2-type model, the 4-type HSM is considerably larger. The full state space is spanned by the 19D vector

$$
\begin{aligned}
&\left(\pi_t, \pi_{t-1}, \pi_{t-2}, \pi_{t-3}, y_t, y_{t-1}, y_{t-2}, y_{t-3}, d_t, \pi^{AD}_{t+1}, y^{AD}_{t+1},\right. \\
&\left. U^{\pi,AD}_t, U^{\pi,WT}_t, U^{\pi,ST}_t, U^{\pi,FD}_t, U^{y,AD}_t, U^{y,WT}_t, U^{y,ST}_t, U^{y,FD}_t\right)',
\end{aligned}
$$

where the additional dimensions cover i) the three lags of inflation and output gap, necessary for the trend following rules and their attractions, ii) two adaptive expectations lags, and iii) the memory of the four attraction measures. The full display of the model can be found in Appendix D in equation set (D.1).

4 Steady state

4.1 Steady state consistency and uniqueness

The PLT rule provides the NK model (8) with an interesting property, namely that the fundamental steady state is in fact the unique fixed point for a wide range of models of expectation formation, including RE and popular variants of HSM.

Lemma 1. *Any steady states requires that the inflation is zero with* $\pi_t = \pi_{t-1} = \ldots = \pi = 0$.

Proof. In any steady state, equation (6) becomes

$$d = d + \pi \quad \rightarrow \quad \pi = 0.$$

□

Remark that Lemma 1 does not depend on any other assumption regarding the calibration, interest rate or expectations. It has a straightforward interpretation – any non-fundamental (excess to the target) inflation will cause the price level deviation d_t to accumulate, forcing the central bank to strengthen its reaction ever more. In spite of its simplicity, Lemma 1 has an important consequence.

To start with, define a model of expectations $\hat{E}_t(\cdot)$ to be *steady state consistent* iff in any steady state that it admits, sans shocks expectations must coincide with realized variables, that is $\hat{E}(x) = x$. In the case of the NK model (8), this condition implies that $\pi^e = \pi$ and $y^e = y$.

Remark that in deterministic economies Rational Expectations simplify to perfect foresight with $\pi^e_{t+1} = \pi_{t+1}$ and $y^e_{t+1} = y_{t+1}$ for any period t, and thus also in any steady state. The concept of the steady state consistency is thus weaker than RE: it allows the agents to commit forecasting errors outside of a steady state, but requires them to achieve "as-if" RE once the economy converges (compare with the notion of eductive stability, for example in Evans et al., 2019). An obvious example of such a model are naive expectations.

Proposition 1. *Consider the NK model with the PLT rule* (8) *and some mechanism of expectation formation* $\left(\pi^e_{t+1}, y^e_{t+1}\right)' = \hat{E}_t(\cdot)$. *If* $\hat{E}(\cdot)_t$ *is steady state consistent, then the fundamental* *steady state $\pi = y = d = 0$ is the unique fixed point of the model.*

Proof. By Lemma 1 $\pi = 0$ and hence by the steady state consistency $\pi^e = 0$, which by the Philips Curve equation (2) implies that $y = 0$ and hence $y^e = 0$. It follows from the Euler equation (1) that the steady state interest rate is also equal to $r = 0$, which together implies that $d = 0$. □

This Proposition has consequences for all the four models of expectations that are considered in this study. As mentioned above, RE and naive expectations are trivially steady state consistent and thus admit the fundamental steady state as their unique fixed point.

4.2 HSM and steady state consistency

HSM is not steady state consistent in general, since it often allows for "biased" fixed points (see Brock and Hommes, 1998, for examples). Nevertheless, the PLT rule can still restrain the set of HSM's attractors in an important fashion.

Proposition 2. *Consider the NK model with the PLT rule* (8)*, and expectations which are formed by the HSM with a set of heuristics H. Suppose that H is such that it i) contains the fundamental rule $F:\ x^F_{t+1} = 0$ and ii) any other rule $h \in H\backslash\{F\}$ is steady state consistent, for $x \in \{\pi, y\}$. Then the fundamental steady state is the unique fixed point of the model.*

Proof. Remark that if an expectation model is individually steady state consistent, it remains so as a part of the heuristic set H of any HSM. Hence, in any steady state in the NK model with the PLT rule, the average forecast for variable $x \in \{\pi, y\}$ is given by

$$x^e = n^{x,F} \times x^F + \sum_{h \in H\backslash\{F\}} n^{x,h} \times x^h = \left(\sum_{h \in H\backslash\{F\}} n^{x,h} \right) x = \left(1 - n^F\right) x.$$

By Lemma 1, in any steady state $\pi = 0$, hence $\pi^h = 0$ for any h and $\pi^e = 0$. The Philips Curve (2) implies thus that $y = 0$, hence for $y^h = 0$ any h and so $y^e = 0$ and $r = d = 0$. □

The three non-fundamental heuristics from the 2- and 4-type HSM – adaptive, naive and trend rules – are all individually steady state consistent.[4] By Proposition 4.2, the 2- and

[4]In particular, in any steady state: $x^N = x$, $x^{VT} = x + \tau_V(x - x) = x$ for $V \in \{W, S\}$ and $x^{AD} = x^{AD} + \alpha\left(x - x^{AD}\right)$, hence $x^{AD} = x$.

4-type HSM variants behave as if they were steady state consistent under the PLT rule and admit only the fundamental steady state as their fixed points. This is in contrast with the standard inflation targeting, which yields pitchfork bifurcations for certain model and policy parameter constellations (see Hommes and Lustenhouwer, 2019, for details). Hence, in the next Section we will focus on the local stability of the fundamental steady state.

5 Stability results

5.1 Stability conditions

Appendix B shows the model in the form necessary for evaluating the Blanchard-Kahn condition under RE, which yields a cubic characteristic polynomial, itself a complicated function of the model's parameters. The same holds for the Jacobian matrices under naive expectations and both variants of the HSM (in which case the dimensionality of the system increases). Therefore, the following analysis is based on numerical evaluation of the respective Jacobian matrices.[5]

For the two variants of the HSM, the following two Propositions summarize their respective stability conditions.

Proposition 3. *In the NK economy* (8) *with expectations formed by the 2-type HSM (with fundamentalists and naive agents) with zero memory* $\rho = 0$*, the fundamental steady state (the unique fixed point) is locally stable if and only if matrix*

$$J_{2T} \equiv \begin{pmatrix} \frac{\nu-\mu\phi_p}{2(1+\mu\phi_p)} & \frac{\kappa-\mu\phi_y}{2(1+\mu\phi_p)} & -\frac{\mu\phi_p}{1+\mu\phi_p} \\ \frac{1-\phi_p(1+\beta)}{2\sigma(1+\mu\phi_p)} & \frac{\sigma-\phi_y}{2\sigma(1+\mu\phi_p)} & \frac{\phi_p}{\sigma(1+\mu\phi_p)} \\ \frac{\nu-\mu\phi_p}{2(1+\mu\phi_p)} & \frac{\kappa-\mu\phi_y}{2(1+\mu\phi_p)} & \frac{1}{1+\mu\phi_p} \end{pmatrix}$$

has eigenvalues within the unit circle.

The proof can be found in the Appendix C. Remark that this 7D system has four trivial eigenvalues $\lambda = 0$ at the steady state, hence only the reduced matrix J_{2T} is relevant. Interest-

[5] The code for the numerical analysis was written in Ox matrix language by Doornik (2009) and is available on request.

ingly, eigenvalues of this matrix do not depend on the intensity of choice parameter γ, which is an uncommon feature for a HSM.

Proposition 4. *In the NK economy* (8) *with expectations formed by the 4-type HSM (with fundamentalists, adaptive and trend chasing agents) with non-negative memory $\rho \geq 0$, the fundamental steady state (the unique fixed point) is locally stable if and only if matrix J_{4T} (presented in Appendix D in equation* (D.2)*) has eigenvalues within the unit circle.*

See Appendix D for the full proof. The 4-type model has 19 dimensions, but the fundamental steady state has four stable eigenvalues $\lambda = 0$ and further eight stable $\lambda = \rho$, hence its stability depends on the remaining seven eigenvalues of the reduced matrix J_{4T}. Like for the 2-type HSM, stability of the 4-type model does not depend on the intensity of choice parameter γ. Furthermore, that stability depends on the sum of the two trend coefficients $\tau = \tau_1 + \tau_2$, but not on their separate values.

5.2 Model comparison

The first stability exercise is conducted, following Hommes and Lustenhouwer (2019), for three popular DSGE model calibrations, which were introduced by Clarida et al. (2000), McCallum and Nelson (1999) and Woodford (1999); henceforth abbreviated as **CGG**, **MN** and **W** respectively (see Table 1). In this exercise we will focus on the stability of different monetary policy portfolios, as measured by the strength of reaction to the price deviation and output, over a fine grid over policy parameter space $\phi_p, \phi_y \in [0, 2]$.[6] Higher parametrizations were disregarded as politically impractical. We will compare the performance of the four models, where the 4-type HSM is based on $\alpha = 0.65$ and $\tau_1 + \tau_2 = 1.7$, as in Anufriev and Hommes (2012b).

Figure 1 presents stability results for the RE and naive benchmark, while Figure 2 shows these for the two HSM variants. Interestingly, naive expectations always render the NK economy unstable, regardless of the policy parameters or calibration of the NK economy (cf. Hommes and Makarewicz, 2021). RE offer a wide range of policies for which the model has

[6] This parameter space was chosen because large coefficients imply aggressive policy that may be politically controversial or unfeasible, whereas this space easily contains parametrizations which are determinate under RE and stable under typical behavioral models for the standard inflation targeting (Hommes and Lustenhouwer, 2019).

Calibration	β	κ	σ
CGG	0.99	0.3	1
MN	0.99	0.3	1/0.164
W	0.99	0.157	0.024

Table 1: Benchmark calibrations of the NK economy: Clarida et al. (2000) (**CGG**), McCallum and Nelson (1999) (**MN**) and Woodford (1999) (**W**).

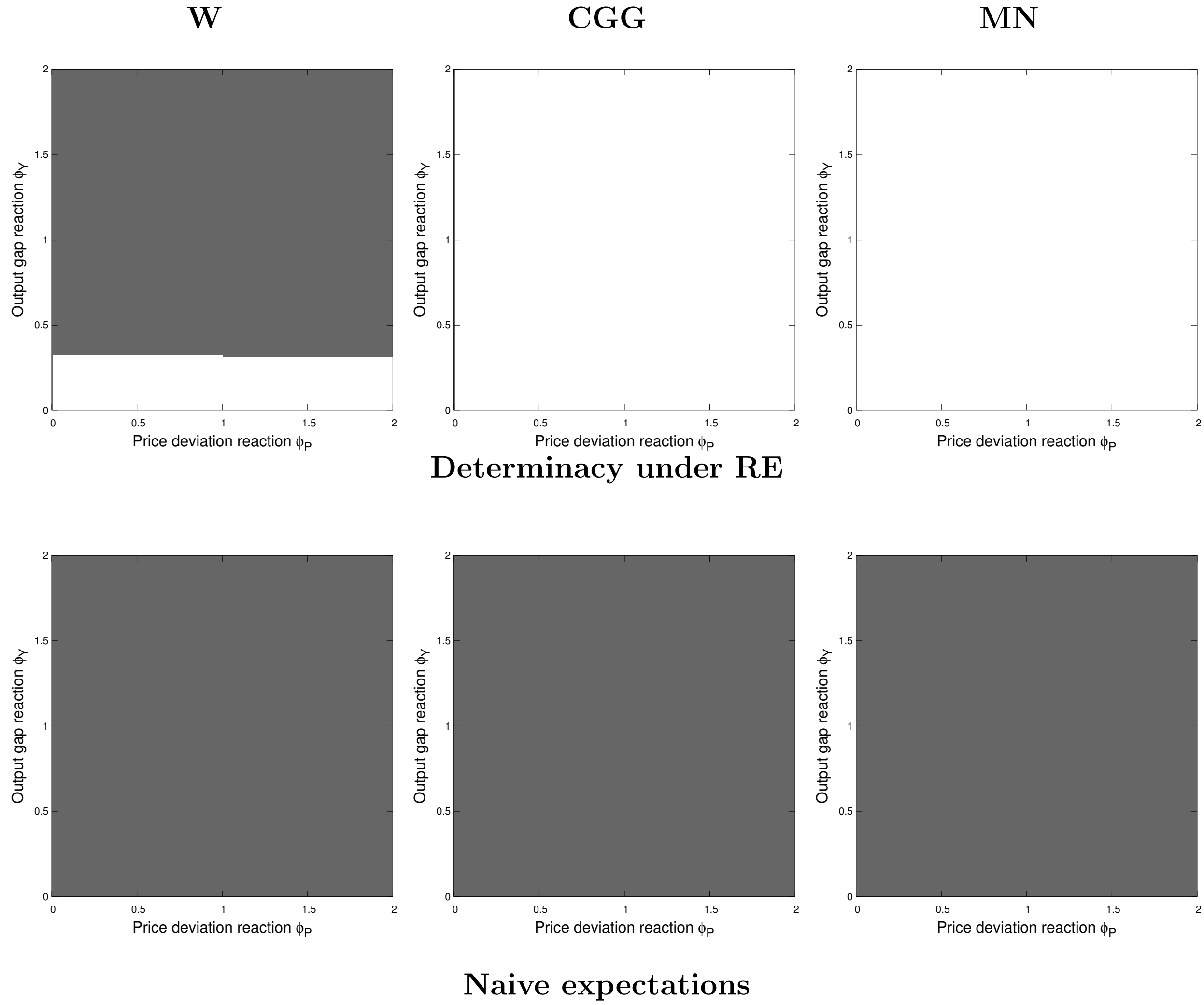


Figure 1: Heatmap: Determinacy under RE and stability under naive expectations of the NK economy with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote respectively determinate/stable and indeterminate/unstable solution under RE/naive benchmark.

a unique solution, in fact for **CGG** and **MN** the solution becomes indeterminate only for unreasonably high ϕ_p and ϕ_y, which are not presented in this paper.

