

Sebastian Sienknecht

Friedrich-Schiller-University Jena, Germany

Inflation persistence amplification in the Rotemberg model**Abstract:**

This paper estimates a Dynamic Stochastic General Equilibrium (DSGE) model for the European Monetary Union using Bayesian techniques. A salient feature of the model is an extension of the typically postulated quadratic adjustment cost structure to amplify the persistence of inflation. The enlargement of the original formulation by Rotemberg (1982) and Hairault and Portier (1993) leads to a structurally more sophisticated hybrid inflation schedule than in the staggering environment of Calvo (1983). In particular, a desired lagged inflation term arises together with a two-period-ahead inflation expectation. The two terms are linked by a structural parameter. We confront this inflation schedule with European data and conclude that it has the potential to outperform its hybrid Calvo counterpart in terms of marginal likelihoods.

Keywords: Inflation Dynamics, New Keynesian Phillips Curve, Business Fluctuations Bayesian, Indexation, Model Comparison.

JEL Classification: E30, E31, E52, E58

1 Introduction

The staggered price-setting according to Calvo (1983) is the most common approach to rule out the neutrality of monetary policy. Similarly to the nominal wage-contracting environment suggested by Taylor (1979), an inertial adjustment of prices emerges as a salient outcome. Equivalent conclusions are derived by assuming the existence of menu costs in nominal adjustment (Rotemberg (1982) and Hairault and Portier (1993)). However, none of these rigidity specifications is able to replicate the highly amplified inflation persistence typically observed in empirical data. Most importantly, the theoretical fit towards observed fundamentals of inflation always involves an extension of the non-neutrality at work. For instance, the Calvo version is usually modified as in Galí et al. (2001) or Christiano et al. (2005) to include rule-of-thumb setters. This means that a fraction of monopolistic agents is forced to index their nominal choice variable to nominal conditions in the past, as they are disallowed to decide optimally. Therefore, the induced endogeneity in inflation variables is due to frictions in the nominal variable only. In contrast, inflation persistence amplification under Taylor contracting and under adjustment costs typically requires real frictions. A standard procedure in the Taylor-type setting is to introduce real-valued wage contracts as done by Fuhrer and Moore (1995), while the adjustment cost environment usually requires additional adjustment costs in input utilization or rule-of-thumb indexation.¹

¹ See for example Lechthaler and Snower (2008) for a very simple specification of labor adjustment costs. While these are embedded in a Calvo setting, the same inflation inertia can be generated by assuming quadratic adjustment costs in the pricing decision instead.

The purpose of our analysis is to show that inflation persistence amplification can be generated under adjustment costs by assuming quadratic rigidities in price adjustment only. To this end, the adjustment cost structure of Rotemberg (1982) and Hairault and Portier (1993) is extended. We show that the resulting New Keynesian Phillips curve displays a higher degree of complexity in terms of expectations than the rule-of-thumb Calvo setting. On the aggregate level, our modified cost structure leads to a lagged inflation term connected with an additional and unavoidable two-period-ahead inflation expectation. This connection is direct, through a structural parameter. Naturally, the question about the plausibility of this structure arises because an additional two-period-ahead inflation expectation is rather unusual. We argue that the importance of extended intertemporal adjustment cost amounts could be taken into consideration, along with a quarterly parameter calibration. The latter makes the two-period-ahead expectational time horizon not questionable on theoretical grounds.

However, these theoretical foundations have to be complemented by sensible econometric results in order to verify the practical usefulness of the implied model. We pursue two main objectives by taking selected time series for the European Monetary Union into account. In a first step, a New Keynesian model containing our extended inflation curve is estimated using Bayesian methods. The Bayesian estimation methodology allows for an estimation of model parameters taking all economic relationships simultaneously into consideration. Moreover, it allows for an intuitive comparison between different models with respect to their empirical fit. This leads us to our second step: two further models are estimated as a point of reference (as in Rabanal and Rubio-Ramírez (2005)). One of them is the framework with a purely forward-looking New Keynesian Phillips curve as postulated by Rotemberg (1982) and Hairault and Portier (1993). The remaining reference model entails a hybrid inflation schedule resulting from a standard Calvo environment with rule-of-thumb setters. We estimate the reference models in order to rank our extended adjustment cost model against them. The ranking criterion is the ability of a model to fit the observed data series in terms of marginal likelihoods.

The remainder is as follows. Section 2 motivates and introduces the extended adjustment cost structure. Moreover, an implementation in the microeconomic optimization of intermediate firms is presented. Section 3 presents a standard internal habit formation setting in order to generate the empirically relevant output persistence amplification. Section 4 presents the abovementioned linearized New Keynesian models differing in their inflation curves. Section 5 describes the dynamics in the model containing the novel New Keynesian Phillips curve. Section 6 presents the data and compares estimation results under the three models. Section 7 summarizes and concludes.

2 Adjustment Cost Structure Revisited

The starting point of our analysis is the adjustment cost structure by Rotemberg (1982) and Hairault and Portier (1993), which is essential for the standard New Keynesian Phillips curve. This adjustment cost structure leads ultimately to a log-linear New Keynesian Phillips curve of the form $\hat{\pi}_t^p = \beta E_t [\hat{\pi}_{t+1}^p] + \gamma (\widehat{MC}_t^f + \hat{\mu}_{p,t})$, with $0 < \beta < 1$, $\gamma > 0$, and $\hat{\pi}_t^p$ as the price inflation rate, $\widehat{MC}_t^f$ as a real marginal cost measure of the firm sector f , and $\hat{\mu}_{p,t}$ as a cost-push variable. This equation is a purely *forward-looking* representation of price inflation dynamics. In the absence of any other *backward-looking* (i.e. predetermined) model element, the only source of persistent deviations from

the steady state is the autoregressive process following a shock.² The resulting impulse response of inflation then shows a monotonous time path, which is completely determined by the monotonous AR(1) development of the exogenous shock variable. This is precisely a major critical point of the basic New Keynesian model: a monotonous impulse response of price inflation is highly at odds with the empirical evidence. As shown by the large number of estimated VAR models, the persistence of price inflation is highly amplified.³ Graphically, this amplification effect materializes in a hump-shaped impulse response after a shock. Formally, a lagged inflation component of a VAR model is found to be significant, and therefore important for the determination of price inflation. This has led theoretical contributions to explore for microeconomic foundations that result in a “hybrid” New Keynesian Phillips curve, which is an inflation schedule of the general form $\hat{\pi}_t^p = \phi \hat{\pi}_{t-1}^p + (1 - \phi) E_t [\hat{\pi}_{t+1}^p] + \tilde{\gamma} (\widehat{MC}_t^f + \hat{\mu}_{p,t})$, $0 < \phi < 1$. It represents an inflation schedule that emphasizes not only a backward-looking inflation term $\hat{\pi}_{t-1}^p$, but also maintains the notion of forward-looking firms with rational expectations.⁴

Our efforts are targeted to a microeconomic reformulation of the adjustment cost structure by Rotemberg (1982) and Hairault and Portier (1993), such that a hybrid New Keynesian Phillips curve arises on the aggregate level. In other words, we search for an economically sensible way to generate a backward-looking term that does *not* rely on additional assumptions, such as capital utilization costs (Lechthaler and Snower (2008)) or price indexation (Galí et al. (2001) and Christiano et al. (2005)). We restrict the context for this search by the set of requirements presented in table 1. They ensure that no other property of the basic structure other than generating the backward-looking inflation term is destroyed. In a strict sense, these requirements have to be fulfilled in order to obtain a hybrid New Keynesian Phillips curve that can be embedded in a standard New Keynesian framework.

Table 1: Requirements for an extension of the basic adjustment cost structure.

-
1. Adjustment costs vanish at the steady state.
 2. Inflation weights add up to one.
 3. The respective first-order condition can be expressed in stationary change rates.
-

The first requirement in table 1 states that adjustment costs should constitute only short run inefficiencies that allow for the short-run non-neutrality of monetary policy (i.e. rigid prices). The second requirement implies that the purely forward-looking schedule should be embedded in the hybrid form. The third requirement demands for aggregate change rates (e.g., price inflation) and the avoidance of aggregate levels (e.g., price level). We first present a general formulation which fulfills these criteria. The inertial development in the change rate of a nominal variable X_t is driven by real adjustment costs of the form:

² The term “persistence” can be defined as the number of periods a variable needs to return to its steady state. In case of log-deviations (in log-linearized models), “persistence” is the time needed until these deviations are zero.

³ This “inertia” of inflation is a well-known stylized fact, see, e.g., Christiano et al. (2005), pp. 4-8 and Peersman and Smets (2001).

⁴ For an overview of the microeconomic extensions undertaken so far, see the references given in section 1.

$$Q_{t+k}^x = \frac{\psi_x}{2} \left(\frac{X_{t+k}}{X_{t+k-1}} - \frac{X}{X} \right)^2 + \frac{v_x}{2} \left(\frac{X_{t+k}}{X_{t+k-1}} - \frac{X_{t+k-1}}{X_{t+k-2}} \right)^2 \quad \text{for } k = 0, 1, \dots; \psi_x, v_x > 0, \quad (1)$$

where X_{t+k} is the choice variable of a monopolistic economic agent not explicitly indexed and X denotes the steady state level of this variable. The suppressed agent indexation makes clear that (1) is also an aggregate formulation due to symmetry assumptions. Notice that setting $\psi_x > 0$ and $v_x = 0$ leads to the basic adjustment cost structure by Rotemberg (1982), pp. 521-523 and Hairault and Portier (1993), p. 1540. That is, our extension is represented by $v_x > 0$, in conjunction with $\psi_x > 0$. When choosing X_t , the monopolistic economic agent optimizes its target function subject to (1) and subject to its own demand function. However, the basic structure ($\psi_x > 0$, $v_x = 0$) already forces the agent to adopt an intertemporal view. This can be visualized by expressing (1) as a rationally expected, asymptotic, and discounted stream of adjustment costs:

$$E_t \sum_{k=0}^{\infty} \Delta_{t,t+k} Q_{t+k}^x, \quad (2)$$

where $\Delta_{t,t+k}$ is the stochastic discount factor. When choosing X_t under $\psi_x > 0$ and $v_x = 0$, the agent takes $k = 0$ and $k = 1$ into account:

$$\Delta_{t,t} Q_t^x + E_t [\Delta_{t,t+1} Q_{t+1}^x] = \Delta_{t,t} \frac{\psi_x}{2} \left(\frac{X_t}{X_{t-1}} - 1 \right)^2 + E_t \left[\Delta_{t,t+1} \frac{\psi_x}{2} \left(\frac{X_{t+1}}{X_t} - 1 \right)^2 \right]. \quad (3)$$

That is, adjustment costs arise in case that X_t differs from X_{t-1} and the (expected and discounted) level X_{t+1} . Therefore, the choice of X_t will be such that the costs dispersing over $k = 0$ and $k = 1$ are minimized *intertemporally*. An important insight from (3) can be gained by assuming $X_t \equiv P_t$: the gross change rates $\frac{X_t}{X_{t-1}}$ and $\frac{X_{t+1}}{X_t}$ correspond to the inflation terms of the purely forward-looking New Keynesian Phillips curve $\ln(P_t) - \ln(P_{t-1}) \equiv \hat{\pi}_t^p$ and $E_t [\ln(P_{t+1}) - \ln(P_t)] = E_t [\hat{\pi}_{t+1}^p]$. Our extension ($\psi_x > 0$, $v_x > 0$) expands the adjustment cost by $k = 2$:

$$\begin{aligned} & \Delta_{t,t} Q_t^x + E_t [\Delta_{t,t+1} Q_{t+1}^x] + E_t [\Delta_{t,t+2} Q_{t+2}^x] \\ &= \Delta_{t,t} \frac{\psi_x}{2} \left(\frac{X_t}{X_{t-1}} - 1 \right)^2 + \Delta_{t,t} \frac{v_x}{2} \left(\frac{X_t}{X_{t-1}} - \frac{X_{t-1}}{X_{t-2}} \right)^2 \\ &+ E_t \left[\Delta_{t,t+1} \frac{\psi_x}{2} \left(\frac{X_{t+1}}{X_t} - 1 \right)^2 \right] + E_t \left[\Delta_{t,t+1} \frac{v_x}{2} \left(\frac{X_{t+1}}{X_t} - \frac{X_t}{X_{t-1}} \right)^2 \right] \\ &+ E_t \left[\Delta_{t,t+2} \frac{\psi_x}{2} \left(\frac{X_{t+2}}{X_{t+1}} - 1 \right)^2 \right] + E_t \left[\Delta_{t,t+2} \frac{v_x}{2} \left(\frac{X_{t+2}}{X_{t+1}} - \frac{X_{t+1}}{X_t} \right)^2 \right], \end{aligned} \quad (4)$$

where $\frac{X_{t-1}}{X_{t-2}}$ would correspond to the desired backward-looking term $\hat{\pi}_{t-1}^p$. However, note that also a term $\frac{X_{t+2}}{X_{t+1}}$ linked to $E_t [\hat{\pi}_{t+2}^p]$ arises. This is simply because adjustment costs now also arise when X_{t-1} differs from X_{t-2} and X_{t+1} differs from X_{t+2} . Moreover, the agent will have to form rational expectations with respect to $t + 2$, given the information set at period t .

Of course, one could ask why our specific adjustment cost extension in (1) should be chosen in the light of possible alternative formulations. We find that $\frac{X_{t+k-1}}{X_{t+k-2}}, \forall k \geq 0$ is the critical component that is needed to induce a backward inflation term, i.e. a term that fulfills the third requirement of table

1. Note that an inclusion of this term inside the bracket of the baseline expression $\frac{\psi_x}{2} \left(\frac{X_{t+k}}{X_{t+k-1}} - 1 \right)^2$ is not possible as $\psi_x = 0$ would eliminate *all* inflation terms. This would contradict the second requirement of table 1. The problem cannot be solved by including a weighting parameter v_x inside the bracket. Ultimately, the expression containing $\frac{X_{t+k-1}}{X_{t+k-2}}, \forall k \geq 0$ has to be *independent* from the baseline term. In other words, the extension will have to be introduced additively outside the baseline term and with a neutralizing parameter $v_x = 0$. Another striking realization is that this separate formulation cannot be analogous to the baseline term, e.g. $\frac{v_x}{2} \left(\frac{X_{t+k-1}}{X_{t+k-2}} - 1 \right)^2$. This extension would also not be compatible with the requirements of table 1. After exploring all potential extensions, we find that the formulation (1) is the *only* expression that is in line with table 1.

