UvA-DARE (Digital Academic Repository)

Optimal policy under model uncertainty: A structural-Bayesian estimation approach

Kriwoluzky, A.; Stoltenberg, C.A.

Publication date
2008

Link to publication

Citation for published version (APA):
Kriwoluzky, A., & Stoltenberg, C. A. (2008). *Optimal policy under model uncertainty: A structural-Bayesian estimation approach*. Afdeling Algemene Economie. http://www.mwpweb.eu/fellows/AlexanderKriwoluzky/resources/document_22.pdf

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (https://dare.uva.nl)

Download date:19 Apr 2024

Optimal Policy under Model Uncertainty: A Structural-Bayesian Estimation Approach*

Alexander Kriwoluzky,† Christian A. Stoltenberg‡

February, 2008

Abstract

In this paper we assess the relevant sources of model uncertainty for the optimal conduct of monetary policy within (parameter uncertainty) and across (specification uncertainty) a set of nested models by applying Bayesian model estimation techniques to EU 13 data. While parameter uncertainty is found to play a minor role, optimal monetary policy is highly sensitive with respect to specification uncertainty. As our main result, we find that following policy prescriptions that are optimal in the model, which includes all features and frictions, does not guard against the risk of model uncertainty.

JEL classification: E32, C51, E52.
Keywords: Optimal monetary policy, model uncertainty, Bayesian model estimation.

*We are especially thankful to Harald Uhlig, Wouter Den Haan, Tom Krebs and Bartosz Maćkowiak. We further thank seminar participants at the Sveriges Riksbank, at the annual conferences of the Royal Economic Society and the European Economic Society as well as at the European workshop on Macroeconomics in Glasgow.

†Department of Economics, European University Institute, Max Weber Programme, Via delle Fontanelle, 10, 50014 San Domenico, Italy, email: Alexander.Kriwoluzky@eui.eu, tel: +39 055 4685 851.

‡Department of Economics, Universiteit van Amsterdam, Roeterstraat 11, 1018 WB Amsterdam, The Netherlands, email: C.A.Stoltenberg@uva.nl, tel: +31 20 525 3913.

1 Introduction

Recently, the empirical evaluation of Dynamic Stochastic General Equilibrium (DSGE) models employing Bayesian methods has made substantial progress (Smets and Wouters, 2003, 2007; An and Schorfheide, 2007; Lubik and Schorfheide, 2004). Policymakers – in particular central banks – nowadays correspondingly employ estimated medium-scale DSGE models, including various features and frictions, in their policy analysis more and more. This practice is based on the implicit idea that by capturing many aspects of the economy, policy prescriptions derived from these models should guard against the risks of an uncertain economic environment.

We set up a structural Bayesian framework to investigate how to conduct optimal monetary policy under model uncertainty for the Euro-13 area. To represent the uncertainty about the true data-generating process, we start with a small benchmark model and subsequently estimate a set of competing and nested models. This bottom-up approach puts us into a position to quantitatively evaluate the gain in explanatory power of each extension separately. As our main result, we find that following the optimal policy derived from the model that nests all features and frictions does not protect against model uncertainty.

The intuition for this result is as follows: A medium-scale model is likely to contain features and frictions which help and some that do not help to explain a given set of time series. Correspondingly, there exists a nested smaller model that comprises only the relevant model components and consequently exhibits a higher model probability. In a micro-founded approach, the policy objective changes according to the features and frictions included, and – in the case of conflicting stabilization aims – optimal policies may differ substantially across models: an optimal policy in one model may be harmful in another environment. A policymaker with a concern for robustness takes into account both, model probabilities and losses, and averages welfare across a set of models to determine the optimal policy under model uncertainty. It follows that the policy found to be optimal in the largest model is, in general, not an optimal policy device, if the policymaker faces uncertainty over the true model.

In a nutshell, our main result is based on the following constellation of nested models: There is a simple benchmark economy and there exist at least two additional features, one that helps to explain the data and another that does not. The inclusion of features of the former type is necessary to demonstrate the fact that there are indeed useful model extensions. Furthermore, a conflicting policy aim is introduced by an additional friction. We illustrate our main result by choosing as benchmark model one of the most popular models employed in monetary analysis nowadays: a standard cashless New Keynesian economy with staggered price-setting without indexation (Woodford, 2003). As examples for specification uncertainty, we subsequently allow for more lags in endogenous variables (indexation and habit formation), for omitted variables (money) and for one large model that nests all these features. While the cashless models are characterized by price stability as the predominant principle, a demand for cash introduces the stabilization of the nominal interest rate as a conflicting aim.

Lucas (2003) highlights the finding that the gains of stabilization policies are limited in representative consumers economies with standard preferences. In particular, he points out that the maximum welfare gain due to stabilization policies is bounded above by half of the product of the unconditional variance of consumption and the degree of relative risk aversion. Using log utility and data for the U.S., he computes this bound as one-twentieth of one percent of annual consumption. Surprisingly, welfare costs typically found employing Bayesian estimation techniques to compute optimal policies (see Levin, Onatski, Williams, and Williams, 2005; Levine, McAdam, Pearlman, and Pierse, 2008) overshoot the Lucas benchmark by a factor of approximately one hundred. The reason for this can be traced to an overestimation of the welfare-relevant moments of interest. To counter this issue, we follow Del Negro and Schorfheide (2008) and choose prior distributions of the parameters characterizing the exogenous disturbances such that the welfare-relevant theoretical unconditional variances at the posterior mean are in line with the ones found in the data. As a result of this procedure, we obtain business cycle costs consistent with Lucas (2003).

In recent papers Cogley, Colacito, and Sargent (2007) and Cogley, Colacito, Hansen, and Sargent (2008) propose another approach to analyzing optimal policy under model uncertainty. In their setup the central bank faces uncertainty over two competing aggregate macro models, of which one is assumed to be the true data generating process. The central bank seeks to maximize a quadratic loss function, which is weighted with the two model probabilities. To serve this final goal, the policymaker may employ his policy instrument to experiment, to learn and therefore to update his belief about the true model over time. By experimenting systematically the central banker learns faster about the true model and reduces losses due to model uncertainty - even if this leads to transitory suboptimal policies. The authors find it is optimal for the policymaker to follow this avenue. This differs from our approach. We do not take a stand on the true data-generating process. On the contrary, we assume that after some process of theorizing and data analysis the policymaker has arrived at a set of competing models. Since uncertainty about the true model cannot be completely resolved by learning, we focus on how to conduct optimal policy when uncertainty about the true model persists. The advantage of our approach is that we can explicitly model the interaction of optimal decisions by private agents and the monetary authority in a general equilibrium setting. However, a combination of both approaches would require a specification of the interaction of central bank learning with optimal private decisions. This opens up an interesting field for future research.

Related Literature

The literature that employs Bayesian model estimation techniques to assess optimal policies under model uncertainty is relatively new. Our paper is related to Levin, Onatski, Williams and Williams (2005, henceforth LOWW) and to Levine, McAdam, Pearlman and Pierse (2008, henceforth LMPP). LOWW estimate a medium-scale New Keynesian Model with staggered price setting using US data and determine the optimal monetary policy in that model across the posterior distribution of the estimated parameters. Their optimal simple interest rate feedback rule is shown to be robust with respect to parameter uncertainty in structural parameters, including the coefficients of the shock processes. However, as a side aspect, they find the optimal rule not to be robust to different extensions of the model, including a demand for cash and different price-and wage-setting algorithms.

In their recent paper, LMPP also use a medium-scale New Keynesian Model to assess the importance of uncertainty over the degree of indexation in wages and prices on the optimal conduct of policy for the Euro area. They compute optimal simple rules that are robust to this source of specification uncertainty and like LOWW find that monetary policy should respond to wage inflation.

While LOWW and LMPP both start with a medium-scale model, our benchmark is a stripped-to-bare-bones New Keynesian model. This modeling and estimation strategy allows us to quantify the importance of each model component in explaining the data and in the optimal conduct of monetary policy separately. In particular, it puts us in a position to investigate whether optimal policies derived using large models can be recommended if the policymaker is uncertain about the economic environment.

The remainder of the paper is organized as follows. In the next section we present our general framework to analyze the optimal conduct of policy in a broad range of micro-founded macroeconomic models. In the following sections we apply our methodology to optimal monetary policy under model uncertainty. The last section concludes.

2 Analyzing optimal policy under model uncertainty

In this section we describe how we analyze the optimal conduct of policy if the decision-maker faces model uncertainty. We assume that after some process of theorizing and data analysis the policymaker has arrived at a set of competing models. We propose to estimate these models by Bayesian model estimation techniques to explain a set of macroeconomic time series. Next, the relevant source of parameter uncertainty is described by the joint posterior distribution of the structural parameters within each particular model. The relevant degree of uncertainty over the true model (specification uncertainty) enters our analysis by the marginal density of each model. After a short description of the general setup, we show how we derive optimal simple feedback rules that are robust with respect to parameter uncertainty, specification uncertainty and a combination of both. The analysis is based on a reasonable assessment of business cycle costs arising under optimal and sub-optimal policies in and, in particular, across models. For that purpose, at the end of the section we propose a procedure such that Bayesian-estimation results empty into reasonable welfare costs of business cycle fluctuations in line with Lucas (2003).

2.1 General setup

Consider a system of linear equations that represent log-linear approximations to the non-linear equilibrium conditions under rational expectations around a deterministic steady state of a particular model i. Let x_t be the vector of state variables, z_t the vector of structural shocks and y_t the vector of observable variables. Furthermore, based on Bayesian model estimation techniques, let Θ denote the random vector of deep parameters and θ a particular realization from the joint posterior distribution.[1] Policy influences the equilibrium outcome through simple feedback rules. The link between the set of policy instruments as a subset of x is characterized by the vector of constant policy coefficients ϕ, i.e. by definition we consider steady state invariant policies. The state space form of the fundamental solution of model i is given by[2]:

$$\widehat{x}_t = T(\theta_i, \phi)\widehat{x}_{t-1} + R(\theta_i, \phi) z_t \tag{1}$$

$$\widehat{y}_t = G\widehat{x}_t, \tag{2}$$

[1] An and Schorfheide (2007) give an overview of Bayesian estimation procedures in DSGE models.

[2] $\widehat{x}_t$ denotes the percentage deviation of the generic variable x_t from a deterministic steady state x chosen as approximation point.

where $T(\theta_i, \phi)$ and $R(\theta_i, \phi)$ are matrices one obtained after solving a DSGE model with standard solution techniques. However, the entries of the matrixes T and R may differ across models since Θ varies. The benchmark model is the model that restricts some entries of Θ and correspondingly T to zero. Perturbations of the benchmark model are defined as dissolving zero restrictions. Suppose one perturbation of a benchmark model is characterized by converting the forward-looking variable consumption from the benchmark model into a history-dependent variable by assuming habit persistence. Then, the column of T that is associated with lagged consumption has zeros in the benchmark model but at least one non-zero entry in the perturbation. In addition, the vector Θ exhibits a non-zero value for the habit parameter in the perturbation but a zero value in the benchmark model. This formulation implies that all perturbations and the benchmark are nested in the largest model. The matrix G is a picking matrix that equates observable and state variables. Using this equation and the solution given by (1) we estimate model i using data Y with Bayesian model estimation techniques.

Given a certain realization of Θ_i, we assess the performance of a particular policy ϕ with two different criteria: a welfare and an implementability measure. As welfare measure we use a particular quadratic loss function $L(\theta_i, \widehat{x})$, which represents the unconditional expectation of the steady-state invariant part that belongs to a valid second-order approximation of representative households' utility in model i:[3]

$$E \sum_{t=t_0}^{\infty} \beta^t U(x_t, \theta_i) \approx \frac{U(\bar{x}, \theta_i)}{1-\beta} - E \sum_{t=t_0}^{\infty} \beta^t A(\theta_i) \widehat{x}_t \widehat{x}_t' = \frac{U(\bar{x}, \theta_i) - L(\theta_i, \widehat{x})}{1-\beta}. \tag{3}$$

We assume that the policymaker can credibly commit to a policy rule ϕ: if a policymaker decides to follow a certain policy rule ϕ once and forever, agents believe indeed that the policymaker will. Thus, the standard task of optimal policy is to find a particular steady-state invariant policy $\phi_i^\star$, such that the implied solution for $\widehat{x}_t$ minimizes the loss of the

[3]Benigno and Woodford show that a purely quadratic micro-founded loss function can be derived for a wide range of models (2006b). Applications of their method include monetary economics (2005) and the classical Ramsey optimal taxation problem (2006a).

representative agent subject to (1). A steady state-invariant policy is a policy which affects the dynamic evolution of the endogenous variables around a steady state, but not the steady state itself. In any case, we only consider fundamental solutions and require the set of equilibrium sequences to be locally stable and unique. We now turn to the problem of how to determine optimal policy under parameter and specification uncertainty.