For the **W** parametrization, the stability condition under RE, 2-type and 4-type HSM look relatively similar: PLT has determinate/stable dynamics under any positive ϕ_p and moderately high ϕ_y. In other words, the central banks is quite free in its response to the price deviation, however, it should neither under- nor over-react to the output gap. In the 2-type HSM, the

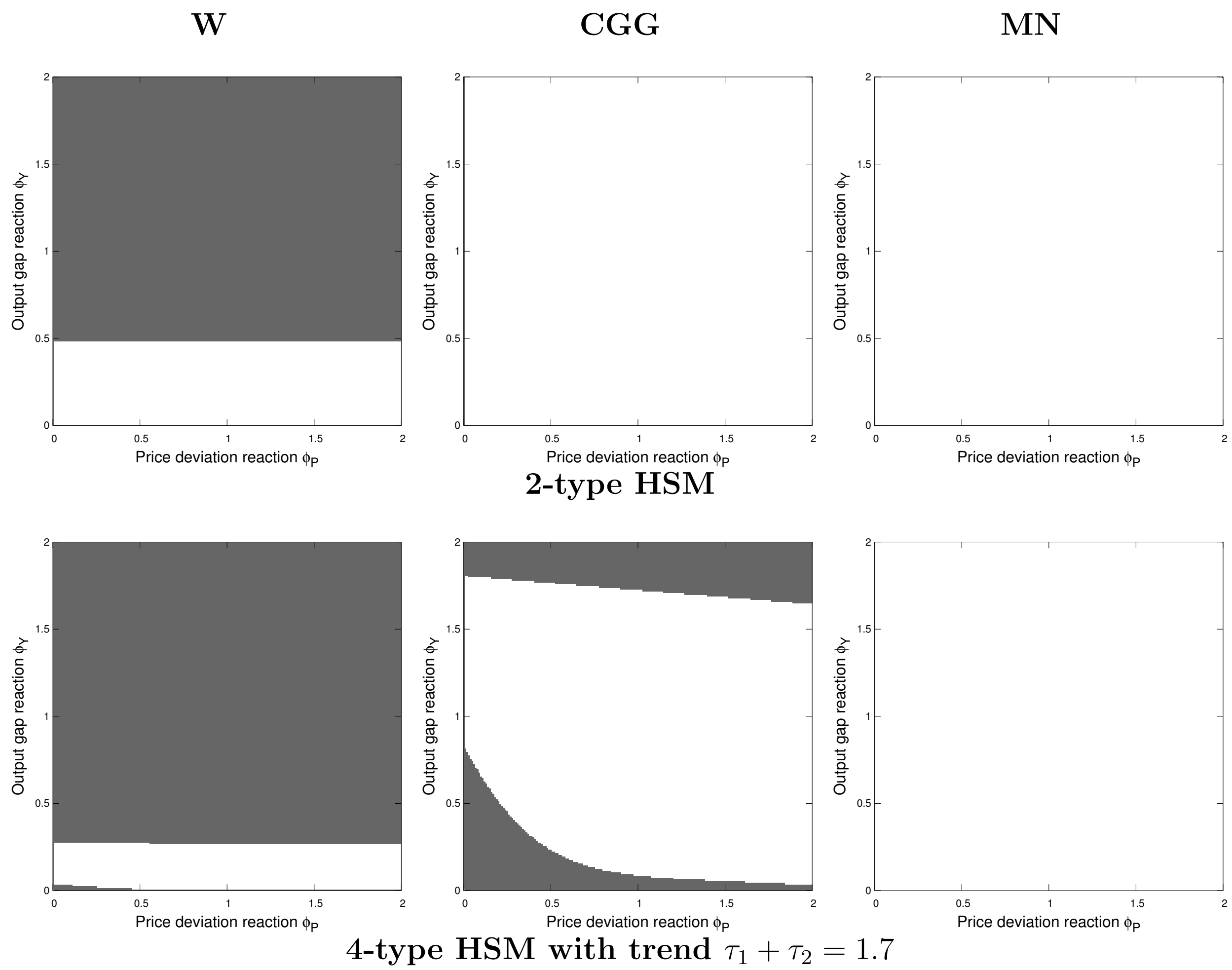


Figure 2: Heatmap: Stability of the NK economy under two variants of HSM with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote stable and unstable eigenvalues respectively.

stable window for the output gap reaction ϕ_y becomes narrower and further away from zero, as the price deviation reaction ϕ_p increases, suggesting a moderate PLT response being the safest. Finally, only the 2-type HSM can have a Jacobian matrix with real eigenvalues, which implies that oscillatory dynamics, regardless of their stability, dominate under PLT and behavioral models of expectations (again, cf. Hommes and Makarewicz, 2021).

The "agreement" between RE and the two variants of the HSM also holds for the **MN** calibration, when in fact *any* policy rule yields stable dynamics under the behavioral model. However, this relationship is less clear once we move to the **CGG** model calibration. Here, RE and the 2-type HSM support any policy rule, while the 4-type HSM has some areas of instability for high reaction to the output gap, as well as low reaction to either output or price deviation, suggesting a more aggressive rule than the rational benchmark.

Another important observation is that the stability regions, regardless which model we consider, are visibly different between the three considered parametrization of the NK economy. This highlights an important channel of model uncertainty – even if we knew exactly the underlying DSGE structure of the economy, as well as the specific mechanism of expectations, minor mistakes in calibration could yield a wrong picture about the viability of a certain PLT rule. It should be noted that in practice such mistakes are simply inevitable, since these parameters need to be estimated or calibrated, in either case with a well understood statistical measurement error. To highlight this issue, let us now consider how the stability of the 4-type HSM depends on two parameters from the NK economy, σ and κ.

5.3 Model calibration and stability of the 4-type HSM

Figure 3 shows how the stability of the 4-type HSM depends on the κ and σ parameters (horizontal and vertical axis respectively), with the three **CGG**, **MN** and **W** calibrations marked by black squares. We consider nine policy constellations, for $\phi_p \times \phi_y \in \{0.1, 0.5, 1.1\}^2$, which represent "weak", "medium" and "strong" reactions to either the price deviation or the output gap. The first interesting observation is that the Jacobian matrix of the steady state always has non-trivial complex eigenvalues, as was the case for the three benchmark calibrations. It should be noted that this may be considered to be a practical disadvantage of PLT, as the communication between the central bank and some economic agents may be obfuscated by the fact that even a one-time shock causes cyclical dynamics.

Both parameters of the NK model have an interesting impact on the overall model dynamics under the 4-type HSM. Recall that κ determines the relationship between realized inflation and output in the Philips Curve, whereas σ represents the utility elasticity of the household. If the latter is relatively small, strong reaction to the output gap renders the model unstable, regardless of the central bank's reaction to the price deviation. One can see this effect in particular in the bottom tiles of Figure 3, ie. for $\phi_p = 1.1$. This suggests that with behavioral expectations and certain types of preferences, restrictive reaction to the output gap can in fact distort the consumer's intertemporal decision-making and thus become self-defeating.

Recall that in the standard NK model under Calvo pricing mechanism, $\lim_{\theta \to 1} \kappa = 0$ and $\lim_{\theta \to 0} \kappa = +\infty$, where θ represents the share of firms that are unable to update their price in a given period (or the stickiness of prices in the model). For the "weak" PLT rule with $\phi_p = 0.1$, κ has initially minute effect, but once it crosses certain threshold, the model remains

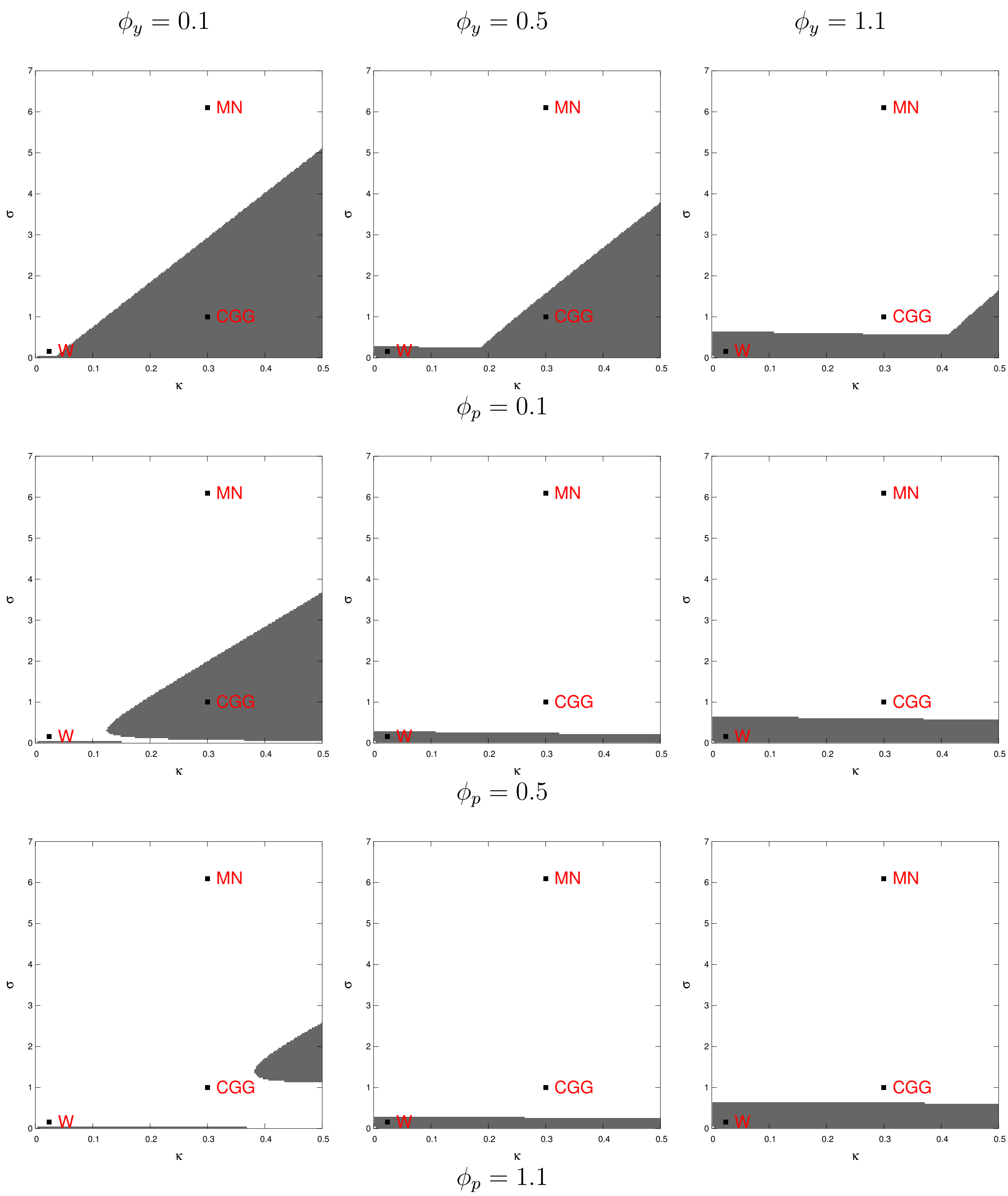


Figure 3: Heatmap: Stability of the NK economy under the 4-type HSM with $\tau = 1.7$ with PLT Taylor rule as a function of model calibration (κ and σ coefficients). White and gray colors denote stable and unstable eigenvalues respectively. Three benchmark calibrations, **CGG**, **MN** and **W**, are shown with black squares.

stable only if σ grows as well. This is represented by dark triangular shapes on the right side of the three top and three left-most panels of Figure 3.

I suggest the following interpretation of this result. In the behavioral model of expectations, agents (including firms) are myopic. They do not realize that certain past trends could be unsustainable under rational solution and thus give a higher weight to deviations from the steady state than they should. If the prices become less sticky, it is easier for such firms to overreact and adjust their production plan more quickly than optimal under RE, which under weak PLT rule allows for run-away dynamics, unless this is mitigated by the consumers being more "conservative" in how they wish to smooth their consumption over time.

