A Bayesian Approach to Optimal Monetary Policy with Parameter and Model Uncertainty[*]

Timothy Cogley,[†]Bianca De Paoli,[‡]Christian Matthes,[§]

Kalin Nikolov,[¶]and Tony Yates[‖]

January 2010

Abstract

This paper undertakes a Bayesian analysis of optimal monetary policy for the U.K. We estimate a suite of monetary-policy models that include both forward- and backward-looking representations as well as large- and small-scale models. We find an optimal simple Taylor-type rule that accounts for both model and parameter uncertainty. For the most part, backward-looking models are highly fault tolerant with respect to policies optimized for forward-looking representations, while forward-looking models have low fault tolerance with respect to policies optimized for backward-looking representations. In addition, backward-looking models often have lower posterior probabilities than forward-looking models. Bayesian policies therefore have characteristics suitable for inflation and output stabilization in forward-looking models.

1   Introduction

Central bankers frequently emphasize the importance of uncertainty in shaping monetary policy (e.g., see Greenspan 2004 and M. King 2004). Uncertainty takes many forms. The central bank must act in anticipation of future conditions, which are affected by shocks that are currently unknown. In addition, because economists have not formed a consensus about the best way to model the monetary transmission

---

[*]The opinions expressed here are our own and do not necessarily reflect those of the Bank of England or members of the Monetary Policy Committee.

[†]New York University. Email: tim.cogley@nyu.edu

[‡]Bank of England. Email: bianca.depaoli@bankofengland.co.uk.

[§]New York University. Email: cm1518@nyu.edu.

[¶]Bank of England. Email: kalin.nikolov@bankofengland.co.uk.

[‖]Bank of England. Email: tony.yates@bankofengland.co.uk.

mechanism, policy makers must also contemplate alternative theories with distinctive operating characteristics. Finally, even economists who agree on a modeling strategy sometimes disagree about the values of key parameters. Central bankers must therefore also confront parameter uncertainty within macroeconomic models.

A natural way to address these issues is to regard monetary policy as a Bayesian decision problem. As noted by Brock, Durlauf, and West (2003), a Bayesian approach is promising because it seamlessly integrates econometrics and decision theory. Thus, we can use Bayesian econometric methods to assess various sources of uncertainty and incorporate the results as an input to a decision problem.

Our aim in this paper is to consider how monetary policy should be conducted in the face of multiple sources of uncertainty, including model and parameter uncertainty as well as uncertainty about future shocks. We apply Bayesian methods root and branch to a suite of macroeconomic models estimated on U.K. data, and we use the results to devise a simple, optimal monetary-policy rule.

1.1   The method in more detail

Just to be clear, we take two shortcuts relative to a complete Bayesian implementation. First, we neglect experimentation. Under model and/or parameter uncertainty, a Bayesian policy maker has an incentive to vary the policy instrument in order to generate information about unknown parameters and model probabilities. In the context of monetary policy, however, a number of recent studies suggest that experimental motives are weak and that 'adaptive optimal policies' (in the language of Svensson and Williams 2008a) well approximate fully-optimal, experimental policies.[1] Because of that, and also because many central bankers are averse to experimentation, our goal is to formulate an optimal non-experimental rule.

We also restrict attention to a simple rule, i.e. one involving a relatively small number of arguments as opposed to the complete state vector. This is for tractability as well as for transparency. For a Bayesian decision problem with multiple models, the fully-optimal decision rule would involve the complete state vector for all the models under consideration. That would complicate our calculations a great deal. Some economists also argue that simple rules constitute more useful communication tools. For example, Woodford (1999) writes that "a simple feedback rule would make it easy to describe the central bank's likely future conduct with considerable precision, and verification by the private sector of whether such a rule is actually being followed should be straightforward as well." Thus, we restrict policy to follow Taylor-like rules.

With those simplifications in mind, our goal is to choose the parameters of a Taylor rule to minimize expected posterior loss. Suppose ϕ represents the policy-rule parameters and that $l_i(\phi, \theta_i)$ represents expected loss conditional on a particular

---

[1]E.g., see Cogley, Colacito, and Sargent (2008); Cogley, Colacito, Hansen, and Sargent (2008); and Svensson and Williams (2007a, 2007b, 2008a, and 2008b).

model i and a calibration of its parameters θ_i. Typically $l_i(\phi, \theta_i)$ is a discounted quadratic loss function that evaluates uncertainty about future shocks. One common approach in the literature is to choose ϕ to minimize $l_i(\phi, \theta_i)$. This delivers a simple optimal rule for a particular model and calibration, but it neglects parameter and model uncertainty.

To incorporate parameter uncertainty within model i, we must first assess how much uncertainty there is. This can be done by simulating the model's posterior distribution, $p(\theta_i|Y, M_i)$, where M_i indexes model i, and Y represents current and past data on variables relevant for that model. Methods for Bayesian estimation of DSGE models were pioneered by Schorfheide (2000) and Smets and Wouters (2003) and are reviewed by An and Schorfheide (2007). If model i were the only model under consideration, expected loss would be

$$l_i(\phi) = \int l_i(\phi, \theta_i)p(\theta_i|Y, M_i)d\theta_i. \tag{1}$$

This integral might seem daunting, but it can be approximated by averaging across draws from the posterior simulation. Assuming evenly weighted draws from the posterior, expected loss is

$$l_i(\phi) \approx N^{-1} \sum_{j=1}^{N} l_i(\phi, \theta_{ij}), \tag{2}$$

where N represents the number of Monte Carlo draws and θ_{ij} is the jth draw for model i. A policy rule robust to parameter uncertainty within model i can be found by choosing ϕ to minimize $l_i(\phi)$.

This is a step in the right direction, but it still neglects model uncertainty. To incorporate multiple models, we attach probabilities to each and weigh their implications in accordance with those probabilities. Posterior model probabilities depend on prior beliefs and on their fit to the data. Suppose that $p(M_i)$ is the policy-makers prior probability on model i, that $p(\theta_i|M_i)$ summarizes his prior beliefs about the parameters of that model, and that $p(Y|\theta_i, M_i)$ is the model's likelihood function.[2] According to Bayes' theorem, the posterior model probability is

$$p(M_i|Y) \propto p(Y|M_i)p(M_i), \tag{3}$$

where

$$p(Y|M_i) = \int p(Y|\theta_i, M_i)p(\theta_i|M_i)d\theta_i \tag{4}$$

is the marginal likelihood or marginal data density. The latter can also be approximated numerically using output of the posterior simulation; see An and Schorfheide

---

[2]For simplicity, we assume that Y is common across models, but that is unnecessary. A technical appendix posted online at http://homepages.nyu.edu/~tc60 describes the more realistic case in which the list of variables differs across model.

for details. To account for model uncertainty, we average $l_i(\phi)$ across models using posterior model probabilities as weights,

$$l(\phi) = \sum\nolimits_{i=1}^{m} l_i(\phi)p(M_i|Y). \tag{5}$$

A policy rule robust to both model and parameter uncertainty can be found by choosing ϕ to minimize $l(\phi)$.

This decision problem might seem complicated, but because the problem is modular it can be solved numerically without much trouble. The main simplification follows from the fact that the econometrics can be done separately for each model and also separately from the decision problem.

1.2   Sketch of previous literature

Our work follows and builds on many previous contributions. As mentioned above, one is the body of work estimating dynamic general equilibrium models using Bayesian methods. This literature has exploded in recent years and includes numerous applications to monetary policy.

A second closely-related literature concerns forecast model averaging. This research was initiated by Bates and Granger (1969) and is now widely regarded as representing best practice in forecasting. Amongst others, recent contributions to the frequentist literature include Clements and Hendry (1998, 2002) and Newbold and Harvey (2002), while examples of Bayesian forecast averaging include Diebold and Pauly (1990), Jacobson and Karlsson (2004), and Kapetanios, Labhard, and Price (2008). Our work is distinct from this in that we are interested not only in forecasting but also in solving a decision problem. Of course, forecasting is an input to our decision problem, but it is not an end in itself. For that reason, we concern ourselves with structural macroeconomic models.

Another important precursor is Brock, Durlauf, and West (2003 and 2007). They also emphasize the importance of accounting for model and parameter uncertainty in policy design, and they describe a variety of Bayesian and frequentist approaches for integrating econometrics and policy design. Our framework follows directly from one of their proposals.[3] They also investigate the robustness of Taylor rules within a class of backward-looking models a la Rudebusch and Svensson (1999). Cogley and Sargent (2005a) apply the ideas of Brock, et al. to investigate how model uncertainty affected monetary policy during the Great Inflation. For tractability, Cogley and Sargent adopt two shortcuts, restricting the model set to a trio of very simple Phillips-curve models and neglecting parameter uncertainty within each model. In our application, we expand the model set to include forward-looking new Keynesian models, and we

---

[3]Among other things, they also discuss frequentist and Waldean approaches to econometrics as well as models in which the decision maker is averse to ambiguity.

explicitly account for parameter uncertainty. Cogley and Aargent's (2005a) work is a positive exercise: our paper follows Brock *et al* (2003,2007) and concentrates on normative questions.

Other routes to robustness include those of McCallum (1988) and Hansen and Sargent (2008). McCallum pioneered an informal version of model averaging, deprecating policy rules optimized with respect to a single model and advocating rules that work well across a spectrum of models. Much of Taylor's (1999) volume on monetary-policy rules can be read as an application of McCallum's ideas. Recent applications include Levin and Williams (2003), Levin, Wieland, and Williams (2003), and Levin, Onatski, Williams, and Williams (2005).[4] We embrace McCallum's approach and extend it by providing Bayesian underpinnings. We want to forge a tighter link between this literature and the literature on Bayesian estimation of DSGE models. Our hope is that a more formal assessment of uncertainties will pay off in policy design.