2.2 Optimal policy under parameter uncertainty

We treat the mean vector $E[\Theta_i] = \bar{\theta}_i$ of the joint posterior distribution as the true value for the deep parameters of model i. Thus, we think of the joint posterior distribution in model i, $f(\theta_i)$, as the relevant uncertainty that a policymaker faces when he makes his decision about ϕ. The optimal policy problem under parameter uncertainty given the posterior distribution for Θ in model i can be stated as follows:

$$\begin{aligned} &\min_{\phi} E_{\Theta_i} L(\Theta_i, \widehat{x}) = \min_{\phi} \int_{\theta_i} L(\theta_i, \widehat{x}) f(\theta_i) d\theta_i \\ &s.t.\ \widehat{x}_t = T(\theta_i, \phi)\widehat{x}_{t-1} + R(\theta_i, \phi) z_t, \quad \forall \theta_i, \end{aligned} \tag{4}$$

where $E_{\Theta_i} L(\Theta_i, \widehat{x})$ is the expected loss when the structural parameters are a random vector – conditional on model i being the true model. Under parameter uncertainty, the policymaker has to average the loss over all possible realizations of Θ with positive probability to find the optimal vector of constant policy coefficients in model i, $\phi^{\star}_{ipu}$.

If $\phi^{\star}_{ipu}$ differs from $\phi^{\star}_{i}$, and pursuing the latter rather than the former leads to high average welfare losses across the parameter space, then parameter uncertainty matters for the optimal conduct of policy even if the true model is known.

2.3 Specification uncertainty

While under parameter uncertainty the policymaker believes that the given model is the true model, under specification uncertainty he has to deal with a situation where the parameter vector is known to him but the true model is not. We consider and estimate a discrete set of models $\mathcal{M} = \{\mathcal{M}_1, ..., \mathcal{M}_n\}$, where each is characterized by the number and position of zero restrictions in the vectors for deep parameters and variables. By subsequently adding zero restrictions to the largest model $\mathcal{M}_n$, one finally arrives at the benchmark model $\mathcal{M}_1$. Since we estimate all models separately using the same data and the same number of shocks, we can reasonably calculate marginal data densities $p(Y|\mathcal{M}_i)$ and posterior probabilities to assess the probability that model i is the true model. The latter are defined as:

$$P(M_i|Y) = \frac{P(M_i)p(Y|\mathcal{M}_i)}{\sum_{j=1}^{n} P(M_j)p(Y|\mathcal{M}_j)}, \tag{5}$$

where $P(M_i)$ denotes the prior probability for each model. Note that this formulation is not restrictive. It captures the possibility of different assumptions about the policymaker's own beliefs – unless he assigns equal prior weights for all models: avoiding the worst outcome in a particular model i (without ignoring information on how likely or unlikely this is) corresponds to setting a high prior probability for that model. [4]

Suppose that the policymaker knows the exact values for all parameters, $\bar{\theta}_i$, but is uncertain about the true model. Then the optimal policy problem can be formulated in the following way:

$$\begin{aligned} &\min_{\phi} E_{\mathcal{M}} L(\bar{\theta}_i, \widehat{x}) = P(M_1|Y)L(\bar{\theta}_1, \widehat{x}) + ... + P(M_n|Y)L(\bar{\theta}_n, \widehat{x}) \\ &s.t.\ \widehat{x}_t = T(\bar{\theta}_i, \phi)\widehat{x}_{t-1} + R(\bar{\theta}_i, \phi)z_t, \quad i = 1, ..., n, \end{aligned} \tag{6}$$

[4] We employ Geweke's(1999) harmonic mean estimator to compute the data likelihood in a certain model. It takes into account both how well the model fits the data and how many parameters are used to achieve this.

where $\sum_{i=1}^{n} P(M_i|Y) = 1$. The gains from this optimal policy rule $\phi^{\star}_{su}$ can be large if the optimal policy prescriptions in the set of models $\phi^{\star}_1, ..., \phi^{\star}_n$ differ substantially due to the presence of different and possibly conflicting stabilization aims. Therefore the inclusion of a micro-founded loss function in each model, instead of using a single arbitrary quadratic function, plays an important role in adequately accounting for these differences across models.

2.4 Parameter and specification uncertainty

A more realistic case is that the decision maker does not know exactly either the true model or the exact values of all the deep parameters. Uncertainty is then jointly composed of parameter and specification uncertainty. Then the optimal policy problem to determine the overall robustly-optimal policy rule can be stated as follows:

$$\begin{aligned} &\min_{\phi} E_{\mathcal{M},\Theta} L(\Theta_i, \widehat{x}) = \pi_1 E_{\Theta_1} L(\Theta_1, \widehat{x}_t) + ... + \pi_n E_{\Theta_n} L(\Theta_n, \widehat{x}_t) \\ &s.t.\ \widehat{x}_t = T(\theta_i, \phi)\widehat{x}_{t-1} + R(\theta_i, \phi) z_t, \quad \forall \theta_i, \quad i = 1, ..., n. \end{aligned} \tag{7}$$

Comparing this optimal device for policy $\phi^{\star}_{psu}$ with the results of (6) allows us to examine the joint role of specification and parameter uncertainty: should policy makers pay special attention to the interaction of these sources of uncertainty?

2.5 Assessing policy performance within and across models

Throughout the paper we express the resulting business cycle costs as the percentage loss in certainty (steady state) equivalent consumption. First we compute the loss of a certain policy $\tilde{\phi}$ given a particular parameter vector $\tilde{\theta}$ in model i to derive overall utility:

$$U\left(c(\tilde{\theta}_i), x_{\backslash c}(\tilde{\theta}_i), \tilde{\theta}_i\right) - L(\tilde{\theta}_i, \tilde{\phi}),$$

where the first term is steady state utility and $x_{\backslash c}$ denotes the variables vector excluding consumption. Since we want to express utility as reduction in certainty consumption equivalents we set this expression to be equal to:

$$U\left(c(\tilde{\theta}_i)(1-\mathcal{BC}), x_{\backslash c}(\tilde{\theta}_i), \tilde{\theta}_i\right)$$

and solve for $\mathcal{BC}$ in percentage terms. Under parameter uncertainty this results in a distribution for $\mathcal{BC}(\tilde{\theta}_i, \tilde{\phi})$ over Θ_i. Taking the expectation of this expression yields a measure of the average losses in certainty consumption equivalents under a particular policy $\tilde{\phi}$. In the absence of parameter uncertainty we set $\tilde{\theta}_i = E[\Theta_i]$.

As can be seen from (3), theoretical unconditional second moments are relevant for households' utility losses due to short fluctuations – and thus for the computation of business-cycle costs under different policies. As Del Negro and Schorfheide (2008) point out, whether the theoretical unconditional moments relevant for policy assessment derived from the DSGE model and the ones observed in the data coincide depends in particular on the specification of the prior distribution of standard deviations and autoregressive coefficients for the driving exogenous disturbances. We choose the prior distribution for the standard deviations of the *i.i.d.* terms in z_t and the autoregressive coefficients of the shocks contained in $T(\bullet)$ such that the relevant theoretical unconditional second moments at the posterior mean in each model are in line the ones computed directly from the stationary times series. This in turn yields welfare costs of short run fluctuations consistent with the limit put forward by Lucas (2003).

3 Optimal monetary policy: the economic environment

To demonstrate our main result, we create a set of monetary models including one model that nests all features and frictions. Starting with a plain-vanilla cashless new Keynesian economy as a benchmark, following Woodford (2003a), we subsequently introduce two additional features (indexation and habit formation) – that may help or not help to explain the data, and a demand for cash (money in the utility function). While the benchmark economy and the first two extensions are characterized by inflation stabilization as the predominant principle, a transaction friction adds the stabilization of the nominal interest rate as an additional policy aim, which is in conflict with achieving price stability. In this section we describe the models, derive the equations characterizing the equilibrium and the relevant policy objectives as the unconditional expectation of households' utility for each model. In the last subsection we describe how we counter the issue of the zero bound requirement for interest rates.

3.1 The benchmark economy

The benchmark economy consists of a continuum of infinitely-lived households indexed with $j \in [0, 1]$. It is assumed that households have identical initial asset endowments and identical preferences. Household j acts as a monopolistic supplier of labor services l_j. Lower (upper) case letters denote real (nominal) variables. At the beginning of period t, households' financial wealth comprises a portfolio of state contingent claims on other households yielding a (random) payment Z_{jt}, and one-period nominally non-state contingent government bonds B_{jt-1} carried over from the previous period. Assume that financial markets are complete, and let $q_{t,t+1}$ denote the period t price of one unit of currency in a particular state of period $t+1$ normalized by the probability of occurrence of that state, conditional on the information available in period t. Then, the price of a random payoff Z_{t+1} in period $t+1$ is given by $E_t[q_{t,t+1}Z_{jt+1}]$. The budget constraint of the representative household reads

$$B_{jt} + E_t[q_{t,t+1}Z_{jt+1}] + P_t c_{jt} \leq R_{t-1}B_{jt-1} + Z_{jt} + P_t w_{jt} l_{jt} + \int_0^1 D_{jit} di - P_t T_t, \qquad (8)$$

where c_t denotes a Dixit-Stiglitz aggregate of consumption with elasticity of substitution ζ, P_t the aggregate price level, w_{jt} the real wage rate for labor services l_{jt} of type j, T_t a lump-sum tax, R_t the gross nominal interest rate on government bonds, and D_{it} dividends from monopolistically competitive firms. Further, households have to fulfil the no-Ponzi game condition: $\lim_{i\rightarrow\infty} E_t q_{t,t+i}(B_{jt+i} + Z_{jt+1+i}) \geq 0$. The objective of the representative household is

$$E_{t_0} \sum_{t=t_0}^{\infty} \beta^t \{u(c_{jt}) - v(l_{jt})\}, \quad \beta \in (0,1), \tag{9}$$

where β denotes the subjective discount factor. The instantaneous utility function is assumed to be non-decreasing in consumption, decreasing in labor time, strictly concave, twice continuously differentiable, and to fulfill the Inada conditions. Households are wage-setters supplying differentiated types of labor l_j, which are transformed into aggregate labor l_t with $l_t^{(\epsilon_t-1)/\epsilon_t} = \int_0^1 l_{jt}^{(\epsilon_t-1)/\epsilon_t} dj$. We assume that the elasticity of substitution between different types of labor, $\epsilon_t > 1$, varies exogenously over time. The time variation in this markup parameter introduces a so called cost-push shock into the model that gives rise to a stabilization problem for the central bank. Cost minimization implies that the demand for differentiated labor services l_{jt}, is given by $l_{jt} = (w_{jt}/w_t)^{-\epsilon_t} l_t$, where the aggregate real wage rate w_t is given by $w_t^{1-\epsilon_t} = \int_0^1 w_{jt}^{1-\epsilon_t} dj$. Maximizing (9) subject to (8) and the no-Ponzi game condition for given initial values Z_0, B_{t_0-1}, and $R_{t_0-1} \geq 0$ leads to the following first-order conditions for consumption, the real wage rate for labor type j, government bonds, and contingent claims:

$$\lambda_{jt} = u_c(c_{jt}), \quad v_l(l_{jt}) = w_{jt}\lambda_{jt}/\mu_t, \tag{10}$$

$$q_{t,t+1} = \frac{\beta\lambda_{jt+1}}{\pi_{t+1}\lambda_{jt}}, \quad \lambda_{jt} = \beta R_t E_t \frac{\lambda_{jt+1}}{\pi_{t+1}} \tag{11}$$

where λ_{jt} denotes a Lagrange multiplier, π_t the inflation rate $\pi_t = P_t/P_{t-1}$, and $\mu_t = \epsilon_t/(\epsilon_t - 1)$ the stochastic wage mark-up with mean $\bar{\mu} > 1$. The first order condition for contingent claims holds for each state in period $t+1$, and determines the price of one unit of currency for a particular state at time $t+1$ normalized by the conditional probability of occurrence of that state in units of currency in period t. Arbitrage-freeness between government bonds and contingent claims requires $R_t = 1/E_t q_{t,t+1}$. The optimum is further characterized by the budget constraint (8) holding with equality and by the transversality condition $\lim_{i\to\infty} E_t\beta^i\lambda_{jt+i}(B_{jt+i} + Z_{jt+1+i})/P_{jt+i} = 0$.