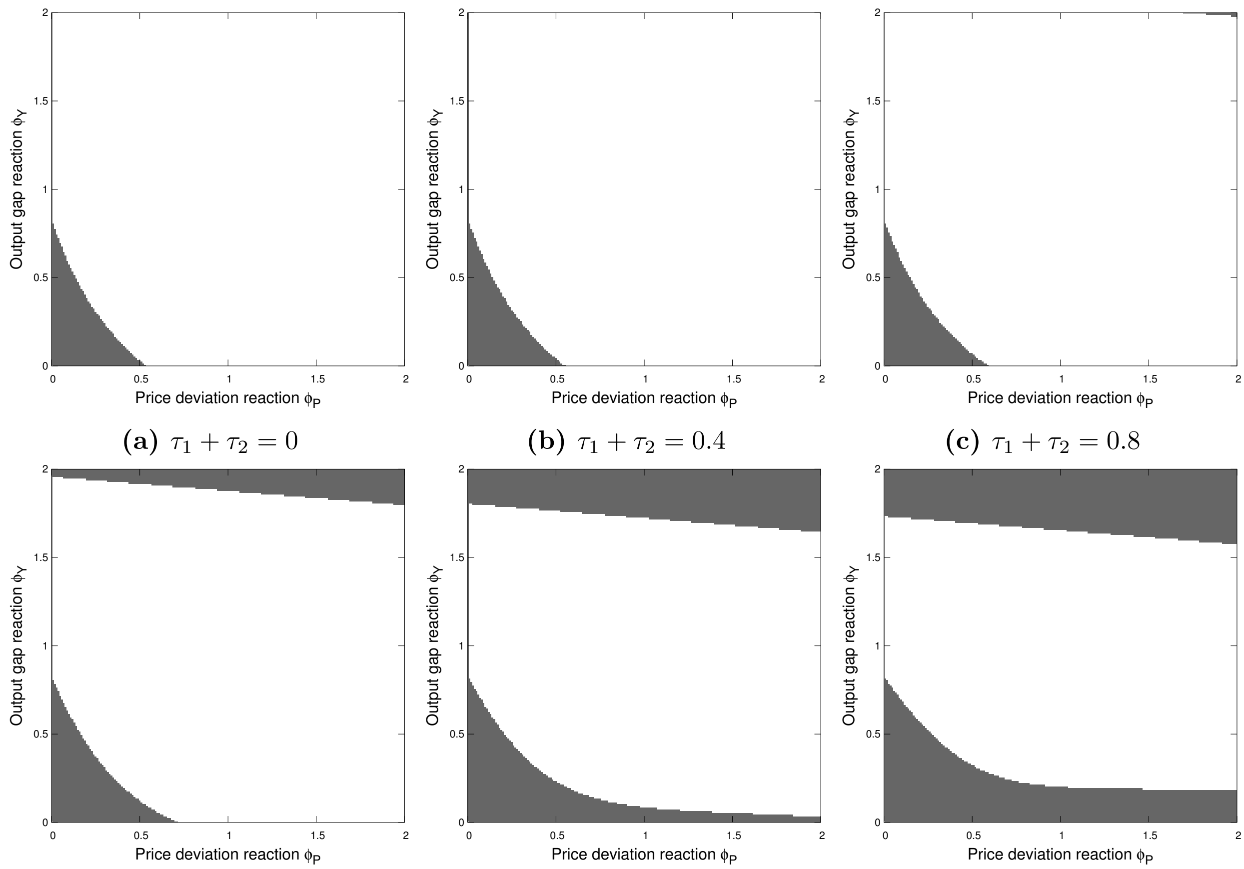
The overall area of unstable model calibrations shrinks with the decisiveness of the PLT rule. I will explore this result in detail in the next part of this section, but before, one should notice that the stability of the PLT rule under the 4-type HSM is indeed highly sensitive to the model calibration. In particular, Figure 3 suggests a combination of a strong reaction to the price deviation and a weak one to the output (with $\phi_p \gg 0$ and $\phi_y \approx 0$). From the nine presented parameter combinations, this is the only one under which all the three benchmark calibrations **CGG**, **MN** and **W** yield stable dynamics. In addition, such a policy is also supported by the 2-type HSM. However, even this policy constellation has a weak spot under the 4-type model for medium values of $\sigma \approx 2$ and high values of $\kappa > 0.4$. If the central bank believes that σ may be large, it may be safer to focus on rules with higher output gap reaction $\phi_y \gg 0$.

This calibration uncertainty holds already for the simplest version of the NK economy. In the actual policy practice, central banks – in particular given their bitter experience with the 2008 financial crisis – may want to use more sophisticated DSGE variants, in particular models with explicit financial frictions or income inequalities. Under behavioral expectations, each of these models is likely to yield very different policy recommendations and "problematic" calibrations for the PLT rule. This highlights the practical limitation of the PLT rule: it is unrobust to model uncertainty and thus should be used with caution.

5.4 Role of the trend chasing

As mentioned in the introduction, an important finding of the behavioral literature on expectations is the role of the so called positive feedback: a case when the realized variable depends positively on its (aggregate) forecast. The NK model (8) can be understood as an example of such a positive feedback, since the direct effect (sans the effect of the interest rate) of inflation and output expectations on their respective variables is given by $\partial y_t / \partial y^e_{t+1} = 1$

and $\partial \pi_t / \partial \pi^e_{t+1} = \beta$. Experimental literature demonstrates that such positive feedback systems often result in persistent, oscillatory dynamics due to the role of chartist strategies: subjects extrapolate past trend, which due to the mathematical properties of the setup results in a continuation of that trend. This result has been replicated by the theoretical literature on HSM, where trend following heuristics perform well under positive feedback and often result in off-equilibrium oscillations (see Anufriev et al., 2019, for a detailed analysis). In this part of the section I argue that this is also true for the NK model under the PLT rule.



(a) $\tau_1 + \tau_2 = 0$ **(b)** $\tau_1 + \tau_2 = 0.4$ **(c)** $\tau_1 + \tau_2 = 0.8$

(d) $\tau_1 + \tau_2 = 1.2$ **(e)** $\tau_1 + \tau_2 = 1.7$ **(f)** $\tau_1 + \tau_2 = 2$

Figure 4: Heatmap: Stability of the NK economy under the 4-type HSM and **CGG** calibration with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_p for different strengths of trend chasing behavior $\tau = \tau_1 + \tau_2$. White and gray colors denote stable and unstable eigenvalues respectively.

In the 4-type HSM, agents can switch between two trend chasing rules (15) and (16), which are parametrized by trend coefficients τ_1 and $\tau_2 \geqslant \tau_1$ respectively. Interestingly, Proposition 4 demonstrates that the stability of the (8) system under the 4-type model depends only on $\tau \equiv \tau_1 + \tau_2$. The intuition behind this result is that in the steady state all four rules have the same share with $n^{ST} = n^{WT} = 0.25$, and hence in the immediate aftermath of a shock the average forecast is a simple average of the four rules with equal weights. The rules themselves are linear in nature, hence in the first period the two trend coefficients simply add up to one term.

Figure 4 shows the stability results for six parametrizations of the 4-type HSM (under **CGG** calibration): α is always set to 0.65 like in Anufriev and Hommes (2012b), whereas the total trend following is gradually increased with $\tau \in \{0, 0.4, 0.8, 1.2, 1.7, 2\}$ (where $\beta = 1.7$ was used by Anufriev and Hommes, 2012a). Stability is again analyzed in terms of the policy rule, on the Cartesian plain $\phi_p \times \phi_y$. A clear pattern can be observed. As the total trend τ increases from 0 to 2, the space of stable PLT rules slowly decreases. Rules, in which both coefficients are low, are always unstable, and even harsher reaction is required if the agents use a stronger trend rule. What is interesting for this calibration is that the output gap reaction is restricted from the top as well: as the total trend coefficient τ increases, ϕ_y should be chosen from a narrower intermediate interval. This again shows that when agents use strong backward-looking forecasting rules, harsh policy rules may in fact be self-defeating, and the central bank is generally more restricted in its behavior than under RE.

To illustrate this finding better, Figures 5 presents sample simulations for the **CGG** model calibration with $\tau = 1.7$ and three different policies: $\phi_y \in \{1, 1.687, 1.68761\}$ and ϕ_p always set to unity. The first policy specification lies well within the stability region, whereas the other two are just inside of this stable region, close to its border, according to Proposition 4. In all three simulations, the economy starts in the fundamental steady when it is subjected to a one-time inflation shock $\varepsilon_1^\pi = 0.1$ and $\varepsilon_s^\pi = \varepsilon_{s'}^y 0$ for any $s > 1$ and $s' > 0$. The first policy swiftly deals with the initial shock, and the population is split evenly between fundamentalists and weak and strong chartists (the adaptive rule requires more time to converge back to the steady state, thus it remains unpopular). In the medium rule with $\phi_y = 1.687$, the dynamics will eventually settle on the fundamental, but this will take a long spell of time. Within the first 100 periods after the initial shock, we can observe two long cycles of inflation and output gap oscillating quickly around zero, although these two cycles have a dampening amplitude. What is interesting, each cycle itself has a cyclical structure, when the oscillations amplify and dampen in a rhythmic manner, which corresponds with periods of higher popularity of the strong trend rule for the output gap forecast. Once the policy parameter ϕ_y approaches

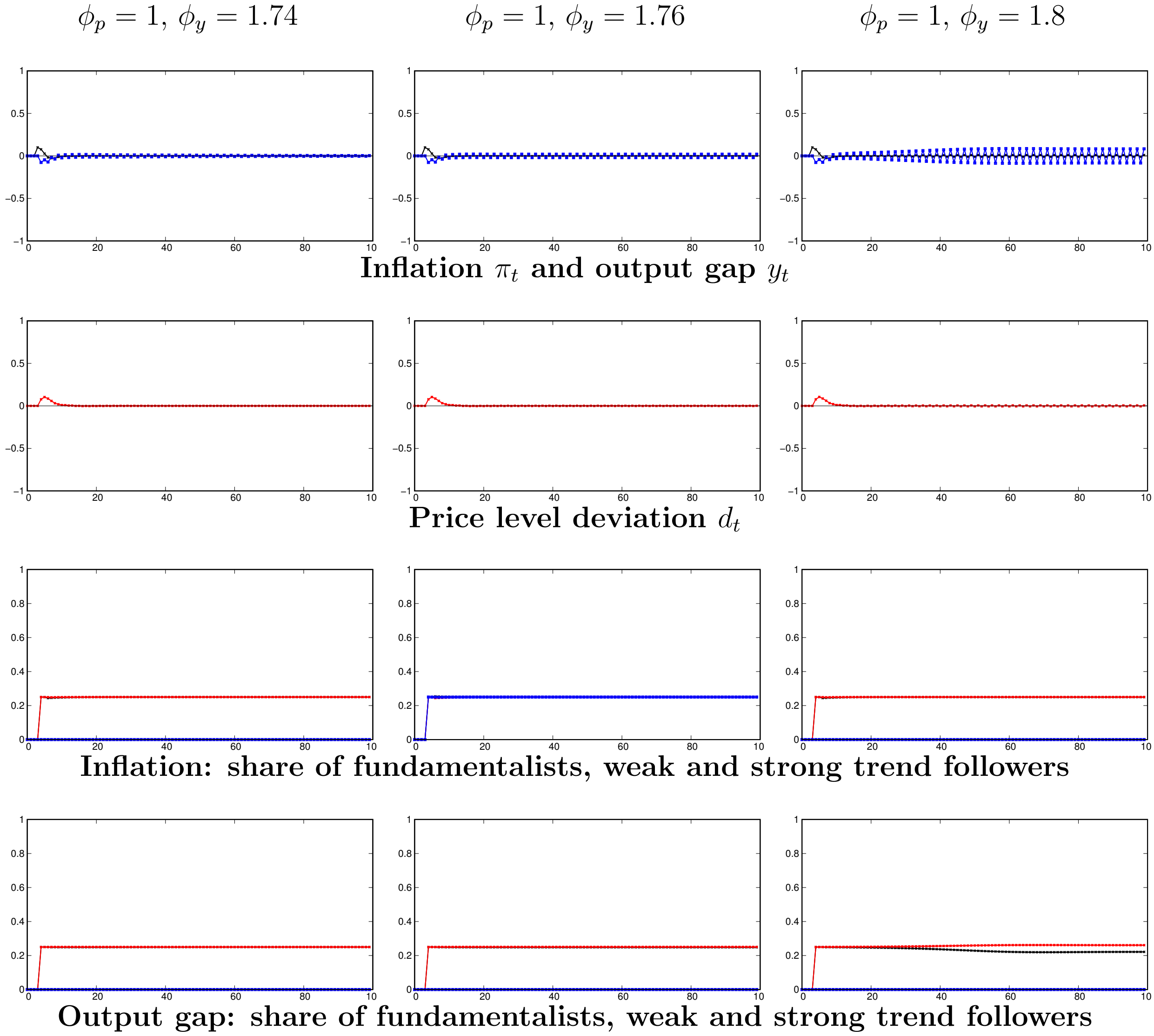
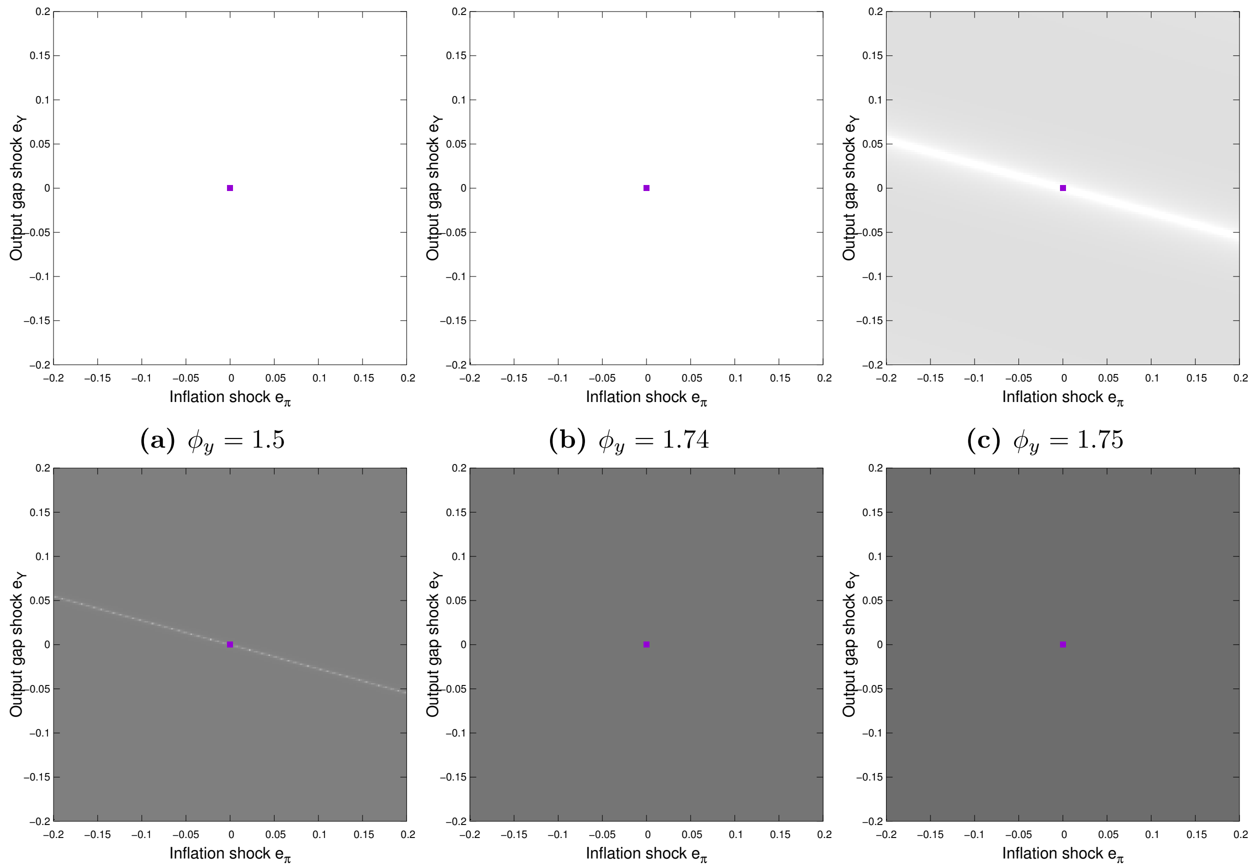


Figure 5: Sample simulations: NK economy under the 4-type HSM, **CGG** calibration, $\tau = 1.7$ and $\phi_p = 1$ with PLT Taylor rule for various output gap reaction parameters ϕ_y. Top panels present inflation π_t (black lines), output gap y_t (blue lines) and price level deviations d_t. Bottom panels present the shares of weak trend followers $n^{x,WT}$ (black lines), strong trend followers $n^{x,ST}$ (blue lines) and fundamentalists $n^{x,F}$ (red lines).

the instability threshold, the model explodes. This demonstrates that the local stability does not imply the global one and sufficiently high shocks can still destabilize the NK model under a poorly chosen PLT rule.