Hansen and Sargent (2008) develop yet another approach to model uncertainty. They specify a single, explicit benchmark model, surround it with an uncountable cloud of alternative models whose entropy relative to the benchmark model is bounded, and find an optimal rule by solving a minimax problem over that set of models. In contrast, we work with a small number of explicit models and assume that policy makers entertain no other possibilities. Our approach no doubt understates the true degree of model uncertainty by excluding a priori a large number of potential alternatives. Despite this shortcoming, we think the Bayesian approach is useful because it is more explicit about the relative probabilities of models within the suite.[5]

1.3 Outline

The paper is organized as follows. Section 2 describes our suite of models, emphasizing their distinctive characteristics and features of the posterior that are most salient for monetary policy. Section 3 reports posterior model weights, and section 4 presents our main results. There we describe an optimal Taylor rule and illustrate how it works in the various submodels.

2 The Suite of Models

We focus on models that command some attention in the monetary-policy literature. Within that class, our intention is to span a variety of approaches to modelling

---

[4]Levin and Williams (2003) cite Patrick Minford as drawing the analogy with a committee of decision makers each with their own model of the inflation process, who would opt for a policy rule provided it does not perform disastrously in any of the individual members' models.

[5]Sometimes the two approaches are distinguished by saying that one explores structured model uncertainty and the other unstructured uncertainty.

inflation dynamics. For instance, we compare microfounded and non-microfounded models, RE vs. non RE models, and small models that offer parsimony at the expense of a rich account of macrodynamics vs. larger models that fit better but involve many more parameters.

Our suite currently comprises four models, all of which are estimated using UK quarterly time series on nominal interest rates, inflation and real GDP growth. As the UK has undergone numerous monetary regime shifts during the post-war period, we only use data from the inflation-targeting period, Q1 1993 - Q3 2006. Our data definitions are as follows. For the nominal interest rate we use the Bank of England's policy rate (source: Bank of England). For inflation we use the quarterly change in the logarithm of the GDP deflator (source: Office of National Statistics). For output growth we use the quarterly change in real GDP at market prices (source: Office of National Statistics). All variables are demeaned prior to estimation.

In what follows, we briefly describe the salient features of each model. A complete presentation can be found in the appendices.

2.1 A traditional backward-looking Keynesian model

We begin with a traditional, backward-looking Keynesian model in the spirit of Rudebusch and Svensson (1999). We include this model for two reasons. As Rudebusch and Svensson emphasize, traditional Keynesian models represent the thinking of many central bankers. Thus, at least one member of a monetary policy committee is likely to place positive probability on a model in this class.

In addition, in studies of robust monetary policy for the U.S., one of the main challenges has been finding a rule that works well both for forward- and backward-looking models. When estimated with U.S. data, backward-looking models typically imply a high degree of intrinsic inflation persistence. In contrast, in forward-looking models, decision rules and the equilibrium law of motion adapt to the policy rule. Because inflation persistence is intrinsic in backward-looking models and endogenous in forward-looking models, rules that succeed in stabilizing inflation in the latter often result in excessive output variability in the former, while gradualist rules well adapted to a backward-looking environment frequently permit more inflation variability in forward-looking models than one might like. Finding a rule well adapted to both environments can be challenging. Furthermore, this can happen even when the probability weight on backward-looking models is small. For example, in a study of the Great Inflation, Cogley and Sargent (2005) report that traditional Keynesian models dominate Bayesian policy despite having probability weights close to zero.[6]

Our version of the Rudebusch-Svensson model is detailed in appendix A, and the priors and posteriors are listed in tables A1-A6. The model consists of three equations

---

[6]This can happen when the period loss function is unbounded.

- a backward-looking Phillips curve, a backward-looking IS curve, and a Taylor rule for monetary policy. Since the model is entirely backward-looking, inflation and output persistence are hard-wired into the structural equations, and expectations of future policy have no effect on current outcomes.

The model is estimated for three sets of priors. Our benchmark prior (table A1) is centered on Rudebusch and Svensson's original estimates[7] but has large prior variances, reflecting the considerable degree of uncertainty about the relevance of U.S. estimates for models of the U.K. At the prior mode, the Phillips and IS curves encode a high degree of intrinsic inflation and output persistence, but because the prior variances are large the data remain influential.

Indeed, the posterior differs from the prior in two respects that are important for monetary policy (see table A4). First, inflation and output turn out to be considerably less persistent than under the prior. In this respect, our estimates confirm studies such as Benati (2008) and Levin and Piger (2006), who also report a marked decline in inflation persistence during the Great Moderation. Second, the slope of the IS curve with respect to the real interest rate is smaller than under the prior, and the lower end of a 95 percent credible set is only slightly above zero. If this coefficient were equal to zero, the central bank would not be able to influence output or inflation via an interest-rate rule, and realizations of output and inflation would be independent of policy-rule parameters.

The finding that there is less instrinsic persistence in U.K. data for the inflation-targeting period than in Rudebusch and Svensson's sample is important for policy design. To assess the robustness of this finding, we re-estimate the model using two alternative priors: (1) a tighter prior based on the original Rudebusch-Svensson estimates (table A2) and (2) a tighter prior centred on simple AR(1) specifications involving a low degree of inflation and output persistence (table A3). The first alternative represents an attempt to force high intrinsic persistence onto the data. Somewhat to our surprise, we found that the estimates are broadly similar to those for the benchmark prior (see table A5). Although this prior is tighter than the benchmark, it is not so tight as to dominate the likelihood function, and since the tight RS prior is centered far from the maximum likelihood estimate the data remain influential. Thus, even when we try, we struggle to force high intrinsic persistence onto the data.

The second alternative prior explores robustness in a different direction. An alternative interpretation of the data is that high inflation persistence arises not from variation within a stable monetary regime, but rather from variation across policy regimes. For instance, Benati (2008), Cogley and Sbordone (2008), and Ireland (2007) argue that shifts in target inflation account for much of the persistence in inflation. Thus, as another robustness check, we re-estimate the model with an informative prior involving low degrees of intrinsic persistence. Once again, we find that model's

---

[7]This is permissable because they use data for the U.S., and we study data for the U.K.

characteristics are qualitatively robust to changes in the prior (see table A6).

The alternative priors will matter for posterior model probabilities. Especially important is the difference between our baseline and the low-persistence prior. The consequences of this modeling choice for optimal policy are discussed below.

2.2 A medium-scale dynamic new Keynesian model

The second member of our suite is a medium-scale new Keynesian model similar to that of Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). This model features a variety of real and nominal rigidities, including habit persistence, sticky wages, sticky prices, variable capital utilisation, and investment-adjustment costs. The Smets-Wouters model fits US and Euro Area data in a way that is competitive to a BVAR, and it has become a workhorse for monetary-policy analysis.

Relative to their specification, we streamline the model to make it more parsimonious. Nevertheless, for a DSGE model it is still heavily parameterized. For Bayesian model averaging, a dense parameterization is both an advantage and a disadvantage. Introducing a rich variety of shocks and frictions improves the model's fit, but the additional parameters are penalized when weighing models. One of our objectives is to explore this tradeoff.

Our version of the Smets-Wouters model is described in detail in appendix B, and the prior and posterior are summarized in tables A7 and A8, respectively. Two features of the posterior are important for policy design. First, in a number of dimensions, the posterior closely resembles the prior. Amongst others, parameters governing the degree of nominal rigidity are weakly identified. This identification problem arises because of the large size of the SW model combined with our limited number of data series and short sample. Thus, along several dimensions, parameter values are effectively set via the priors. The large number of parameters and the fact that some are weakly identified will count against the model when calculating posterior model probabilities.

One exception to this general result concerns the degree of price indexation, which we are able to estimate precisely, in the sense that the mode of the posterior turns out to be considerably lower than in the prior. At the posterior mode, the price-indexation parameter is 0.162. In contrast, Christiano, et al. (2005) calibrate their price-indexation parameter at unity. Thus, as for the RS model, our version of the SW model involves substantially less intrinsic inflation persistence than in versions for the US.

2.3   A small-scale new Keynesian model with credit frictions

Since the optimal tradeoff between fit and parsimony is an open question, we also study a small-scale dynamic new Keynesian model that fits fewer features of the data but which is more parsimonious. Among many candidates, we chose a model similar to that of Bernanke, Gertler, and Gilchrist (1999), which builds a financial accelerator into an otherwise standard new Keynesian model. The BGG model is much smaller in scale than the Smets-Wouters model, and we include it in the suite not only because we want to compare small- and medium-scale models but also because we are interested in the financial accelerator.

This model is presented in appendix C, and its prior and posterior are recorded in tables A9 and A10. In many ways, the estimates agree with those for the SW model. They differ, however, in one important respect, viz. that nominal prices are considerably more flexible in the BGG model than in the SW model. At the posterior mode, the estimates imply that prices are re-optimized once every 1.5 quarters. Because there is so little nominal rigidity, inflation is volatile but not persistent. This feature of the BGG model will matter later when designing a Bayesian policy.

2.4   A small-open-economy model

Last but not least, for the UK we feel that the suite should include an open-economy model in order to take into account international dimensions of monetary policy. Accordingly, we consider a small-open-economy model in the style of Gali and Monacelli (2005). The model assumes that home price setting follows a Calvo-type contract and features complete pass-through, as prices are set in the producer's currency. Moreover, even though the law of one price holds, deviations from purchasing power parity arise from the existence of home bias in consumption. Finally, markets are complete, and domestic and foreign agents optimally share risk.

The model is presented in appendix D, with the prior and posterior shown in tables A11 and A12, respectively. Concentrating on features that are important for monetary policy, the estimates imply a degree of price flexibility in between those found for the BGG and SW models. Unlike the other models, the model implies that the terms of trade enter both the IS and Phillips curves. Because no international data are used for estimation, however, it is difficult to obtain sharp estimates of the parameters governing these channels. This handicaps the model in comparison with the others.

3   Posterior model probabilities

For each model, we estimate the marginal data density using Geweke's (1999) modified harmonic-mean estimator. Then we combine marginal data densities with

9

prior model probabilities to compute posterior model probabilities. In every case, we assume equal prior odds on the four models, as shown in table 1. Posterior model probabilities are reported in table 2.

Table 1: Prior model probabilities

Model	RS1	RS2	RS3
Rudebusch-Svensson	1/4	1/4	1/4
Smets-Wouters	1/4	1/4	1/4
BGG	1/4	1/4	1/4
SOE	1/4	1/4	1/4

Table 2: Posterior model probabilities

Model	RS1	RS2	RS3
Rudebusch-Svensson	0.0204	0.0000	0.8010
Smets-Wouters	0.8008	0.8175	0.1627
BGG	0.1757	0.1793	0.0357
SOE	0.0031	0.0033	0.0006

Note: RS1 refers to the baseline prior, RS2 to the prior tightly centered on Rudebusch and Svensson's estimates, and RS3 to the prior involving weak persistence.