The final consumption good Y_t is an aggregate of differentiated goods produced by monopolistically competitive firms indexed with $i \in [0,1]$ and defined as $y_t^{\frac{\zeta-1}{\zeta}} = \int_0^1 y_{it}^{\frac{\zeta-1}{\zeta}} di$, with $\zeta > 1$. Let P_{it} and P_t denote the price of good i set by firm i and the price index for the final good. The demand for each differentiated good is $y_{it}^d = (P_{it}/P_t)^{-\zeta} y_t$, with $P_t^{1-\zeta} = \int_0^1 P_{it}^{1-\zeta} di$. A firm i produces good y_i using a technology that is linear in the labor bundle $l_{it} = [\int_0^1 l_{jit}^{(\epsilon_t-1)/\epsilon_t} dj]^{\epsilon_t/(\epsilon_t-1)}$: $y_{it} = a_t l_{it}$, where $l_t = \int_0^1 l_{it} di$ and a_t is a productivity shock with mean 1. Labor demand satisfies: $mc_{it} = w_t/a_t$, where $mc_{it} = mc_t$ denotes real marginal costs independent of the quantity that is produced by the firm. We allow for a nominal rigidity in form of a staggered price setting as developed by Calvo (1983). Each period firms may reset their prices with probability $1-\alpha$ independently of the time elapsed since the last price setting. A fraction $\alpha \in [0,1)$ of firms are assumed to keep their previous period's prices, $P_{it} = P_{it-1}$, i.e. indexation is absent. Firms are assumed to maximize their market value, which equals the expected sum of discounted dividends $E_t \sum_{T=t}^{\infty} q_{t,T} D_{iT}$, where $D_{it} \equiv P_{it}y_{it}(1-\tau) - P_t mc_t y_{it}$ and firms also have access to contingent claims. Here, τ denotes an exogenous sales tax introduced to offset the inefficiency of steady-state output due to markup pricing following Rotemberg and Woodford (1999). In each period a measure $1-\alpha$ of randomly selected firms set new prices $\widetilde{P}_{it}$ as the solution to $\max_{\widetilde{P}_{it}} E_t \sum_{T=t}^{\infty} \alpha^{T-t} q_{t,T}(\widetilde{P}_{it} y_{iT}(1-\tau) - P_T mc_T y_{iT})$, s.t. $y_{iT} = (\widetilde{P}_{it})^{-\zeta} P_T^{\zeta} y_T$. The first-order condition for the price of re-optimizing producers for $\alpha > 0$ is given by

$$\frac{\widetilde{P}_{it}}{P_t} = \frac{\zeta}{\zeta - 1}\frac{F_t}{K_t}, \tag{12}$$

where K_t and F_t are given by the following expressions:

$$F_t = E_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} u_c(c_T) y_T \left(\frac{P_T}{P_t}\right)^{\zeta} mc_T \tag{13}$$

and

$$K_t = E_t \sum_{T=t}^{\infty} (\alpha\beta)^{T-t} u_c(c_T)(1-\tau) y_T \left(\frac{P_T}{P_t}\right)^{\zeta-1}. \tag{14}$$

Aggregate output is given by $y_t = a_t l_t / \Delta_t$, where $\Delta_t = \int_0^1 (P_{it}/P_t)^{-\zeta} di \geq 1$ and thus $\Delta_t = (1-\alpha)(\widetilde{P}_t/P_t)^{-\zeta} + \alpha \pi_t^{\zeta} \Delta_{t-1}$. The dispersion measure Δ_t captures the welfare decreasing effects of staggered price setting. If prices are flexible, $\alpha = 0$ and the first order condition for the optimal price of the differentiated good reads: $mc_t = (1-\tau)\frac{\zeta-1}{\zeta}$.

The public sector consists of a fiscal and a monetary authority. The central bank as the monetary authority is assumed to control the short-term interest rate R_t with a simple feedback rule contingent on past interest rates, inflation and output:

$$R_t = f(R_{t-1}, \pi_t, y_t). \tag{15}$$

The fiscal authority issues risk-free one period bonds, has to finance exogenous government expenditures $P_t G_t$, receives lump-sum taxes from households and tax-income from an exogenous given constant sales tax τ, such that the consolidated budget constraint reads: $R_{t-1}B_{t-1} + P_t G_t = B_t + P_t T_t + \int_0^1 P_{it} y_{it} \tau di$. The exogenous government expenditures G_t evolve around a mean $\bar{G}$, which is restricted to be a constant fraction of output, $\bar{G} = \bar{y}(1-sc)$. We assume that tax policy guarantees government solvency, i.e., ensures $\lim_{i\to\infty} (B_{t+i}) \prod_{v=1}^{i} R_{t+v}^{-1} = 0$. Due to the existence of the lump-sum tax, we consider only the demand effect of government expenditures and focus exclusively on optimal monetary policy.

We collect the exogenous disturbances in the vector $\xi_t = [a_t, G_t, \mu_t]$. It is assumed that the percentage deviations of the first two elements of the vector from their means evolve according to autonomous AR(1)-processes with autocorrelation coefficients $\rho_a, \rho_G \in [0,1)$. The process for $\log(\mu_t/\bar{\mu})$ and all innovations, $z_t = [\epsilon_t^a, \epsilon_t^g, \epsilon_t^\mu]$, are assumed to be i.i.d..

The recursive equilibrium is defined as follows:

Definition 1 *Given initial values $P_{t_0-1} > 0$ and $\Delta_{t_0-1} \geq 1$, a monetary policy and a ricardian fiscal policy $T_t\ \forall\, t \geq t_0$, and a sales tax τ, a rational expectations equilibrium (REE) for $R_t \geq 1$, is a set of sequences $\{y_t,\ c_t,\ l_t,\ mc_t,\ w_t,\ \Delta_t,\ P_t,\ \widetilde{P}_{it}, R_t\}_{t=t_0}^\infty$ satisfying the following conditions: firms' first order condition $mc_t = w_t/a_t$, (12) with $\widetilde{P}_{it} = \widetilde{P}_t$, and $P_t^{1-\zeta} = \alpha P_{t-1}^{1-\zeta} + (1-\alpha)\widetilde{P}_t^{1-\zeta}$, the households' first order conditions $u_c(y_t - G_t)w_t = v_l(l_t)\mu_t$, $u_c(y_t - G_t)/P_t = \beta R_t E_t u_c(y_{t+1} - G_{t+1})/P_{t+1}$, the aggregate resource constraint $y_t = a_t l_t/\Delta_t$, where $\Delta_t = (1-\alpha)(\widetilde{P}_t/P_t)^{-\zeta} + \alpha(P_t/P_{t-1})^\zeta \Delta_{t-1}$, clearing of the goods market $c_t + G_t = y_t$, and the transversality condition, for $\{\xi_t\}_{t=t_0}^\infty$.*

In the next step, we seek to estimate the model by employing Bayesian methods. To do so, we log-linearize the structural equations around the deterministic steady state under zero inflation. Thus, the dynamics in the benchmark economy are described by the following two structural equations:

$$\sigma(E_t\widehat{y}_{t+1} - E_t\widehat{y}_{t+1}^n) = \sigma(\widehat{y}_t - \widehat{y}_t^n) + \widehat{R}_t - E_t\widehat{\pi}_{t+1} - \widehat{R}_t^n \tag{16}$$

$$\widehat{\pi}_t = \beta E_t\widehat{\pi}_{t+1} + \kappa(\widehat{y}_t - \widehat{y}_t^n), \tag{17}$$

where $\sigma = -u_{cc}c/(u_c sc)$, $\omega = v_{ll}l/v_l$ and $\kappa = (1-\alpha)(1-\alpha\beta)(\omega+\sigma)/\alpha$. Furthermore, $\widehat{k}_t$ denotes the percentage deviation of a generic variable k_t from its steady-state value k. The natural rates of output and interest, i.e the values for output and real interest under flexible prices, are given by the following expressions

$$\widehat{y}_t^n = \frac{(1+\omega)\widehat{a}_t + \sigma g_t - \widehat{\mu}_t}{\omega + \sigma}, \tag{18}$$

where $g_t = (G_t - G)/y$ and

$$\widehat{R}_t^n = \sigma[(g_t - \widehat{y}_t^n) - E_t(g_{t+1} - \widehat{y}_{t+1}^n)]. \tag{19}$$

The model is closed by a simple interest rate feedback rule:

$$\widehat{R}_t = \rho_R \widehat{R}_{t-1} + \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t. \tag{20}$$

The general system (1) in the benchmark model then is the fundamental locally stable and unique solution that satisfies (16)-(20) for a certain vector of constant policy coefficients $\phi = (\rho_R, \phi_\pi, \phi_y)$.

Our welfare measure is the unconditional expectation of representative households' utility. Building on Woodford (2003a), after averaging over all households, the purely quadratic approximation to (9) in our benchmark model can be computed as[5]:

$$\frac{1}{1-\beta}[u - \frac{u_c y \zeta(\omega + \sigma)}{2\kappa}\{var(\widehat{\pi}_t) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e)\}], \tag{21}$$

where $\lambda_d = \kappa/\zeta$ and the efficient rate of output is given by

$$\widehat{y}_t^e = \widehat{y}_t^n + \widehat{\mu}_t/(\omega + \sigma). \tag{22}$$

Note that from the perspective of the central bank the vector of structural parameters may be random. Therefore the objective of the central bank (for a given realization θ_i) reads:[6]

$$L(\theta_i, \widehat{x}) = \frac{u_c y \zeta(\omega + \sigma)}{2\kappa}\{var(\widehat{\pi}_t) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e)\}. \tag{23}$$

[5]Throughout we assume that the steady state is rendered efficient by an appropriate setting of the sales tax rate.

[6]Note that the argument $\widehat{x}$ does not depend on time since we compute unconditional variances, i.e. $var(\widehat{x}_t) = var(\widehat{x})$, $\forall t$.

In the next subsection we consider habit formation and indexation to past inflation as examples of missing lags in the nuclear variables output, and inflation.

3.2 Additional features: habit formation and indexation

One example of a missing lag in an endogenous variable is to allow for an internal habit (e.g. Boivin and Giannoni (2006); Woodford (2003a)) in households' total consumption. In that case, the marginal utility of consumption λ_t is no longer exclusively a function of current consumption. Instead, the amount consumed in the last period affects households' behavior:

$$\lambda_t = u_c(c_t - \eta c_{t-1}) - \beta\eta E_t u_c(c_{t+1} - \eta c_t), \quad 0 \leq \eta \leq 1, \tag{24}$$

with η as the habit parameter. Correspondingly, the constituting equations for (1) are the policy rule (20) and the modified versions of the Euler equation and the New Keynesian Philips curve:

$$\begin{aligned} \varphi[d_t - \eta d_{t-1}] - \varphi\beta\eta E_t[d_{t+1} - \eta d_t] &= E_t\widehat{\pi}_{t+1} + \widehat{R}_t^n - \widehat{R}_t ... \\ &+ E_t\varphi[d_{t+1} - \eta d_t] - \varphi\beta\eta E_t[d_{t+2} - \eta d_{t+1}] \end{aligned} \tag{25}$$

$$\widehat{\pi}_t = \kappa_h[(d_t - \delta^* d_{t-1}) - \beta\delta^* E_t(d_{t+1} - \delta^* d_t)] + \beta E_t\widehat{\pi}_{t+1}, \tag{26}$$

where $d_t = \widehat{y}_t - \widehat{y}_t^n$, $\kappa_h = \eta\varphi\kappa[\delta^*(\omega + \sigma)]^{-1}$, $\varphi = \sigma/(1 - \eta\beta)$, and the natural rate of output follows[7]

$$\begin{aligned} [\omega + \varphi(1 + \beta\eta^2)]\widehat{y}_t^n - \varphi\eta\widehat{y}_{t-1}^n - \varphi\eta\beta E_t\widehat{y}_{t+1}^n &= \varphi(1 + \beta\eta^2)g_t - \varphi\eta g_{t-1} - \varphi\eta\beta E_t g_{t+1} ... \\ &+ (1 + \omega)\widehat{a}_t - \widehat{\mu}_t. \end{aligned} \tag{27}$$

[7]The parameter δ^*, $0 \leq \delta^* \leq \eta$, is the smaller root of the quadratic equation $\eta\varphi(1 + \beta\delta^2) = [\omega + \varphi(1 + \beta\eta^2)]\delta$. This root is assigned to past values of the natural and efficient rate of output in their stationary solutions.