In the following, the extended adjustment cost structure (1) is embedded in a New Keynesian framework. However, we aim to keep the model as small as possible, in order to demonstrate its potential explanatory power for *price* inflation.

3 Amplification of Price Inflation Persistence

Each commodity $Y_t(i)$ is produced by a single firm with a standard decreasing return to labor technology. Moreover, there is a technology variable A_t following a first-order autoregressive process. The real marginal costs of a firm i are given by⁵

$$MC_t^f(i) = \left(\frac{1}{1-\alpha} \right) \frac{W_t}{P_t} Y_t(i)^{\frac{\alpha}{1-\alpha}} A_t^{-\frac{1}{1-\alpha}}. \quad (5)$$

A firm chooses its own price $P_t(i)$ in order to maximize monopolistic real profits, subject to its own demand schedule and to the appearance of adjustment costs as formulated under section 1. Specifically, the adjustment cost structure is

$$Q_t^p(i) = \frac{\psi_p}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} - \pi^p \right)^2 + \frac{v_p}{2} \left(\frac{P_t(i)}{P_{t-1}(i)} - \frac{P_{t-1}(i)}{P_{t-2}(i)} \right)^2, \quad \psi_p, v_p > 0, \quad (6)$$

where $v_p \neq 0$ implies that the extension of the traditional adjustment cost structure is at work. The expected sum of real profits of an intermediate firm is written in a standard way. The first-order derivative with respect to $P_t(i)$ leads to the optimal price setting from the viewpoint of firm i :

⁵ This follows from the production function $Y_t(i) = A_t N_t(i)^{1-\alpha}$ and the total real costs $\frac{W_t}{P_t} N_t(i)$, where $\frac{W_t}{P_t}$ is the real wage, $N_t(i)$ represents labor hours, and A_t is a technology parameter. See Galí (1999), p. 252 for further details.

$$\begin{aligned}
 \left(\frac{P_t(i)}{P_{t-1}(i)} - \pi^p \right) \frac{P_t(i)}{P_{t-1}(i)} = E_t \left[\frac{\Delta_{t,t+1}}{\Delta_{t,t}} \left(\left(\frac{P_{t+1}(i)}{P_t(i)} - \pi_p \right) \frac{P_{t+1}(i)}{P_t(i)} \right. \right. \\
 \left. \left. + \frac{v_p}{\psi_p} \left(\frac{P_{t+1}(i)}{P_t(i)} - \frac{P_t(i)}{P_{t-1}(i)} \right) \left(\frac{P_{t+1}(i)}{P_t(i)} + \frac{P_t(i)}{P_{t-1}(i)} \right) \right) \right] \\
 - E_t \left[\frac{\Delta_{t,t+2}}{\Delta_{t,t}} \frac{v_p}{\psi_p} \left(\frac{P_{t+2}(i)}{P_{t+1}(i)} - \frac{P_{t+1}(i)}{P_t(i)} \right) \frac{P_{t+1}(i)}{P_t(i)} \right] \\
 - \frac{v_p}{\psi_p} \left(\frac{P_t(i)}{P_{t-1}(i)} - \frac{P_{t-1}(i)}{P_{t-2}(i)} \right) \frac{P_t(i)}{P_{t-1}(i)} \\
 + \frac{\epsilon_{p,t} Y_t(i)}{\psi_p} \left(MC_t^f(i) - \frac{1}{\mu_{p,t}} \frac{P_t(i)}{P_t} \right). \quad (7)
 \end{aligned}$$

Note that under full price flexibility ($\psi_p = v_p = 0$), an intermediate firm sets its price as a mark-up over its nominal marginal costs. By adopting the symmetry conditions $P_t(i) = P_t$ and $Y_t(i) = Y_t$:

$$\begin{aligned}
 \left(\frac{P_t}{P_{t-1}} - \pi^p \right) \frac{P_t}{P_{t-1}} = E_t \left[\frac{\Delta_{t,t+1}}{\Delta_{t,t}} \left(\left(\frac{P_{t+1}}{P_t} - \pi_p \right) \frac{P_{t+1}}{P_t} + \frac{v_p}{\psi_p} \left(\frac{P_{t+1}}{P_t} - \frac{P_t}{P_{t-1}} \right) \left(\frac{P_{t+1}}{P_t} + \frac{P_t}{P_{t-1}} \right) \right) \right] \\
 - E_t \left[\frac{\Delta_{t,t+2}}{\Delta_{t,t}} \frac{v_p}{\psi_p} \left(\frac{P_{t+2}}{P_{t+1}} - \frac{P_{t+1}}{P_t} \right) \frac{P_{t+1}}{P_t} \right] \\
 - \frac{v_p}{\psi_p} \left(\frac{P_t}{P_{t-1}} - \frac{P_{t-1}}{P_{t-2}} \right) \frac{P_t}{P_{t-1}} + \frac{\epsilon_{p,t} Y_t}{\psi_p} \left(MC_t^f - \frac{1}{\mu_{p,t}} \right). \quad (8)
 \end{aligned}$$

By the definition of the gross price inflation rate $\pi_{t+k}^p \equiv \left(1 + \frac{P_{t+k} - P_{t+k-1}}{P_{t+k-1}} \right) = \frac{P_{t+k}}{P_{t+k-1}}$ (for $k = 0, 1, \dots$), we arrive at the following equation:

$$\begin{aligned}
 (\pi_t^p - \pi^p) \pi_t^p = E_t \left[\frac{\Delta_{t,t+1}}{\Delta_{t,t}} \left((\pi_{t+1}^p - \pi_p) \pi_{t+1}^p + \frac{v_p}{\psi_p} (\pi_{t+1}^p - \pi_t^p) (\pi_{t+1}^p + \pi_t^p) \right) \right] \\
 - E_t \left[\frac{\Delta_{t,t+2}}{\Delta_{t,t}} \frac{v_p}{\psi_p} (\pi_{t+2}^p - \pi_{t+1}^p) \pi_{t+1}^p \right] - \frac{v_p}{\psi_p} (\pi_t^p - \pi_{t-1}^p) \pi_t^p + \frac{\epsilon_{p,t} Y_t}{\psi_p} \left(MC_t^f - \frac{1}{\mu_{p,t}} \right). \quad (9)
 \end{aligned}$$

This aggregate relationship is our *extended* purely nonlinear New Keynesian Phillips curve, which will be log-linearized in further sections.

4 Habit-Driven Output Inertia

The lack of persistence described above is not only present in price inflation, but also in real output. In order to replicate the amplified behavior of real output found in the data, we assume that household utility depends positively on a consumption level that is relative to an internal habit formation term. The internal habit formation behavior is formulated following Constantinides (1990), p. 522-523, Casares (2006), p. 877 and Erceg et al. (2000), p. 287:

$$U_{t+k} = \frac{(C_{t+k}(j) - h C_{t+k-1}(j))^{1-\sigma}}{1-\sigma} - \frac{(N_{t+k}(j))^{1+\eta}}{1+\eta} \quad \text{for } k = 0, 1, \dots, \quad (10)$$

where the parameter $0 < h < 1$ gives the degree of habit formation.⁶ The household chooses its optimal demand for consumption $C_t(j)$ and riskless bonds $B_t(j)$ by maximizing the expected discounted sum of (10) subject to the expected discounted sum of the following real budget constraint:

$$C_{t+k}(j) + \frac{B_{t+k}(j)}{P_{t+k}} = \frac{W_{t+k}}{P_{t+k}} N_{t+k}(j) + R_{t+k-1} \frac{B_{t+k-1}(j)}{P_{t+k}} + Div_{t+k}^f(j) \quad \text{for } k = 0, 1, \dots \quad (11)$$

Combining the resulting first-order conditions with respect to $C_t(j)$ and $B_t(j)$ leads to the consumption Euler equation of household j with internal habit formation. Aggregation delivers:

$$(C_t - hC_{t-1})^{-\sigma} - \beta E_t \left[((C_{t+1} - hC_t)^{-\sigma} - \beta h (C_{t+2} - hC_{t+1})^{-\sigma}) \frac{R_t}{\pi_{t+1}^p} + h (C_{t+1} - hC_t)^{-\sigma} \right] = 0, \quad (12)$$

which is a purely nonlinear Euler consumption equation with internal habits to be log-linearized in further sections. Note that setting $h = 0$ gives the purely forward-looking nonlinear Euler equation for consumption. Moreover, the discount factor ratios entering the nonlinear Phillips curve (9) are given by

$$\frac{\Delta_{t,t+1}}{\Delta_{t,t}} = \beta \left(\frac{\partial U_{t+1}/\partial C_{t+1}}{\partial U_t/\partial C_t} \right) = \beta \frac{(C_{t+1} - hC_t)^{-\sigma} - \beta h (C_{t+2} - hC_{t+1})^{-\sigma}}{(C_t - hC_{t-1})^{-\sigma} - \beta h (C_{t+1} - hC_t)^{-\sigma}} \quad (13)$$

and

$$\frac{\Delta_{t,t+2}}{\Delta_{t,t}} = \beta^2 \left(\frac{\partial U_{t+2}/\partial C_{t+2}}{\partial U_t/\partial C_t} \right) = \beta^2 \frac{(C_{t+2} - hC_{t+1})^{-\sigma} - \beta h (C_{t+3} - hC_{t+2})^{-\sigma}}{(C_t - hC_{t-1})^{-\sigma} - \beta h (C_{t+1} - hC_t)^{-\sigma}}. \quad (14)$$

Households also choose a labor supply $N_t(j)$, which maximizes (10) subject to (11). In aggregate terms, we obtain

$$\frac{W_t}{P_t} = \frac{N_t^\eta}{(C_t - hC_{t-1})^{-\sigma} - \beta h (C_{t+1} - hC_t)^{-\sigma}}, \quad (15)$$

which is another relationship to be log-linearized later on. Again, notice that setting $h = 0$ leads to a standard labor supply schedule $\frac{W_t}{P_t} = MRS_t = C_t^\sigma N_t^\eta$.

5 Linearized DSGE Models

This section outlines three different log-linear equation systems. Each collection of equations is characterized by a specific New Keynesian Phillips curve, while all relationships other than the inflation schedule remain identical across the models (as in Rabanal and Rubio-Ramírez (2005)). The baseline (BAC) model contains a purely forward-looking Phillips curve following Rotemberg (1982) and Hairault and Portier (1993). We compare the theoretical dynamics and the empirical

⁶ Notice that $h = 0$ implies utility streams from consumption levels. Further, habit formation is due to the household's own consumption, i.e. $hC_{t+k-1}(j)$ for $k = 0, 1, \dots$. An alternative assumption is the behavior with *external* habit formation, i.e. hC_{t+k-1} for $k = 0, 1, \dots$ (see Abel (1990), p. 38 and Campbell and Cochrane (1999), pp. 208-210). The main distinction is that the first-order derivative with respect to $C_t(j)$ and therefore the Euler consumption equations will differ across both assumptions. This is because the past consumption term is irrelevant for the choice of $C_t(j)$ in the case of external habits. The assumption of habit formation was mainly intended to remedy issues in asset pricing, see Grishchenko (2010), pp. 177-178 for a recent overview.

performance of this baseline specification to the hybrid (HAC) alternative postulated above (we set $v_p \neq 0$). Most importantly, we pursue a comparison of this newly introduced hybrid inflation curve against the standard Calvo (1983) schedule with rule-of-thumb setters (HPS). The latter is derived according to Galí et al. (2001).