Three scenarios are considered, corresponding to the three priors on the Rudebusch-Svensson model. The first column records the outcome for our baseline RS prior. Recall that the benchmark prior has its mode at the original RS estimates but is fairly loose. For this scenario, the SW model is the most probable with a weight of over 80%, the BGG model comes second with a weight of 17.6%, the Rudebusch-Svensson model comes third with a weight of 2.0%, while the weight on the SOE model is less than 1%. Thus, at least in this comparison, fit seems to trump parsimony, as the densely parameterized SW model is assigned a probability 4 times that of the more parsimonoius BGG model. The BGG model weight is nontrivial, however.

Alas, the other two models are assigned low probability weights. Two factors explain the low weight on the SOE model. One is that the model assumes a purely-forward Phillips curve and abstracts from inflation indexation. Although the estimated degree of indexation in the SW and BGG models is not large, it is not zero, and a little bit of indexation seems to help fit the data. In addition, although the structural model is of an open economy, it resembles a closed economy with: (A) a different definition for potential output that includes foreign output; and (B) an unobserved endogenous variable (namely, the exchange rate) driving differences between the consumer and producer prices. Neither feature is well captured in our estimation

because no data on exchange rates or foreign variables are used. For model averaging, the models must be conditioned on the same variables, and since the others have nothing to say about exchange rates or foreign variables, we cannot condition on these variables.

Turning to the Rudenbusch-Svensson model, there are two reasons why it has a low posterior probability. One is that the posterior mean for the U.K. is very different from the prior, which is based on estimates for the U.S. As noted above, the U.S. estimates imply high instrinsic persistence, while those for the U.K. imply very little persistence. The marginal data density is the prior expectation of the likelihood function, and it tends to be low when the maximum likelihood estimate is far from the prior mode. A second reason is that the baseline RS prior is loose. We adopted a loose prior because the RS model is not microfounded; hence it was difficult for us to formulate an informative prior. A loose prior spreads probability mass throughout the parameter space and can put a lot of weight on regions in which the likelihood is small. Other things equal, that also reduces the prior expectation of the likelihood.

Thus, the baseline prior may put the RS model at a disadvantage in posterior model comparisons. Since posterior model probabilities can be sensitive to the choice of prior when the prior is weakly informative, we perform two sensitivity analyses on the location and tightness of the RS prior.

Our first sensitivity check involves tightenening the baseline RS prior, while still centering it on Rudebusch and Svensson's estimates. The estimates still suggest a low degree of persistence despite the tight prior. However, the marginal data density actually falls very sharply because the priors now are located even further from the maximum likelihood estimate. The second column of Table 2 shows this very clearly, as the weight on the RS model declines to zero in this case.

Our second sensitivity check involves using information from a growing literature on the Great Moderation which suggests that the persistence of all economic series in the US has declined dramatically over the past 25 years. This would suggest that the prior degree of persistence should be considerably lower compared to what Rudebusch and Svensson originally estimated. Thus, we specify a tight prior centred on a very low degree of persistence for output and inflation. Again, this change of prior matters more for the marginal data density than for the posterior mean of the parameters. Now prior and likelihood agree to a much greater extent and, as the final column of Table 2 shows, the posterior probability of the Rudebusch-Svensson model jumps dramatically to over 80%, making it the likeliest model in our suite. The posterior probabilities of the other models correspondingly fall: Smets-Wouters now has a probability of 16.3%, BGG has a probability of 3.6%, while the probability of SOE model falls below 0.1%.

The next section devises Bayesian policy rules for each scenario shown in table 2.

4 The Policy Problem

To implement our method, we must specify the function l which maps the values that a model generates for a set of variables, under a given policy rule, into losses, or welfare. Strictly speaking, we should use the welfare of households in each model. So for each model and parameterization there would be a different l function. However, we will, at least initially, abstract from this step, and choose one that we know from previous work[8] will not be too bad an approximation to the 'microfounded' loss function.

4.1 The Period Loss Function

For a given policy ϕ, and a given model j with parameterization θ_{jk}, the period loss function is

$$l_j(\phi, \theta_{jk}) = E\left[var\left(4\pi_t\right) + \lambda_y var\left(y_t - y_t^*\right) + \lambda_i var(4i_t) | \phi, \theta_{jk}\right]. \qquad (6)$$

The loss function depends on the unconditional variance of annualised inflation, the output gap and the annualised nominal interest rate, where $\lambda_y = 1$ is the relative weight on the output gap while $\lambda_i = 0.1$ is the weight on nominal interest rate variability.

The output gap used in computing the loss function differs between the models. In the micro-founded models SW, BGG and SOE, a model consistent measure of 'natural output' y_t^* can be computed and this is what we use in order to compute the output gap $y_t - y_t^*$. In the RS model, the concept of 'natural output' is undefined and therefore we use detrended output y_t.

A small weight on interest-rate variability is included in order to avoid extreme volatility of the policy rate. Woodford (2003) motivates such a term in the objective function by appealing to the desirability of damping variation in the tax on the liquidity services of money. He also argues that an interest-smoothing term helps central banks avoid hitting the zero lower bound on nominal interest rates.

To find the expected loss, we integrate across models and parameterizations, as described in section (1.1). For particular parameter values, some policy settings generate indeterminate equilibria in some of the models. In such cases, we set $l_j(\phi, \theta_{jk}) = \infty$, thus ensuring that the Bayesian policy rule guarantees determinacy.

---

[8] An example is Amato and Laubach (2003) who note that the optimal criterion for monetary poilcy is not much changed by varying inherent backward-lookingness in the inflation process deriving from rule-of thumb behaviour on the part of price-setters.

4.2 Optimal simple rules

We choose the coefficients on our simple rule

$$i_t = \phi_i i_{t-1} + (1 - \phi_i)\left(\phi_\pi \pi_t + \phi_y y_t\right) + \phi_{dy}\left(y_t - y_{t-1}\right) \tag{7}$$

in order to minimise the loss function (5). Notice that the central bank responds to detrended output y_t instead of the model-consistent output gap. This is likely to involve some welfare loss. However, responding to the output gap in a multiple-model world involves considerable complications, and we leave this for future work. For now we take the simple approach of picking a rule which only responds to detrended output, output growth and price inflation.

We begin by examining optimal simple rules in each model. Then we study the optimal policy in the suite as a whole under the assumption that each model has been given its Bayesian weight. We also consider a scenario in which each model receives an equal weight.

When numerically searching for the optimal coefficients in individual models we found loss functions to be rather flat around the optimum. As a result, coefficients often moved a very long way without a significant change in the value of the loss function. For this reason we imposed a limit of 100 on the long-run response coefficients to inflation and output.

4.2.1 Optimal simple rules in individual models

Tables 3 records the policy-rule parameters for optimal simple rules in each of the individual models, and table 4 summarizes the volatility of inflation, output, and nominal interest under those policies. The rules differ in interesting ways.

Table 3: Optimal policy coefficients in the individual models

Coefficients	SW	BGG	SOE	RS1	RS2	RS3
Smoothing	0.99	0.03	0.61	0.06	0.81	0.05
Inflation	65.3	100.0	42.19	0.01	1.01	0.01
Output	7.71	-0.06	-0.20	0.03	0.08	0.05
Output Growth	1.71	-0.20	4.10	0.00	0.10	0.00
Loss	5.62	0.035	0.83	3.45	6.75	3.28

Table 4: Volatility under Model-Specific Policies

	Inflation	Output	Nominal Interest
SW	4.31	1.18	1.34
BGG	0.0002	0.003	0.30
SOE	0.20	1.05	2.64
RS1	3.33	0.10	0.17
RS2	6.17	0.27	3.19
RS3	3.14	0.12	0.17

Note: RS1 refers to the baseline prior, RS2 to the prior tightly centered on Rudebusch and Svensson's estimates, and RS3 to the prior involving weak persistence.

For the BGG model, the optimal policy approximates pure inflation tareting. The BGG model has little price stickiness or price indexation and consequently, both inflation and output-gap volatility can be reduced almost to zero by responding aggressively to inflation deviations from target. The threat of an aggressive response keeps the inflation gap close to zero, and since there are no cost-push shocks and wages are flexible, this also keeps the output gap close to zero. The inflation response coefficient hits the upper bound of 100. Increasing the inflation response further led to very small additional reductions in expected loss, which is why we were happy to cap the coefficient at this level. The response coefficients to output and interest-smoothing parameters are close to zero.

In the Smets-Wouters model, the monetary-policy tradeoff is more challenging because there are large and persistent shocks to the markup and sticky wages as well as sticky prices. Thus, even under the optimized rule, a substantial amount of inflation and output variability remains. Nevertheless, the optimal rule calls for aggressive long-run responses to inflation and output but with an extremely high degree of interest rate smoothing (almost a unit root). This is similar to the optimal simple rules which performed well in the paper by Levin, Wieland and Williams (2003). Accordingly, they advocate 'first-difference' rules, i.e. rules where the *change* in the nominal interest rate is a linear function of inflation and the output gap.[9]

In terms of monetary-policy challenges, the SOE model lies between the BGG and SW models. In this model, the central bank can simultaneously stabilize the output gap and producer prices. The welfare loss cannot be driven to zero, however, because the period loss function (6) depends on consumer price inflation – which is also affected by movements in international relative prices – and movements in interest rates. Like the other forward-looking models, the optimal rule calls for a high long-run coefficient on inflation ($\phi_\pi = 41$), with more interest smoothing than in the SW model and less than in the BGG model. The response to output is also weaker than in the SW model, but the response to output growth is stronger. The

---

[9]Orphanides and Williams (2007) report that first-difference rules also perform well under learning.

minimized welfare loss is significantly lower than in the SW model but higher than in the BGG model.