To further highlight this issue, consider Figure 6, which shows basins of attraction for a model with fixed trend rules ($\tau_1 = 1.2$ and $\tau_2 = 0.5$) and price deviation reaction parameter $\phi_p = 1$, but with different reactions to the output gap, with $\phi_y \in \{1.5, 1.74, 1.75, 1.76, 1.77, 1.8\}$. Only in the of the last two cases the fundamental steady state is locally unstable, whereas in the other cases it is in fact locally stable (with $\phi_y = 1.76$ at the edge of the stability region). Each basin is constructed in the following way. For a given initial shock vector $(e_\pi, e_y)' \in$



(a) $\phi_y = 1.5$ **(b)** $\phi_y = 1.74$ **(c)** $\phi_y = 1.75$

(d) $\phi_y = 1.76$ **(e)** $\phi_y = 1.77$ **(f)** $\phi_y = 1.8$

Figure 6: Basin of attraction: response to inflation/output shock of the NK model under 4-type HSM and **CGG** calibration with $\tau = 1.7$, $\phi_p = 1$ and various values of ϕ_y. Heatmap shows the value of the natural logarithm of the inflation variance in periods $t \in \{1001, \ldots, 1200\}$, in the range between -32 (white color) to 4 (black color).

$[-0.2, 0.2] \times [-0.2, 0.2]$ the simulation runs for 1200 periods. Next, $S = \ln\left(Var\left(\pi_{1001:1200}\right)\right)$ is computed as the stability measure. Figure 6 reports heat maps for various model variants, for S constrained to the interval between $-32 \leqslant S$ (white color, corresponding with standard deviation $\approx 10^{-7}$) and $S \leqslant 4$ (black color, corresponding with standard deviation ≈ 7.389, which is up to ≈ 24 times the magnitude of the initial shock).[7]

The results highlight the role of local stability. The PLT parametrizations that yield eigenvalues of the Jacobian matrix closest to zero, with $\phi_p = 1$ and $\phi_y = 1.5$ or $\phi_y = 1.74$, attract the whole shock space with ease and simulations quickly converge in a manner of few periods. However, we observe residual dynamics already for $\phi_y = 1.75$, where for majority the initial shocks the inflation's standard deviation after 1000 periods $SD^{\pi}_{1001:1200}$ is in the order of magnitude of approximately 10^{-5}. For the $\phi_y = 1.76$ policy parametrization, this standard deviation jumps to the magnitude of 10^{-3}, even for relatively small shocks. Once the policy

[7] The largest considered shock is for $|e_\pi| = |e_y| = 0.2$, for which the norm is $\sqrt{0.2^2 + 0.2^2} \approx 0.283$.

parameter ϕ_y further increases, the simulations start to explode to infinity. In addition, the stability measure does not depend on the size of the shock itself, suggesting that for a poorly chosen rule, persistent dynamics are essentially inevitable.

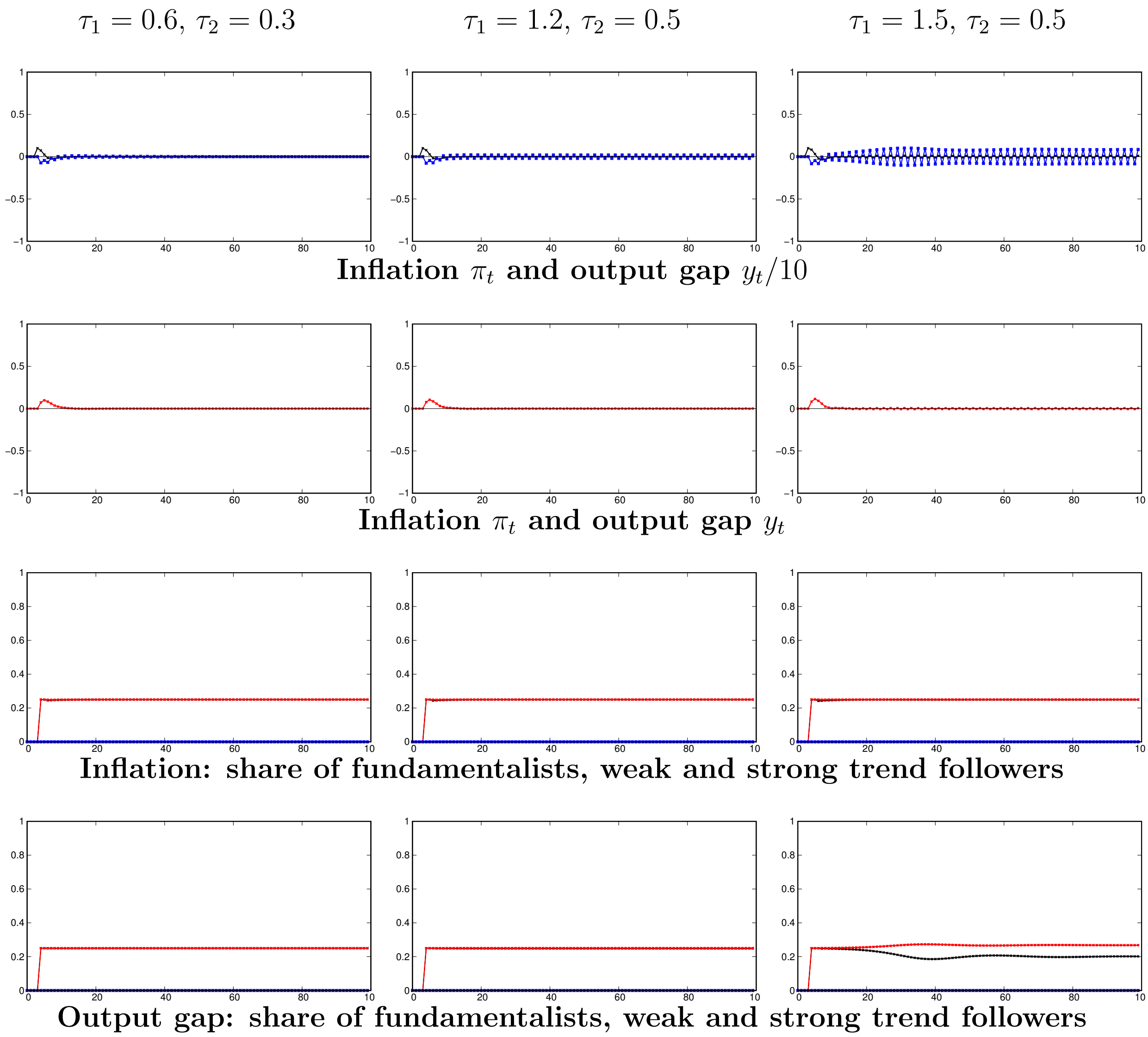
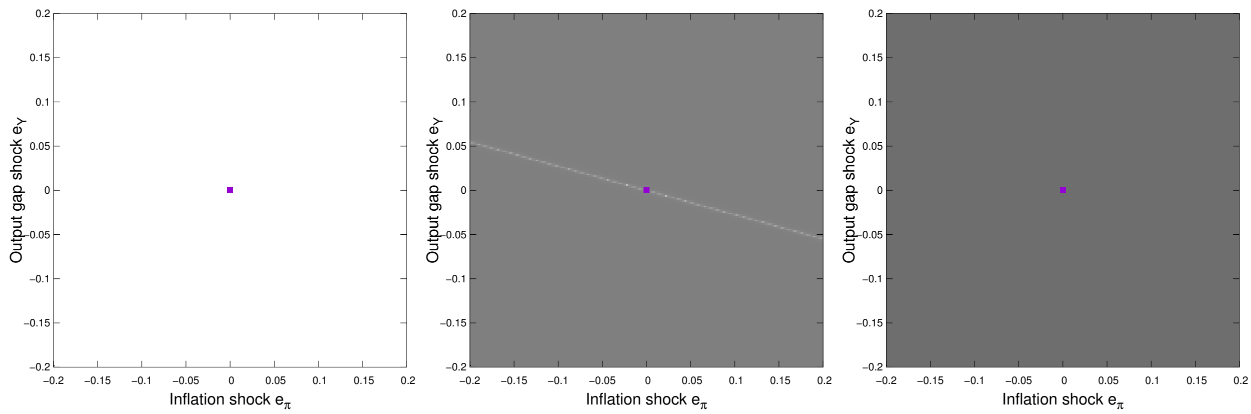


Figure 7: Sample simulations: NK economy under the 4-type HSM, **CGG** calibration, $\phi_p = 1$ and $\phi_y = 1.76$, with PLT Taylor rule for various strengths of trend following τ. Top panels present inflation π_t (black lines), output gap y_t (blue lines) and price level deviations d_t. Bottom panels present the shares of weak trend followers $n^{x,WT}$ (black lines), strong trend followers $n^{x,ST}$ (blue lines) and fundamentalists $n^{x,F}$ (red lines)

Figure 7 presents sample simulations for a fixed, mild PLT policy (with $\phi_p = 1$ and $\phi_y = 1.76$) and **CGG** calibration. The Figure focuses on three different overall strengths of the trend following: $\tau = 0.9$, $\tau = 1.7$ and $\tau = 2$. These three cases are respectively in the stable, just within the stable, and just outside of the stable parametrization region. The stable simulation quickly settles as expected, but the one at the edge of stability region keeps mildly oscillating, driven by the popularity of the weak trend following rule. Similar dynamics are visible for the unstable simulation, where the the weak trend rule is slightly more popular and oscillations gradually widen up, which will eventually lead to explosive dynamics. These results are further confirmed by the basins of attractions which are presented in Figure 8.



(a) $\tau_1 = 0.6, \tau_2 = 0.3$ **(b)** $\tau_1 = 1.2, \tau_2 = 0.5$ **(c)** $\tau_1 = 1.5, \tau_2 = 0.5$

Figure 8: Basin of attraction: response to inflation/output shock of the NK model under 4-type HSM and **CGG** calibration with $\phi_p = 1$ and $\phi_y = 1.76$, for various values of trend following strength τ. Heatmap shows the value of the natural logarithm of the inflation variance in periods $t \in \{1001, \ldots, 1200\}$, in the range between -32 (white color) to 4 (black color).

The highlighted relationship between the harshness of the monetary policy and the strength of the trend chasing behavior has an important implication. Suppose that the central bank is convinced that the expectations are driven by the 4-type HSM, but is unsure about the degree of trend following. In such a case, it may be difficult to ascertain a "safe" policy rule, since "restrictive" rules can be equally unstable as the "soft" ones, depending on the actual empirical value of the overall trend chasing coefficient τ and other model parameters. Moreover, the central bank may try one specific rule and at first find it to be successful, however, a larger shock may then initialize persistent oscillations, even without any further shocks – a case when PLT provides an initial sense of false safety. Finally, in this analysis we assume that τ_1 and τ_2 are simple constants, whereas Anufriev et al. (2019) suggest that there exists a separate feedback between the structure of the economy and the chartist strategies of the economic agents. In our case, poorly chosen rule may gradually teach the agents to focus on more aggressive trend chasing behavior, exaggerating the issue. This again shows that the PLT rule can be used successfully under behavioral expectations, but only if the central bank can understand the economic agents well.

6 Intensity of choice

The unusual feature of the PLT NK model is that the intensity of choice parameter γ plays no role in the stability conditions. This is the consequence of the Lemma 1, which constrains the set of the steady states in an unusual way the. In this section we will address the question of what is then the effect of this parameter on the PLT model. For the sake of brevity, we will again focus on the 4-type HSM.

Figure 9 presents the basins of attraction as defined in the previous Section, for three parameter choices for the intensity of choice: $\gamma \in \{0.01, 1, 20\}$, representing low, medium and high levels of intensity of choice, for three policy constellations that under the **CGG** calibration render the model stable, at the edge of stability and unstable. As we can see from the Figure, figures in each vertical panel (representing three levels of γ for one policy specification) are virtually identical. This suggests that, due to the uniqueness of the steady state, i) the intensity of choice plays little to no role in the long-run dynamics, whereas ii) the policy strength and the trend chasing behavior are crucial.

7 Conclusions

In the wake of the Great Recession, macroeconomists and policy makers have become increasingly interested in unconventional monetary policy tools, such as Price Level Targeting. At the same time, many have expressed doubts in the traditional toolbox of the monetary policy, in particular the modeling framework which is built upon Rational Expectations.

This paper investigates the dynamic properties of the Price Level Targeting in a simple New Keynesian model, where the agents are unable to form Rational Expectations and instead use Heuristic Switching Model to forecast future inflation and output gap. The Heuristic Switching Model is a popular alternative to the Rational Expectations and is used widely in behavioral and experimental literature. This paper focuses on two variants of this model, two-type (where agents switch between the fundamental and naive forecast) and four-type (where agents switch between the fundamental, adaptive, weak trend following and strong trend following forecasts).