Several relationships are common to the three models considered here. We log-linearize the nonlinear Euler equation for intertemporal consumption (equation (12)). Since aggregate demand is driven by consumption only ($\hat{Y}_t = \hat{C}_t$), we obtain the following IS curve:

$$\hat{Y}_t = \Theta_1 \hat{Y}_{t-1} + \Theta_2 E_t [\hat{Y}_{t+1}] - \Theta_3 E_t [\hat{Y}_{t+2}] - \Theta_4 (\hat{R}_t - E_t [\hat{\pi}_{t+1}^p]). \quad (16)$$

The variables $\Theta_1, \dots, \Theta_7$ are functions of deep model parameters:

$$\Theta_1 = \frac{h}{1 + h(1 + \beta h)}, \quad (17)$$

$$\Theta_2 = \frac{1 + \beta h(1 + h)}{1 + h(1 + \beta h)}, \quad (18)$$

$$\Theta_3 = \frac{\beta h}{1 + h(1 + \beta h)}, \quad (19)$$

$$\Theta_4 = \frac{(1 - h)(1 - \beta h)}{\sigma(1 + h(1 + \beta h))}. \quad (20)$$

Aggregate employment responds to the real wage following the first-order condition (15). In log-linear form, we obtain

$$\hat{W}_t - \hat{P}_t = \eta \hat{N}_t + \iota_1 \hat{Y}_t - \iota_2 \hat{Y}_{t-1} - \iota_3 E_t [\hat{Y}_{t+1}], \quad (21)$$

where

$$\iota_1 = \frac{\sigma(1 + \beta h^2)}{(1 - h)(1 - \beta h)}, \quad (22)$$

$$\iota_2 = \frac{\sigma h}{(1 - h)(1 - \beta h)}, \quad (23)$$

$$\iota_3 = \frac{\sigma \beta h}{(1 - h)(1 - \beta h)}. \quad (24)$$

Aggregate employment and real output are linked to one another through the production technology

$$\hat{N}_t = \frac{1}{1 - \alpha} (\hat{Y}_t - \hat{A}_t). \quad (25)$$

The log-linear real marginal cost schedule can be derived from equation (5) as

$$\widehat{MC}_t^f = \hat{W}_t - \hat{P}_t + \left(\frac{\alpha}{1 - \alpha} \right) \hat{Y}_t - \left(\frac{1}{1 - \alpha} \right) \hat{A}_t. \quad (26)$$

The monetary policy authority is assumed to follow a modified version of the interest rate rule by Smets and Wouters (2003), p. 1136 and Smets and Wouters (2007), p. 591:

$$\hat{R}_t = (1 - \phi) \left(\delta_\pi \hat{\pi}_t^p + \delta_y \left(\hat{Y}_t - \hat{Y}_t^n \right) \right) + \phi \hat{R}_{t-1} + \hat{e}_{R,t} + \delta_{\Delta\pi} \hat{\pi}_{t-1}^p + \delta_{\Delta y} \left(\hat{Y}_{t-1} - \hat{Y}_{t-1}^n \right). \quad (27)$$

Under $\delta_{\Delta\pi} = \delta_{\Delta y} = 0$, (27) collapses to a simple instrument rule with interest rate smoothing. The choice of (27) is driven by the fact that our estimation results are improved in comparison to the rules by Smets and Wouters (2003, 2007).⁷ The natural output level entering the output gap $\hat{Y}_t - \hat{Y}_t^n$ can be derived as

$$\hat{Y}_t^n = \frac{1 + \eta}{\sigma(1 - \alpha) + \eta + \alpha} \hat{A}_t. \quad (28)$$

The technology shock variable follows an AR(1) process:

$$\hat{A}_t = \rho_A \hat{A}_{t-1} + e_{a,t}, \quad \rho_A \in [0, 1), \quad e_{a,t} \sim N(0, \sigma_a^2). \quad (29)$$

The time-varying monopolistic mark-up entering the Phillips curves is inversely related to the price elasticity of demand $\hat{\epsilon}_{p,t}$, with ϵ_p as its steady state counterpart:

$$\hat{\mu}_{p,t} = -\frac{1}{\epsilon_p - 1} \hat{\epsilon}_{p,t}. \quad (30)$$

The price elasticity entering (30) follows a white noise process:

$$\hat{\epsilon}_{p,t} = e_{p,t}, \quad e_{p,t} \sim N(0, \sigma_p^2), \quad (31)$$

which is assumed because persistent movements in the observed inflation rate should be captured by the price rigidity parameters, rather than by an autoregressive shock persistence parameter. The interest rate shock variable entering (27) also follows a white noise process:

$$\hat{e}_{R,t} = e_{R,t}, \quad e_{R,t} \sim N(0, \sigma_R^2), \quad (32)$$

which is assumed because persistent movements in the observed instrument rate should be captured by the interest rate smoothing parameter ϕ , rather than by an autoregressive persistence parameter. The equations presented in this subsection are the common components of the three models. Following, e.g., Smets and Wouters (2003, 2007) and others, the empirical data involved in the estimations will contain a *detrended* series of real output which corresponds to the theoretical output gap entering the interest rate rule (27). We define the output gap as

$$\hat{X}_t \equiv \hat{Y}_t - \hat{Y}_t^n, \quad (33)$$

and eliminate $\hat{Y}_t$ from the common equations. Table 2 presents the common model block. This system needs to be closed by one of the inflation equations presented below.

⁷ If d is the first-difference operator, the formulation by Smets and Wouters (2003), p. 1136 reads

$\hat{R}_t = (1 - \phi) \left(\delta_\pi \hat{\pi}_t^p + \delta_y \left(\hat{Y}_t - \hat{Y}_t^n \right) \right) + \phi \hat{R}_{t-1} + \hat{e}_{R,t} + \delta_{\Delta\pi} d\hat{\pi}_t^p + \delta_{\Delta y} d \left(\hat{Y}_t - \hat{Y}_t^n \right)$, while Smets and Wouters (2007), p. 591 assume the same rule with $\delta_{\Delta\pi} = 0$. A much more sophisticated rule is assumed by Henzel et al. (2009), p. 274 and Hülsewig et al. (2009), p. 1314, who argue that the specific formulation of the interest rate rule is driven by the dynamics observed in the data. In our case, the rule (27) delivered slightly lower standard deviations of parameter estimates.

Table 2: The log-linear common model block in terms of the output gap. The equation system has to be closed by a New Keynesian Phillips curve.

1. IS curve	$\hat{X}_t = \Theta_1 \hat{X}_{t-1} + \Theta_2 E_t [\hat{X}_{t+1}] - \Theta_3 E_t [\hat{X}_{t+2}] - \Theta_4 (\hat{R}_t - E_t [\hat{\pi}_{t+1}^p]) - \hat{Y}_t^n + \Theta_1 \hat{Y}_{t-1}^n + \Theta_2 E_t [\hat{Y}_{t+1}^n] - \Theta_3 E_t [\hat{Y}_{t+2}^n]$	(16)
2. Labor supply	$\hat{W}_t - \hat{P}_t = \eta \hat{N}_t + \iota_1 \hat{X}_t - \iota_2 \hat{X}_{t-1} - \iota_3 E_t [\hat{X}_{t+1}] + \iota_1 \hat{Y}_t^n - \iota_2 \hat{Y}_{t-1}^n - \iota_3 E_t [\hat{Y}_{t+1}^n]$	(21)
3. Technology	$\hat{N}_t = \frac{1}{1-\alpha} \left(\hat{X}_t - \frac{\sigma(1-\alpha)+\alpha-1}{1+\eta} \hat{Y}_t^n \right)$	(25)
4. Real marginal costs	$\widehat{MC}_t^f = \hat{W}_t - \hat{P}_t + \left(\frac{\alpha}{1-\alpha} \right) \hat{X}_t - \frac{\sigma(1-\alpha)+\eta+\alpha-\alpha(1+\eta)}{(1-\alpha)(1+\eta)} \hat{Y}_t^n$	(26)
5. Interest rate rule	$\hat{R}_t = (1-\phi) \left(\delta_\pi \hat{\pi}_t^p + \delta_y \hat{X}_t \right) + \phi \hat{R}_{t-1} + \delta_{\Delta\pi} \hat{\pi}_{t-1}^p + \delta_{\Delta y} \hat{X}_{t-1} + \hat{\epsilon}_{R,t}$	(27)
6. Natural output	$\hat{Y}_t^n = \frac{1+\eta}{\sigma(1-\alpha)+\eta+\alpha} \hat{A}_t$	(28)
7. Technology process	$\hat{A}_t = \rho_A \hat{A}_{t-1} + e_{A,t} \quad , \quad \rho_A \in [0, 1) \quad , \quad e_{A,t} \sim N(0, \sigma_A^2)$	(29)
8. Price mark-up	$\hat{\mu}_{p,t} = -\frac{1}{\epsilon_p - 1} \hat{\epsilon}_{p,t}$	(30)
9. Cost-push shock	$\hat{\epsilon}_{p,t} = e_{p,t} \quad , \quad e_{p,t} \sim N(0, \sigma_p^2)$	(31)
10. Interest rate shock	$\hat{\epsilon}_{R,t} = e_{R,t} \quad , \quad e_{R,t} \sim N(0, \sigma_R^2)$	(32)

6 Baseline Adjustment Cost (BAC) Model

The first model version is given by the equations of table 2, together with the purely forward-looking New Keynesian Phillips curve due to costs of price adjustment for the monopolistic agent (Rotemberg (1982) and Hairault and Portier (1993)). This is simply the log-linearized version of the nonlinear Phillips curve (9) under the assumption $v_p = 0$. We obtain the following inflation schedule:

$$\hat{\pi}_t^p = \beta E_t [\hat{\pi}_{t+1}^p] + \gamma \left(\widehat{MC}_t^f + \hat{\mu}_{p,t} \right). \quad (34)$$

The degree of inflation reactivity γ is a function of deep parameters:

$$\gamma = \frac{(\epsilon_p - 1) Y}{\psi_p}, \quad (35)$$

where the steady state level of real output can be derived as

$$Y = \left(\frac{(1-\alpha)(1-\beta h)}{\mu_p (1-h)^\sigma} \right)^{\frac{1-\alpha}{\sigma(1-\alpha)+\alpha+\eta}}. \quad (36)$$

7 Hybrid Adjustment Cost (HAC) Model

The log-linear version of the nonlinear Phillips curve (9) is the core expression in our analysis. The inflation curve to be included into the system of table 2 is

$$\hat{\pi}_t^p = \gamma_1 \hat{\pi}_{t-1}^p + \gamma_2 E_t [\hat{\pi}_{t+1}^p] - \gamma_3 E_t [\hat{\pi}_{t+2}^p] + \gamma_4 \left(\widehat{MC}_t^f + \hat{\mu}_{p,t} \right), \quad (37)$$

where $\gamma_1, \dots, \gamma_4$ are functions of deep parameters:

$$\gamma_1 = \frac{v_p}{v_p(1+2\beta) + \psi_p}, \quad (38)$$

$$\gamma_2 = \frac{v_p(2+\beta)\beta + \psi_p\beta}{v_p(1+2\beta) + \psi_p}, \quad (39)$$

$$\gamma_3 = \frac{v_p\beta^2}{v_p(1+2\beta) + \psi_p}, \quad (40)$$

$$\gamma_4 = \frac{(\epsilon_p - 1)Y}{v_p(1+2\beta) + \psi_p}. \quad (41)$$

Note that this New Keynesian Phillips curve is not only hybrid, but also features a negative inflation expectation term two periods ahead. This is a direct result of the adjustment cost extension presented above. Moreover, the parameter v_p relates the last period's inflation rate $\hat{\pi}_{t-1}^p$ to the two-period-ahead expectation $E_t[\hat{\pi}_{t+2}^p]$ because $v_p = 0$ eliminates both terms and our inflation curve collapses to the traditional forward-looking New Keynesian Phillips curve (BAC). We also obtain the desirable property for the inflation weights $\gamma_1 + \gamma_2 + (-\gamma_3) = 1$. Overall, our expression (37) fulfills all requirements stated in table 1.

8 Hybrid Price Staggering (HPS) Model

The most widespread approach of nominal rigidity modeling is the staggered price-setting environment of Calvo (1983). As a common practice, purely forward-looking New Keynesian Phillips curves are transformed into hybrid versions by assuming a fraction $\omega_p \in (0, 1)$ of rule-of-thumb setters. We formulate this behavior in the spirit of Galí et al. (2001), p. 1247. In our notation, the following hybrid New Keynesian Phillips curve is obtained:

$$\hat{\pi}_t^p = \gamma_1^c \hat{\pi}_{t-1}^p + \gamma_2^c E_t[\hat{\pi}_{t+1}^p] - \gamma_3^c E_t[\hat{\pi}_{t+2}^p] + \gamma_4^c (\widehat{MC}_t^f + \hat{\mu}_{p,t}), \quad (42)$$

where the superscript ‘ c ’ denotes reaction parameters under the Calvo pricing assumption. Again, they depend on deep model parameters:

$$\gamma_1^c = \frac{\omega_p}{\omega_p(1 - \theta_p(1 - \beta)) + \theta_p}, \quad (43)$$

$$\gamma_2^c = \frac{\beta\theta_p}{\omega_p(1 - \theta_p(1 - \beta)) + \theta_p}, \quad (44)$$

$$\gamma_3^c = 0, \quad (45)$$

$$\gamma_4^c = \frac{(1 - \alpha)(1 - \beta\theta_p)(1 - \theta_p)(1 - \omega_p)}{(1 + \alpha(\epsilon_p - 1))(\theta_p + \omega_p(1 - \theta_p(1 - \beta)))}. \quad (46)$$

Note that in this well-known approach, an expectation of two periods ahead does not arise. We included the term $\gamma_3^c E_t[\hat{\pi}_{t+2}^p]$ with $\gamma_3^c = 0$ merely for the sake of comparability against the hybrid adjustment cost (HAC) case.

9 Dynamic Properties

In the following, we investigate the dynamic responses resulting in the HAC model after a shock. In particular, we investigate how the amplification of inflation persistence is influenced by the

parameter v_p . Note that because of $v_p = 0$, the BAC model is implicitly included and serves as a point of reference. In contrast, we will describe only analogies of the HPS model to the BAC/HAC setting, as the dynamics of the Calvo approach with rule-of-thumb behavior are well-known. The analysis is based on assigned (“calibrated”) parameter values which are usually assumed on a quarterly basis. We mostly adopt values widely found in, e.g., Smets and Wouters (2003), Casares (2006), and Taylor (1979). Table 3 gives an overview of the deep parameter values. They determine steady state levels and the composite parameter variables $\Theta_1, \dots, \Theta_4$, $\iota_1, \dots, \iota_3$, and $\gamma_1, \dots, \gamma_4$.⁸

Table 3: Calibrated parameters underlying theoretical responses in the HAC model.