Approximating households' utility to second order results in the following expression for the objective of the central bank:

$$L(\theta_2, \widehat{x}) = \frac{(1-\beta\eta)\eta\varphi u_c^h y^h \zeta}{2\kappa_h \delta^*}\{var(\widehat{\pi}_t) + \lambda_{d,h} var(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e))\}, \tag{28}$$

where $\lambda_{d,h} = \kappa_h/\zeta$ and the efficient rate of output is characterized by

$$\begin{aligned} [\omega + \varphi(1+\beta\eta^2)]\widehat{y}_t^e - \varphi\eta\widehat{y}_{t-1}^e - \varphi\eta\beta E_t\widehat{y}_{t+1}^e &= \varphi(1+\beta\eta^2)g_t - \varphi\eta g_{t-1} - \varphi\eta\beta E_t g_{t+1} ... \\ &+ (1+\omega)\widehat{a}_t. \end{aligned} \tag{29}$$

Like habit formation, the indexation of prices to past inflation induces the economy to evolve in a history-dependent way. We assume that the fraction of prices that are not reconsidered, α, adjusts according to the following simple rule:

$$\log P_{it} = \log P_{it-1} + \gamma \log \pi_{t-1}, \tag{30}$$

with $0 \leq \gamma \leq 1$ as the degree of indexation. This implies that the dispersion measure is given by

$$\Delta_t = (1-\alpha)(\frac{\widetilde{P}_t}{P_t})^{-\zeta} + \alpha\pi_{t-1}^{-\zeta\gamma}\Delta_{t-1}\pi_t^{\zeta}.$$

Correspondingly, the economy with indexation is characterized by a modified aggregate supply curve

$$\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1} = \beta E_t(\widehat{\pi}_{t+1} - \gamma\widehat{\pi}_t) + \kappa(\widehat{y}_t - \widehat{y}_t^n), \tag{31}$$

(16) and (20). The corresponding loss function of the central bank reads (Woodford, 2003):

$$L(\theta_3, \widehat{x}) = \frac{u_c y\zeta(\omega+\sigma)}{2\kappa}\{var(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1}) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e)\}, \tag{32}$$

where λ_d and the efficient rate of output are defined as in the benchmark economy.

3.3 Additional friction: money in utility function

While the extensions discussed above modify preferences or technology, they do not add other frictions to the economy. The primary aim – in principle – is to stabilize inflation. In order to allow for the possibility of conflicting stabilization aims, we introduce a transaction friction by letting real money balances enter households' utility in a separable way. Notably, this change does not affect the dynamics of inflation and output for a given monetary policy.[8] More precisely, an euler equation or demand equation for real money balances enters the set of equilibrium conditions:

$$\frac{z_m(m_t)}{\lambda_t} = \frac{R_t - 1}{R_t}. \tag{33}$$

We assume that $z(m_t)$ implies satiation in real money balances at a finite positive level. The derivatives z_m, z_{mm} have finite limiting values as m approaches the satiation level from below. In particular, the limiting value of z_{mm} from below is negative (see Woodford, 2003a, Assumption 6.1). Log-linearizing (33) results in:

$$\widehat{m}_t = -\frac{1}{\sigma_m(R-1)}\widehat{R}_t - \frac{1}{\sigma_m}\widehat{\lambda}_t, \tag{34}$$

where $\sigma_m = -z_{mm}m/z_m$. In this case the stabilization loss that results is given by:

$$L(\theta_4, \widehat{x}) = \frac{u_c y\zeta(\omega+\sigma)}{2\kappa}\{var(\widehat{\pi}_t) + \lambda_d var(\widehat{y}_t - \widehat{y}_t^e) + \lambda_{1R} var(\widehat{R}_t)\}, \tag{35}$$

where $\lambda_d = \kappa/\zeta$, $\lambda_{1R} = \lambda_d\beta[v(\omega+\sigma)(1-\beta)\sigma_m]^{-1}$ and $v = y/m$. The general form (1) has to satisfy the interest rate rule, (16)-(17) and (34).

In the following proposition we derive a quadratic to households' utility when we combine the features habit formation, indexation and a demand for cash. Later on we refer to this

[8] However, we assume that this does not lead to an optimal steady state that corresponds to Friedman's rule, i.e. the approximation point is still characterized by flexible prices and zero inflation (see Paustian and Stoltenberg (2008)). Note that our specification of utility is consistent with recent findings by Andrés, López-Salido, and Vallés (2006) for the Euro area. They estimate the role of money for the business cycle of the Euro area and the US and find that preferences are separable between consumption and real money balances.

model as the large model.

Proposition 1 *If the fluctuations in y_t around y, R_t around R, ξ_t around ξ, π_t around π are small enough, $(R-1)/R$ is small enough, and if the steady state distortions ϕ vanish due to the existence of an appropriate subsidy τ, the utility of the average household can be approximated by:*

$$U_{t_0} = -E_{t_0} \sum_{t=t_0}^{\infty} \beta^{t-t_0} L(\theta_5, \widehat{x}) + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t, (R-1)/R\|^3), \tag{36}$$

where t.i.s.p. indicate terms independent of stabilization policy,

$$L(\theta_5, \widehat{x}) = \frac{(1-\beta\eta)\eta\varphi u_c^h y^h \zeta}{2\kappa_h \delta^*}\{var(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1}) + \lambda_{d,h} var(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e)) + \lambda_{2R} var(\widehat{R}_t)\}, \tag{37}$$

$$\lambda_x = \lambda_{d,h} = \kappa_h/\zeta, \tag{38}$$

$$\lambda_{2R} = \frac{\lambda_{d,h}\beta\delta^*}{v\sigma_m(1-\beta)\eta\varphi}, \tag{39}$$

and $v = y/m > 0$.

Proof: see appendix A.1. The equilibrium conditions in this case are (25), (34) and

$$\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1} = \beta E_t(\widehat{\pi}_{t+1} - \gamma\widehat{\pi}_t) + \kappa_h[(d_t - \delta^* d_{t-1}) - \beta\delta^* E_t(d_{t+1} - \delta^* d_t)] \tag{40}$$

for a monetary policy (20).

3.4 Monetary policy and the lower bound on interest

The question of implementability is of particular interest in monetary policy: the zero bound on interest rates, $R_t \geq 1$, imposes a natural restriction on the set of feasible policies.[9] We approximate this constraint by analyzing implementable policies.

[9]As Sims (2001) points out by the zero bound on interest should be taken especially seriously under model uncertainty.

Definition 2 *Consider the locally stable and unique fundamental solution for model i under a particular policy ϕ that satisfies (1) and the transversality condition. The policy ϕ is implementable if the unconditional standard deviation of the net nominal interest rate σ_i under the fundamental solution is at least k- times smaller than the steady state value for interest:*

$$k\sigma_i \leq R - 1, \quad k > 0.$$

Notably, the higher is k, the lower is the probability that the zero bound on interest rates in fact becomes binding. We stress that this approach ignores certain feedback channels.[10] Nevertheless, computing the standard deviation of the nominal interest rate is one way to gauge the severity of the lower bound constraint in linear models (see Woodford (2003a) or more recently Schmitt-Grohé and Uribe (2007)).

Throughout our optimal policy analysis we compute two vectors $\phi^{\star}$: either requiring optimal policy to be implementable or not. While the corresponding results for the latter case will be discussed in the next section, the analysis of the zero bound requirement can be found in appendix A.3.

4 Results

In this section we first present and interpret the estimation results. These results will be key for the assessment of the relevant model uncertainty faced by the policymaker. In the second part we compute optimal simple rules along with the procedures laid out in section 2. As a standard, we determine optimal monetary policy at the posterior mean, i.e. optimal policy in the absence of any model uncertainty. Then we analyze optimal policy in the presence of parameter and specification uncertainty to deliver our main result: when the policymaker is uncertain about the true model, the optimal policy in the model that comprises all features and frictions cannot be recommended as a policy device. By comparing these results to the

[10] Adam and Billi (2006) and Eggertsson and Woodford (2003) explicitly account for the non-linear zero bound constraint and show how the possibility of a binding constraint affects agents' decisions.

optimal policy at the posterior mean we explore the pitfalls of ignoring the issue of model uncertainty.

4.1 Data and estimation results

We treat the variables real wage, output and consumer price inflation as observables. The data consists of HP filtered[11] quarterly values of these variables for the EU 13 countries from 1970-2006 [12].

We do not estimate the discount factor $\beta = 0.99$, the fraction of private consumption relative to GDP $c/y = 0.8$ and the elasticity of substitution between differentiated goods $\zeta = 6$. While we assume the disturbances g_t and $\widehat{a}_t$ to follow stationary $AR(1)$ processes, $\widehat{\mu}_t$ is supposed to be $i.i.d.$(see appendix A.2 for the assumed prior distribution of the estimated parameters). Since we are interested in evaluating the explanatory power of each extension of the benchmark model separately, this requires common parameters in the set of models to exhibit the same sufficient prior statistics. In particular, the marginal prior distributions for the set of coefficients that describe the shock processes, ψ_g, ψ_a and $\sigma_g, \sigma_a, \sigma_\mu$, do not change across models, and they are specified according to the procedure explained in Section 2.5.

We approximate the joint posterior distribution of structural parameters by drawing $100,000$ times employing an MCMC-algorithm. To insure convergence, we discard the first $80,000$ draws. In general, all deep parameters except the relative risk aversion with respect to real money balances σ_m are identified (see Figure 1-5 and Table 9 in appendix A.2). The fact that we are not able to identify σ_m comes as no surprise since it only appears in equations that have no influence on predicting output, inflation or the real wage.

In order to assess the relevance of each perturbation, we compute marginal likelihoods and the corresponding posterior odds as their weighted average.[13] The results are presented in Table 1. Note that adding features to the model does not necessarily increase the marginal

[11] We set $\lambda_{HP} = 1600$ as suggested by Hodrick and Prescott (1997).

[12] The data-set we use was kindly provided by the Euro Area Business Cycle Network (EABCN). For a description of how this data is constructed see Fagan, Henry, and Mestre (2001).

[13] We employ equal prior weights for each model.

likelihood. This is supported by the following observations: First, enriching the benchmark model with a demand for cash lowers the marginal likelihood, since real money balances do not help to predict the observable variables. On the contrary, the additional parameter to estimate (σ_m) blurs the precision of the remaining parameters slightly (1683.57 < 1683.98). Second, while habit formation does modify the dynamics of the model it does not improve the fit of the model. This points to a well-known problem in Bayesian model estimation, namely that the informative prior on the habit parameter introduces curvature into the posterior density surface, which tilts it away from being zero at the posterior, which would increase the predictive density. Third, history dependence in inflation seems to improve the predictive power of the model, and Model 3 exhibits with approximately 81% the highest posterior probability. Thus, the largest model incorporates features that helps to predict the data (indexation) and others that do not (habit and money). It therefore exhibits a higher marginal likelihood than the benchmark model but lower than the best one.

Table 1: Posterior probabilities and marginal data densities

	M_1	M_2	M_3	M_4	M_5
$p(Y\|\mathcal{M}_i)$	1683.98	1682.69	1696.83	1683.57	1695.39
$P(M_i\|Y)$	0.00	0.00	0.81	0.00	0.19

The welfare-assessment of optimal and sub-optimal policies in and across models depends on the magnitude of the resulting stabilization losses, i.e. the welfare relevant unconditional variances or standard deviations. In our context, these are the unconditional fluctuations in inflation and consumption (expressed in terms of a welfare-relevant output gap) for the models without a transaction friction (see e.g. (23)), and additionally fluctuations in interest rates, when money enters the utility function (see e.g (37)). As can be verified in Table 2, our estimated theoretical moments at the posterior mean are consistent with the corresponding ones directly estimated from the stationary times series. Due to this we obtain reasonable business-cycle costs under optimal policies.

Table 2: Welfare-relevant standard deviations: models vs. data

	M_1	M_2	M_3	M_4	M_5	*data*
$std(c, \bar{\theta}_i)$	0.0070	0.0090	0.0068	0.0070	0.0078	0.0073
$std(\pi, \bar{\theta}_i)$	0.0020	0.0020	0.0023	0.0020	0.0022	0.0020
$std(R, \bar{\theta}_i)$	0.0028	0.0027	0.0031	0.0028	0.0031	0.0028

In the next section we begin the analysis of optimal policies in and across models.