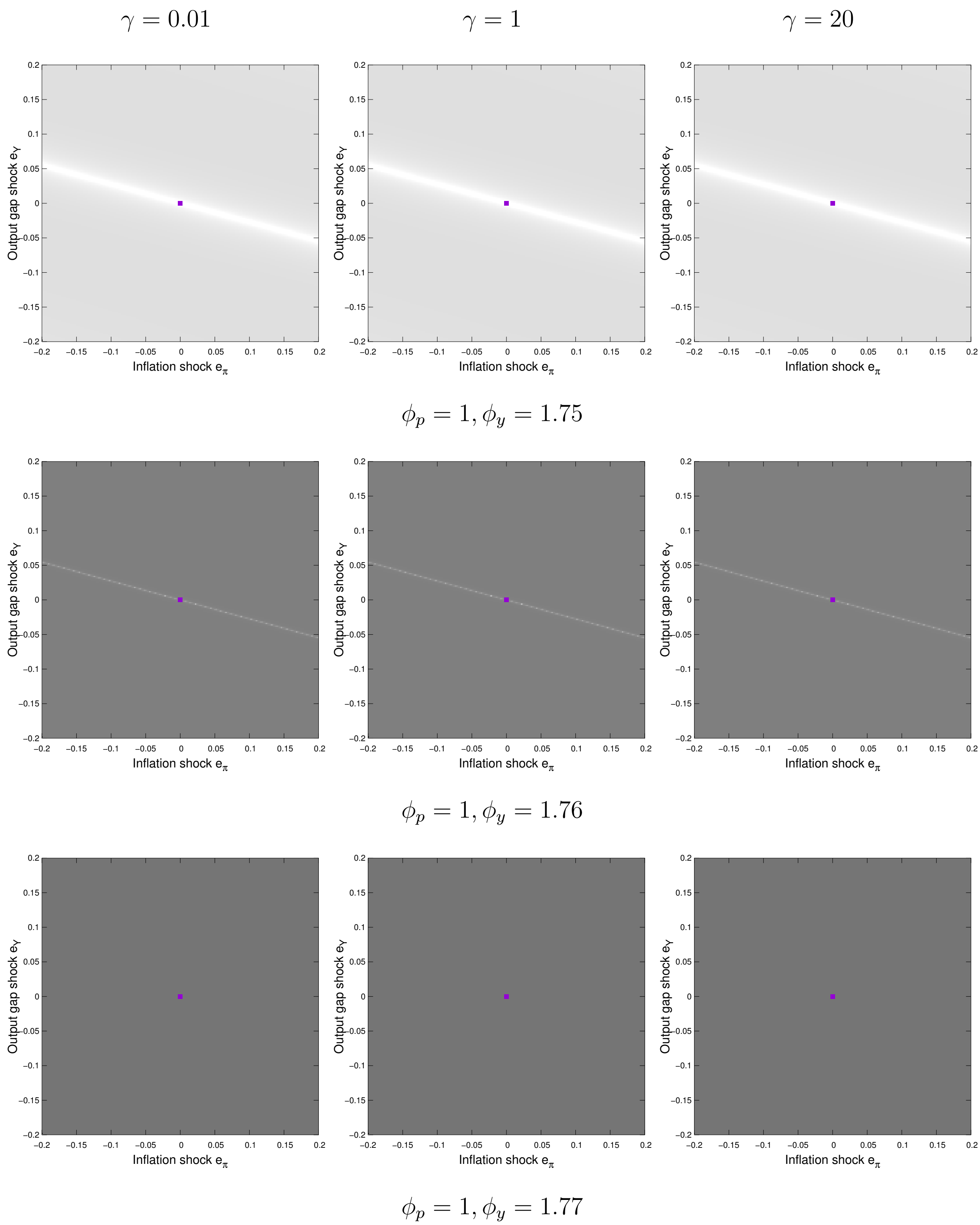


Figure 9: Basin of attraction: response to inflation/output shock of the NK model under 4-type HSM and **CGG** calibration with $\tau_p = 1.7$ and $\phi_p = 1$, for various values of intensity of choice γ and output gap reaction parameter ϕ_y. Heatmap shows the value of the natural logarithm of the inflation variance in periods $t \in \{1001, \ldots, 1200\}$, in the range between -32 (white color) to 4 (black color).

The results are mixed. An advantage of the Price Level Targeting rule is that for many important models of expectation formation, it admits only the fundamental equilibrium as the steady state. However, the dynamic properties of this fixed point are unclear. The same Taylor rule may lead to stable or unstable dynamics, depending on the specific model of expectations and the calibration of the economy. In particular the behavioral models suggest a much narrower area of safe policy rules than is the case for the Rational Expectations. It seems that the safest rule is one with a strong reaction to the price deviation and mild to the output gap. Nevertheless, there exists no single rule that is robust against all the variants of the studied economy. This suggests caution on the side of policy makers, who can use the Price Level Target only if they remain confident in their macroeconomic model. Otherwise, the traditional inflation targeting may be simply safer to use.

These findings confirm and give a good intuition for the inconsistent experimental results on the efficacy of the Price Level Target – relatively small changes to the experimental economy or framing may lead to surprisingly different behavior of the experimental subjects. Finally, this study confirms one important insight of the Learning-to-Forecast literature, namely the destabilizing role of trend following forecasting strategies.

This study has a number of limitations. The model is based on the simplest possible New Keynesian economy and thus cannot be used as a direct proof for more complicated models that are employed by many policy makers. For the sake of simplicity I have also abstracted from such issues as the Zero-Lower bound, guidance and communication of the central bank. All these relevant policy issues are left for future studies.

References

Ahrens, S., Lustenhouwer, J., and Tettamanzi, M. (2017). The stabilizing role of forward guidance: A macro experiment.

Ambler, S. (2009). Price-level targeting and stabilisation policy: a survey. *Journal of Economic Surveys*, 23(5):974–997.

Anufriev, M., Chernulich, A., and Tuinstra, J. (2022). Asset price volatility and investment horizons: An experimental investigation. *Journal of Economic Behavior & Organization*, 193:19–48.

Anufriev, M. and Hommes, C. (2012a). Evolution of market heuristics. *The Knowledge Engineering Review*, 27(2):255–271.

Anufriev, M. and Hommes, C. (2012b). Evolutionary selection of individual expectations and aggregate outcomes in asset pricing experiments. *American Economic Journal: Microeconomics*, 4(4):35–64.

Anufriev, M., Hommes, C., and Makarewicz, T. (2019). Simple forecasting heuristics that make us smart: Evidence from different market experiments. *Journal of the European Economic Association*, 17(5):1538–1584.

Arifovic, J. and Petersen, L. (2017). Stabilizing expectations at the zero lower bound: Experimental evidence. *Journal of Economic Dynamics and Control*, 82:21–43.

Armantier, O., Bruine de Bruin, W., Potter, S., Topa, G., Van Der Klaauw, W., and Zafar, B. (2013). Measuring inflation expectations. *Annu. Rev. Econ.*, 5(1):273–301.

Assenza, T., Bao, T., Hommes, C., and Massaro, D. (2014). Experiments on expectations in macroeconomics and finance. In *Experiments in macroeconomics*. Emerald Group Publishing Limited.

Assenza, T., Heemeijer, P., Hommes, C. H., and Massaro, D. (2021). Managing self-organization of expectations through monetary policy: A macro experiment. *Journal of Monetary Economics*, 117:170–186.

Bao, T. and Hommes, C. (2019). When speculators meet suppliers: Positive versus negative feedback in experimental housing markets. *Journal of Economic Dynamics and Control*, 107:103730.

Bao, T., Hommes, C., and Pei, J. (2021). Expectation formation in finance and macroeconomics: A review of new experimental evidence. *Journal of Behavioral and Experimental Finance*, 32:100591.

Berg, C. and Jonung, L. (1999). Pioneering price level targeting: the swedish experience 1931–1937. *Journal of Monetary Economics*, 43(3):525–551.

Bernanke, B. S. and Woodford, M. (2005). *The inflation-targeting debate*, volume 32. University of Chicago Press.

Bolt, W., Demertzis, M., Diks, C., Hommes, C., and Van Der Leij, M. (2019). Identifying booms and busts in house prices under heterogeneous expectations. *Journal of Economic Dynamics and Control*, 103:234–259.

Boswijk, H. P., Hommes, C. H., and Manzan, S. (2007). Behavioral heterogeneity in stock prices. *Journal of Economic dynamics and control*, 31(6):1938–1970.

Brázdik, F. and Marsal, A. (2011). Survey of research on financial sector modeling within dsge models: What central banks can learn from it. *Available at SSRN 2274689*.

Brock, W. A. and Hommes, C. H. (1998). Heterogeneous beliefs and routes to chaos in a simple asset pricing model. *Journal of Economic dynamics and Control*, 22(8-9):1235–1274.

Brock, W. A. and Hommes, C. H. (2001). A rational route to randomness. In *Growth Theory, Nonlinear Dynamics and Economic Modelling*, pages 402–438. Edward Elgar Publishing.

Brzoza-Brzezina, M. and Kolasa, M. (2013). Bayesian evaluation of dsge models with financial frictions. *Journal of Money, Credit and Banking*, 45(8):1451–1476.

Clarida, R., Gali, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: evidence and some theory. *The Quarterly journal of economics*, 115(1):147–180.

Colasante, A., Palestrini, A., Russo, A., and Gallegati, M. (2017). Adaptive expectations versus rational expectations: Evidence from the lab. *International Journal of Forecasting*, 33(4):988–1006.

Cornand, C. and Hubert, P. (2020). On the external validity of experimental inflation forecasts: A comparison with five categories of field expectations. *Journal of Economic Dynamics and Control*, 110:103746.

Cornea-Madeira, A., Hommes, C., and Massaro, D. (2019). Behavioral heterogeneity in us inflation dynamics. *Journal of Business & Economic Statistics*, 37(2):288–300.

Covas, F. and Zhang, Y. (2010). Price-level versus inflation targeting with financial market imperfections. *Canadian Journal of Economics/Revue canadienne d'économique*, 43(4):1302–1332.

Deak, S., Levine, P., Mirza, A., Pearlman, J., et al. (2020). Is price level targeting a robust monetary rule? Technical report.

Doornik, J. A. (2009). An object-oriented matrix programming language ox 6.

Evans, C. L. (2012). Monetary policy in a low-inflation environment: Developing a state-contingent price-level target. *Journal of Money, Credit and Banking*, 44:147–155.

Evans, G. W., Guesnerie, R., and McGough, B. (2019). Eductive stability in real business cycle models. *The Economic Journal*, 129(618):821–852.

Evans, G. W. and Honkapohja, S. (2012). Learning and expectations in macroeconomics. In *Learning and Expectations in Macroeconomics*. Princeton University Press.

Franke, R. and Westerhoff, F. (2012). Structural stochastic volatility in asset pricing dynamics: Estimation and model contest. *Journal of Economic Dynamics and Control*, 36(8):1193–1211.

Galí, J., Giusti, G., and Noussair, C. N. (2021). Monetary policy and asset price bubbles: a laboratory experiment. *Journal of Economic Dynamics and Control*, 130:104184.

Gambacorta, L. and Signoretti, F. M. (2014). Should monetary policy lean against the wind?: An analysis based on a dsge model with banking. *Journal of Economic Dynamics and Control*, 43:146–174.

Gerali, A., Neri, S., Sessa, L., and Signoretti, F. M. (2010). Credit and banking in a dsge model of the euro area. *Journal of money, Credit and Banking*, 42:107–141.

Grazzini, J., Richiardi, M. G., and Tsionas, M. (2017). Bayesian estimation of agent-based models. *Journal of Economic Dynamics and Control*, 77:26–47.

Heemeijer, P., Hommes, C., Sonnemans, J., and Tuinstra, J. (2009). Price stability and volatility in markets with positive and negative expectations feedback: An experimental investigation. *Journal of Economic dynamics and control*, 33(5):1052–1072.

Hommes, C. (2011). The heterogeneous expectations hypothesis: Some evidence from the lab. *Journal of Economic dynamics and control*, 35(1):1–24.

Hommes, C. and Lustenhouwer, J. (2019). Inflation targeting and liquidity traps under endogenous credibility. *Journal of Monetary Economics*, 107:48–62.

Hommes, C. and Makarewicz, T. (2021). Price level versus inflation targeting under heterogeneous expectations: a laboratory experiment. *Journal of Economic Behavior & Organization*, 182:39–82.

Hommes, C., Massaro, D., and Weber, M. (2019). Monetary policy under behavioral expectations: Theory and experiment. *European Economic Review*, 118:193–212.

Honkapohja, S. and Mitra, K. (2020). Price level targeting with evolving credibility. *Journal of Monetary Economics*, 116:88–103.

Jääskelä, J. P. (2005). Inflation, price level and hybrid rules under inflation uncertainty. *Scandinavian Journal of Economics*, 107(1):141–156.

Jang, T.-S. and Sacht, S. (2021). Forecast heuristics, consumer expectations, and new-keynesian macroeconomics: A horse race. *Journal of Economic Behavior & Organization*, 182:493–511.

Jump, R. C. and Levine, P. (2019). Behavioural new keynesian models. *Journal of Macroeconomics*, 59:59–77.

Kirman, A. (2014). Is it rational to have rational expectations? *Mind & Society*, 13(1):29–48.

Kopányi, D., Rabanal, J. P., Rud, O. A., and Tuinstra, J. (2019). Can competition between forecasters stabilize asset prices in learning to forecast experiments? *Journal of Economic Dynamics and Control*, 109:103770.

Kostyshyna, O., Petersen, L., and Yang, J. (2022). A horse race of monetary policy regimes: An experimental investigation. Technical report, National Bureau of Economic Research.

Kouwenberg, R. and Zwinkels, R. (2014). Forecasting the us housing market. *International Journal of Forecasting*, 30(3):415–425.

Kryvtsov, O. and Petersen, L. (2013). Expectations and monetary policy: experimental evidence. Technical report, Bank of Canada.

Kryvtsov, O. and Petersen, L. (2021). Central bank communication that works: Lessons from lab experiments. *Journal of Monetary Economics*, 117:760–780.

Kukacka, J. and Barunik, J. (2017). Estimation of financial agent-based models with simulated maximum likelihood. *Journal of Economic Dynamics and Control*, 85:21–45.

Kukacka, J., Jang, T.-S., and Sacht, S. (2018). On the estimation of behavioral macroeconomic models via simulated maximum likelihood. *Available at SSRN 3299450*.