The Rudebusch and Svensson model variants were perhaps the biggest surprise of all, featuring extremely weak responses to all endogenous variables. Under the baseline prior and the prior which places most weight on low inflation and output persistence, the optimal rule fails to satisfy the Taylor principle. Even when we imposed a tight prior on high inflation and output persistence, the optimal rule remains relatively unresponsive although it just satisfies the Taylor principle.

The reasons for this are simple. First of all, since the model is entirely backward-looking, indeterminacy is not an issue. Secondly, the estimated persistence of both inflation and output are low even in the case of a tight prior on Rudebusch and Svensson's original estimates. This implies that shocks die out relatively quickly regardless of the policy response. Finally, the slope of the IS curve is robustly estimated to be very low. This implies that, due to the penalty on interest-rate variability, responding aggressively to shocks which policy will find it hard to control anyway is not worthwhile. Under priors 1 and 3, the central bank essentially ignores inflation and output and tries to minimize nominal interest volatility. Those rules approximate a pure nominal interest peg.

4.2.2 Fault tolerance of model-specific optimal policies

Next, following McCallum (1988), we consider how the model-specific optimal rules perform in other models. The purpose is to develop intuition about pitfalls central bankers face because of model uncertainty and about the nature of the policies that are optimal across models.

The results are presented in table 5. Each column shows how loss increases as we replace the optimal rule for that model with the optimal rule for another model. A relative loss of unity implies that the alternative rule peforms just as well as the model-specific optimal policy, and a large number indicates that the alternative rule delivers a much inferior performance. We assign an infinite loss whenever a policy rule results in instability or indeterminacy. So, for example, the first column shows how the optimal rule for the Smets-Wouters model performs across the suite. The rule delivers a slight deterioration relative to model-specific optimal rules in the SOE model and the three variants of the RS model. Its relative performance is poor in the BGG model, which dislikes the SW rule's strong response to the output gap.

Table 5: Relative Loss in Model i (rows) Under a Policy Optimized for Model j (columns)

15

	SW	BGG	SOE	RS1	RS2	RS3
SW	1	∞	5.37	∞	1.35	∞
BGG	334	1	3.82	∞	3339	∞
SOE	5.98	1.40	1	∞	50	∞
RS1	2.77	∞	45	1	1.02	1.00
RS2	3.30	∞	∞	∞	1	∞
RS3	1.94	∞	44	1.00	1.02	1

Note: RS1 refers to the baseline prior, RS2 to the prior tightly centered on Rudebusch and Svensson's estimates, and RS3 to the prior involving weak persistence. We report an infinite loss (denoted by ∞) when the model is unstable or indeterminate.

The first and most important lesson that emerges from this table is that rules optimized for variants of the Rudebusch-Svensson model are dangerous for forward-looking economies. The rules optimized for RS1 and RS3 fail to satisfy the Taylor principle and result in indeterminacy in forward-looking models. The rule optimized for RS2 satisfies the Taylor principle and delivers a unique solution in all models in the suite, but since the long-run inflation response is only slighly above unity this rule does badly in models with little nominal inertia (BGG and SOE). Such models have the implication that a strong response to inflation can drive losses almost to zero, and relative losses rise rapidly as the inflation response weakens. Thus, the forward-looking models in our suite have low fault tolerance with respect to policies devised for the backward-looking models. The BGG and SOE models are less fault tolerant than the SW model.

A second important lesson is that inflation-only Taylor rules can be dangerous. The BGG optimal policy, which responds little to variables other than inflation, works well when there is little nominal inertia (BGG and SOE), but it works poorly in the other models, generating explosive outcomes in the backward-looking models. Somewhat to our surprise, the SW model also becomes unstable when subjected to the BGG-optimal rule. The SW model has considerable nominal inertia, and backward-looking indexation makes price and wage inflation partly pre-determined. The SW-optimal policy calls for a high degree of interest-rate smoothing in order to stabilise inflation and the output gap without excessive volatility in the short-term interest rate. The highly inertial response to current conditions allows long real rates to fluctuate substantially while the short rate (which enters the loss function) remains stable. In contrast, the BGG-optimal rule calls for an enormous short-run response to price inflation with essentially no interest smoothing. This very strong response to a partly predetermined variable makes outcomes unstable.

A third lesson is that RS1 and RS3 have high fault tolerance as long as the policy rule does not have an enormous short-run response coefficient on inflation. Except for the BGG-optimal rule, all the policies deliver acceptable performance in these models. This follows from that fact that intrinsic persistence is weak, that the slope

of the IS curve is close to zero, and that the weight on nominal interest volatility is small.

Finally, the SW-optimal policy performs reasonably well in all models. The relative loss in the BGG model is 334, but the absolute loss under the BGG-optimal policy is small, and 334 times that small number comes to 11.7. This is approximately twice the absolute loss in the SW model under the SW policy.

Table 6 below presents additional information on the behavior of the models under alternative policies. Each cell contains three numbers - the standard deviations of annualised inflation, the output gap and the annualised nominal interest rate. This allows us to trace the exact sources of fault intolerance in the various models.

The first three rows describe the performance of the three forward-looking models. The SW model has high fault tolerance with respect to the SOE- and RS2-optimal policies. Like the SW-optimal policy, the RS2-optimal rule also involves high interest smoothing, but with weak responses to inflation and output. This results in lower nominal interest volatility and only slightly higher inflation volatility, but output volatility increases by a factor of 2.5. The SOE-optimal policy involves less interest smoothing and a more aggressive short-term response to inflation. The more aggressive short-term response to inflation reduces inflation volatility by about 20 percent, but output volatility increases by a factor of 12 and there is an enormouse increase in nominal interest volatility. This happens because the cost-push shocks and nominal wage inertia makes inflation stabilisation much more costly in the SW model than in the SOE model.

Table 6: Volatility in Model i (rows) Under a Policy Optimized for Model j (columns)

	SW	BGG	SOE	RS1	RS2	RS3
SW	4.31 1.18 1.34	–	3.45 13.1 136	–	4.57 2.93 0.77	–
BGG	11.09 0.03 0.96	0.0002 0.003 0.3	0.09 0.003 0.35	–	108 0.23 38.3	–
SOE	4.43 0.88 1.58	0.001 1.14 6.82	0.20 1.05 2.64	–	38.7 1.11 15.12	–
RS1	3.32 0.17 60.5	–	3.28 1.60 1507	3.33 0.10 0.17	3.32 0.10 1.051	3.33 0.10 0.17
RS2	5.90 1.13 153	–	–	–	6.17 0.27 3.19	–
RS3	3.14 0.19 30.3	–	3.17 2.36 1385	3.14 0.12 0.17	3.14 0.12 0.95	3.14 0.12 0.17

Note: The entries in each cell represent the standard deviation of inflation, output, and nominal interest, respectively. Empty cells refer to indeterminate or explosive outcomes. RS1 refers to the baseline prior, RS2 to the prior tightly centered on Rudebusch and Svensson's estimates, and RS3 to the prior involving weak persistence.

The BGG model has high fault tolerance with respect to the SOE-optimal policy but low fault tolerance with respect to the SW and RS2 optimal rules. Because prices are estimated to be almost flexible in the BGG model, fluctuations in detrended output are efficient and do not correspond to movements in the output gap (which is approximately equal to zero at all times). Therefore an aggressive response to detrended output such as under the SW-optimal policy leads to enormous fluctuation in inflation. Equally, the BGG model behave poorly under with a weak long-run inflation response such as the RS2-optimal policy because these policies fail to stabilise inflation and also lead to volatile nominal interest rates. The BGG model performs well under policies that respond strongly to inflation and weakly to output. Among the other model-specific rules, the SOE-optimal policy comes closest to this description.

The SOE behaves similarly (see the third row of the table). Prices are again fairly flexible and inflation has no intrinsic persistence. Consequently welfare under the SOE model deteriorates either when a rule responds strongly to output (the SW rule) or when it responds insufficiently to inflation (the RS2 rule). The model performs well under the BGG-optimal policty, albeit with a substantial increase in

nominal interest volatility.

The last three rows of Table 6 describe the performance of the RS model variants under alternative policies. RS1 and RS3 are highly fault tolerant. One remarkable feature of these models is that, with the exception of the BGG-optimal rule, inflation and output volatility is approximately invariant to changes in policy. The difference in welfare under alternative policies is due almost entirely to changes interest rate variability. For example, under the RS1- and RS3-optimal rules, nominal interest volatility is 0.17. This increases to about 1 percent under the RS2-optimal policy, but rises enormously under the SW or SOE-optimal rules. Other components of the loss function hardly change. This demonstrates yet again what kinds of policies are optimal in the RS model. Because the slope of the IS curve is small and inflation has a little intrinsic persistence, policy cannot do much to stabilise the economy. Responding aggressively carries little benefit in terms of lower output gap and inflation volatility but plenty of costs in terms of higher nominal interest rate variability.

4.2.3   Optimal simple rules under Bayesian model weights

Table 7 describe Bayesian policies for three versions of our suite. The first column shows the optimal simple rule formed by combining SW, BGG, SOE, and RS1. The second and third columns retain the forward-looking models and replace RS1 with RS2 and RS3, respectively. Within each suite, the models are weighed in accordance with the posterior probabilities shown in table 2.

Table 7: Optimal policy coefficients with Bayesian model weights

Coefficients	Bayes 1	Bayes 2	Bayes 3
Smoothing	0.97	0.97	0.51
Inflation	39.5	48.81	1.53
Output	4.60	4.92	0.07
Output growth	1.60	1.85	-0.01
Loss	5.59	5.42	4.09

Note: Bayes 1, 2, and 3 refer to suites formed by combining the forward-looking models with RS1, RS2, and RS3, respectively.

The optimal simple rules in the first and second columns are similar in that both feature a high degree of interest smoothing along with large long-run responses to inflation and real activity. Indeed, these policy rules differ only slightly from the SW-optimal rule. Bayesian policies 1 and 2 call for a bit less interest smoothing than in the SW-optimal policy, and the long-run inflation and output responses are slightly lower. This results in stronger short-run responses to inflation and output, hedging slightly in the direction of the BGG- and SOE-optimal rules. This outcome reflects

19

the high probability weight on the SW model (greater than 80 percent in both suites), the high fault tolerance of the other models with respect to the SW-optimal rule, and the low fault tolerance of various models with respect to other model-specific optimal policies. Notice in particular that the backward-looking models in these suites have little influence on Bayesian policy because they have low probability weight and high fault tolerance.