4.2 Optimal policy at the posterior mean

To establish a standard, we determine the optimal policy $\phi_i^\star = (\rho_R^\star, \phi_\pi^\star, \phi_y^\star)_i$ at the posterior mean $\bar{\theta}_i$ for each model i, $i = 1, 2, .., 5$, i.e. if the policymaker is neither uncertain about the true model nor about the corresponding parameters. To ease the numerical computation and to exclude unreasonably high policy responses, we assume the following bounds for the policy coefficients of the simple interest rate rule:

$$\rho_R \in [0, 20], \quad \rho_\pi \in [0, 20], \text{ and } \rho_y \in [0, 20].$$

The optimal coefficients and the resulting business cycles costs ($\mathcal{BC}$) expressed as equivalent reductions in steady-state consumption are displayed in Table 3.

Table 3: Optimal policy at the posterior mean ($\phi_i^\star$)

$$\widehat{R}_t = \rho_R \widehat{R}_{t-1} + \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t$$

	M_1	M_2	M_3	M_4	M_5
ρ_R	0.81	1.05	0.62	1.26	1.36
ϕ_π	20.00	20.00	20.00	2.42	1.01
ϕ_y	0.00	0.02	0.00	0.00	0.00
$\mathcal{BC}(\bar{\theta}_i, \phi_i^\star)$	0.0014%	0.0014%	0.0020%	0.0194%	0.0178%

Optimal policies are characterized by drawing on past interest rates. Put differently, optimal policy is history-dependent (Woodford, 2003a,b). In the first three models inflation stabilization is the predominant aim. Correspondingly, optimal policies feature a strong reaction on inflation. In models 4 and 5, households value real money balances as a medium for transactions. This introduces stabilization of the nominal interest rate as a conflicting aim to price stability (see (35) and (36)) in the presence of fluctuations in the natural rate of interest. For intuition on this, suppose that ϕ_y is small and that the economy in Model 1 is hit by a cost-push shock. To fight inflationary tendencies the output gap must decrease according to the aggregate supply curve (17). This in turn requires a strong increase in the nominal interest rate to fulfil the euler equation (16), since the cost-push shock affects the natural rate of interest. Therefore, optimal policies in models with a demand for cash exhibit a higher coefficient ρ_R to smooth interest rates and a less aggressive response to inflation.[14]

According to this, welfare costs in models that feature a transaction friction are substantially higher. This increase is due to two effects. First, the stabilization of the interest rate adds a new component to the welfare-relevant stabilization loss, which accounts for over fifty percent of the increase in business cycle costs in Model 4 relative to Model 1. The second effect relates to the conflict of stabilizing interest rates, inflation and the output gap simultaneously, as apparent in the muted response to inflation in the optimal rules for models 4 and 5. The resulting increase in the unconditional weighted variances of inflation and the output gap accounts for the remaining increase in costs of business cycle fluctuations.

Table 4: The weights λ_d and λ_R at the posterior mean

Weights	M_1	M_2	M_3	M_4	M_5
λ_d	0.0063	0.0231	0.0079	0.0057	0.0328
λ_R	-	-	-	0.0602	0.0728

Table 4 shows how the importance of stabilization aims relative to inflation for households

[14]Note that even if $\rho_\pi < 1$ equilibrium sequences are uniquely determined and converge to the steady state. The condition for local stability and uniqueness of equilibrium sequences reads: $\rho_\pi + (1-\beta)/\kappa\phi_y > 1 - \rho_R$ (Woodford, 2003).

changes across models. For example, the stabilization of the output gap is five times more important in Model 5 than in Model 1. In addition, the exact gap that policy should stabilize to maximize welfare differs (see (23) and (36)). Furthermore, comparing the two models that feature a demand for cash reveals that the optimal response to changes in inflation is larger in Model 4 than in Model 5. Although both specifications incorporate stabilizing the nominal interest rate as a policy aim, this aim is relatively more important in Model 5 than in Model 4.

4.3 Optimal monetary policy under model uncertainty

In this section we analyze how a central bank should act when it is uncertain about either the true data generating process, about the parameters of a given model, or about a combination of both. In Table 5 we list average business cycle costs in the presence of model uncertainty. For example, row one displays the resulting welfare losses from following optimal policy prescriptions derived at the posterior mean in Model 1 ($\phi_1^\star$) across all draws in models 1, 3 and 5: the first column of this row captures parameter uncertainty, the second and third both sources of uncertainty jointly.

Table 5: Welfare losses due to model uncertainty expressed as consumption equivalents

	M_1	M_3	M_5
$\mathcal{BC}(\Theta_i, \phi_1^\star) - \mathcal{BC}(\Theta_i, \phi_{ipu}^\star)$	0.00%	0.00%	0.07%
$\mathcal{BC}(\Theta_i, \phi_3^\star) - \mathcal{BC}(\Theta_i, \phi_{ipu}^\star)$	0.00%	0.00%	0.08%
$\mathcal{BC}(\Theta_i, \phi_5^\star) - \mathcal{BC}(\Theta_i, \phi_{ipu}^\star)$	0.02%	0.01%	0.00%

As can be seen from column 3, the highest additional stabilization losses due to model uncertainty arise in the model that nests all features and frictions. Thus, at a first glance, following an interest-rate rule that is optimal in this model may not only guard against model uncertainty by including all features and frictions but can also be motivated as a minmax approach to model uncertainty. Employing the optimal policy rule in Model 5 as a guard against specification uncertainty, but taking into account parameter uncertainty in that model as described in (5), yields the following feedback rule:

$$\phi^{\star}_{5pu}: \quad \rho_R = 1.34; \quad \phi_\pi = 1.17; \quad \phi_y = 0.00.$$

However, Model 5 is not the most likeliest model since it also contains features which do not help to explain the given time series of GDP, inflation and the real wage (see Table 1). A policymaker pursuing a Bayesian approach to model uncertainty weights welfare losses in a particular model with its posterior probability, i.e. derives an optimal policy over all draws and models according to (7):

$$\phi^{\star}_{psu}: \quad \rho_R = 1.39; \quad \phi_\pi = 3.36; \quad \phi_y = 0.00.$$

Comparing the characteristics of the two rules reveals two similarities and one difference. Both rules draw heavily on past interest rates to avoid welfare-reducing fluctuations in the interest rate in models 4 and 5, and put no emphasize on stabilizing the output gap. The main difference between both rules is the preference to stabilize inflation. While there is a conflict in stabilizing inflation and the nominal interest rate jointly in Model 5, this trade-off is absent in the most likeliest model, Model 3. Comparing the differences in business cycle cost in row one from Table 6 with the absolute values listed in Table 3, the optimal rule $\phi^{\star}_{psu}$ performs substantially better in Models 1, 2 and 3 where inflation stabilization is the predominant principle. Nevertheless, by reacting less harshly to inflation than the optimal rules from those models, it avoids high welfare losses in Model 5. On average, the optimal rule leads to 41% lower business cycle costs relative to the optimal policy rule derived from the large model. Alternatively, the policymaker could be tempted to follow policy prescriptions derived in the most likeliest model (Model 3). But this cannot be recommended either: this policy rule fails to prevent high stabilization losses in Model 5, and it leads to an increases in business cycle costs of 74% on average compared to the optimal policy under model uncertainty. In other words, even if a certain friction does not add any explanatory power it may be relevant for welfare considerations and for the assessment of optimal policies. Summing up, optimal policy should take into account both, possible high stabilization losses and their probability of occurrence.

Table 6: Relative performance of policy under model uncertainty

	M_1	M_2	M_3	M_4	M_5	WM_{rel}
$\mathcal{BC}(\Theta_i, \phi^\star_{5pu}) - \mathcal{BC}(\Theta_i, \phi^\star_{psu})$	0.01%	0.01%	0.01%	0.01%	−0.01%	1.41
$\mathcal{BC}(\Theta_i, \phi^\star_{3pu}) - \mathcal{BC}(\Theta_i, \phi^\star_{psu})$	-0.01%	-0.01%	-0.01%	0.03%	0.07%	1.74

WM_{rel} indicates the relation of weighted business cycle costs over the modelspace. For model j, $j = 3, 5$, this is defined as $WM_{rel} = E_{\mathcal{M},\Theta}\mathcal{BC}(\Theta_i, \phi^\star_{jpu})/E_{\mathcal{M},\Theta}\mathcal{BC}(\Theta_i, \phi^\star_{psu})$.

Our analysis so far stresses the importance of both sources of model uncertainty, implying substantial welfare gains of a policy that incorporates this risk. Consequently, a natural question arises as to whether these gains are mainly due to parameter or specification uncertainty. We find that specification uncertainty is more important than parameter uncertainty. First, as can be seen from the entries on the diagonal of Table 5, the mistake of ignoring parameter uncertainty is negligible if the policymaker knows the true model. Moreover, in Table 7 we compare optimal policy rules that focus on specification uncertainty only (computed according to (6)), $\phi^\star_{su}$, to optimal policy under specification and parameter uncertainty, $\phi^\star_{psu}$. To disentangle both sources of uncertainty, we calculate the expected welfare costs over all draws and models under both policy rules and the difference in their performance (rows four and five). The gains from taking parameter uncertainty into consideration are found to be minor.

Table 7: Disentangling specification and parameter uncertainty

Coefficients	$\phi^{\star}_{psu}$	$\phi^{\star}_{su}$
ρ_R	1.39	1.43
ϕ_π	3.36	3.07
ϕ_y	0.00	0.00
$E_{\mathcal{M},\Theta}\mathcal{BC}(\Theta_i, \phi^{\star}_{\bullet})$	0.0114%	0.0117%
$E_{\mathcal{M},\Theta}\mathcal{BC}(\Theta_i, \phi^{\star}_{su}) - E_{\mathcal{M},\Theta}\mathcal{BC}(\Theta_i, \phi^{\star}_{psu})$	0.0003%	

Remarkably, the optimal rule over all draws and models reacts more aggressively to changes in inflation. This is due to the characteristics of the distribution of the standard deviations of the interest rate and inflation. Those distributions affect the policy trade-off between stabilizing inflation and the nominal interest rate in Model 5. While the median of the distribution of the standard deviation of inflation is literally the same as the standard deviation of inflation at the posterior mean, this is not the case for the nominal interest rate. There, the median of the distribution is smaller than the standard deviation at the posterior mean. This eases the trade-off faced by the central bank, thereby allowing a more aggressive response to inflation.

5 Conclusion

In this paper we have analyzed how to optimally conduct policy from a Bayesian perspective when the policymaker faces uncertainty about the parameters and the structure of the economy. In particular, we investigated whether the recent tendency of central banks to employ medium-scale DSGE models including a wide variety of features and frictions in their policy analysis can provide insurance against the risk of model uncertainty. Using a bottom-up approach, we subsequently estimated a set of nested monetary representative consumer models for the EU 13 area. Consequently, parameter uncertainty was characterized by the posterior distribution of the structural parameters, and specification uncertainty by each model's posterior probability. Eventually, we computed optimal policy to maximize representative households' expected utility across the model and parameter space.

We have identified uncertainty about the true structure of the economy as the main source of model uncertainty. For the central question of the paper, this implies that conducting policy according to one model only, even if this captures many features and frictions of the economy, leads to substantial welfare losses across the model space. A policymaker with a concern for robustness takes this into account and optimizes policy across a range of models weighted with their posterior probability.

References

Adam, K., and R. M. Billi (2006): "Optimal Monetary Policy under Commitment with a Zero Bound on Nominal Interest Rates," *Journal of Money, Credit and Banking*, 38(7), 1877–1905.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models by S. An and F. Schorfheide," *Econometric Reviews*, 26(2), 187–192.

Andrés, J., D. López-Salido, and J. Vallés (2006): "Money in an estimated business cycle model of the euro area," *Economic Journal*, 116, 457–477.

Boivin, J., and M. Giannoni (2006): "Has monetary policy become more effective?," *Review of Economics and Statistics*, 3(88), 445–462.

Calvo, G. A. (1983): "Staggered price setting in a utility-maximizing framework.," *Journal of Monetary Economics*, 12, 383–398.

Cogley, T., R. Colacito, L. P. Hansen, and T. J. Sargent (2008): "Robustness and U.S. Monetary Policy Experimentation," mimeo, New York University.

Cogley, T., R. Colacito, and T. J. Sargent (2007): "Benefits from U.S. Monetary Policy Experimentation in the Days of Samuelson and Solow and Lucas," *Journal of Money, Credit and Banking*, 39(s1), 67–99.

Del Negro, M., and F. Schorfheide (2008): "Forming Priors for DSGE Models (and How it Affects the Assessment of Nominal Rigidities)," NBER Working Papers 13741, National Bureau of Economic Research.