$\alpha = 0.3$	$\beta = 0.99$	$\sigma = 1$	$\eta = 2$	$h = 0 \vee 0.7$	$\phi = 0 \vee 0.8$	$\rho_a = 0.8$
$\epsilon_p = 6$	$\psi_p = 170$	$v_p \in [0, 15000]$	$\delta_\pi = 1.7$	$\delta_y = 0.5$	$\delta_{\Delta\pi} = 0 \vee 0.1$	$\delta_{\Delta y} = 0 \vee 0.16$

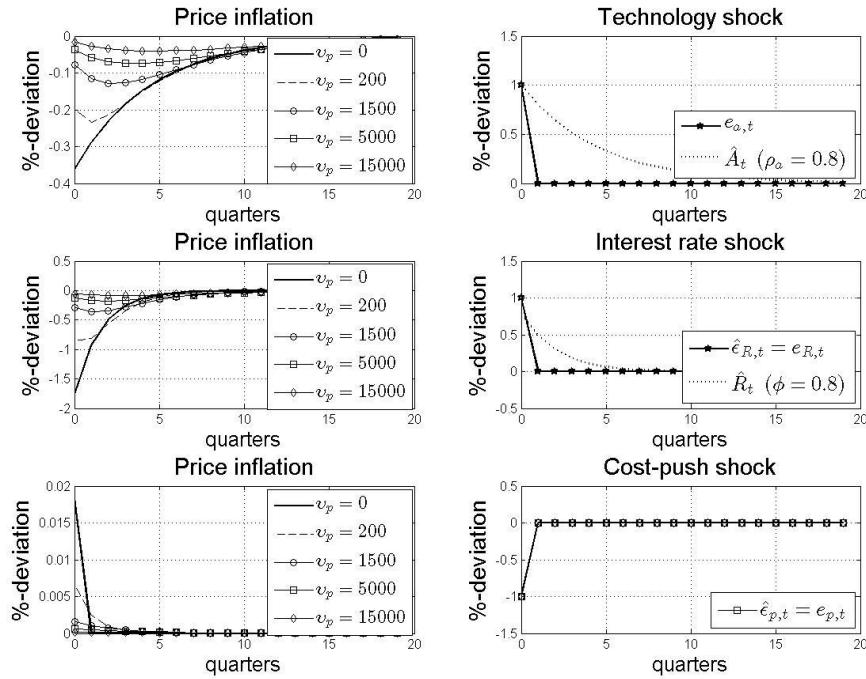
Note that the sources of persistence affecting *price inflation* after a shock are $h > 0$ (habit formation), $0 \leq \phi < 1$ (interest rate smoothing), $0 \leq \delta_{\Delta\pi} < 1$ (inflation reaction smoothing), $0 \leq \delta_{\Delta y} < 1$ (output gap reaction smoothing), $0 \leq \rho_a < 1$ (technology shock persistence), and the backward-looking inflation term, which is determined by $v_p \geq 0$. In order to isolate the role of this parameter, we simulate the shocks (items 7-10 in table 2) under $v_p \in [0, 1500]$ and $h = \delta_{\Delta\pi} = \delta_{\Delta y} = 0$. In the case of the technology and the cost-push shock, we eliminate an amplification of inflation through interest rate effects by setting $\phi = 0$.

The impulse responses after a positive *technology shock* are depicted in the first two panels of figure 1. We set $\sigma_a = 0.01$, which implies a one percent shock. This is visualized in the right subfigure by the fact that the shock variable $e_{a,t}$ increases by one percent, leading to an impact increase of the technology level $\hat{A}_t$ in the same magnitude. However, $\rho_a = 0.8$ implies that the one percent shock is persistent: since $\hat{A}_t$ is driven by the AR(1) process $\hat{A}_t = \rho_a \hat{A}_{t-1} + e_{a,t}$, the shock on the variable $\hat{A}_t$ lasts longer than the shock itself. Even if the shock $e_{a,t}$ is fully reverted in the following period $t = 1$, the variable $\hat{A}_t$ decreases only gradually towards the zero-deviation line. The impact increase (and the following dynamics) of $\hat{A}_t$ translate to price inflation (37) through an impact increase in natural output $\hat{Y}_t^n$ and, therefore, through an impact decrease in real marginal costs $\widehat{MC}_t^f$ (items 6 and 4 of table 2).

In the case that $v_p = 0$, the HAC model is converted into the purely forward-looking BAC setting. The monotonous AR(1) dynamics of $\hat{A}_t$ are in this case the only dynamics driving inflation. This is visualized by the left subfigure: price inflation decreases on impact and returns to the zero-deviation line monotonously from below. This is a major weakness of the BAC model because empirical impulse responses show a considerably damped response when the shock occurs and an amplified development of nominal and real variables in the following periods.

⁸ Note that the reaction parameter with respect to the output gap is $\delta_\pi = 0.5$, instead of $\delta_\pi = 0.5/4 = 0.125$. This is because we annualize the inflation rate such that $\hat{\pi}_t^{p,ann.} = \hat{\pi}_t^p * 4$. This is in line with Taylor (1993), p. 202 and with the New Keynesian empirical literature, see, e.g., Henzel et al. (2009) and Hülsewig et al. (2009).

Figure 1: Impulse responses of annualized price inflation in the HAC model under different values of the novel inflation persistence parameter.



The assumption $v_p > 0$ implies that $\hat{\pi}_{t-1}^p$, and $E_t [\hat{\pi}_{t+2}^p]$ are now relevant in the New Keynesian Phillips curve (37). The backward-looking inflation term $\hat{\pi}_{t-1}^p$ now becomes a source of endogenous inflation dynamics: even if the shock variable $\hat{A}_t$ entails no persistence (i.e. $\rho_a = 0$), inflation would show a persistent monotonic behavior. In other words, inflation is able to be persistent without any other exogenous persistence source. However, the characteristic behavior of inflation observed in VAR studies is generated by the *combination* of the exogenous and endogenous persistence sources. In the left top panel of figure 1, the exogenous transmission of monotonous persistence ($\rho_a = 0.8$) to inflation through $\hat{A}_t$ is overlaid by the endogenous dynamics of the backward-looking inflation term whenever $v_p > 0$. The impulse response of price inflation now shows typical properties observed in VAR studies, namely a damped jump reaction (decrease) in the shock period $t = 0$ and an amplified long-lasting reversal to the zero-deviation line. This is the distinctive “hump shape” widely observed in the empirical macroeconomics literature. Since $\frac{\partial \gamma_1}{\partial v_p} = \frac{\psi_p/v_p^2}{(1+2\beta+\psi_p/v_p)^2} > 0$, an increasing value of v_p implies that the weight γ_1 in (37) becomes larger, the term $\hat{\pi}_{t-1}^p$ more important, and the overlaying of the exogenous persistence source stronger. From the left top panel in figure 1, it should become clear that an increasing value of v_p increases the damping of the jump reaction and accentuates persistence (i.e. the number of periods until the zero-deviation line is reached increases). At the same time, the hump becomes flatter and wider. The effects connected with an increasing value of v_p are *qualitatively* the same as those connected with an increasing value of the respective parameter in the HPS model (42)-(46). In that setting, the parameter driving the persistence amplification is ω_p , which can be seen as a counterpart of v_p . However, in our setting an increasing value of v_p also increases (with a negative sign) the weight γ_3 on $E_t [\hat{\pi}_{t+2}^p]$, which is equal to zero in the HPS model. Therefore, we expect *quantitative* differences, which we intend to clarify in the estimations presented below.

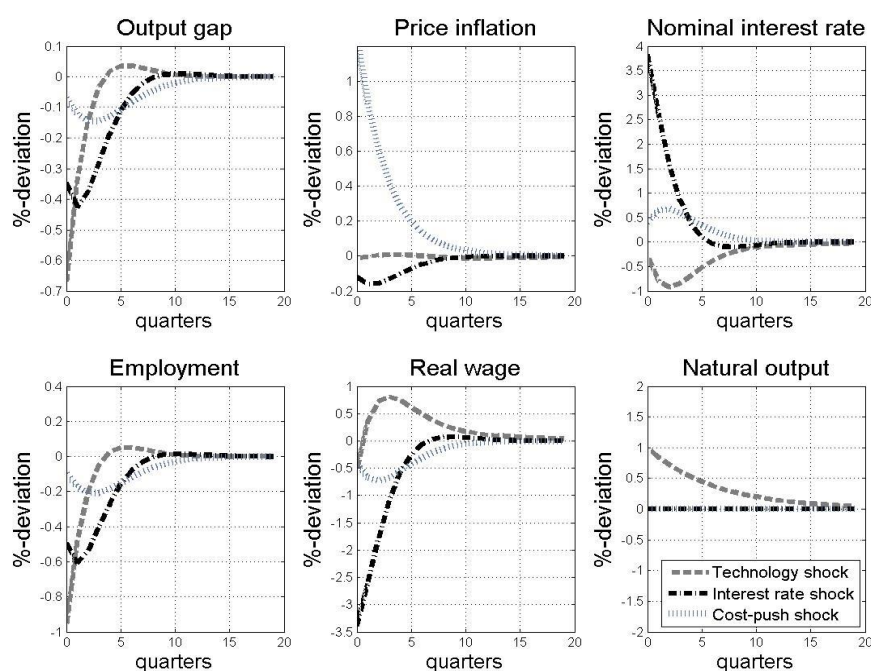
The middle panels of figure 1 show impulse responses after a one percent ($\sigma_R = 0.01$) positive *interest rate shock*. In this case, the positive smoothing parameter $\phi = 0.8$ induces a persistent increase of the interest rate $\hat{R}_t$ that dissipates monotonously. This, in turn, translates to inflation through an impact decrease of the output gap $\hat{X}_t$ which decreases real marginal costs $\widehat{MC}_t^f$ on impact (see items 1 and 4 of table 2). Again, $v_p = 0$ (BAC model) causes inflation to reflect completely the monotonous behavior of the interest rate, while $v_p > 0$ is connected to a hump-shaped impulse response.

The bottom panels of figure 1 show impulse responses under a negative one percent cost-push shock, i.e. $\sigma_p = -0.01$. This leads to an impact increase of inflation (37) via an impact increase of the mark-up $\hat{\mu}_{p,t}$ (see item 8 of table 2). Note that, as there is no exogenous source of persistence in the shock variable $\hat{\epsilon}_{p,t}$, the only possibility of inflation to be persistent is $v_p > 0$.⁹ Here, the isolated workings of persistence endogeneity are clearly visible: a positive value of v_p leads to a monotonous impulse response of inflation. Moreover, an increasingly higher v_p increases the damping of the impact and accentuates persistence.

In summary, two ingredients are needed in order to generate empirically relevant responses in our model: a persistent autoregressive process driving the exogenous variable and an endogenous persistence source. The computational results shown so far underline the potential ability of the outlined adjustment cost specification (1) to generate inflation persistence. In order to gain an insight about the relative impact of the shocks on endogenous variables, we plot the impulse responses *together* after the shocks and under the *positive* calibration values of table 3. Figure 2 is generated under the assumptions $\sigma_a = \sigma_R = -\sigma_p = 0.01$ (the qualitative responses are standard; see Galí (1999) and Christiano et al. (2005)), and $v_p = 2000$. However, the effect of the cost-push shock on endogenous variables is small in comparison to the other shocks. In order to plot the impulse responses of all three shocks together, the responses after the cost-push have to be magnified by the factor 100000 (or 1000 if responses are in percentage points). This large difference in shock magnitude will be crucial when defining our prior shock values in the estimation exercises presented below.

⁹ The assumption of non-persistent cost-push shocks is widely applied; see in particular Smets and Wouters (2003), p. 1140 and Rabanal and Rubio-Ramírez (2005), p. 1154. Moreover, the empirical evidence of hump-shaped impulse responses mostly assumes interest rate and/or technology shocks, see, e.g., Christiano et al. (2005), pp. 4-8.

Figure 2: Impulse responses of selected variables in the HAC model after one percent shocks and with internal habit formation, interest rate smoothing, and technology persistence.



10 Estimation

We proceed to estimate the presented theoretical models and to assess their empirical fit on the basis of European data. More precisely, we estimate the BAC (table 2 and (34)), the HAC (table 2 and (37)), and the HPS (table 2 and (42)) models using Bayesian estimation techniques. The Bayesian estimation methodology is the most common choice when it comes to evaluate a DSGE model with empirical data (see An and Schorfheide (2007), Schorfheide (2000), and Fernández-Villaverde (2010)). It allows for an estimation of model parameters taking all economic relationships simultaneously into consideration (as in Smets and Wouters (2003, 2007) and Almeida (2009)).

Our first target is focused on the HAC model and its new parameter v_p . In contrast to other parameters (such as ψ_p), there is no hint in the existing literature about the magnitude of v_p . In the context of Bayesian estimations, the significance of this parameter is easily assessed: an estimated parameter $v_p = 0$ would imply its insignificance and, consequently, that the HAC model is misspecified. This question is of special interest because $v_p > 0$ does not only generate a backward-looking inflation term, but also a forward-looking term two periods ahead. Therefore, the empirical plausibility of $v_p > 0$ is not clear in advance.

Our second target is conditional on the first target: if the HAC model is not misspecified, we ask about its ability to describe the data, compared to the BAC and the HPS models. In Bayesian estimation, the measure suitable for this comparison is the marginal density of the data conditional on the respective model. Using this criterion, we are able to rank the three models (as in Fernández-Villaverde and Rubio-Ramírez (2004) and Rabanal and Rubio-Ramírez (2005)).

Our ultimate target is to show that the adjustment cost extension (1) and the resulting inflation curve (37) have the potential to compete with the popular assumption of rule-of-thumb setters under Calvo pricing (equation (42)).

Apart from presenting the set of time series used, the following sections explain the rationale behind our fixed parameter values and the parameter priors. They are crucial for the Bayesian estimation procedure, which is also briefly reviewed.

11 The Data

The underlying set of raw data includes four quarterly and seasonally adjusted series. All series are from the Area-Wide-Model (AWM) for the time period 1970Q1-2005Q4 and can be inspected in appendix A:

1. Real consumption PCR_t , seasonally adjusted (AWM code: PCR).
2. Number of employed persons LEN_t , seasonally adjusted (AWM code: LEN).
3. Consumption deflator PCD_t , seasonally adjusted (AWM code: PCD).
4. Short-term nominal interest rate STN_t (p.a.), nominal in percent (AWM code: STN).