Levine, P. (2020). The state of dsge modeling. In *Oxford Research Encyclopedia of Economics and Finance*.

Lindé, J., Smets, F., and Wouters, R. (2016). Challenges for central banks' macro models. In *Handbook of macroeconomics*, volume 2, pages 2185–2262. Elsevier.

Lustenhouwer, J. and Salle, I. (2022). Forecast revisions in the presence of news: a lab investigation. AWI discussion paper series No. 714.

Lux, T. and Zwinkels, R. C. (2018). Empirical validation of agent-based models. In *Handbook of computational economics*, volume 4, pages 437–488. Elsevier.

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

Mauersberger, F. (2021). Monetary policy rules in a non-rational world: A macroeconomic experiment. *Journal of Economic Theory*, 197:105203.

McCallum, B. T. and Nelson, E. (1999). Performance of operational policy rules in an estimated semiclassical structural model. In *Monetary policy rules*, pages 15–56. University of Chicago Press.

Merola, R. (2015). The role of financial frictions during the crisis: An estimated dsge model. *Economic Modelling*, 48:70–82.

Nuzzo, S. and Morone, A. (2017). Asset markets in the lab: A literature review. *Journal of Behavioral and Experimental Finance*, 13:42–50.

Plott, C. R. and Sunder, S. (1982). Efficiency of experimental security markets with insider information: An application of rational-expectations models. *Journal of political economy*, 90(4):663–698.

Poledna, S., Miess, M. G., Hommes, C., and Rabitsch, K. (2023). Economic forecasting with an agent-based model. *European Economic Review*, 151:104306.

Salle, I. (2021). What to target. *Insights from theory and lab experiments. Bank of Canada Staff Working Paper.*

Sargent, T. J. (2008). Evolution and intelligent design. *American Economic Review*, 98(1):5–37.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, pages 1–48.

Svensson, L. E. (1996). Price level targeting vs. inflation targeting: a free lunch?

Trichet, J.-C. (2018). Unconventional monetary policy measures: Principles-conditions-raison d'être. *29th issue (January 2013) of the International Journal of Central Banking*.

Vestin, D. (2006). Price-level versus inflation targeting. *Journal of Monetary Economics*, 53(7):1361–1376.

Weizsäcker, G. (2010). Do we follow others when we should? a simple test of rational expectations. *American Economic Review*, 100(5):2340–60.

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

Woodford, M. (2012). Inflation targeting and financial stability. Technical report, National Bureau of Economic Research.

Woodford, M. (2013). Macroeconomic analysis without the rational expectations hypothesis.

A Contemporaneous and backward rules

The three possible Taylor rules are given as:

$$\text{(A.1)} \qquad \textbf{Forward TR:} \quad r_t = \phi_p d_{t+1}^e + \phi_y y_{t+1}^e = \phi_p \left(\pi_{t+1}^e + \pi_t + d_{t-1}\right) + \phi_y y_{t+1}^e,$$

$$\text{(A.2)} \qquad \textbf{Contemporaneous TR:} \quad r_t = \phi_p d_t + \phi_y y_t = \phi_p \left(\pi_t + d_{t-1}\right) + \phi_y y_t,$$

$$\text{(A.3)} \qquad \textbf{Backward TR:} \quad r_t = \phi_p d_{t-1} + \phi_y y_{t-1}.$$

With straightforward derivations, one can show that the NK economy (1–2), together with the three PLT, rules can be represented as

$$\text{(A.4)} \qquad \begin{aligned} \pi_t &= \omega_1 \pi_{t+1}^e + \omega_2 y_{t+1}^e + \omega_3 y_{t-1} + \omega_4 d_{t-1} \\ y_t &= \zeta_1 \pi_{t+1}^e + \zeta_2 y_{t+1}^e + \zeta_3 y_{t-1} + \zeta_4 d_{t-1} \\ d_t &= \pi_t + d_{t-1} = \omega_1 \pi_{t+1}^e + \omega_2 y_{t+1}^e + \omega_3 y_{t-1} + \left(\omega_4 + 1\right) d_{t-1}, \end{aligned}$$

with the specific parametrization given by

Forward TR:

$$\text{(A.5)} \qquad \pi_t = \frac{\nu - \mu\phi_p}{1 + \mu\phi_p} \pi_{t+1}^e + \frac{\kappa - \mu\phi_y}{1 + \mu\phi_p} y_{t+1}^e - \frac{\mu\phi_p}{1 + \mu\phi_p} d_{t-1},$$

$$\text{(A.6)} \qquad y_t = \frac{1 - \phi_p \left(1 + \beta\right)}{\sigma \left(1 + \mu\phi_p\right)} \pi_{t+1}^e + \frac{\sigma - \phi_y}{\sigma \left(1 + \mu\phi_p\right)} y_{t+1}^e - \frac{\phi_p}{\sigma \left(1 + \mu\phi_p\right)} d_{t-1},$$

Contemporaneous TR:

$$\text{(A.7)} \qquad \pi_t = \frac{\nu\sigma + \beta\phi_y}{\sigma + \kappa\phi_p + \phi_y} \pi_{t+1}^e + \frac{\kappa\sigma}{\sigma + \kappa\phi_p + \phi_y} y_{t+1}^e - \frac{\kappa\phi_p}{\sigma + \kappa\phi_p + \phi_y} d_{t-1},$$

$$\text{(A.8)} \qquad y_t = \frac{1 - \beta\phi_p}{\sigma + \kappa\phi_p + \phi_y} \pi_{t+1}^e + \frac{\sigma}{\sigma + \kappa\phi_p + \phi_y} y_{t+1}^e - \frac{\phi_p}{\sigma + \kappa\phi_p + \phi_y} d_{t-1},$$

Backward TR:

$$\text{(A.9)} \qquad \pi_t = \nu\pi_{t+1}^e + \kappa y_{t+1}^e - \mu\phi_y y_{t-1} - \mu\phi_p d_{t-1},$$

$$\text{(A.10)} \qquad y_t = \sigma^{-1}\pi_{t+1}^e + y_{t+1}^e - \sigma^{-1}\phi_y y_{t-1} - \sigma^{-1}\phi_p d_{t-1},$$

each with the price deviation equation (6).

B Blanchard-Kahn conditions

The system can be rewritten into the following form for the three Taylor rules.

Forward TR:

$$\begin{pmatrix} \pi^e_{t+1} \\ y^e_{t+1} \\ d_t \end{pmatrix} = \begin{bmatrix} 1/\beta & -\kappa/\beta & 0 \\ \frac{\phi_p(1+\beta)-1}{\beta(\sigma-\phi_y)} & \frac{\beta\sigma+\kappa(1-\phi_p)}{\beta(\sigma-\phi_y)} & \frac{\phi_p}{\sigma-\phi_y} \\ 1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \pi_t \\ y_t \\ d_{t-1} \end{pmatrix}$$

Contemporaneous TR:

$$\begin{pmatrix} \pi^e_{t+1} \\ y^e_{t+1} \\ d_t \end{pmatrix} = \begin{bmatrix} 1/\beta & -\kappa/\beta & 0 \\ \frac{\phi_p\beta-1}{\beta\sigma} & \frac{\beta\sigma+\beta\phi_y+\kappa}{\beta\sigma} & \frac{\phi_p}{\sigma} \\ 1 & 0 & 1 \end{bmatrix} \begin{pmatrix} \pi_t \\ y_t \\ d_{t-1} \end{pmatrix}$$

Backward TR:

$$\begin{pmatrix} \pi^e_{t+1} \\ y^e_{t+1} \\ d_t \\ l_t \end{pmatrix} = \begin{bmatrix} 1/\beta & -\kappa/\beta & 0 & 0 \\ -1/(\beta\sigma) & \frac{\beta\sigma+\kappa}{\beta\sigma} & \frac{\phi_p}{\sigma} & \frac{\phi_y}{\sigma} \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{pmatrix} \pi_t \\ y_t \\ d_{t-1} \\ l_{t-1} \end{pmatrix}$$

Remark that given that the formulation of the Phillips Curve (2) and price level deviation equation (6), changes to the Taylor rule influence only the Euler equation (1) (second row in the BK matrices). The backward TR introduces an additional output gap lag y_{t-1} into the system, hence the trick with $l_t = y_t$ as an additional jump variable.

C Derivations for the 2-type HSM

Proof for Proposition 3.

Proof. For the sake of simplifying the notation, define

$$\begin{aligned} G^{w,z}_t \equiv & \tanh\left[\gamma\left(w_{t-1}z_{t-1} - 0.5z^2_{t-1}\right)\right], \\ H^{w,z} \equiv & \gamma w_{t-1}\left(1 - (G^{w,z}_t)^2\right). \end{aligned}$$

It is easy to see that when $w_{t-1} = z_{t-1} = 0$ then $G_t^{0,0} = H_t^{0,0} = 0$. Then

$$\begin{aligned}
\frac{\partial G_t^{w,z} w_{t-1}}{\partial w_{t-1}} =& G_t^{w,z} + w_{t-1}\left(1 - (G_t^{w,z})^2\right)\gamma z_{t-1} = G_t^{w,z} + z_{t-1} H_t^{w,z} \\
&=_{|w_{t-1}=z_{t-1}=0} 0, \\
\frac{\partial G_t^{w,z} w_{t-1}}{\partial z_{t-1}} =& w_{t-1}\left(1 - (G_t^{w,z})^2\right)\gamma\left(w_{t-1} - z_{t-1}\right) = H_t^{w,z}\left(w_{t-1} - z_{t-1}\right) \\
&=_{|w_{t-1}=z_{t-1}=0} 0.
\end{aligned}$$

Define the state vector of the NK model (A.4) as $x_t \equiv (\pi_t, k_{1,t}, k_{2,t}, y_t, m_{1,t}, m_{2,t}, d_t)'$, with vector $x^* = \mathbf{0}$ representing the fundamental steady state. The system can be written down as the following set of difference equations:

(C.1a) $$\begin{aligned}
\pi_t =& \omega_1 \pi_{t+1}^e + \omega_2 y_{t+1}^e + \omega_3 y_{t-1} + \omega_4 d_{t-1} \\
=& 0.5\omega_1 \pi_{t-1} + 0.5\omega_1 \tanh\left[\gamma\left(\pi_{t-1} k_{2,t-1} - 0.5 k_{2,t-1}^2\right)\right]\pi_{t-1} + (0.5\omega_2 + \omega_3) y_{t-1} \\
&+ 0.5\omega_2 \tanh\left[\gamma\left(y_{t-1} m_{2,t-1} - 0.5 m_{2,t-1}^2\right)\right] y_{t-1} + \omega_4 d_{t-1} \\
=& 0.5\omega_1 \pi_{t-1} + 0.5\omega_1 G_t^{\pi,k2} \pi_{t-1} + (0.5\omega_2 + \omega_3) y_{t-1} + 0.5\omega_2 G_t^{y,m2} y_{t-1} + \omega_4 d_{t-1},
\end{aligned}$$

(C.1b)

$$k_{,1t} = \pi_{t-1},$$

(C.1c) $k_{2,t} = k_{1,t-1},$

(C.1d) $y_t = 0.5\zeta_1 \pi_{t-1} + 0.5\zeta_1 G_t^{\pi,k2} \pi_{t-1} + (0.5\zeta_2 + \zeta_3) y_{t-1} + 0.5\zeta_2 G_t^{y,m2} y_{t-1} + \zeta_4 d_{t-1},$

(C.1e)

$$m_{1,t} = y_{t-1},$$

(C.1f)

$$m_{2,t} = m_{1,t-1},$$

(C.1g) $$\begin{aligned}
d_t =& \pi_t + d_{t-1} \\
=& 0.5\omega_1 \pi_{t-1} + 0.5\omega_1 G_t^{\pi,k2} \pi_{t-1} + (0.5\omega_2 + \omega_3) y_{t-1} + 0.5\omega_2 G_t^{y,m2} y_{t-1} + (\omega_4 + 1) d_{t-1}.
\end{aligned}$$

The specific values for the ω and ζ parameters depend on the Taylor rule.