Eggertsson, G. B., and M. Woodford (2003): "The Zero Bound on Interest Rates and Optimal Monetary Policy," *Brookings Papers on Economic Activity*, 34(2003-1), 139–235.

Fagan, G., J. Henry, and R. Mestre (2001): "An area-wide model (AWM) for the euro area," Discussion Paper 42, ECB.

Hodrick, R. J., and E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1), 1–16.

Levin, A., A. Onatski, J. Williams, and N. Williams (2005): "Monetary Policy under Uncertainty in Microfounded Macroeconometric Models," *NBER Macroeconomics Annual*.

Levine, P., P. McAdam, J. Pearlman, and R. Pierse (2008): "Risk Management in Action. Robust monetary policy rules under structured uncertainty," Working Paper Series 870, European Central Bank.

Lubik, T. A., and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94(1), 190–217.

Lucas, R. E. (2003): "Macroeconomic Priorities," *American Economic Review*, 93(1), 1–14.

Paustian, M., and C. A. Stoltenberg (2008): "Optimal interest rate stabilization in a basic sticky-price model," *Journal of Economic Dynamics and Control*, 32, 3166–3191.

Rotemberg, J., and M. Woodford (1999): "Interest-Rate-Rules in an Estimated Sticky-Price-Model," in *Monetary Policy Rules*, ed. by J. Taylor. University of Chicago Press, Chicago.

Schmitt-Grohé, S., and M. Uribe (2007): "Optimal simple and implementable monetary and fiscal rules," *Journal of Monetary Economics*, 54(6), 1702–1725.

Sims, C. A. (2001): "Pitfalls of a Minimax Approach to Model Uncertainty," *American Economic Review*, 91(2), 51–54.

Smets, F., and R. Wouters (2003): "An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area," *Journal of the European Economic Association*, 1(5), 1123–1175.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.

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

——— (2003b): "Optimal interest rate smoothing," *Review of Economic Studies*, 70, 861–886.

A Appendix

A.1 Proof of proposition 1

The period utility function of the average household in equilibrium is given by:

$$\int_0^1 [u(\bullet) - v(l_{jt}) + z(m_t)]dj = u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) - \int_0^1 v(l_{jt})dj + z(m_t).$$

To derive (36) we need to impose that, in the optimal steady state, real money balances are sufficiently close to being satiated (see Woodford, 2003a, Assumption 6.1) such that we can treat $(R-1)/R$ as an expansion parameter.

The first summand can be approximated to second order by:

$$\begin{aligned} u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) = u_c y(1-\beta\eta)[\widehat{y}_t + \frac{1}{2}(1 - \varphi(1+\eta^2\beta))\widehat{y}_t^2 + \varphi\eta\widehat{y}_t\widehat{y}_{t-1} \\ + \varphi\widehat{y}_t(-\eta g_{t-1} - \beta\eta g_{t+1} + (1+\eta^2\beta)g_t)] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3), \quad (41) \end{aligned}$$

where we used $(x_t - x) = x(\widehat{x}_t + 0.5\widehat{x}_t^2) + \mathcal{O}(\|\widehat{x}_t\|^3)$, $\varphi = \frac{\sigma}{1-\beta\eta}$, t.i.s.p denotes terms independent of stabilization policy, $\sigma = -yu_{11}/u_1$, and $g_t = (G_t - G)/y$.

Since $y_t = a_t l_t/\Delta_t$, the second term can be approximated by

$$v(l_t) = u_c(1-\beta\eta)[\widehat{y}_t + \frac{1+\omega}{2}\widehat{y}_t^2 - (1+\omega)\widehat{a}_t\widehat{y}_t + \widehat{\Delta}_t] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3), \quad (42)$$

where we posited that in the equilibrium under flexible wages each household supplies the same amount of labor, $l = y$, $\omega = \frac{v_{ll}}{v_l}l$, and that due to the existence of an output subsidy the steady state is rendered efficient with $v_l = u_c(1-\beta\eta)$. In the next step we combine (41) and (42), employ $\tilde{g}_t = -\eta g_{t-1} - \beta\eta E_t g_{t+1} + (1+\eta^2\beta)g_t$, and obtain:

$$u(y_t - g_t - \eta(y_{t-1} - g_{t-1})) - \int_0^1 v(l_{jt})dj = u_c y(1-\beta\eta)[\frac{1}{2}(-\varphi(1+\eta^2\beta) - \omega)\widehat{y}_t^2$$
$$+ \varphi\eta\widehat{y}_t\widehat{y}_{t-1} + \varphi\widehat{y}_t\tilde{g}_t + (1+\omega)\widehat{a}_t\widehat{y}_t - \widehat{\Delta}_t] + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3). \quad (43)$$

The efficient rate of output is defined by the following difference equation:

$$[\omega + \varphi(1+\beta\eta^2)]\widehat{y}_t^e - \varphi\eta\widehat{y}_{t-1}^e - \varphi\eta\beta E_t\widehat{y}_{t+1}^e = \varphi\tilde{g}_t + (1+\omega)\widehat{a}_t + \mathcal{O}(\|\widehat{\xi}_t\|^2).$$

If we use this expression to rewrite (43), we obtain the following:

$$E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\{u(\bullet) - \int_0^1 v(l_{jt})dj\} = -E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}u_c y(1-\beta\eta)\{\frac{1}{2}(\varphi(1+\eta^2\beta)+\omega)\widehat{y}_t^2$$
$$-\varphi\eta\widehat{y}_t\widehat{y}_{t-1} - [\omega + \varphi(1+\beta\eta^2)]\widehat{y}_t^e\widehat{y}_t + \varphi\eta\widehat{y}_{t-1}^e\widehat{y}_t + \varphi\eta\beta E_t\widehat{y}_{t+1}^e\widehat{y}_t + \widehat{\Delta}_t\} + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3). \quad (44)$$

We seek to rewrite this expression in purely quadratic terms of the welfare-relevant gaps for inflation and output. To do so we apply the method of undetermined coefficients to reformulate the first part (all but $\widehat{\Delta}_t$), i.e. we seek to find the coefficient δ_0, such that (44) and

$$
\begin{aligned}
& -\frac{1}{2}\delta_0(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e))^2 \\
= & -\frac{1}{2}\delta_0[\widehat{y}_t^2 - 2\widehat{y}_t\widehat{y}_t^e + (\widehat{y}_t^e)^2 - 2\delta^*(\widehat{y}_t - \widehat{y}_t^e)(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e) + (\delta^*)^2(\widehat{y}_{t-1}^2 - 2\widehat{y}_{t-1}\widehat{y}_{t-1}^e + (\widehat{y}_{t-1}^e)^2)] \\
= & -\frac{1}{2}\delta_0[\widehat{y}_t^2 - 2\widehat{y}_t\widehat{y}_t^e - 2\delta^*\widehat{y}_t\widehat{y}_{t-1} + 2\delta^*\widehat{y}_t\widehat{y}_{t-1}^e + 2\delta^*\widehat{y}_t^e\widehat{y}_{t-1} + (\delta^*)^2\widehat{y}_{t-1}^2 - 2(\delta^*)^2\widehat{y}_{t-1}\widehat{y}_{t-1}^e] \\
= & -\frac{1}{2}\delta_0[((\delta^*)^2\beta + 1)\widehat{y}_t^2 - 2\delta^*\widehat{y}_t\widehat{y}_{t-1} + 2\delta^*\widehat{y}_t\widehat{y}_{t-1}^e + 2\delta^*\beta\widehat{y}_{t+1}^e\widehat{y}_t - (2(\delta^*)^2\beta + 2)\widehat{y}_t\widehat{y}_t^e] \\
= & -\frac{1}{2}\delta_0((\delta^*)^2\beta + 1)\widehat{y}_t^2 + \delta_0\delta^*\widehat{y}_t\widehat{y}_{t-1} - \delta_0\delta^*\widehat{y}_t\widehat{y}_{t-1}^e - \delta_0\delta^*\beta\widehat{y}_{t+1}^e\widehat{y}_t + \delta_0((\delta^*)^2\beta + 1)\widehat{y}_t\widehat{y}_t^e
\end{aligned}
$$

are consistent. We use that $\widehat{y}_{t0-1}$ is $t.i.s.p.$. The parameter δ^*, $0 \leq \delta^* \leq \eta$, is the smaller root of this quadratic equation: $\eta\varphi(1+\beta\delta^2) = [\omega+\varphi(1+\beta\eta^2)]\delta$. This root is assigned to past values of the natural and efficient rate of output in their stationary solutions. Comparing coefficients, δ_0 is

$$\delta_0 = \frac{u_c y(1-\beta\eta)\eta\varphi}{\delta^*}.$$

If firms are allowed to index with past inflation, such that

$$E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}2\widehat{\Delta}_t = E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\frac{\zeta\alpha}{(1-\alpha)(1-\alpha\beta)}(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1})^2 + t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3),$$

the quadratic approximation in (44) can be written as:

$$
\begin{aligned}
- E_{t_0}\sum_{t=t_0}^{\infty}\beta^{t-t_0}\frac{u_c y(1-\beta\eta)}{2}[\frac{\eta\varphi}{\delta^*}(\widehat{y}_t - \widehat{y}_t^e - \delta^*(\widehat{y}_{t-1} - \widehat{y}_{t-1}^e))^2 + \frac{\zeta\alpha}{(1-\alpha)(1-\alpha\beta)}(\widehat{\pi}_t - \gamma\widehat{\pi}_{t-1})^2] \\
+ t.i.s.p. + \mathcal{O}(\|\widehat{\xi}_t\|^3).
\end{aligned}
$$

The last approximation needed is that involving the utility of real money balances. Applying similar techniques we get

$$z(m_t) = z + yu_c(s_m(\widehat{m}_t + \frac{1}{2}s_m(1-\sigma_m)\widehat{m}_t^2) + t.i.s.p + \mathcal{O}(\|\widehat{\xi}_t\|^3), \tag{45}$$

where we employ $s_m = z_m m/(u_c y) = (R-1)(1-\beta\eta)R$ and $\sigma_m = -z_{mm}m/z_m$. Since we treat $(R-1)/R$ as an expansion parameter, s_m and $1/\sigma_m$ are of first order. However, $s_m\sigma_m$ approaches a finite limit for $(R-1)/R \rightarrow 0$, which is given by

$$s_m\sigma_m = \frac{z_{mm}m^2}{yu_c}.$$

The interest elasticity of money demand is given by the following expression:

$$\eta_i = -\frac{u_c(1-\beta_n)}{z_{mm}}\frac{1-\frac{R-1}{R}}{m} = \frac{1}{\sigma_m(R-1)}.$$

At the limit for $(R-1)/R \rightarrow 0$, it follows that $\eta_i = -u_c(1-\beta\eta)/(z_{mm}m)$ and therefore $s_m\sigma_m = (1-\beta\eta)/(v\eta_i)$, with $v = y/m$. A first-order approximation of the money demand equation (33) yields

$$\widehat{m}_t = -\eta_i\widehat{R}_t - \frac{1}{\sigma_m}\widehat{\lambda}_t + \mathcal{O}(\|\widehat{\xi}_t|^2),$$

where

$$\widehat{\lambda}_t = -\varphi(\widehat{y}_t - \eta\widehat{y}_{t-1}) + \beta\eta\varphi(\widehat{y}_{t+1} - \eta\widehat{y}_t) + \varphi(g_t - \eta g_{t-1}) - \beta\eta\varphi(g_{t+1} - \eta g_t) + \mathcal{O}(\|\widehat{\xi}_t|^2).$$

Using all the above we can rewrite $z(m_t)$ in the following way:

$$z(m_t) = -\frac{\eta_i y u_c}{2v}(1-\beta\eta)(\widehat{R}_t^2 + 2\frac{R-1}{R}\widehat{R}_t) + t.i.s.p + \mathcal{O}(\|\widehat{\xi}_t, (R-1)/R\|^3). \qquad (46)$$

We assume for simplicity that $[(R-1)/R - 0]$ is of second order, and sum the results in expression (36) in the text.