Tables 8 and 9 provide more intuition about the Bayesian policies. Table 8 shows relative loss in each model under the various Bayesian rules. As before, a value of unity means that the Bayesian rule performs just as well as the model-specific optimal policy, and a large value indicates a substantial deterioration in performance. Table 9 breaks down expected loss into its components, viz. the standard deviations of inflation, output, and nominal interest.

Table 8: Relative Loss Under Bayesian Policies

	Bayes 1	Bayes 2	Bayes 3
SW	1.04	1.07	1.11
BGG	106	84.5	142.1
SOE	3.01	2.66	7.67
RS1	3.57	–	–
RS2	–	5.69	–
RS3	–	–	1.09

Note: Losses are reported relative to the policy that is optimal in each model.

Table 9: Volatility Under Bayesian Policies

	Bayes 1			Bayes 2			Bayes 3		
SW	4.21	1.39	2.70	4.17	1.49	3.36	4.47	1.35	4.49
BGG	3.57	0.01	0.79	2.76	0.01	0.71	4.56	0.01	2.04
SOE	2.06	0.93	1.39	1.80	0.93	1.26	5.34	0.99	8.05
RS1	3.31	0.21	87.7		–			–	
RS2		–		5.66	1.94	308		–	
RS3		–			–		3.14	0.12	3.06

Note: The entries in each cell represent the standard deviation of inflation, output, and nominal interest, respectively. Bayes 1, 2, and 3 refer to suites formed by combining the forward-looking models with RS1, RS2, or RS3.

By design, Bayesian policies rule out indeterminate and explosive outcomes. Therefore, in contrast with table 5, the losses reported in table 8 are all finite.[10] Indeed, except for the BGG model, relative loss never exceeds 10. Also, for the BGG model,

---

[10]Empty cells refer to models not in the suite, not to infinite or undefined losses.

absolute loss never exceeds 5. As shown in table 3, this approximates the loss in the SW model under the SW-optimal policy. Thus, although the Bayesian policies result in large relative losses in the BGG model, they do not result in large absolute losses.

By hedging slightly in the direction of BGG- and SOE-optimal policies, the Bayesian policy-maker tries to mitigate losses in the BGG and SOE models while still achieving good performance in the SW model. Relative to the SW-optimal rule, these policies reduce expected loss by two-thirds to three-quarters in the BGG model and by about half in the SOE model. For both models, most of the improvement is due to a reduction in inflation volatility. These gains are accomplished at the expense of a slight rise in expected loss in the SW model, which increases by 4 and 7 percent, respectively, in the two suites. Inflation volatility in the SW model is about the same under the Bayesian policies as under the SW-optimal policy, but the standard deviation of output is about 20-25 percent higher, and the standard deviation of the nominal interest rate increases by a factor of 2 or 3.

On the other hand, outcomes in the RS models are worse under the Bayesian policies than under the SW-optimal rule. As explained above, alternative policies have little influence on inflation and output volatility in the RS models. Their main influence is on nominal interest volatility. In the RS models, the nominal interest rate is enormously volatile under the SW-optimal policy, and it is even more volatile under Bayesian policies 1 and 2. The Bayesian decision maker is content with this outcome because the RS models have low probability weights in suites 1 and 2.

Matters are different in suite 3, which combines the forward-looking models with RS3. Recall that this version of the RS model was estimated under a tight prior featuring low output and inflation persistence. In this suite, the RS model has the highest probability weight – approximately 80 percent – and the forward-looking models have low probability weights. Because the probability weight on the backward-looking model is much greater than in the other suites, the Bayesian policy differs substantially from those in the first two columns, involving a modest degree of interest rate smoothing, a long-run inflation response coefficient around 1.5, and small response coefficients to real activity (see the final column of table 7). Except for the small output coefficients, this resembles a conventional Taylor rule with interest smoothing.

Interestingly, for this suite the Bayesian policy differs significantly from the optimal policy of its most probable member. Recall from the subsection 4.2.1 that the RS3-optimal rule has response coefficients close to zero on all arguments, approximating a pure nominal interest peg. According to RS3, monetary policy has little influence on output because the slope of the IS curve is close to zero and shocks die out quickly on their own. Roughly speaking, since movements in nominal interest are penalized and have little influence on inflation or output, the best a central bank can do is to minimize nominal interest volatility.

The RS3-optimal policy cannot be optimal for the suite, however, because it violates the Taylor principle and generates indeterminacy in the forward-looking models. Because policy rules that generate indeterminate outcomes are heavily penalized, our Bayesian central banker shies away from the RS3-optimal rule, as well as from anything close to it. First and foremost, s/he seeks a rule that guarantees determinacy in all the models. Within that family, s/he strikes a balance between performance in the various models. The RS model is more influential here than in suites 1 and 2 because of its higher probability weight, but it cannot be dominant for policy because its recommended policy generates infinite loss in the other models.[11]

In this case, a Bayesian decision maker must balance concerns about nominal interest volatility against requirements for determinacy in other members of the suite. S/he strikes this balance by moving the long-run response coefficient on inflation into the determinacy region while leaving the response coefficients on output close to zero and smoothing interest rates to a modest extent.

The RS3 model performs well under this policy. Inflation and output volatility are about the same as under the RS3-optimal policy, but the standard deviation of the nominal interest rate is about 20 times higher. Although this increase in volatility is costly, it pales in comparison with the expected cost of indeterminacy in the other models. The Bayesian policy maker is content with this compromise, despite the high probability weight on RS3.

The SW model also performs well under the Bayesian rule. No comparison is possible with outcomes under the RS3-optimal policy because that rule generates indeterminacy in the SW model. Relative to Bayesian policies 1 and 2, however, inflation and output volatility are about the same, and the standard deviation of the nominal interest rate is one- to two-thirds higher. With less interest smoothing, more vigorous movements in the short-term interest rate are needed to stabilize output and inflation. Although more interest smoothing and a higher long-run inflation response would be desirable for this model, it would be counterproductive for the suite because of its implications for nominal interest volatility in the RS3 model.

The BGG and SOE models also perform reasonably well under Bayesian rule 3. Output volatility is roughly the same as under Bayesian policies 1 and 2, but inflation and nominal interest volatility are higher. A more aggressive long-run response to inflation and less interest smoothing would be desirable for these models, but again would be counterproductive for RS3. Since the BGG and SOE models have low probability weight in this suite, the policy maker is content with ensuring determinacy and does not attempt to fine-tune outcomes in these models.

---

[11]Cogley and Sargent (2005) report the opposite finding for the U.S. During the Great Inflation, backward-looking models were dominant for policy despite having low probability weights because the policy designed for a more probable forward-looking model would have been disastrous for backward-looking economies. In this suite, the backward-looking model has the highest probability weight, yet its policy would be disastrous for forward-looking economies.

4.2.4   Optimal simple rules under equal model weights

The forecasting literature has found that model averaging sometimes works better when models are assigned simple (usually equal) weights as opposed to Bayesian probability weights. With this result in mind, we repeat the policy design exercise with equal model weights. Our loss function now becomes the simple average of expected losses conditional on each individual model,

$$l(\phi) = \frac{1}{m} \sum\nolimits_{i=1}^{m} l_i(\phi). \tag{8}$$

Like Bayesian model averaging, this approach rules out indeterminate and explosive outcomes by design, preserving that important aspect of robustness.

The results are presented in tables 10 and 11. Table 10 records the optimal policy-rule coefficients for each suite, and table 11 describes the volatility of inflation, output, and nominal interest under these policies.

Table 10: Optimal policy coefficients with equal model weights

Coefficients	Suite 1	Suite 2	Suite 3
Smoothing	0.37	0.27	0.37
Inflation	2.55	2.17	2.57
Output	0.04	0.01	0.03
Output growth	0.63	0.57	0.64
Loss	3.58	4.68	3.53

Table 11: Volatility under equal-weight policies

	Suite 1			Suite 2			Suite 3		
SW	4.14	1.92	8.09	4.18	1.91	8.14	4.13	2.00	8.10
BGG	0.99	0.01	0.84	1.15	0.01	0.99	1.05	0.01	0.82
SOE	1.21	1.02	2.66	1.36	1.02	3.20	1.17	1.02	2.58
RS1	3.32	0.10	11.8		—			—	
RS2		—		5.82	0.30	21.9		—	
RS3		—			—		3.14	0.12	11.4

Note: the entries in each cell represent the standard deviation of inflation, output, and nominal interest, respectively. Suites 1, 2, and 3 are formed by combining the forward-looking models with RS1, RS2, or RS3.

In all three cases, the policies resemble speed-limit versions of the Taylor rule with a modest degree of interest smoothing. The response coefficients on output are close to zero, those on output growth are around 0.6, the long-run response to inflation

ranges from 2.2 to 2.6, and the interest-smoothing parameter hovers around 0.35. The policies are similar across suites, which shows that differences in the prior over RS-model parameters have almost no impact on optimal policy over and above their effect on model weights.

In suites 1 and 2, the evenly-weighted policies deviate quite a bit more from the SW-optimal rule than the Bayesian policies. There is a lot less interest smoothing, the long-run response coefficients on inflation and output are much smaller, and the coefficient on output growth falls by about two-thirds. This reflects the fact that the BGG, SOE, and RS models are now assigned higher weight at the expense of the SW model. In particular, the weights on SOE and RS rise from close to 0 to 0.25. Thus, while a Bayesian policy maker could essentially ignore the SOE and RS models, that is no longer the case. In particular, the consequences for nominal interest volatility in the RS models must now be taken more seriously. As shown in table 11, outcomes in the BGG, SOE, and RS models improve, while those in the SW model deteriorate. Inflation is less volatile in the BGG and SOE models and nominal interest volatility is lower in the SOE and RS models. The biggest change is a decline in the standard deviation of the nominal interest rate in the RS model, which falls from 87.7 to 11.8 in suite 1 and from 308 to 22 in suite 2. This is purchased at the expense of higher output and nominal interest volatility in the SW model. The standard deviation of output rises by 30 and 40 percent in the two suites and the standard deviation of the nominal interest rate increases by 140 and 200 percent.