The full $J^{2T}(x_t)$ Jacobian matrix of the system is given by equation (C.3) at the end of this Appendix, as it is too large to fit a horizontally aligned page. Evaluating $J(x^*)$ at the fundamental steady state $x^* = \mathbf{0}$, all the $G(\cdot)$ and $H(\cdot)$ terms drop out and the matrix simplifies into

$$J^{2T}\left(x^f\right) = \begin{pmatrix} 0.5\omega_1 & 0 & 0 & \omega_3 + 0.5\omega_2 & 0 & 0 & \omega_4 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5\zeta_1 & 0 & 0 & \zeta_3 + 0.5\zeta_2 & 0 & 0 & \zeta_4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0.5\omega_1 & 0 & 0 & \omega_3 + 0.5\omega_2 & 0 & 0 & \omega_4 + 1 \end{pmatrix}. \tag{C.2}$$

Remark that the third and sixth columns of this matrix (associated with $\partial(\cdot)/\partial k_{2,t-1}$ and $\partial(\cdot)/\partial m_{2,t-1}$) contain only zeros and can be dropped, after which the same is true for the second and fifth column (associated with $\partial(\cdot)/\partial k_{1,t-1}$ and $\partial(\cdot)/\partial m_{1,t-1}$). Hence the characteristic polynomial of the Jacobian at the fundamental steady state simplifies to

$$\mathbf{H}(\lambda) = \lambda^4 \det \left[\begin{pmatrix} 0.5\omega_1 & \omega_3 + 0.5\omega_2 & \omega_4 \\ 0.5\zeta_1 & \zeta_3 + 0.5\zeta_2 & \zeta_4 \\ 0.5\omega_1 & \omega_3 + 0.5\omega_2 & \omega_4 + 1 \end{pmatrix} - \lambda \mathbb{I} \right] \equiv \lambda^4 \det \left[J_{2T} - \lambda \mathbb{I} \right].$$

The four trivial eigenvalues are given by stable $\lambda_{1:4} = 0$ and thus the stability of the system depends on the stability of J_{2T}. $\square$

(C.3)

$$J^{2T}(x_t) = \begin{pmatrix} 0.5\omega_1\left(1+G_t^{\pi,k2}+k_{2,t-1}H_t^{\pi,k2}\right) & 0 & 0.5\omega_1\left(\pi_{t-1}-k_{2,t-1}\right)H_t^{\pi,k2} & \omega_3+0.5\omega_2\left(1+G_t^{y,m2}+m_{2,t-1}H_t^{y,m2}\right) & 0 & 0.5\omega_2\left(y_{t-1}-m_{2,t-1}\right)H_t^{y,m2} & \omega_4 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0.5\zeta_1\left(1+G_t^{\pi,k2}+k_{2,t-1}H_t^{\pi,k2}\right) & 0 & 0.5\zeta_1\left(\pi_{t-1}-k_{2,t-1}\right)H_t^{\pi,k2} & \zeta_3+0.5\zeta_2\left(1+G_t^{y,m2}+m_{2,t-1}H_t^{y,m2}\right) & 0 & 0.5\zeta_2\left(y_{t-1}-m_{2,t-1}\right)H_t^{y,m2} & \zeta_4 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0.5\omega_1\left(1+G_t^{\pi,k2}+k_{2,t-1}H_t^{\pi,k2}\right) & 0 & 0.5\omega_1\left(\pi_{t-1}-k_{2,t-1}\right)H_t^{\pi,k2} & \omega_3+0.5\omega_2\left(1+G_t^{y,m2}+m_{2,t-1}H_t^{y,m2}\right) & 0 & 0.5\omega_2\left(y_{t-1}-m_{2,t-1}\right)H_t^{y,m2} & \omega_4+1 \end{pmatrix}$$

D Derivations for the 4-type HSM

Model representation. The model becomes a 19D map

(D.1a) $$\pi_t = \omega_1 \pi^e_{t+1} + \omega_2 y^e_{t+1} + \omega_3 y_{t-1} + \omega_4 d_{t-1},$$

(D.1b) $$k^1_t = \pi_{t-1},$$

(D.1c) $$k^2_t = k^1_{t-1},$$

(D.1d) $$k^3_t = k^2_{t-1},$$

(D.1e) $$y_t = \zeta_1 \pi^e_{t+1} + \zeta_2 y^e_{t+1} + \zeta_3 y_{t-1} + \zeta_4 d_{t-1},$$

(D.1f) $$l^1_t = y_{t-1},$$

(D.1g) $$l^2_t = l^1_{t-1},$$

(D.1h) $$l^3_t = l^2_{t-1},$$

(D.1i) $$d_t = \omega_1 \pi^e_{t+1} + \omega_2 y^e_{t+1} + \omega_3 y_{t-1} + (\omega_4 + 1) d_{t-1},$$

(D.1j) $$\pi^{AD}_{t+1} = \alpha \pi_{t-1} + (1-\alpha)\pi^{AD}_t,$$

(D.1k) $$y^{AD}_{t+1} = \alpha y_{t-1} + (1-\alpha) y^{AD}_t,$$

together with equations (25–28) and auxiliary definitions of the heuristics' weights $n^{x,h}_t$ for $x \in \{\pi, y\}$ and $h \in \{AD, WT, ST, FD\}$ (see Appendix A for the definition of ω and ζ coefficients).

Proof for Proposition 4. The Jacobian matrix for the map (D.1) is obviously complicated, with many interaction terms between the model coefficients. However, in the fundamental steady state many of them will disappear, since the steady state variables become zero. This allows for a relatively simple Jacobian that can be efficiently evaluated numerically. In the case of a model with multiple fixed points, the picture would in fact become much more complicated.

Proof. The system (D.1) is mostly linear, and in fact the only source of non-linearities in the model are heuristic weights $n^{x,h}_t$ for $x \in \{\pi, y\}$ and $h \in \{AD, WT, ST, FD\}$. Given that each of these weights then depends on all four attractions $U^{x,h}_t$ of its variable x, they have a complicated dependency on the four variable x lags and other variables. However, they simplify greatly at the steady state.

The partial derivatives of the two average expectations $x^e_{t+1} \in \{\pi^e_{t+1}, y^e_{t+1}\}$ in respect to other variables, evaluated at the fundamental steady state $(\pi^e_, y^e)' = (\pi, y)' = (0, 0)$ and $n^{x,h} = 0.25$ (which is denoted as $\overset{FS}{=}$) are as follows:

$$\begin{aligned}
\frac{\partial x^e_{t+1}}{\partial x_{t-1}} =& \alpha n^{x,AD}_t + (1+\tau_1) n^{x,WT}_t + (1+\tau_2) n^{x,ST}_t \\
&+ x_{t-1} \frac{\partial \left(\alpha n^{x,AD}_t + (1+\tau_1) n^{x,WT}_t + (1+\tau_2) n^{x,ST}_t \right)}{\partial x_{t-1}} \\
&- x_{t-2} \frac{\partial (\tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t)}{\partial x_{t-1}} + (1-\alpha) x^{AD}_t \frac{\partial n^{x,AD}_t}{\partial x_{t-1}} \\
\overset{FS}{=}& 0.25 \left(\alpha + 2 + \tau_1 + \tau_2 \right) \equiv \Gamma_1 \\
\frac{\partial x^e_{t+1}}{\partial x_{t-2}} =& x_{t-1} \frac{\partial \left(\alpha n^{x,AD}_t + (1+\tau_1) n^{x,WT}_t + (1+\tau_2) n^{x,ST}_t \right)}{\partial x_{t-2}} - (\tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t) \\
&- x_{t-2} \frac{\partial (\tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t)}{\partial x_{t-1}} + (1-\alpha) x^{AD}_t \frac{\partial n^{x,AD}_t}{\partial x_{t-1}} \\
\overset{FS}{=}& - 0.25 (\tau_1 + \tau_2) \equiv \Gamma_2 \\
\frac{\partial x^e_{t+1}}{\partial x^{AD}_t} =& x_{t-1} \frac{\partial \left(\alpha n^{x,AD}_t + (1+\tau_1) n^{x,WT}_t + (1+\tau_2) n^{x,ST}_t \right)}{\partial x_{t-1}} - x_{t-2} \frac{\partial (\tau_1 n^{x,WT}_t + \tau_2 n^{x,ST}_t)}{\partial x_{t-1}} \\
&+ (1-\alpha) n^{x,AD}_t + (1-\alpha) x^{AD}_t \frac{\partial n^{x,AD}_t}{\partial x^{AD}_t} \overset{FS}{=} 0.25 (1-\alpha) \equiv \Gamma_3,
\end{aligned}$$

and for any other lag variable $\hat{z}$ from the state space

$$\begin{aligned}
\frac{\partial x^e_{t+1}}{\partial \hat{z}} =& x_{t-1} \frac{\partial \left(\alpha n^{AD}_t + (1+\tau_1) n^{WT}_t + (1+\tau_2) n^{ST}_t \right)}{\partial \hat{z}} \\
&- x_{t-2} \frac{\partial (\tau_1 n^{WT}_t + \tau_2 n^{ST}_t)}{\partial \hat{z}} + (1-\alpha) x^{AD}_t \frac{\partial n^{AD}_t}{\partial \hat{z}} \overset{FP}{=} 0.
\end{aligned}$$

Next for any heuristic $h \in \{AD, WT, ST, FD\}$ and any $x \in \{\pi, y\}$

$$
\begin{aligned}
\frac{\partial U_t^{x,h}}{\partial U_{t-1}^{x,h}} =& \rho, \\
\frac{\partial x_{t+1}^{AD}}{\partial x_{t-1}} =& \alpha, \quad \frac{\partial x_{t+1}^{AD}}{\partial x_t^{AD}} = 1-\alpha, \\
\frac{\partial U_t^{x,AD}}{\partial x_{t-1}} =& -2x_{t-1} + 2x_{t-1}^{AD} \overset{FP}{=} 0, \quad \frac{\partial U_t^{x,AD}}{\partial x_t^{AD}} = 2x_{t-1} - 2x_t^{AD} \overset{FP}{=} 0, \\
\frac{\partial U_t^{x,FD}}{\partial x_{t-1}} =& -2x_{t-1} \overset{FP}{=} 0, \\
\frac{\partial U_t^{x,WT}}{\partial x_{t-1}} =& -2(x_{t-1} - x_{t-3} - \tau_1(x_{t-3} - x_{t-4})) \overset{FP}{=} 0 \\
\frac{\partial U_t^{x,WT}}{\partial x_{t-2}} =& 0, \\
\frac{\partial U_t^{x,WT}}{\partial x_{t-3}} =& -2(x_{t-1} - x_{t-3} - \tau_1(x_{t-3} - x_{t-4}))(-1-\tau_1) \overset{FP}{=} 0, \\
\frac{\partial U_t^{x,WT}}{\partial x_{t-4}} =& -2(x_{t-1} - x_{t-3} - \tau_1(x_{t-3} - x_{t-4}))\tau_1 \overset{FP}{=} 0,
\end{aligned}
$$

similarly for the strong trend heuristic and any other lag variable $\hat{z}$ from the state space

$$
\begin{aligned}
\frac{\partial U_t^{x,ST}}{\partial x_{t-s}} =& \overset{FP}{=} 0 \quad \text{for any } s \in \{1,2,3,4\}, \\
\frac{\partial U_t^{x,h}}{\partial \hat{z}} =& 0.
\end{aligned}
$$

Therefore, the Jacobian of the system at the steady state becomes

$$
J^4\left(x^f\right) \equiv \begin{pmatrix} J_{11\times 11}^4 & \mathbf{0}_{8\times 8} \\ \mathbf{0}_{8\times 8} & \rho \mathbb{I}_{8\times 8} \end{pmatrix},
$$

where

$$J^4_{11\times 11} = \begin{pmatrix} \omega_1\Gamma_1 & \omega_1\Gamma_2 & 0 & 0 & \omega_2\Gamma_1+\omega_3 & \omega_2\Gamma_2 & 0 & 0 & \omega_4 & \omega_1\Gamma_3 & \omega_2\Gamma_3 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \zeta_1\Gamma_1 & \zeta_1\Gamma_2 & 0 & 0 & \zeta_2\Gamma_1+\zeta_3 & \zeta_2\Gamma_2 & 0 & 0 & \zeta_4 & \zeta_1\Gamma_3 & \zeta_2\Gamma_3 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \omega_1\Gamma_1 & \omega_1\Gamma_2 & 0 & 0 & \omega_2\Gamma_1+\omega_3 & \omega_2\Gamma_2 & 0 & 0 & \omega_4+1 & \omega_1\Gamma_3 & \omega_2\Gamma_3 \\ \alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-\alpha & 0 \\ 0 & 0 & 0 & 0 & \alpha & 0 & 0 & 0 & 0 & 0 & 1-\alpha \end{pmatrix}.$$

We can eliminate the fourth and eight columns (since they contain only zeros), which in turn allows us to drop the third and the seventh columns. Further more, J^4 has 8 trivial eigenvalues equal to $\lambda_{12:19} = \rho$, hence

$$\det\left(J^4\left(x^f\right) - \lambda\mathbb{I}\right) = \lambda^2\left(\rho-\lambda\right)^8 \det\left[J_{4T} - \lambda\mathbb{I}\right],$$

where

$$J_{4T} = \begin{pmatrix} \omega_1\Gamma_1 & \omega_1\Gamma_2 & \omega_2\Gamma_1+\omega_3 & \omega_2\Gamma_2 & \omega_4 & \omega_1\Gamma_3 & \omega_2\Gamma_3 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \zeta_1\Gamma_1 & \zeta_1\Gamma_2 & \zeta_2\Gamma_1+\zeta_3 & \zeta_2\Gamma_2 & \zeta_4 & \zeta_1\Gamma_3 & \zeta_2\Gamma_3 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \omega_1\Gamma_1 & \omega_1\Gamma_2 & \omega_2\Gamma_1+\omega_3 & \omega_2\Gamma_2 & \omega_4+1 & \omega_1\Gamma_3 & \omega_2\Gamma_3 \\ \alpha & 0 & 0 & 0 & 0 & 1-\alpha & 0 \\ 0 & 0 & \alpha & 0 & 0 & 0 & 1-\alpha \end{pmatrix}. \tag{D.2}$$

□

E Determinacy and stability conditions for forward, backward and contemporaneous PLT rules