In order match the series to the model variables, the data is transformed in several steps. We first compute

- a) $\hat{C}_t^{obs.} = \left(\ln\left(\frac{PCR_t}{LEN_t}\right) - \ln\left(\frac{PCR_t}{LEN_t}\right)^{trend} \right) \times 100$ (per capita consumption in percent)
- b) $\hat{\pi}_t^{p,obs.} = \ln\left(\frac{PCD_t}{PCD_{t-1}}\right) \times 400$ (annualized price inflation in percent)
- c) $\hat{R}_t^{obs.} = \ln\left(1 + \frac{STN_t}{100}\right) \times 100$ (per annum interest rate factor in percent)

The choice of this particular set of series is following Rabanal and Rubio-Ramírez (2005), p. 1156.¹⁰ The series $\hat{C}_t^{obs.}$, $\hat{\pi}_t^{p,obs.}$, $\hat{R}_t^{obs.}$ are visualized (now 1970Q2-2005Q4) in the first three panels of figure 3. Our transformations are explained as follows. Since our model is cyclical and abstracts from economic growth, we compute and extract the trend $\ln\left(\frac{PCR_t}{LEN_t}\right)^{trend}$ from the per capita consumption series $\ln\left(\frac{PCR_t}{LEN_t}\right)$. As a standard procedure, we interpret the trend series $\ln\left(\frac{PCR_t}{LEN_t}\right)^{trend}$ as being driven by technology shocks only. That is, the trend series is approximately represented in the model by the natural output level $\hat{Y}_t^n$. We compute this trend with a Hodrick-Prescott filter on a quarterly time basis.¹¹ The extraction of the trend then gives $\hat{C}_t^{obs.}$, which is a purely cyclical series corresponding to the theoretical output gap $\hat{X}_t$.¹² Inflation and interest rate rates do not follow any (growth-) trend.

¹⁰ Note that these authors include a series of real wages. However, this is only necessary when estimating the inverse elasticity of labor supply η , which we will fix at a constant value in section 13. Moreover, each series has a corresponding shock in the model. In our case, the technology shock corresponds to the explanation of the detrended consumption series, the interest rate shock is in line with the interest rate series, and the cost-push shock corresponds with the inflation series. In contrast to our setting, Rabanal and Rubio-Ramírez (2005) include a preference shock which corresponds to their real wage series.

¹¹ The Hodrick-Prescott filter decomposes the series into a trend and a cyclical component, where the cyclical component is given by $\hat{C}_t^{obs.}$, see e.g., Dejong and Dave (2011), pp. 36-38.

¹² For example, Taylor (1993), p. 202 defines the output gap as $y = 100(Y - Y^*)/Y^*$, with Y as real GDP and Y^* as trend real GDP. Moreover, it is true that $\ln(Y) - \ln(Y^*) \approx (Y - Y^*)/Y^*$. This is completely consistent to our output

In order visualize well-known European inflation rates (see the upper middle panel in figure 3), we annualized the series of inflation by multiplying by four. We interpret the log-deviation of a model variable in percentages. To achieve consistency of the data, all series are multiplied by 100.

The estimation exercises demand that all series are stationary. While $\hat{C}_t^{obs.}$ is already (trend-) stationary, the series for $\hat{\pi}_t^{p,obs.}$ and $\hat{R}_t^{obs.}$ are not. We achieve stationarity for these series by computing first differences, which results in stationary series centered around zero (i.e. the series are $I(1)$). Additionally, we compute the first-difference of the series $\hat{C}_t^{obs.}$, which can be interpreted as a quarterly growth rate. If d is the first-difference operator, taking the first difference of the variables $\hat{C}_t^{obs.}$, $\hat{\pi}_t^{p,obs.}$, and $\hat{R}_t^{obs.}$ involves

- a) $d\hat{C}_t^{obs.} = \hat{C}_t^{obs.} - \hat{C}_{t-1}^{obs.}$ (per capita consumption, first difference).
- b) $d\hat{\pi}_t^{p,obs.} = \hat{\pi}_t^{p,obs.} - \hat{\pi}_{t-1}^{p,obs.}$ (annualized price inflation, first difference).
- c) $d\hat{R}_t^{obs.} = \hat{R}_t^{obs.} - \hat{R}_{t-1}^{obs.}$ (per annum interest rate factor, first difference),

where the first-differenced data is for the time period 1970Q3-2005Q4. The last step is standard and involves the demeaning of the three differenced series. This is related to the fact that the differenced series still have a positive mean, which contradicts the definition of the model variables. In the model, it is true that (differenced) log-deviations are zero in mean values. In order to account for this in the data, we first compute the mean of each of the differenced series in order to obtain the mean values $d\hat{C}_t^{obs.,mean}$, $d\hat{\pi}_t^{p,obs.,mean}$, and $d\hat{R}_t^{obs.,mean}$. In a second step, we extract this mean value from the respective series in order to obtain a demeaned series. Note that DSGE models are quarterly frameworks. Therefore, the variables $\hat{C}_t^{obs.}$, $\frac{1}{4}\hat{\pi}_t^{p,obs.}$, $\frac{1}{4}\hat{R}_t^{obs.}$ from the observed data are *direct* correspondences to the model variables $\hat{X}_t$, $\hat{\pi}_t^p$, and $\hat{R}_t$.

Consistently, we divide the demeaned first differences of inflation (which was annualized for checking purposes) and interest rates (which are per annum) by four:¹³

- a) $d\hat{C}_t^{obs.,dem.} = d\hat{C}_t^{obs.} - d\hat{C}_t^{obs.,mean}$ (consumption, dem. first difference).
- b) $d\hat{\pi}_t^{p,obs.,dem.} = \frac{1}{4} \left(d\hat{\pi}_t^{p,obs.} - d\hat{\pi}_t^{p,obs.,mean} \right)$ (price inflation, dem. first difference).
- c) $d\hat{R}_t^{obs.,dem.} = \frac{1}{4} \left(d\hat{R}_t^{obs.} - d\hat{R}_t^{obs.,mean} \right)$ (interest rate factor, dem. first difference).

The series $d\hat{C}_t^{obs.,dem.}$, $d\hat{\pi}_t^{p,obs.,dem.}$, and $d\hat{R}_t^{obs.,dem.}$ in figure 3 represent the transformed dataset that is used in our estimation exercises presented below. However, the theoretical models presented so far do *not* contain first-differenced variables (except for inflation as a first-difference of

gap $\hat{X}_t$, where the steady state levels do not matter: $\hat{X}_t = \hat{Y}_t - \hat{Y}_t^n = \ln(Y_t) - \ln(Y_t^n)$, since $\ln(Y) = \ln(Y^n)$. Therefore, $\hat{X}_t$ corresponds exactly to the way $\hat{C}_t^{obs.}$ is constructed.

¹³ We confirmed the correctness of these transformations by noting that a subset of our parameter estimates would otherwise deviate consistently from standard results (Smets and Wouters (2003, 2007) and Taylor (1993)). For example, if the interest rate and/or inflation series were not divided by four, we estimated some reaction parameters in the interest rate rule which were approximately four times higher than standard estimates in the literature.

logarithmical prices). This is accounted for by adding equations into the model that resemble the demeaned first-differenced data. These equations and their empirical correspondences read

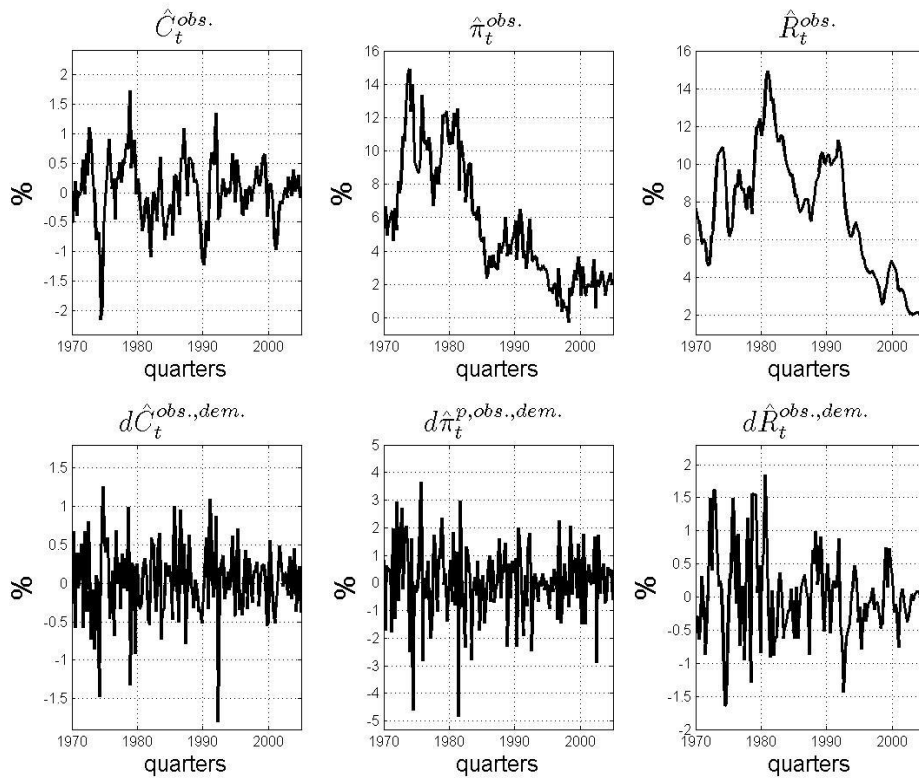
$$d\hat{X}_t = (\hat{X}_t - \hat{X}_{t-1}) \times 100, \quad \text{which counterparts} \quad d\hat{C}_t^{obs.,dem.}, \quad (47)$$

$$d\hat{\pi}_t^p = (\hat{\pi}_t^p - \hat{\pi}_{t-1}^p) \times 100, \quad \text{which counterparts} \quad d\hat{\pi}_t^{p,obs.,dem.}, \quad (48)$$

and

$$d\hat{R}_t = (\hat{R}_t - \hat{R}_{t-1}) \times 100, \quad \text{which counterparts} \quad d\hat{R}_t^{obs.,dem.}. \quad (49)$$

Figure 3: Transformation of the time series used for the Bayesian estimations.



12 Estimation Methodology

Instead of reviewing the vast amount of literature on Bayesian estimation methods, we provide this section with the very essential idea. The interested reader is referred to core contributions such as An and Schorfheide (2007), Schorfheide (2000), and Fernández-Villaverde (2010). Since we use the Matlab preprocessor Dynare for solving and estimating our three DSGE models, the following explanations rely heavily on the pertinent instruction manuals.¹⁴

The Bayesian procedure estimates a subset of model parameters with a weighted Maximum Likelihood approach. More precisely, priors for the parameters to be estimated (mean, variance, and type of distribution) are pre-specified and combined with the model-specific likelihood function.

¹⁴ We use Matlab version 7.11.0.584 (R2010b) and Dynare version 4.2.3. See for a reference manual Adjemian et al. (2011). Our explanations are mainly based on Griffoli (2007), pp. 77-87. See also Dejong and Dave (2011), pp. 217-264.

This gives the target function to be maximized with an optimization routine. As an outcome, one obtains the combination of posterior parameter estimates that renders the dataset and the imposed a-priori beliefs “most likely”. The following lines should give a sensible idea of the Bayesian estimation procedure.

All steps explained below follow Griffoli (2007), pp. 77-87 and are ultimately targeted towards a posterior density function of the form $p(\theta|Y_T)$, where θ denotes the vector of parameters to be estimated and Y_T is a vector of observables up to period T . In the first step, the log-likelihood (i.e. the logarithmical density of the observed data, given the model parameters) $\ln p(Y_T|\theta)$ is retrieved by using the Kalman filter recursion.¹⁵ Given this likelihood function and having defined a prior density $p(\theta)$, the computation of the posterior density $p(\theta|Y_T)$ as an update of the priors is straightforward. Using the Bayes theorem, the posterior density is written as

$$p(\theta|Y_T) = \frac{p(\theta; Y_T)}{p(Y_T)}, \quad (50)$$

where $p(\theta; Y_T)$ is the joint density of the parameters and the data and $p(Y_T)$ is the marginal density of the data. Similarly, one can write for the density of the data conditional on the parameters (the likelihood function):

$$p(Y_T|\theta) = \frac{p(\theta; Y_T)}{p(\theta)} \Leftrightarrow p(\theta; Y_T) = p(Y_T|\theta)p(\theta). \quad (51)$$

A combination of (50) and (51) gives the posterior density as

$$p(\theta|Y_T) = \frac{p(Y_T|\theta)p(\theta)}{p(Y_T)}. \quad (52)$$

Since the marginal density $p(Y_T)$ is independent from the parameter vector θ , it can be treated as a constant. The posterior kernel $\mathcal{K}(\theta|Y_T)$ or the unnormalized posterior density corresponds to the numerator of (52):

$$p(\theta|Y_T) \propto p(Y_T|\theta)p(\theta) \equiv \mathcal{K}(\theta|Y_T). \quad (53)$$

Taking logs of the kernel expression delivers:

$$\ln \mathcal{K}(\theta|Y_T) = \ln p(Y_T|\theta) + \ln p(\theta), \quad (54)$$

where $\ln p(Y_T|\theta)$ is a known term from the Kalman recursion procedure and $\ln p(\theta)$ contains the (also known) priors. The next step is to maximize the log kernel (54) with respect to θ in order to find estimates for the posterior mode θ^m . However, the analytical intractability of the log-likelihood $\ln p(Y_T|\theta)$ as a nonlinear and complicated function of the deep parameters θ requires the use of a numerical optimization algorithm. We call in Dynare the Matlab routine “newrat” for the purpose of obtaining the mode θ^m and the Hessian matrix $H(\theta^m)$ evaluated at the mode.¹⁶ After computing the

¹⁵ For details on the Kalman filter recursion, see for example Dejong and Dave (2011), pp. 80-86, Fernández-Villaverde (2010), p. 17-18 and Hamilton (1994), Ch. 13. See also Griffoli (2007), p. 82, and Almeida (2009), p. 16.