The contrast between equal- and Bayesian-weighted policies is less stark for suite 3. In this case, the main difference is that the equal-weight policy has a higher speed-limit coefficient (0.64 v. -0.01) and also a higher long-run inflation response coeffient (2.57 v. 1.53). This reflects the fact that RS3 loses influence relative to the forward-looking models. That reduces concerns about nominal interest volatility in RS3 and increases the weight placed on obtaining good outcomes in the forward-looking models, especially in the BGG and SOE models. Thus, in the RS3 model, the standard deviation of the interest rate rises from 3.06 under Bayes policy 3 to 11.4 under the evenly-weighted policy. Performance in the SW model again deteriorates under the evenly-weighted policy, with a 50 percent increase in the standard deviation of output and an 80 percent increase in the standard deviation of the nominal interest rate. Balanced against that deterioration are substantial improvements in outcomes in the BGG and SOE models. In these models, the standard deviation of inflation falls by almost 80 percent, and the standard deviation of the nominal interest rate falls by 40 to 70 percent.

Whether these changes represent improvements relative to Bayesian policies depends on one's attitudes about model weights, which is subjective. There are cogent arguments both for and against model averaging with equal weights. Arguments in favor stress difficulties associated with managing the model set and estimating Bayesian model probabilities. Arguments against stress that assigning equal weights

favors poor-fitting models at the expense of good-fitting models. As this is an open area of research, we are content to present both sets of policies and to leave questions about how best to assign model weights to future research.

5   Conclusions

This paper executes a Bayesian analysis of optimal monetary policy for the U.K. Our method takes into account model and parameter uncertainty as well as uncertainty about future shocks and outcomes. We examine a suite of models that have received a lot of attention in the monetary policy, including versions of the Rudebusch-Svensson (1999) model, the Smets-Wouters (2007) model, the Bernanke, Gertler and Gilchrist (1999) model, and the small-open-economy model of Gali and Monacelli (2005). We estimate each model using Bayesian methods and calculate posterior model probabilities. Then we compute the coefficients of a simple rule that minimises expected losses, where expectations are taken across uncertainty about shocks, parameters, and models, and where losses are defined as a weighted sum of the unconditional variance of inflation, the output gap and the change in the interest rate. Since our methods are modular, adding new models to the suite is straightforward. Indeed, because of its modular nature, it would be possible to extend this research through a network of decentralized modeling groups.

Several conclusions emerge from our analysis. First, the rule which is optimal within each model differs substantially across models. Our best estimates of the RS model suggest there is little intrinsic inflation inertia. Since that model is backward looking and shocks dissipate quickly on their own, the optimal RS rule is passive and seeks mainly to minimize interest-rate volatility. Indeed, for two versions of the RS model, the model-specific optimal policy approximates a pure nominal-interest peg. At the other end of the spectrum, the policy optimal for the BGG model is approximately equivalent to an inflation-only Taylor rule. Our estimates of the BGG model find little evidence of inflation inertia. Because this is a forward-looking model, the optimal BGG rule responds very aggressively to deviations of inflation from its target, with little response to other variables. The SW-optimal rule approximates a first-difference rule for the nominal interest rate with high long-run response coefficients on inflation and output. This follows from the fact that the SW model features sticker prices and both sticky wages and large and persistent cost-push shocks, thus presenting a more challenging policy tradeoff. Finally, in the small open economy model, the central bank can simultaneously stabilize the output gap and producer prices (though not consumer prices). As a result welfare loss is significantly lower than in the SW model. Like the other forward-looking models, the optimal rule calls for a high long-run coefficient on inflation.

Second, the forward-looking models have low fault tolerance with respect to poli-

cies designed for the backward-looking models. Those policies either violate the Taylor principle or barely satisfy the Taylor principle with long-run inflation response coefficients just above 1. Outcomes in the forward-looking models are poor in either case.

In contrast, the backward-looking models have high fault tolerance with respect to policies designed for forward-looking models. In this respect, results for the U.K. contrast sharply with those for the U.S. One of the main challenges for the U.S. is to find a rule that works well both for forward- and backward-looking models. Backward-looking models typically imply a high degree of intrinsic inflation persistence when estimated with U.S. data. Policy rules that succeed in stabilizing inflation in forward-looking models often result in excessive output variability in backward-looking models, while gradualist rules well adapted to a backward-looking environment permit more inflation variabity in forward-looking models than one might like. Finding a rule well adapted to both environments is difficult. For the U.K., this turns out not to be an issue because backward-looking models estimated with U.K. data for the inflation-targeting period involve little instrinsic persistence. Thus, rules that work well for forward-looking models also work well in our backward-looking models. Hence optimal rules bear a closer resemblance to those for forward-looking models than would be the case for the U.S.

In two of the three suites, the backward-looking model has a low probability weight. Since it is also highly fault tolerant, it has virtually no influence on the optimal Bayesian policy. In those suites, the SW model has a high probability weight, and the optimal Bayesian policy resembles the SW-optimal policy, with a slight hedge in the direction of policies appropriate for the other forward-looking models. Relative to the SW-optimal policy, the Bayesian policy improves outcomes substantially in the other forward-looking models at the cost of a slight deterioration in outcomes in the SW model.

In the third suite, the backward-looking model has a probability weight of 0.8, and the forward-looking models collectively have weight of 0.2. Despite that, the optimal Bayesian policy differs subsantially from the policy that is optimal for the backward-looking model, which violates the Taylor principle. Since we assign an infinite loss to indeterminate outcomes, our Bayesian policy maker shies away from the RS-optimal rule, seeking first and foremost a rule that guarantees determinacy in all the models. Within that family, s/he strikes a balance between performance in the various models. The optimal Bayesian policy in this case is a Taylor rule with modest interest smoothing, a long-run inflation response around 1.5, and virtually no reaction to output or output growth.

6   References

Amato, J and Laubach, T (2003) 'Rule-of-thumb behaviour and monetary policy', *European Economic Review*, vol 47(4), October, pp 791-831.

An, S. and F. Schorfheide, 2007, Bayesian analysis of DSGE models, *Econometric Reviews* 26, 113-172.

Bates, and Granger (1969) 'The combination of forecasts', *Operations Research Quarterly* 20, 451-68.

Benati, L., 2008, Investigating inflation persistence across monetary regimes, Quarterly Journal of Economics 123, 1005-1060.

Bernanke, B., M. Gertler, and S. Gilchrist, 1999, The financial accelerator in a quantitative business cycle framework, in J.B. Taylor and M. Woodford, eds., *Handbook of Macroeconomics,* North Holland.

Brock, W., S. Durlauf, and K. West, 2003, Policy evaluation in uncertain environments, *Brookings Papers on Economic Activity*, 235-322.

Brock, W, S. Durlauf, K. West, 2007, Model uncertainty and policy evaluation: some theory and empirics, *Journal of Econometrics* 136 (2), 629-664.

Chari, V.V., P.J. Kehoe, and E.R. McGrattan, 2002, Can sticky price models generate volatile and persistent real exchange rates?, *Review of Economic Studies* 69, 533-563.

Christiano, L.J., M. Eichenbaum, and C.L. Evans, Nominal Rigidities and the dynamic effects of a shock to monetary policy, *Journal of Political Economy* 113, 1-45.

Clements, M. and D. Hendry, 1998, *Forecasting economic time series*, Cambridge, CUP.

Clements, M. and D. Hendry, 2002, Pooling of forecasts, *Econometrics Journal*,5,1-26.

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*, supplement to Vol. 39, pp. 67–99.

Cogley, T., R. Colacito, L.P. Hansen, and T.J. Sargent, 2008, Robustness and U.S. monetary policy experimentation, *Journal of Money, Credit and Banking* 40, 1599-1623.

Cogley, T. and T.J. Sargent, 2005, The Conquest of US Inflation: learning and robustness to model uncertainty, *Review of Economic Dynamics* 8, 528-563.

Cogley, T. and A.M. Sbordone, 2008, Trend inflation, indexation, and inflation persistence in the new Keynesian Phillips curve, *American Economic Review* 98, 2101-2126.

De Paoli, B., 2009, Monetary policy and welfare in a small open economy, Journal of International Economics 77, 11-22.

Diebold, F. and P. Pauly, 1987, Structural change and the combination of forecasts, *Journal of Forecasting* 6, 503-508

Gali, J. and T. Monacelli, 2005, Monetary policy and exchange rate volatility in a small open economy, *Review of Economic Studies* 72, 707-734.

Greenspan, A., 2004, Risk and uncertainty in monetary policy, *American Economic Review Papers and Proceedings* 94(2), 33-40.

Geweke 1999 ???.

Hansen, L.P. and T.J. Sargent, 2007, *Robustness*, Princeton, PUP.

Ireland, P.N., 2007, Changes in the federal reserve's inflation target: causes and consequences, *Journal of Money Credit and Banking* 39, 1851-1882.

Jacobson, T. and S. Karlsson, 2004, Finding good predictors for inflation: a Bayesian model averaging approach, *Journal of Forecasting* 23, 479-496.

King, M., 2004, Innovations and issues in monetary policy: panel discussion," *American Economic Review Papers and Proceedings* 94(2), 43-45.

Kapetanios, G., V. Labhard, and S. Price, 2008, Forecasting using Bayesian and information-theoretic model averaging: An application to U.K. inflation, *Journal of Business and Economic Statistics* 26, 33-41.

Kimball, M.S., 1995, The quantitative analytics of the basic neomonetarist model, *Journal of Money Credit and Banking* 27, 1241-1277.

Levin, A.T., A. Onatski, J.C. Williams, N. Williams, 2005, Monetary policy under uncertainty in micro-founded macroeconometric models, *NBER Macroeconomics Annual* 20, 229-287.

Levin, A.T. and J.M. Piger, 2006, Is inflation persistence intrinsic in industrial economies? unbublished manuscript, University of Oregon.

Levin, A.T., V. Wieland, and J.C. Williams, 2003, The performance of forecast-based monetary policy rules under model uncertainty, *American Economic Review* 93, 622-645.

Levin, A.T. and J.C. Williams, 2003, Robust monetary policy with competing reference models, *Journal of Monetary Economics* 50, 945-975.