A.2 Estimation Results

Table 8: Prior distribution of the structural parameters

	Prior distribution		
Parameter	distribution	mean	std
ρ	beta	0.8	0.1
ϕ_π	normal	1.7	0.1
ϕ_y	normal	0.125	0.05
ω	gamma	1	0.5
σ_c	normal	1.5	0.375
α	beta	0.75	0.05
η	beta	0.7	0.1
γ	beta	0.75	0.15
σ_m	normal	1.25	0.375
ψ_g	beta	0.7	0.1
ψ_a	beta	0.7	0.1
σ_g	invgamma	0.04	0.026
σ_a	invgamma	0.04	0.026
σ_μ	invgamma	0.04	0.026

Table 9: Posterior estimates of the structural parameters in each model

	$Model_1$		$Model_2$		$Model_3$		$Model_4$		$Model_5$	
Parameter	mean	std	mean	std	mean	std	mean	std	mean	std
ρ	0.4379	0.0719	0.4582	0.0727	0.3578	0.0742	0.4477	0.0730	0.3756	0.0756
ϕ_π	1.6972	0.0983	1.7188	0.0985	1.6255	0.1009	1.6719	0.0963	1.6772	0.0972
ϕ_y	0.0964	0.0263	0.0701	0.0261	0.1154	0.0324	0.0939	0.0261	0.0821	0.0297
ω	0.4170	0.1104	0.2610	0.0547	0.4156	0.0984	0.3962	0.0903	0.2904	0.0644
σ_c	1.4202	0.1744	1.4296	0.1798	1.4142	0.1879	1.4117	0.1735	1.4158	0.1779
α	0.8810	0.0139	0.8988	0.0126	0.8672	0.0177	0.8855	0.0156	0.8777	0.0168
η	-	-	0.8368	0.0566	-	-	-	-	0.8307	0.0515
γ	-	-	-	-	0.4677	0.0842	-	-	0.4563	0.0796
σ_m	-	-	-	-	-	-	1.2685	0.3866	1.2230	0.3777
ψ_g	0.8126	0.0438	0.7880	0.0579	0.7992	0.0482	0.8099	0.0477	0.7950	0.0571
ψ_a	0.7733	0.0367	0.7929	0.0372	0.6766	0.0545	0.7836	0.0403	0.6780	0.0562
σ_g	0.0095	0.0007	0.0084	0.0005	0.0097	0.0007	0.0095	0.0007	0.0084	0.0005
σ_a	0.0168	0.0030	0.0206	0.0034	0.0178	0.0032	0.0174	0.0026	0.0201	0.0039
σ_μ	0.0092	0.0006	0.0090	0.0006	0.0097	0.0007	0.0092	0.0006	0.0094	0.0006