¹⁶ We also tried the routine “csmnwel”, which is usually the number one choice for computing the mode. However, the routine “newrat” proved more successful in finding the respective mode for our three models. See Adjemian et al. (2011), p. 46-47 for an overview of the Matlab routines available in Dynare.

mode, a version of the random-walk Metropolis-Hastings algorithm is called in order to generate the posterior distribution (mean and variance) of our parameters around their mode.¹⁷ Having estimated each of our three model versions in the way described so far, it is possible to undertake a model comparison using posterior distributions. See for the following explanations An and Schorfheide (2007), pp. 144-149, Griffoli (2007), pp. 86-87, and Schorfheide (2000), p. 651. First, define the prior distribution over two competing models A and B as $p(A)$ and $p(B)$. Using the Bayes rule as in (52), one can compute the posterior probability of each model as

$$p(A|Y_T) = \frac{p(Y_T|A)p(A)}{p(A)p(Y_T|A) + p(B)p(Y_T|B)} \quad (55)$$

and

$$p(B|Y_T) = \frac{p(Y_T|B)p(B)}{p(A)p(Y_T|A) + p(B)p(Y_T|B)}. \quad (56)$$

The expressions above describe the probability of a model being true after observing the data. Therefore, a natural way to compare the empirical fit of two models is given by the posterior odds ratio:

$$PO_{A,B} = \frac{p(A|Y_T)}{p(B|Y_T)} = \frac{p(Y_T|A)p(A)}{p(Y_T|B)p(B)}, \quad (57)$$

where $\frac{p(Y_T|A)}{p(Y_T|B)}$ is the Bayes factor describing the evidence in the data favoring model A over model B and $\frac{p(A)}{p(B)}$ is the prior odds ratio giving the relative probability subjectively assigned to the models. The terms entering the numerator and denominator of the Bayes factor are (corresponding to $p(Y_T)$ in (52)) the marginal densities of the data conditional on the respective model. They can be obtained, at least theoretically, by integrating the parameters out of the posterior kernel (c.f. Griffoli (2007), p. 80):

$$p(Y_T|A) = \int_{\theta_A} p(Y_T, \theta_A|A) d\theta_A = \int_{\theta_A} \underbrace{p(Y_T|\theta_A, A)p(\theta_A|A)}_{=K(\theta_A|Y_T, A)} d\theta_A, \quad (58)$$

$$p(Y_T|B) = \int_{\theta_B} p(Y_T, \theta_B|B) d\theta_B = \int_{\theta_B} \underbrace{p(Y_T|\theta_B, B)p(\theta_B|B)}_{=K(\theta_B|Y_T, B)} d\theta_B. \quad (59)$$

However, the integration is analytically intractable and has to be substituted by the Laplace or the Harmonic Mean approximation (see for details Griffoli (2007), pp. 86-87, An and Schorfheide (2007), pp. 146-147, and Schorfheide (2000), p. 660). As in Smets and Wouters (2003), p. 1150, Dynare gives the logarithmic value of the approximated marginal likelihood of each model $\ln p(Y_T|A)^{Lapl./H.Mean} < 0$ and $\ln p(Y_T|B)^{Lapl./H.Mean} < 0$. After assuming a uniform distribution across two models ($p(A) = p(B)$), we compare them by using the posterior odds ratio

¹⁷ We opted for 1,000,000 draws from each model's posterior distribution with 2 distinct chains. For details on the Dynare version of the Metropolis-Hastings algorithm, see Griffoli (2007), pp. 83-85, An and Schorfheide (2007), p. 131, and Schorfheide (2000), p. 668-669.

$PO_{A,B}^{Lapl./H.Mean} = \exp \left(\ln p(\mathbf{Y}_T | \mathbf{A})^{Lapl./H.Mean} \right) / \exp \left(\ln p(\mathbf{Y}_T | \mathbf{B})^{Lapl./H.Mean} \right)$. A ratio greater than one is interpreted as an outperformance of model $\mathbf{A}$ over model $\mathbf{B}$ in the description of the dataset $\mathbf{Y}_T$ (see Dejong and Dave (2011), p. 242 and Rabanal and Rubio-Ramírez (2005), pp. 1156-1158). In our context, the models BAC, HAC, and HPS correspond to the labels $\mathbf{A}$ and $\mathbf{B}$ following $\mathbf{A} = \mathbf{HAC}$ and $\mathbf{B} = \mathbf{BAC}, \mathbf{HPS}$.¹⁸ Moreover, the dataset is in our case $\mathbf{Y}_{T=2005Q4} = \left(d\hat{C}_t^{obs.,dem.}, d\hat{\pi}_t^{p,obs.,dem.}, d\hat{R}_t^{obs.,dem.} \right)$.

13 Parameters and Priors

This section presents the subset of fixed deep parameter values and the priors. The preliminary definition of some parameters is due to identification problems and resulting difficulties in estimating them. As pointed out by Rabanal and Rubio-Ramírez (2005), p. 1157, the absence of capital services in our model hinders the estimation of the household discount factor β and the capital share of output α . Therefore, we set $\beta = 0.99$ and $\alpha = 0.3$, which are standardly used values. Moreover, identification problems arise between the steady state elasticity ϵ_p and the corresponding rigidity parameter ψ_p (BAC and HAC) or θ_p (HPS). For this reason, we assume $\epsilon_p = 6$. In contrast to Rabanal and Rubio-Ramírez (2005), p. 1156, we do not assume a preference shock and a series for real wages. This results in identification problems with respect to the inverse of the real wage elasticity η . Therefore, we assume $\eta = 2$. Note that all fixed parameter values are the same as in in Rabanal and Rubio-Ramírez (2005), p. 1157. The remaining parameters and shocks to be estimated for each of the three models are $\boldsymbol{\theta}_{BAC} = (\psi_p, \boldsymbol{\theta}^*)'$, $\boldsymbol{\theta}_{HAC} = (v_p, \boldsymbol{\theta}_{BAC})'$, and $\boldsymbol{\theta}_{HPS} = (\theta_p, \omega_p, \boldsymbol{\theta}^*)'$, with $\boldsymbol{\theta}^* = (\sigma, h, \delta_\pi, \delta_y, \phi, \delta_{\Delta\pi}, \delta_{\Delta y}, \rho_a, \sigma_a, \sigma_p, \sigma_R)$. This means that the BAC and HAC models have common priors up to the parameter v_p , which appears with a non-zero value only in the HAC model. Similarly, the HAC and HPS models have common priors up to the parameters ψ_p , v_p , θ_p , and ω_p .

Choosing priors is mostly a matter of subjective beliefs. As stated by Griffoli (2007), p. 48, the researcher should reflect about the domain, its boundedness, and the shape of the prior distribution. Regarding the latter, Rabanal and Rubio-Ramírez (2005), p. 156 choose a completely uninformative prior distribution, i.e., a uniform distribution for the shocks. This is due to their uncertainty with respect to the magnitude of these shocks. However, we prefer the imposition of informative distributions, as in the majority of contributions and as in Smets and Wouters (2003), p. 1142-1143. A more important issue is whether to choose loose or tight priors. The choice of tight priors means that low prior standard deviations are imposed. This implies that the researcher lets the prior densities to concentrate around their prior means. The prior beliefs $\ln p(\boldsymbol{\theta})$ are then intended to dominate over the log-likelihood (the information from the data) $\ln p(\mathbf{Y}_T | \boldsymbol{\theta})$ in the computation of the posterior $\ln \mathcal{K}(\boldsymbol{\theta} | \mathbf{Y}_T) = \ln p(\mathbf{Y}_T | \boldsymbol{\theta}) + \ln p(\boldsymbol{\theta})$, see (54) and Griffoli (2007), p. 80-81. The choice of tight priors makes identification and estimation of large models easier. More importantly, the parameter values are more likely to remain in economically reasonable ranges (see Fernández-

¹⁸ Note that our pairwise comparison is rather unappealing if the models are exclusively intended for forecasting, see Schorfheide (2000), pp. 653-654. However, the pairwise comparison is enough for our exercise of showing the potential of the HAC model to describe the data. In this sense, our procedure is in line with Rabanal and Rubio-Ramírez (2005), p. 1160.

Villaverde (2010), p. 6-37). The downside of tight priors is, of course, that the role of the data is small. In contrast, the choice of loose priors implies that the researcher lets $\ln p(\mathbf{Y}_T|\boldsymbol{\theta})$ dominate in the posterior $\ln \mathcal{K}(\boldsymbol{\theta}|\mathbf{Y}_T)$. In other words, the data plays a dominant role over the researchers beliefs $\ln p(\boldsymbol{\theta})$. Fernández-Villaverde (2010), p. 36-37 argues that the choice of loose or tight priors could be made depending on the application of the model. A policy-oriented model calls for tight priors that avoid odd posterior estimates that would otherwise hinder the communication of results to policy makers. In contrast, a research-oriented model calls for loose priors that let the data drive the posterior results.

Concerning *our* parameter priors, we opt to choose prior mean values that are close to those of Smets and Wouters (2003), p. 1142-1143. Their tightness (prior standard deviation) is set as loosely as possible in order to let the data drive our posterior results. However, the looseness degree is restricted by the fact that most of the priors for the parameters in $\boldsymbol{\theta}_{BAC}$, $\boldsymbol{\theta}_{HAC}$, and $\boldsymbol{\theta}_{HPS}$ are common to the three models. These common priors have to ensure that a posterior mode is found for each of the three different versions of the log-kernel (54). We find that the success in finding the posterior mode for each model is more likely, when the common priors are tight. This can lead to prior densities for some of the common parameters that are somewhat tighter than in Smets and Wouters (2003). We checked with the Dynare command “mode_check” that the the maximum of the posterior log kernel (54) is obtained for each parameter. An overview of the priors is given by the respective columns of table 4. Across the BAC and HAC models, the prior of ψ_p is assumed to be normally distributed with a mean of 170 and a standard deviation (henceforth in brackets) of (80).

By the relationship $\psi_p = \frac{\theta_p(1+\alpha(\epsilon_p-1))Y(\epsilon_p-1)}{(1-\theta_p)(1-\beta\theta_p)(1-\alpha)}$, this corresponds to a prior mean for the Calvo parameter θ_p in the HPS model given by 0.75.¹⁹ Since the value of this parameter lies between zero and one, it is assumed to follow a beta distribution with a standard deviation of (0.01), as in Smets and Wouters (2003), p. 1142-1143. Note that the different structures of the hybrid inflation curves (37) and (42) do not allow to establish a sensible link between v_p and ω_p .²⁰ Moreover, we don't have any prior knowledge about v_p . This means that a loose prior for this parameter is rather desirable: the data should be driving its posterior, rather than our prior beliefs. From the perspective of the log-likelihood maximization (i.e. finding a posterior mode), we find that the choice of a loose prior 2000 (1000) is a reasonable alternative. In contrast, the existing literature enables us to assume standard values for the beta-distributed prior of ω_p . We assume for this parameter 0.75 (0.015), as in Smets and Wouters (2003), p. 1142-1143. Similarly, the interest rate rule coefficients have prior means $\delta_\pi = 1.7$, $\delta_y = 0.125$, $\delta_{\Delta\pi} = 0.1$, $\delta_{\Delta y} = 0.163$, and we impose gamma distributions ensuring a positively bounded density with standard deviations (0.1), (0.05), (0.01), and (0.05), respectively. The smoothing parameter ϕ follows a beta distribution with a mean equal to 0.8 and a standard deviation of (0.1). As in Smets and Wouters (2003), p. 1142-1143, the household parameter σ follows a normal distribution with mean 1. Accordingly, we impose a respective standard deviation (0.375). The prior of the habit formation parameter h is assumed to follow a beta distribution with 0.7 (0.1). Table 4 also gives the priors for the shocks. Their prior mean values (which equal their

¹⁹ This link between ψ_p and ω_p is true under $v_p = \omega_p = 0$ and can be derived by equalizing (37) to (42).

²⁰ The structural difference is given by $\gamma_3 > 0$ and $\gamma_3^c = 0$.

standard deviations) are driven by the fact that the cost-push shock leads to log-deviations that are small compared to the technology and interest rate shocks (see figure 2). In other words, σ_p must have a relatively high value compared to σ_a and σ_R in order to generate meaningful fluctuations. This is in line with the relative importance of cost-push shocks estimated by Rabanal and Rubio-Ramírez (2005), p. 1157.²¹ Therefore, we set mean values (standard deviations) as 0.005, 0.001, and 30, for σ_a , σ_R , and σ_p , respectively. As done by Smets and Wouters (2003), p. 1142-1143, we use the inverse gamma distribution, which ensures a positive support of the density function of the shocks. The prior for the autoregressive shock parameter ρ_a is assumed to follow a beta distribution with 0.85 (0.1).

14 Parameter Estimates

This section reports posterior parameter distributions. An example of graphical diagnostics created by Dynare can be found in appendix B. They ensure that the Bayesian estimation of the three models has been successful and that all parameters in the vectors $\theta_{BAC} = (\psi_p, \theta^*)'$, $\theta_{HAC} = (v_p, \theta_{BAC})'$, $\theta_{HPS} = (\theta_p, \omega_p, \theta^*)'$, and $\theta^* = (\sigma, h, \delta_\pi, \delta_y, \phi, \delta_{\Delta\pi}, \delta_{\Delta y}, \rho_a, \sigma_a, \sigma_p, \sigma_R)$ are identified.²²

The last three columns of table 4 contain the posterior mean and the posterior standard deviation (SD) of the estimated parameters and shocks. Moreover, figure 4 gives a sense of how the prior and posterior densities are shaped. Estimation results for the BAC model are presented in the fourth column of table 4. Accordingly, the degree of price rigidity increases from $\psi_p = 170$ (prior mean) to $\psi_p = 267.77$ (posterior mean) and the degree uncertainty decreases (standard deviation (SD) decreases from (35) to (27.48)). Most of the remaining parameter estimates in the BAC model are close to their priors and have a lower degree of uncertainty. We obtain a posterior for the inverse of the intertemporal elasticity of substitution which is lower than in the standard literature, namely $\sigma = 0.094$.²³ Another glance at the fourth column of table 4 reveals a posterior volatility of the price mark-up $\sigma_p = 6.78$. At the same time, the importance of some shocks for the explanation of fluctuations is marginally decreased ($\sigma_A = 0.004$) or increased ($\sigma_R = 0.0015$). The posterior autoregressive parameter ρ_a shows little deviation from our prior distribution beliefs.