Lubik, T.A. and F. Schorfheide, 2007, Do central banks respond to exchange rate movements? A structural investigation, *Journal of Monetary Economics* 54, 1069-1087.

McCallum, B.M., 1988, Robustness properties of a rule for monetary policy', *Carnegie-Rochester Conference Series on Monetary Policy* 29, 173-203.

Newbold, P. and A. Harvey, 2002, Forecast combination and encompassing' in Clements, M. and Hendry, D. [eds] , *A companion to economic forecasting*, Oxford, Basil Blackwell.

Orphanides. A. and J.C. Williams, 2007, Robust monetary policy with imperfect knowledge, Journal of Monetary Economics 54 (5 Spec. Iss.), 1406-1435.

Rudebusch, G.D. and L.E.O. Svensson, 1999, Policy rules for inflation targeting, in J.B. Taylor, ed., *Monetary Policy Rules,* Chicago, University of Chicago Press.

Schorfheide, F., 2000, Loss function-based analysis of DSGE models, *Journal of Applied Econometrics* 15, 645-670.

Smets, F. and R. Wouters, 2003, An estimated dynamic stochastic general equilibrium model of the euro area, *Journal of the European Economic Association* 1, 1123-1175.

Smets, F. and R. Wouters, 2007, Shocks and frictions in US business cycles: A Bayesian DSGE approach, American Economic Review 97, 586-606.

Svensson, L.E.O. and N. Williams, 2007a, Monetary policy with model uncertainty: distribution forecast targeting, unpublished manuscript, University of Wisconsin.

Svensson, L.E.O. and N. Williams, 2007a, Bayesian and adaptive optimal policy under model uncertainty, unpublished manuscript, University of Wisconsin.

Svensson, L.E.O. and N. Williams, 2008a, Optimal monetary policy under uncertainty: a Markov jump-linear-quadratic approach, *Federal Reserve Bank of St. Louis Review* 90 (4), 275-29.

Svensson, L.E.O. and N. Williams, 2008b, Optimal monetary policy in DSGE models: a Markov jump-linear-quadratic approach, unpublished manuscript, University of Wisconsin.

Taylor, J.B., ed., 1999, *Monetary Policy Rules,* Chicago, University of Chicago Press.

Townsend, R.M., 1979, Optimal contracts and competitive markets with costly state verification, Journal of Economic Theory 21, 265-293.

Woodford, M., 1999, Optimal Monetary Policy Inertia, *Manchester School* 67 (Suppl. 1), 1-35.

Woodford, M., 2003, *Interest and Prices*, Princeton: PUP.

A The Rudebusch-Svensson (1999) model

The Rudebusch-Svensson model consists of three equations - a Phillips curve, an IS curve, and a policy rule. Inflation is determined according to a reduced-form

Phillips curve,

$$\pi_t = \sum_{i=1}^{4} a_{pi}\pi_{t-i} + a_y y_{t-1} + \varepsilon_{pt}, \tag{9}$$

where π_t is inflation and y_t is the output gap. For estimation, the output gap is measured as linearly detrended output. Aggregate demand is governed by an IS curve,

$$y_t = \sum_{i=1}^{2} b_{yi} y_{t-i} + b_r r_{t-1} + \varepsilon_{gt}, \tag{10}$$

where the ex post real interest rate r_t is defined as

$$r_{t-1} = (1/4) \sum_{i=1}^{4} \left(i_{t-i} - \pi_{t-i} \right). \tag{11}$$

Finally, the monetary authorities set the nominal interest rate in accordance with a Taylor-type rule,

$$i_t = \phi_i i_{t-1} + (1 - \phi_i) \left(\phi_\pi \pi_t + \phi_y y_t \right) + \phi_{dy} \left(y_t - y_{t-1} \right) + \varepsilon_{rt}. \tag{12}$$

A.1   Priors for the RS model

Table 1 summarizes our baseline prior. Because the model lacks micro-foundations, it is not easy for us to elicit an informative prior for its parameters. As a benchmark, we therefore choose a prior whose mode lies on Rudebusch and Svensson's original estimates. Its key features can be characterised as follows. First, the Phillips curve encodes a high degree of intrinsic inflation persistence, with the lag coefficients on inflation summing to unity. Inflation is also more responsive to current output than in conventional calibrations of new Keynesian models. The IS curve also encodes instrinsic output persistence, with the lag coefficients on output summing to 0.91. On the other hand, the slope of the IS curve with respect to the real interest rate is relatively small. Finally, the prior variances are large, reflecting the considerable degree of uncertainty about the relevance of U.S. estimates for models of the U.K.

Table A1: Benchmark Prior for the Rudesbusch-Svennson model

Parameter	Distribution	Mean	Standard Deviation
a_{p1}	Beta	0.6	0.22
a_{p2}	Normal	-0.1	0.2
a_{p3}	Normal	0.28	0.2
a_{p4}	Normal	0.12	0.2
a_y	Gamma	0.14	0.1
b_{y1}	Normal	1.16	0.3
b_{y2}	Normal	-0.25	0.2
b_r	Gamma	0.26	0.2
ϕ_r	Beta	0.7	0.05
$\phi_\pi - 1$	Gamma	1.0	0.1
ϕ_y	Normal	0.125	0.05
ϕ_{dy}	Normal	0.125	0.05
σ_p^2	Inverse Gamma	0.25	0.2
σ_g^2	Inverse Gamma	0.25	0.2
σ_r^2	Inverse Gamma	0.25	0.2

We also examine the sensitivity of our results with respect to two alternative priors: (1) tighter priors based on the original Rudebusch-Svensson estimates and (2) tighter priors centred on simple AR(1) specifications involving a low degree of inflation and output persistence. Those priors are summarized in tables 2 and 3.

Table A2: A tight prior around the original RS estimates

Parameter	Distribution	Mean	Standard Deviation
a_{p1}	Normal	0.7	0.1
a_{p2}	Normal	-0.1	0.1
a_{p3}	Normal	0.28	0.1
a_{p4}	Normal	0.12	0.1
a_y	Gamma	0.14	0.1
b_{y1}	Normal	1.16	0.1
b_{y2}	Normal	-0.25	0.1
b_r	Gamma	0.18	0.1
ϕ_r	Beta	0.7	0.05
$\phi_\pi - 1$	Gamma	1.0	0.1
ϕ_y	Normal	0.125	0.05
ϕ_{dy}	Normal	0.125	0.05
σ_p^2	Inverse Gamma	0.25	0.2
σ_g^2	Inverse Gamma	0.25	0.2
σ_r^2	Inverse Gamma	0.25	0.2

Table A3: A tight prior around low inflation and output persistence

Parameter	Distribution	Mean	Standard Deviation
a_{p1}	Normal	0.3	0.1
a_{p2}	Normal	0.0	0.1
a_{p3}	Normal	0.0	0.1
a_{p4}	Normal	0.0	0.1
a_y	Gamma	0.14	0.1
b_{y1}	Normal	0.7	0.1
b_{y2}	Normal	0.0	0.1
b_r	Gamma	0.18	0.1
ϕ_r	Beta	0.7	0.05
$\phi_\pi - 1$	Gamma	1.0	0.1
ϕ_y	Normal	0.125	0.05
ϕ_{dy}	Normal	0.125	0.05
σ_p^2	Inverse Gamma	0.25	0.2
σ_g^2	Inverse Gamma	0.25	0.2
σ_r^2	Inverse Gamma	0.25	0.2

A.2   Posterior for the RS model

Table 4 summarizes the posterior distribution under the baseline prior. Two aspects of the estimates are apparent. First, inflation and output are considerably less persistent than in the prior. For instance, according to posterior mean estimates, the lag coefficients on inflation in the Phillips curve sum to 0.243, while the lag coefficients on output in the IS curve sum to 0.423. Second, the slope of the IS curve with respect to the real interest rate is smaller than in the prior. Indeed, the lower end of a 95 percent credible set is only slightly above zero. If this coefficient were equal to zero, the central bank could not influence output or inflation via an interest-rate rule, and realizations of output and inflation would be independent of policy-rule parameters.

Table A4: Posterior for the Rudebusch-Svensson model (baseline prior)

Parameter	Prior mean	Post. mode	Post. mean	5th %ile	95th %ile
a_{p1}	0.6	0.1212	0.1592	0.0430	0.3171
a_{p2}	-0.1	-0.0758	-0.0619	-0.2484	0.1264
a_{p3}	0.28	-0.0758	-0.0442	-0.2423	0.1570
a_{p4}	0.12	0.1919	0.1900	-0.0113	0.3838
a_y	0.14	0.0506	0.1118	0.0201	0.2635
b_{y1}	1.16	0.4899	0.5069	0.3072	0.7139
b_{y2}	-0.25	-0.0606	-0.0842	-0.2729	0.1031
b_r	0.26	0.0506	0.0953	0.0167	0.2138
ϕ_r	0.7	0.9646	0.9589	0.9358	0.9781
$\phi_\pi - 1$	1.0	0.8636	0.8762	0.5996	1.1815
ϕ_y	0.125	0.1011	0.1275	0.0563	0.2220
ϕ_{dy}	0.125	0.0506	0.0603	0.0290	0.0989
σ_p^2	0.25	0.4142	0.4213	0.3582	0.4953
σ_g^2	0.25	0.2627	0.2699	0.2302	0.3174
σ_r^2	0.25	0.1011	0.1009	0.0853	0.1193

Recall that the benchmark prior was centered on Rudebusch and Svensson's estimates for the U.S. Table 4 therefore documents that there is less instrinsic persistence in U.K. data for the inflation-targeting period than in Rudebusch and Svensson's sample. To assess the robustness of this finding, we re-estimate the model using the tighter RS prior listed in table 2. This represents an attempt to force high intrinsic persistence onto the data. The posterior corresponding to this prior is summarized in table 5. To our surprise, we found that the estimates are broadly similar to those for the benchmark prior. As another robustness check, we re-estimate the model with an informative prior involving low degrees of intrinsic persistence (see table 3). Table 6 summarizes the posterior associated with this prior. Once again, we find that model's characteristics are qualitatively robust to changes in the prior.