Figure 1: Deep parameters prior vs. **posterior** distribution in Model 1

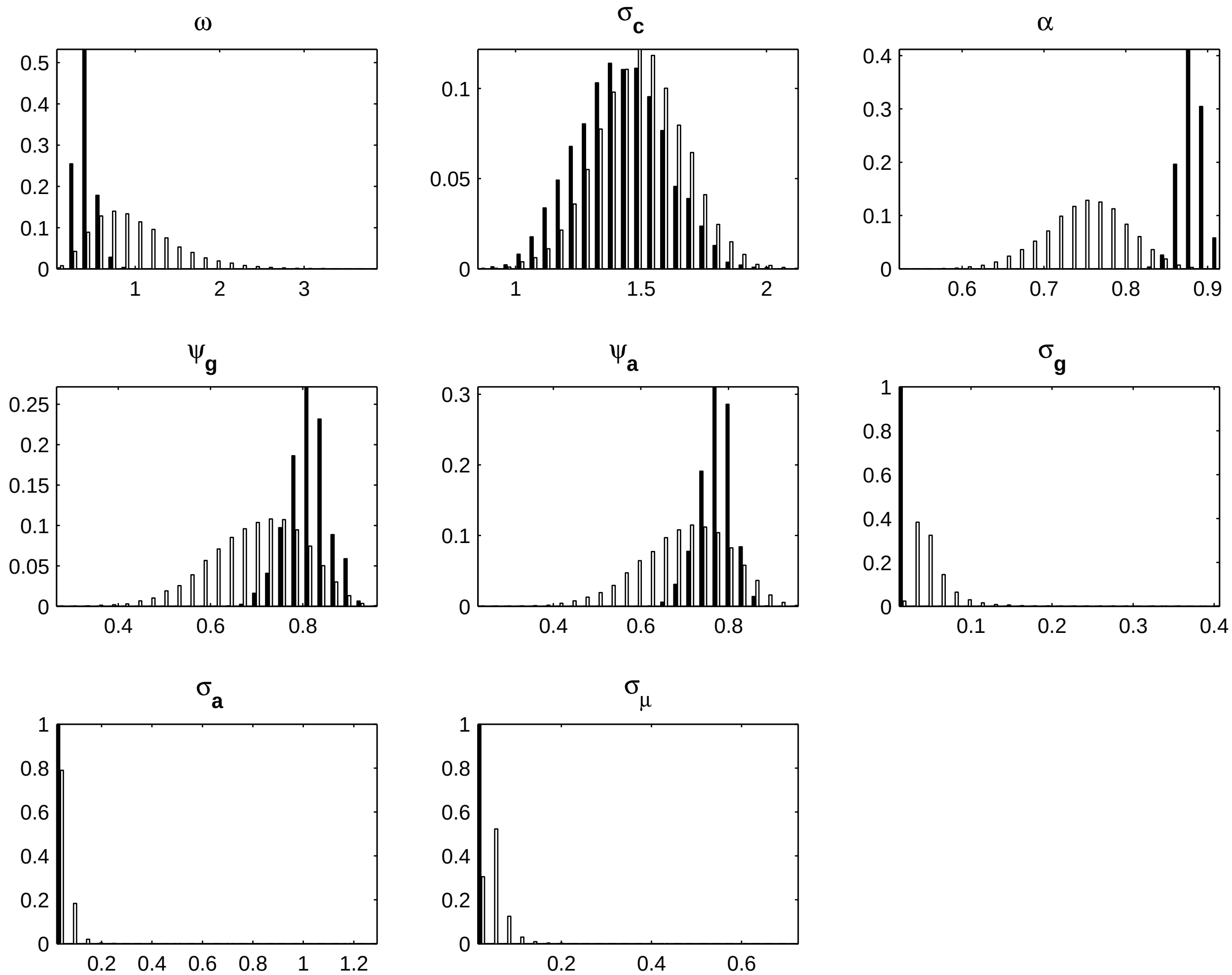


Figure 2: Deep parameters prior vs. **posterior** distribution in Model 2

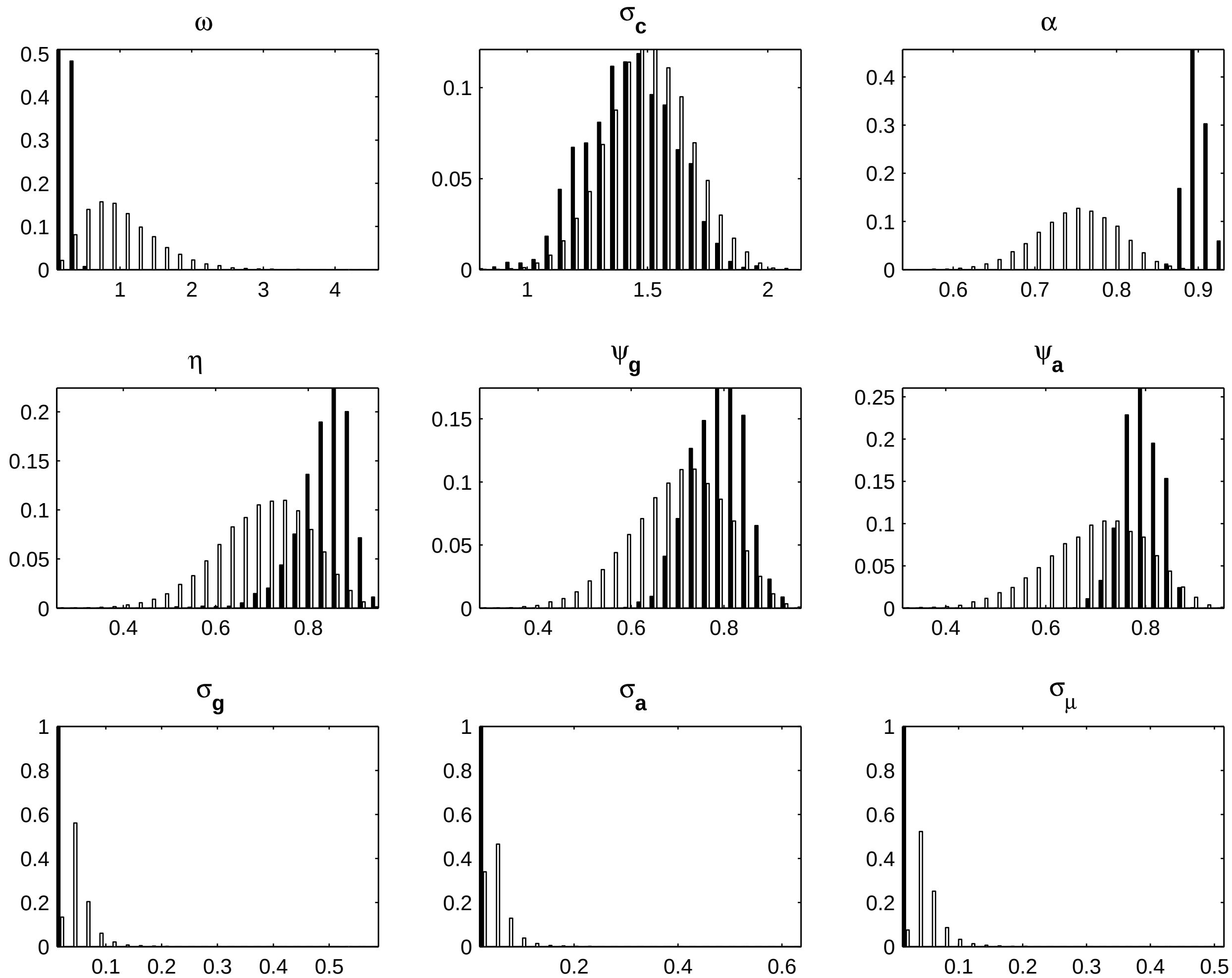


Figure 3: Deep parameters **prior** vs. **posterior** distribution in Model 3

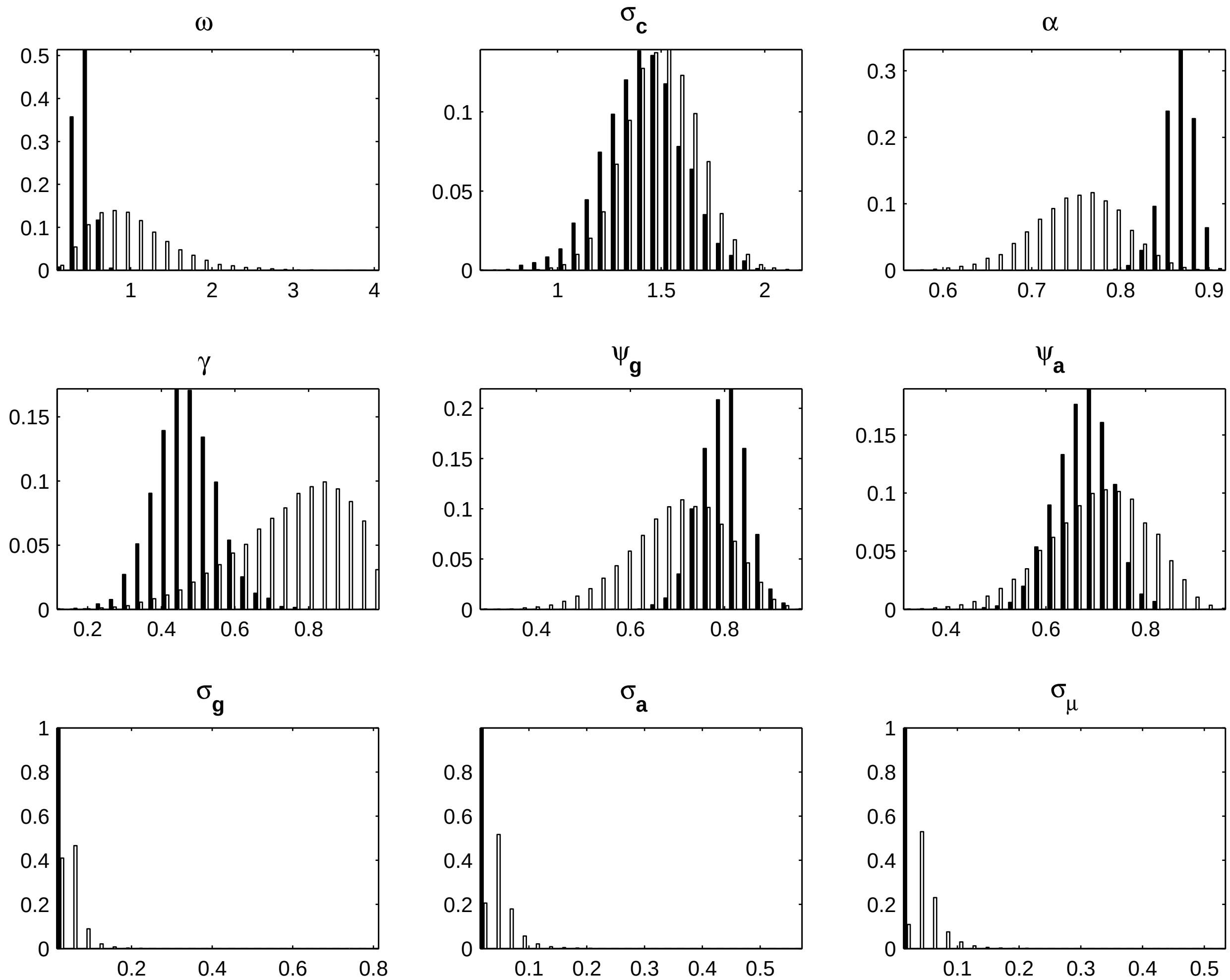


Figure 4: Deep parameters **prior** vs. **posterior** distribution in Model 4

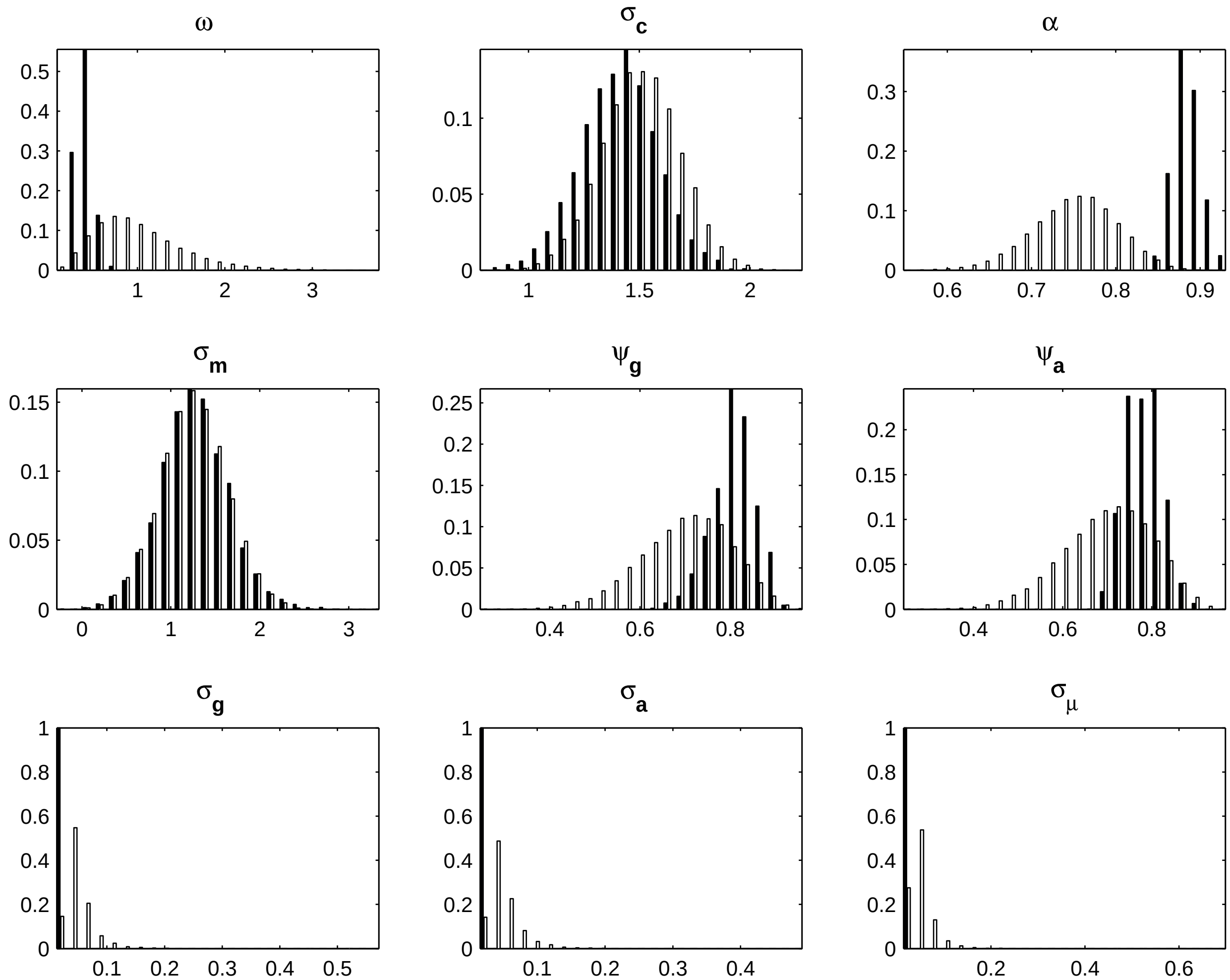
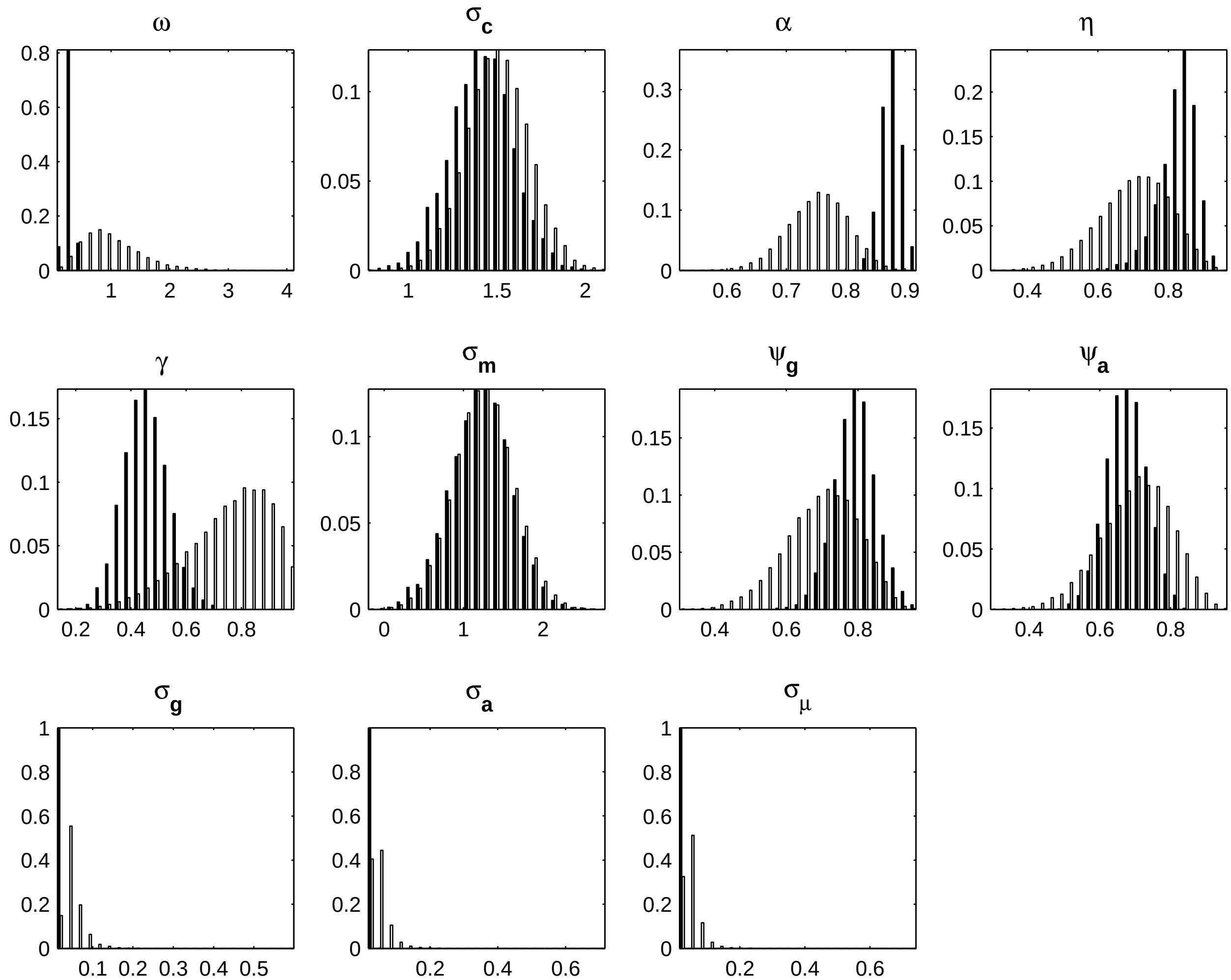


Figure 5: Deep parameters **prior** vs. **posterior** distribution in Model 5



A.3 Optimal monetary policy and the zero bound on interest

In this section we investigate how the zero bound constraint affects the conduct of optimal monetary policy. We approximate this non-linear condition by restraining the fluctuations in interest and require twice the unconditional standard deviation of the nominal interest rate to not exceed the difference between steady-state interest and the zero bound (see definition 2 with $k = 2$).

In general, the zero bound on interest shapes the optimal policy reaction in the following way: optimal simple interest rate rules give less weight to inflation and more to the degree of interest rate inertia to smooth interest rates over time (compare Tables 3 and 10).

Table 10: Optimal policy, parameter uncertainty and the zero bound on interest

$$\widehat{R}_t = \rho_R \widehat{R}_{t-1} + \phi_\pi \widehat{\pi}_t + \phi_y \widehat{y}_t$$

	M_1	M_2	M_3	M_4	M_5
ρ_R	1.13	1.45	1.42	1.26	1.36
ϕ_π	19.99	14.09	12.65	2.42	1.01
ϕ_y	0.00	0.04	0.00	0.00	0.00
$\mathcal{BC}(\bar{\theta}_i, \phi^\star_{izb}) - \mathcal{BC}(\bar{\theta}_i, \phi^\star_i)$	0.0002%	0.0009%	0.0013%	0.0000	0.0000
$ZBV(\phi^\star_i)$	42%	45%	41%	0%	0%

ZBV indicates that the optimal policy computed at the posterior mean violates the zero bound requirement in x percent of all draws.

Requiring policies to be implementable leads to additional welfare losses in models 1-3 relative to monetary policy unconstrained by the zero bound (fourth row of Table 10), where households' utility does not depend on variations in interest rates. In such an environment the implementability constraint implicitly introduces stabilization of the nominal interest rate as an additional policy aim by restricting the set of feasible policies. In particular, policies that aggressively fight inflation drop out, since these policies result in a high variability of interest.

In contrast to previous results, parameter uncertainty is an obstacle for monetary policy when the lower bound on interest is taken into account. Following the optimal policy at the posterior mean is not a good policy recommendation: the central bank risks violating the implementability requirement with high probability over the parameter space. In models where optimal policy is characterized by a binding implentability constraint at the posterior mean (models 1,2 and 3), variations over the parameter vector easily lead to a violation of this constraint (in 41 to 45 percent of all draws, see last row of Table 10).

Not surprisingly, the zero bound also matters when the policymaker is uncertain about the true model. In general, a policy that evades violations of this constraint in one model is not a good device if the true model is another. Exceptions are models in which preferences require stabilization of the nominal interest rate as an additional aim (see Table 11).

Table 11: Optimal policy under specification uncertainty ($\phi^{\star}_{su}$): zero bound

	M_1	M_2	M_3	M_4	M_5
$\mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{1zb}) - \mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{izb})$	0.00%	ZBV	$ZBV*$	0.02%	ZBV
$\mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{2zb}) - \mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{izb})$	0.00%	0.00%	ZBV	0.01	ZBV
$\mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{3zb}) - \mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{izb})$	0.00%	0.00%	0.00%	0.01%	ZBV
$\mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{4zb}) - \mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{izb})$	0.01%	0.01%	0.00%	0.00%	0.04%
$\mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{5zb}) - \mathcal{BC}(\bar{\theta}_i, \phi^{\star}_{izb})$	0.02%	0.02%	0.01%	0.00%	0.00%

ZBV in (i, j) indicates that policy j violates the zero bound in model i.