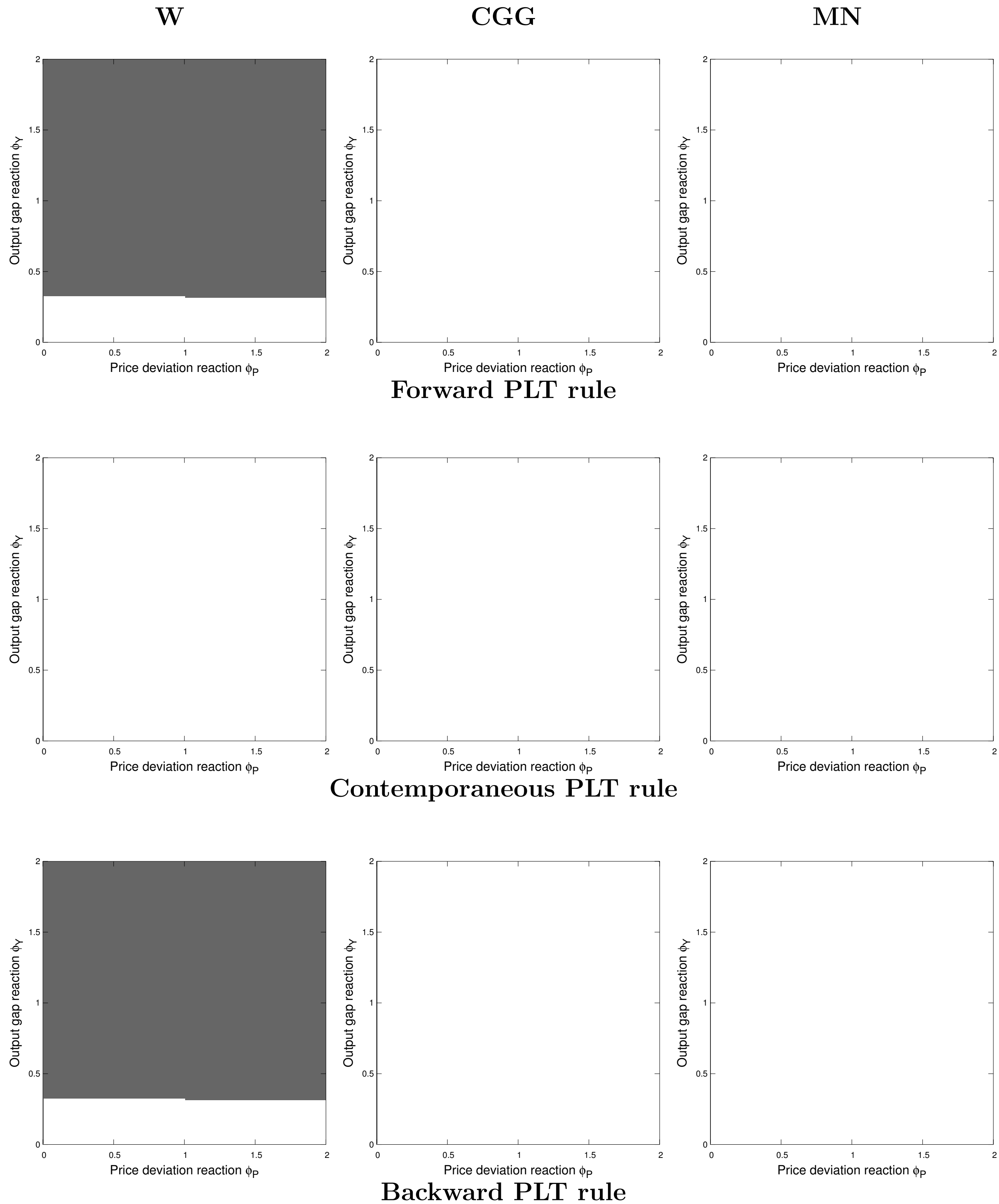


Figure 10: Heatmap: Determinacy under **RE** of the NK economy with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote respectively determinate and indeterminate solution under RE.

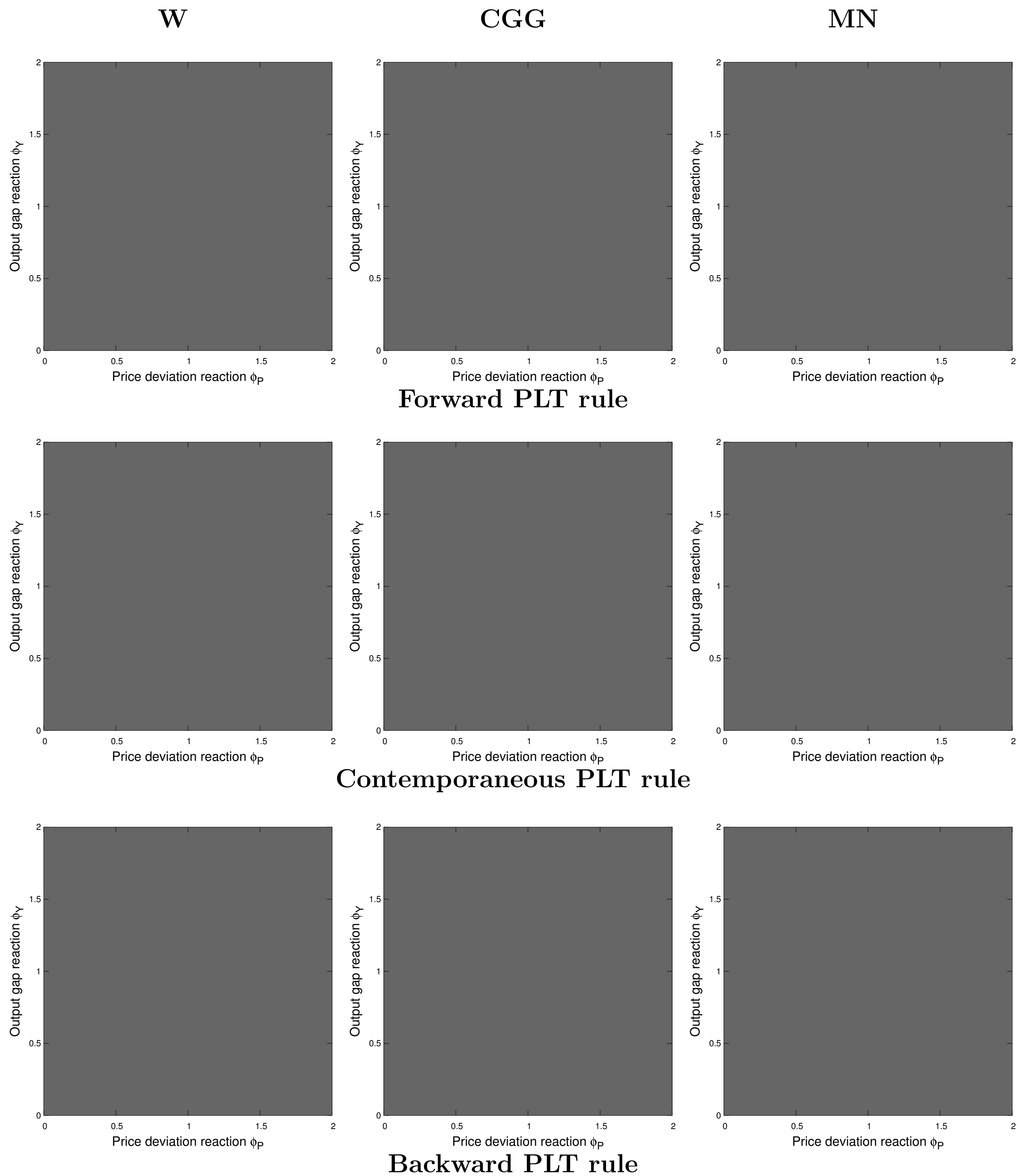


Figure 11: Heatmap: Stability of the NK economy under **naive benchmark** with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote stable and unstable eigenvalues respectively.

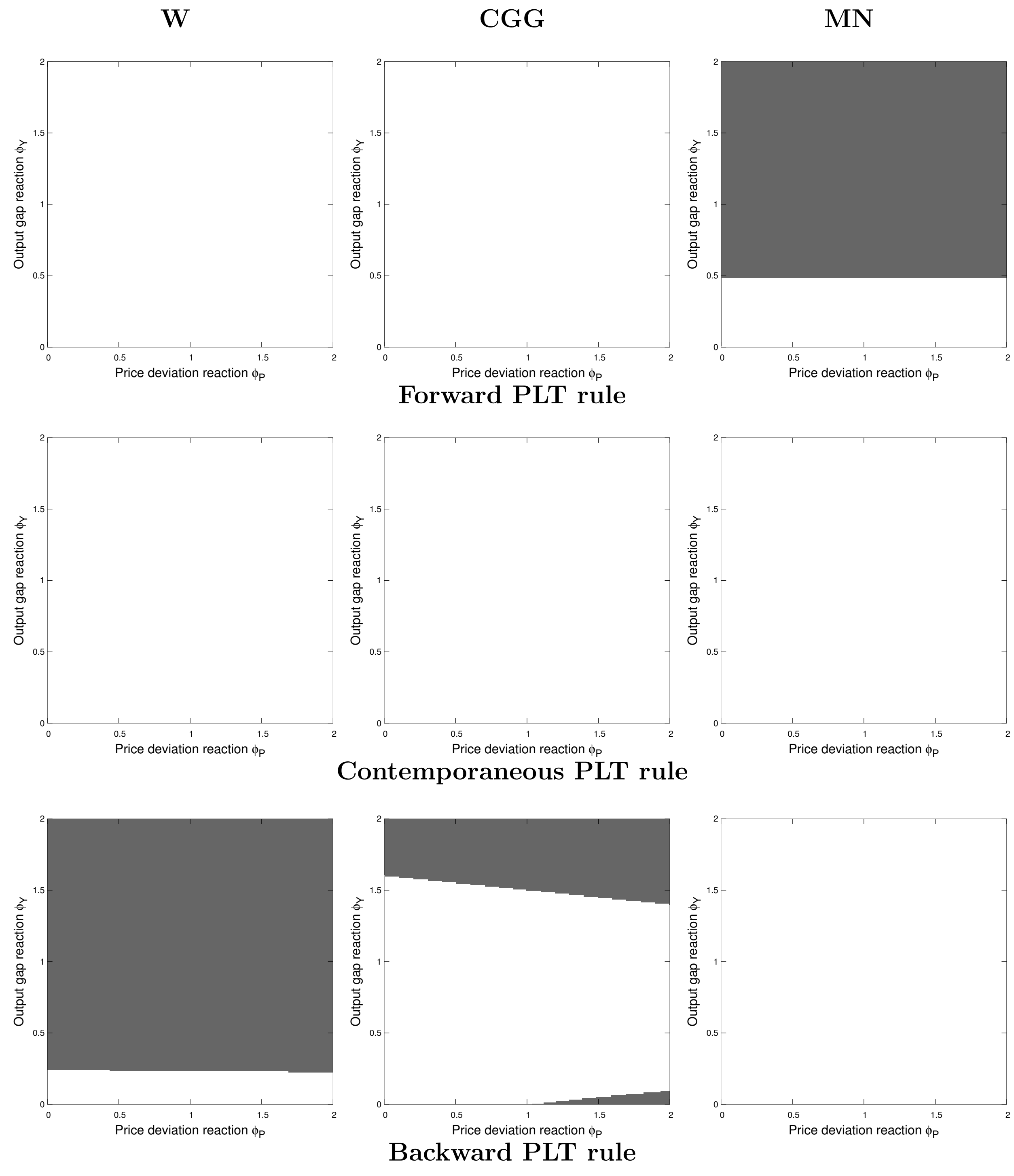


Figure 12: Heatmap: Stability of the NK economy the **2-type HSM** with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote stable and unstable eigenvalues respectively.

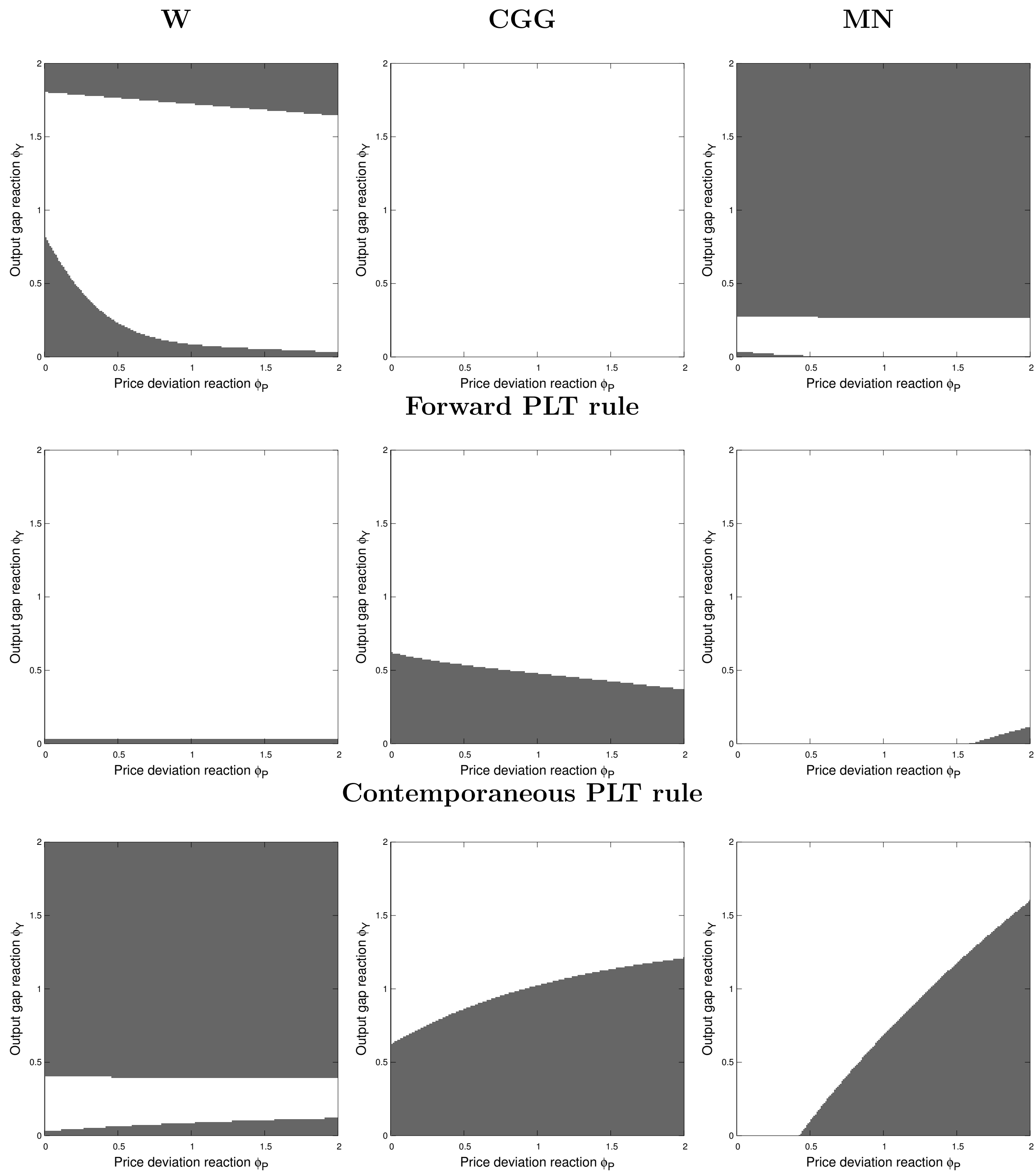


Figure 13: Heatmap: Stability of the NK economy the **4-type HSM** with PLT Taylor rule as a function of policy reaction parameters ϕ_p and ϕ_y. White and gray colors denote stable and unstable eigenvalues respectively.