Posterior parameter estimates for the HAC model can be inspected in the fifth column of table 4. Again, the posterior degree of uncertainty is lower for all parameters than in our prior subjective beliefs. We obtain a lower degree of price rigidity, namely $\psi_p = 146.49$. The most important posterior results are those for the novel parameter in the HAC model. We obtain the posterior estimate for v_p as 4372.15 (714.98). This result clearly indicates that this parameter is relevant for

²¹ However, Smets and Wouters (2003) and others do not find this result. This simply because they assume, e.g., capital services, thus increasing the relative importance of, e.g., technology shocks. In contrast, our models are in the spirit of Rabanal and Rubio-Ramírez (2005).

²² The diagnostics are complemented by brief explanations, which are included in appendix B.

²³ However, Rabanal and Rubio-Ramírez (2005), p. 1157 find in their version of the BAC model a value $\sigma = 6.67$, which is higher compared to standard values $\sigma \in [1, 2]$. This deviation is a result of different data sources, different data samples, and differing prior configurations. Nevertheless, the Bayesian estimation ensures that the posterior parameter values in the BAC, HAC, and HPS models fulfill the Blanchard-Kahn stability condition, see Blanchard and Kahn (1980).

the description of the underlying data. In order to reject the hybrid specification in the HAC model (equation (37)), one would need to obtain a value of the parameter v_p in the neighborhood of zero or a highly increased degree of uncertainty. Note that the posterior positive value of v_p is not driven by the lagged inflation term $\hat{\pi}_{t-1}^p$ alone. The theoretical result is an additional two-period-ahead expectation term that cannot be disentangled from the lagged term (i.e. $v_p = 0$ in (37) eliminates $\hat{\pi}_{t-1}^p$ and $E_t[\hat{\pi}_{t+2}^p]$). Taking this theoretical result into account, our posterior parameter value $v_p > 0$ is driven by *both* inflation terms.

The remaining posterior estimates in the HAC model are similar to those obtained in the BAC environment. The inverse of the intertemporal substitution in consumption is somewhat lower ($\sigma = 0.98$ (0.034)) than its prior. The fifth column of table 4 also shows a higher posterior volatility in the price mark-up shock, whereas the relevance of the remaining shocks is rather limited. Again, the posterior autoregressive parameter shows little deviation from our prior beliefs.

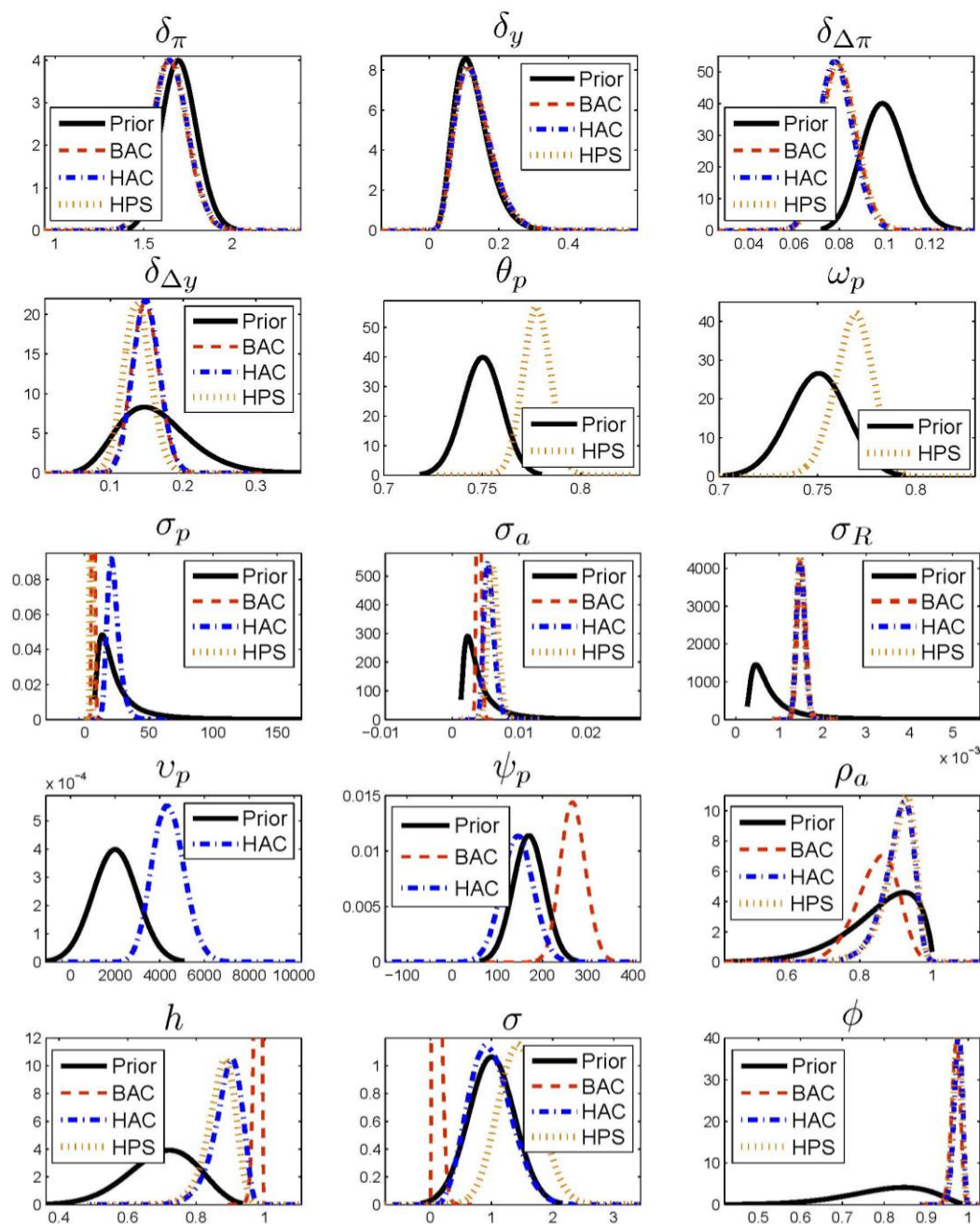
The estimation outcomes for the hybrid Calvo-type model (HPS) are reported in the sixth column of table 4. All posterior estimates are rather close to their priors and are similar to the results obtained by Smets and Wouters (2003), p. 1140. The posterior of σ now resembles the value commonly found in hybrid New Keynesian models with Calvo pricing, namely 1.45 (0.35). Moreover, the cost-push shock drops to a similar value as in the BAC model ($\sigma_p = 5.02$).

Table 4: Priors, posteriors, and logarithmical marginal likelihoods. The entries in brackets give the respective standard deviation (SD).

Parameters, shocks, and marginal likelihoods	Priors		Posteriors		
	Distribution	All models	BAC	HAC	HPS
		Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
ψ_p Rotemberg rigidity	normal	170 (35)	267.77 (27.48)	146.49 (34.86)	(—)
v_p Rotemberg inflation lag	normal	2000 (1000)	(—)	4372.15 (714.98)	(—)
θ_p Calvo rigidity	beta	0.75 (0.01)	(—)	(—)	0.77 (0.0069)
ω_p Calvo inflation lag	beta	0.75 (0.015)	(—)	(—)	0.76 (0.0096)
σ Consumption utility	normal	1 (0.375)	0.094 (0.036)	0.98 (0.034)	1.45 (0.35)
h Consumption habits	beta	0.7 (0.1)	0.97 (0.0072)	0.88 (0.0037)	0.87 (0.036)
δ_π Taylor inflation	gamma	1.7 (0.1)	1.64 (0.099)	1.66 (0.099)	1.65 (0.099)
δ_y Taylor output	gamma	0.125 (0.05)	0.131 (0.047)	0.127 (0.047)	0.130 (0.047)
ϕ Taylor smoothing	gamma	0.8 (0.1)	0.97 (0.01)	0.97 (0.0097)	0.97 (0.0098)
$\delta_{\Delta\pi}$ Taylor inflation lag	gamma	0.1 (0.01)	0.08 (0.0076)	0.077 (0.0074)	0.079 (0.0075)
$\delta_{\Delta y}$ Taylor output lag	gamma	0.163 (0.05)	0.147 (0.018)	0.151 (0.018)	0.136 (0.018)
ρ_a Shock persistence	beta	0.85 (0.1)	0.84 (0.059)	0.89 (0.039)	0.9 (0.035)

σ_a Technology shock	inv. gamma	0.005 (0.005)	0.004 (0.004)	0.0056 (0.0056)	0.0061 (0.0061)
σ_R Interest rate shock	inv. gamma	0.001 (0.001)	0.0015 (0.0015)	0.0015 (0.0015)	0.0015 (0.0015)
σ_p Cost-push shock	inv. gamma	30 (30)	6.78 (6.78)	22.27 (22.27)	5.026 (5.02)
$\ln p(Y_T Model)^{Lapl.}$			-231.42	-99.07	-105.77
$\ln p(Y_T Model)^{H.Mean}$			-231.14	-99.33	-105.62

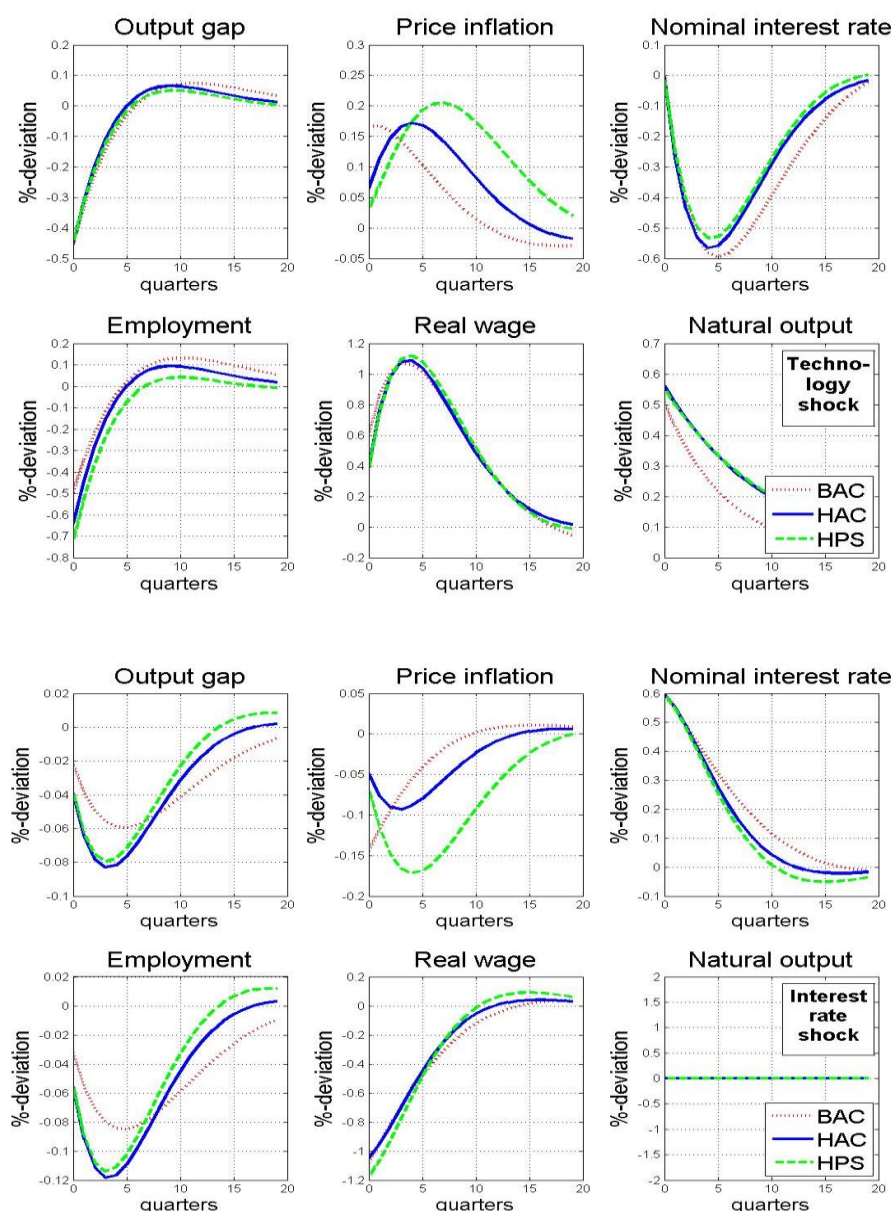
Figure 4: Prior and posterior distributions.

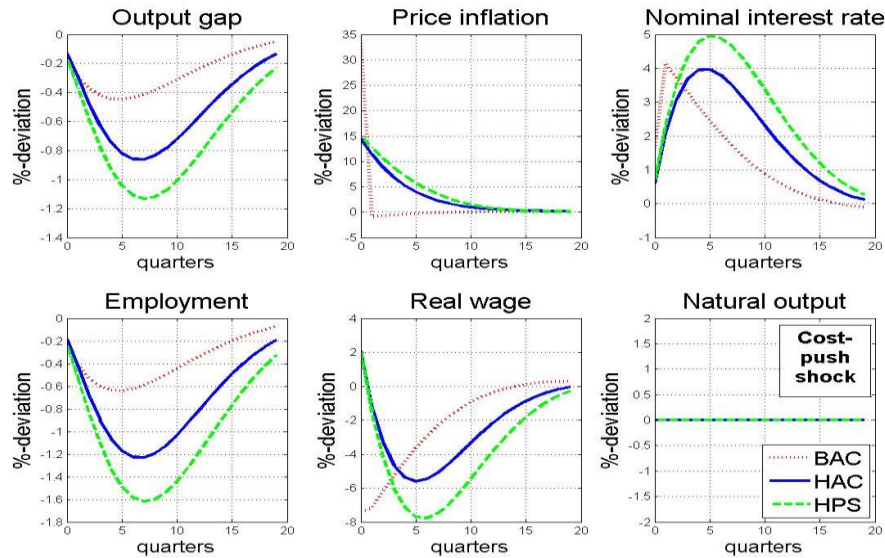


15 Model Comparison

We first generate impulse responses in the BAC, HAC, and HPS models based on the posterior mean parameter values and the posterior standard deviation shocks. That is, we parameterize the models with the posteriors of table 4. Figure 5 shows that the posterior mean parameter values are such that impulse responses can be generated, thus confirming that the Blanchard-Kahn stability condition is met for all three models. The response of inflation in the HAC model displays the desired amplification of inflation persistence across the technology and interest rate shock. Moreover, all responses are qualitatively very similar to those in the HPS model.

Figure 5: Impulse responses based on the posterior estimates of table 4.





As explained in section 12, a model A outperforms a framework B in the description of a dataset Y_T if the posterior odds ratio $PO_{A,B}^{Lapl./H.Mean}$ is greater than one. We follow a well-known comparison guidance, which is reproduced by Dejong and Dave (2011), p. 242. Thus, $PO_{A,B}^{Lapl./H.Mean} \in [1, 3]$: “very slight evidence” in favor of model A . $PO_{A,B}^{Lapl./H.Mean} \in (3, 10]$: “slight evidence” in favor of model A , $PO_{A,B}^{Lapl./H.Mean} \in (10, 100]$: “strong to very strong evidence” in favour of model A , and $PO_{A,B}^{Lapl./H.Mean} > 100$: “decisive evidence” in favor of model A .

As a starting point, we compare the HAC model against the BAC setting. The resulting outperformer is then checked against the HPS model. In each of the two comparisons, a uniform prior probability distribution across the two models involved was assumed, such that $p(HAC) = p(BAC)$ and $p(HAC) = p(HPS)$.

The last two rows of table 4 give the marginal likelihoods as Laplace-/Harmonic Mean-approximations and in natural logs. Note first that the HAC model clearly outperforms the BAC model since $PO_{HAC,BAC}^{Lapl.} = \exp(-99.07) / \exp(-231.42) = \frac{9.4286 \cdot 10^{-44}}{3.1302 \cdot 10^{-101}} = 3.0121 \cdot 10^{57}$ (Laplace) and $PO_{HAC,BAC}^{H.Mean} = \exp(-99.33) / \exp(-231.14) = \frac{7.2699 \cdot 10^{-44}}{4.1417 \cdot 10^{-101}} = 1.7553 \cdot 10^{57}$ (Harmonic mean). Therefore, there is “decisive evidence” in favor of the HAC model when compared to the BAC model. This result is not surprising, as the purely forward-looking BAC model fails to describe the amplified sluggishness of inflation present in the data.

Having established the superiority of the HAC model against the baseline adjustment cost framework, we now ask if there is also an outperformance against the HPS model.²⁴ From the viewpoint of the HAC model, the posterior odds ratios are $PO_{HAC,HPS}^{Lapl.} = \exp(-99.07) / \exp(-105.77) = \frac{9.4286 \cdot 10^{-44}}{1.1606 \cdot 10^{-46}} = 812.4058$ (Laplace) and

²⁴ The results in table 4 also imply that the HPS setting is superior against the BAC model.

$PO_{HAC,HPS}^{H.Mean} = \exp(-99.33) / \exp(-105.62) = \frac{9.4286 \cdot 10^{-44}}{1.3483 \cdot 10^{-46}} = 539.1533$ (Harmonic mean). Therefore, there is “decisive evidence” in favor of the HAC model when compared to the HPS model.

The results let us conclude that the HAC model should be preferred over the BAC model. At the same time, the HAC model seems to have the potential to outperform the HPS model. The higher marginal likelihood of the HAC model against the HPS setting could be interpreted as a potential of the term $E_t[\hat{\pi}_{t+2}^p]$ to capture information in the data that the HPS model is not able to capture.

16 Conclusions

We showed that an amplification of inflation persistence can be achieved by undertaking a slight alteration of the real adjustment cost structure by Rotemberg (1982) and Hairault and Portier (1993). The result of this specification is a New Keynesian Phillips curve with a backward-looking inflation term which is directly linked to an inflation expectation two periods ahead. This link is established through a new parameter which is necessary to generate our hybrid inflation curve. Our theoretical findings were complemented with Bayesian estimations based on historical European data. We showed that the new parameter adopts a significant empirical value unequal to zero. Moreover, we provided evidence in terms of logarithmical and approximated marginal likelihoods that our hybrid inflation schedule performs better than a standard purely forward-looking specification. In addition, the marginal likelihood of our hybrid inflation schedule was found to be higher than the marginal likelihood of the Calvo model with indexation.

We conclude from our results that our specified New Keynesian Phillips curve has the potential to compete with the widespread hybrid New Keynesian Phillips curve resulting from Calvo assumptions. However, this potential relies in the description of historical data and our analysis could be extended along several dimensions. One task not considered here is how our model performs in (out-of-sample) forecasting, which is primarily important for policy makers. Similarly, a VAR model could be taken into account as a competing benchmark. In this way, the dominance of one model over another could be assessed by comparing, e.g., impulse responses and confidence intervals of the VAR model. Finally, our results could be complemented by an extensive sensitivity analysis. Naturally, our estimated (posterior) parameter values depend on the subset of fixed parameter values and on the set of priors. A sensitivity analysis would unveil how sensitive our estimates are with respect to the fixed and prior parameter setting.

References

- Abel, A.B. (1990). Asset pricing under habit formation and catching up with the Joneses. *American Economic Review* 80 (2), 38-42.
- Adjemian, S., Bastani, H., Karamé, F., Juillard, M., Maih, J., Mihoubi, F., Perendia, G., Pfeifer, J., Ratto, M., Villemot, S. (2011). *Dynare: Reference Manual, Version 4*. Dynare Working Paper 1, CEPREMAP.
- Almeida, V. (2009). Bayesian estimation of a DSGE model for the Portuguese economy. *Estudos e Documentos de Trabalho, Working Paper 14/2009*, Banco de Portugal.
- An, S., Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric Reviews, Taylor and Francis Journals* 26 (2), 113-172.
- Blanchard, O., Kahn, C. (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica* 48 (5), 1305-1311.

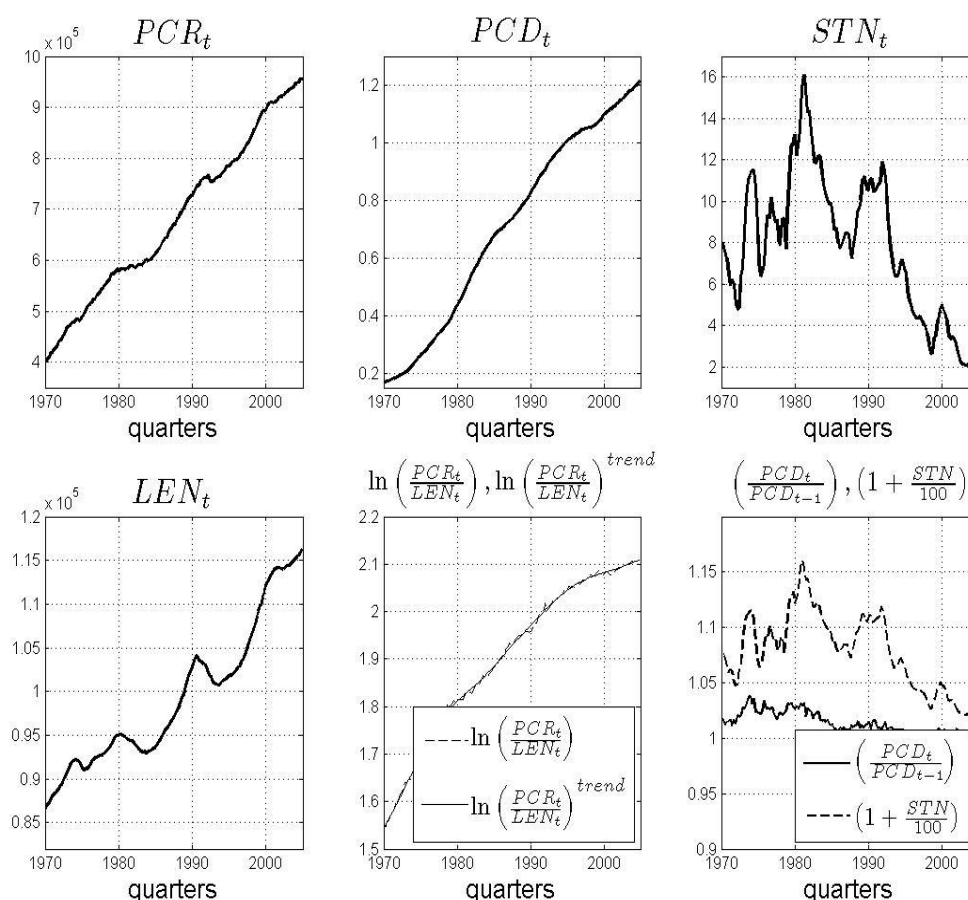
- Brooks, S.P., Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics* 8 (4), 434-455.
- Calvo, G. (1983). Staggered prices in a utility-maximizing framework. *Journal of Monetary Economics* 12 (3), 383-398.
- Campbell, J.Y., Cochrane, J.H. (1999). By force of habit: a consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy* 107 (2), 205-251.
- Casares, M. (2006). Monetary Policy Rules in a New Keynesian Euro Area Model. *Journal of Money, Credit, and Banking* 39 (4), 875-900.
- Constantinides, G.M. (1990). Habit Formation: A Resolution of the Equity Premium Puzzle. *Journal of Political Economy* 98 (3), 519-543.
- Christiano, L.J., Eichenbaum, M., Evans, C. (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* 113 (1), 1-45.
- Dejong, D., Dave, C. (2011). *Structural Macroeconometrics*. 2nd ed., Princeton University Press.
- Erceg, C.J., Henderson, D., Levin, A.T. (2000). Optimal Monetary Policy with Staggered Wage and Price Contracts. *Journal of Monetary Economics* 46 (2), 281-313.
- Fagan, H., Henry, J., Mestre, R. (2005). An area-wide model for the euro area. *Economic Modelling* 22 (1), 39-59.
- Fernández-Villaverde, J. (2010). The Econometrics of DSGE models. *SERIEs* 1 (1), 3-49.
- Fernández-Villaverde, J., Rubio-Ramírez, J.F. (2004). Comparing dynamic equilibrium models to data: a Bayesian approach. *Journal of Econometrics* 123 (1), 153-187.
- Fuhrer, J.C., Moore, G.R. (1995). Inflation Persistence. *Quarterly Journal of Economics* 110 (1), 127-159.
- Galí, J. (1999). Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations? *American Economic Review* 89 (1), 249-271.
- Galí, J., Gertler, M., López-Salido, J.D. (2001). European inflation dynamics. *European Economic Review* 45 (7), 1237-1270.
- Griffoli, T.M., (2007). *Dynare User Guide: An Introduction to the Solution and Estimation of DSGE Models*. CEPREMAP, mimeo.
- Grishchenko, O.V. (2010). Internal vs. external habit formation: The relative importance for asset pricing. *Journal of Economics and Business* 62 (3), 176-194.
- Hairault, J.-O., Portier, F. (1993). Money, New Keynesian macroeconomics and the business cycle. *European Economic Review* 37 (8), 1533-1568.
- Hamilton, J.D. (1994). *Time Series Analysis*. Princeton University Press.
- Henzel, S., Hülsewig, O., Mayer, E., Wollmershäuser, T. (2009). The price puzzle revisited: Can the cost channel explain a rise in inflation after a monetary policy shock? *Journal of Macroeconomics* 31 (2), 268-289.
- Hülsewig, O., Mayer, E., Wollmershäuser, T. (2009). Bank behavior, incomplete interest rate pass-through, and the cost channel of monetary policy transmission. *Economic Modelling* 26 (6), 1310-1327.
- Lechthaler, W., Snower, D. (2008). Quadratic Labor Adjustment Costs and the New-Keynesian Model. Kiel Working Paper 1453, Kiel Institute for the World Economy.
- Peersman, G., Smets, F. (2001). The Monetary Transmission in the Euro Area: More Evidence from VAR Analysis. ECB Working Paper 91, European Central Bank.
- Rabanal, P., Rubio-Ramírez, J.F. (2005). Comparing New Keynesian models of the business cycle: A Bayesian approach. *Journal of Monetary Economics* 52 (6), 1151-1166.

- Rotemberg, J. (1982). Monopolistic Price Adjustment and Aggregate Output. *Review of Economic Studies* 49 (5), 517-531.
- Schorfheide, F. (2000). Loss function-based evaluation of DSGE models. *Journal of Applied Econometrics* 15 (6), 645-670.
- Smets, F., Wouters, R. (2003). An estimated stochastic dynamic general equilibrium model of the euro area. *Journal of the European Economic Association* 1 (5), 1123-1175.
- Smets, F., Wouters, R. (2007). Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review* 97 (3), 586-606.
- Taylor, J.B. (1979). Staggered Wage Setting in a Macro Model. *American Economic Review* 69 (2), 108-113.
- Taylor, J. B. (1993). Discretion versus policy rules in practice. *Carnegie-Rochester Conference Series on Public Policy* 39 (1), 195-214.

Appendix A: Raw Data

The data was obtained from the website <http://www.eabcn.org/area-wide-model>. See Fagan et al. (2005) for an overview of the AWM and the underlying historical series.

Figure 6: Visualized time series of the Area Wide Model (AWM).



Appendix B: Bayesian Estimation Diagnostic

The figures presented in this section are based on the HAC model and exemplify how Dynare allows for a statement on the parametrical convergence of the Metropolis-Hastings sampling algorithm (following Brooks and Gelman (1998), p. 451). According to Almeida (2009), p. 37, the overall convergence information is summarized in three graphs, each representing a global convergence measure. The two distinct lines (red and blue) display the results within and between the chains of the Metropolis-Hastings algorithm. These measures are related to the analysis of the parameters' mean ("Interval"), of the variances ("m2") and of the third moments ("m3"). Ideally, the two lines converge to each other and become stable for each one of the three measures. The figure shows in the example of v_p , ψ_p , and ρ_a that parametrical convergence was obtained.

Figure 7: MCMC univariate diagnostic (Brooks and Gelman (1998)).