Table A5: RS model prosterior (A tight prior around the original RS estimates)

Parameter	Prior mean	Post. mode	Post. mean	5th %ile	95th %ile
a_{p1}	0.7	0.4545	0.4592	0.3160	0.6007
a_{p2}	-0.1	-0.0909	-0.0981	-0.2309	0.0346
a_{p3}	0.28	-0.0758	-0.0442	-0.2423	0.1570
a_{p4}	0.12	0.1515	0.1504	0.0165	0.2812
a_y	0.14	0.0607	0.1223	0.0212	0.2924
b_{y1}	1.16	0.9141	0.9199	0.7784	1.0561
b_{y2}	-0.25	-0.2121	-0.2217	-0.3563	-0.0928
b_r	0.18	0.0910	0.1164	0.0372	0.2243
ϕ_r	0.7	0.9646	0.9599	0.9372	0.9787
$\phi_\pi - 1$	1.0	0.8939	0.8854	0.6096	1.1755
ϕ_y	0.125	0.1011	0.1273	0.0566	0.2220
ϕ_{dy}	0.125	0.0506	0.0604	0.0292	0.0993
σ_p^2	0.25	0.4647	0.4732	0.4010	0.5584
σ_g^2	0.25	0.2930	0.3040	0.2575	0.3593
σ_r^2	0.25	0.1011	0.1008	0.0853	0.1192

Table A6: RS model prosterior (tight prior around low inflation and output persistence)

Parameter	Prior mean	Post. mode	Post. mean	5th %ile	95th %ile
a_{p1}	0.3	0.1616	0.1695	0.0435	0.3053
a_{p2}	0.0	-0.0152	-0.0235	-0.1590	0.1089
a_{p3}	0.0	-0.0606	-0.0647	-0.1986	0.0717
a_{p4}	0.0	0.0707	0.0792	-0.0539	0.2155
a_y	0.14	0.0607	0.1197	0.0205	0.2878
b_{y1}	0.7	0.5707	0.5731	0.4379	0.7060
b_{y2}	0.0	-0.0152	-0.0107	-0.1380	0.1189
b_r	0.18	0.0910	0.1153	0.0377	0.2181
ϕ_r	0.7	0.9646	0.9591	0.9363	0.9782
$\phi_\pi - 1$	1.0	0.8485	0.8842	0.6186	1.1913
ϕ_y	0.125	0.1061	0.1272	0.0556	0.2226
ϕ_{dy}	0.125	0.0556	0.0602	0.0289	0.0988
σ_p^2	0.25	0.4142	0.4237	0.3608	0.4981
σ_g^2	0.25	0.2627	0.2717	0.2316	0.3189
σ_r^2	0.25	0.1011	0.1008	0.0853	0.1192

B   A version of the Smets-Wouters (2007) model

B.1   The final goods sector

The final-goods sector is perfectly competitive and produces a final good Y_t by bundling together a continuum of intermediate goods $Y_t(z)$. Final-good producers

choose inputs and outputs to maximize profits,

$$\max_{Y_t,Y_t(i)} P_t Y_t - \int_0^1 P_t(z) Y_t(z)\, dz \quad \text{s.t.} \quad \left[\int_0^1 G\left(\frac{Y_t(z)}{Y_t}; \varepsilon_p\right) dz \right] = 1, \qquad (13)$$

where P_t and $P_t(z)$ are the price of the final and intermediate goods respectively, and G is a strictly concave and increasing function characterised by $G(1) = 1$. The variable ε_t^p is an exogenous shock that changes the elasticity of demand and therefore the markup. We assume that ε_t^p follows an $ARMA(1,1)$ process,

$$\ln \varepsilon_t^p = \rho_p \ln \varepsilon_{t-1}^p - \theta_p \eta_t^p + \eta_t^p, \ \ \eta_t^p \sim N(0, \sigma_p) \qquad (14)$$

B.2 Intermediate-goods sector

The intermediate goods sector is monopolistically competitive and features sticky prices. There is a continuum of intermediate-goods firms, indexed by z, with technology

$$Y_t(z) = \varepsilon_t^a (K_t^s(z))^\alpha (L_t(z))^{1-\alpha} - \Phi. \qquad (15)$$

The variable K_t^s represents capital services, L_t is labour input, Φ is a fixed cost, and ε_t^a is an exogenous shock to total factor productivity. The technology shock follows an $AR(1)$ process,

$$\ln \varepsilon_t^a = \rho_z \ln \varepsilon_{t-1}^a + \eta_t^a, \ \ \eta_t^a \sim N(0, \sigma_a). \qquad (16)$$

The firm's profit is given by

$$P_t(z) Y_t(z) - W_t L_t(z) - R_t^k K_t(z), \qquad (17)$$

where W_t is the aggregate nominal wage and R_t^k is the rental rate on capital.

Under Calvo pricing with partial indexation, a firm that is allowed to re-optimise its price solves

$$\max \sum_{k=0}^{\infty} E_t \xi_p^k \beta^k \frac{\Xi_{t+k}}{\Xi_t} \frac{P_t}{P_{t+k}} Y_{t+k} \left[\widetilde{P}_t(z)(\Pi_{l=1}^k \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}) - \mathcal{MC}_{t+k} \right] = 0, \qquad (18)$$

$$s.t. Y_{t+k}(z) = G'^{-1}\left(\frac{P_t(z)X_{t.k}}{P_{t+k}} \int_0^1 G'\left(\frac{Y_t(z)}{Y_t}\right) \frac{Y_t(z)}{Y_t} dz \right) Y_{t+k}, \qquad (19)$$

where $\widetilde{P}_t(z)$ is the newly set price, ζ_p is the Calvo probability of being allowed to re-optimise one's price, $\pi_t = P_t/P_{t-1}$ is gross inflation, $\beta^k \frac{\Xi_{t+k}}{\Xi_t} \frac{P_t}{P_{t+k}}$ is the firm's nominal stochastic discount factor (which equals the discount factor for households). Following Kimball (1995), $G(\cdot)$ is specified so that the demand for input $Y_t(z)$ is decreasing in its relative price $P_t(z)/P_t$, with the elasticity of demand being a positive function of its relative price. Finally, $X_{t.k} = \Pi_{l=1}^k \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}$, unless $k = 0$, in which case $X_{t.k} = 1$. The term $\Pi_{l=1}^k \pi_{t+l-1}^{\iota_p}$ captures the fact that prices of firms that do not receive a price signal are indexed to last period's inflation rate, and the term $\pi_*^{1-\iota_p}$ is an adjustment for trend inflation.

B.3 Households

Households are indexed by a and have identical preferences defined over the consumption of a composite good C and hours worked L,

$$E_t \sum_{i=0}^{\infty} \beta^i \left[\frac{(C_{t+i}(a) - \lambda C_{t+i-1}(a))^{1-\sigma_c}}{1 - \sigma_c} \right] \exp\left(\frac{(\sigma_c - 1) L_{t+i}(a)^{1+\sigma_l}}{1 + \sigma_l} \right). \tag{20}$$

The parameter $\beta \in (0,1)$ represents their subjective discount factor, σ_c is the inverse of the intertemporal elasticity of substitution, σ_l is the inverse elasticity of labor supply, and λ governs the degree of external habit formation.

A household's period-by-period budget constraint is given by

$$C_t(a) + I_t(a) + \frac{B_t(a)}{R_t P_t} - T_t(a) \tag{21}$$

$$\leq \frac{B_{t-1}(a)}{P_t} + \frac{W_t^h}{P_t} L_t(a) + \frac{R_t^k Z_t(a) K_{t-1}(a)}{R_t P_t} - a(Z_t(a) K_{t-1}(a)) + \frac{Div_t}{P_t},$$

where I_t represents gross investment, B_t is a nominally riskless discount bond paying gross interest R_t, and T_t is net lump sum taxes. The household earns a nominal wage W_t^h and collects financial income from its bond holdings, from renting capital to firms, and from collecting dividends distributed by the labour unions.

The capital-accumulation identity is

$$K_t(a) = (1 - \delta) K_{t-1}(a) + I_{t+i}(z) \left[1 - S\left(\frac{I_t(a)}{I_{t-1}(a)} \right) \right], \tag{22}$$

where δ is the depreciation rate and $S(.)$ is an adjustment-cost function, with with $S(1) = 0$; $S'(1) = 0$ and $S(.)'' > 0$.

B.4 Intermediate labour union sector

The supply side of the labor market involves three agents: households, unions, and labor packers. Households supply homogenous labour to a labour union which differentiates their labour services and sets wages following a Calvo mechanism. Unions sell differentiated labor to labor packers, who re-package labor services and sell them to intermediate-goods producers.

Working backwards, intermediate-goods producers employ a composite L_t of labor services,

$$L_t = \left[\int_0^1 L_t(l)^{\frac{1}{1-\lambda_{w,t}}} dl \right]^{1-\lambda_{w,t}}. \tag{23}$$

This composite is supplied by labor packers, who maximize profits in a perfectly competitive environment. Demand for variety $L_t(l)$ is therefore given by

$$L_t(l) = \left(\frac{W_t(l)}{W_t} \right)^{-\frac{1-\lambda_{w,t}}{\lambda_{w,t}}} L_t. \tag{24}$$

Labor packers buy variety $L_t(l)$ from labor unions. The unions allocate and differentiate labour services from the households and have market power. In their negotiations with labour packers, unions take the household's marginal rate of substitution between consumption and labour as the cost of labour services. The unions choose the wage subject to the labour demand equation and to nominal rigidities á la Calvo. Specifically, unions can readjust wages with probability $1 - \zeta_w$ in each period. For those that cannot adjust wages, $W_t(l)$ increases at the weighted average of the steady state inflation π_* and of last period's inflation π_{t-1}. For those that can adjust, the problem is to choose a wage $\widetilde{W}_t(l)$ that maximizes wage income in all states where the union is stuck with that wage:

$$\max \sum_{k=0}^{\infty} E_t \zeta_w^k \beta^k \frac{\Xi_{t+k}}{\Xi_t} \frac{P_t}{P_{t+k}} L_{t+k}(l) \left[W_{t+k}(l) - W_{t+k}^h \right] = 0, \tag{25}$$

$$\text{where } W_{t+k}(l) = \widetilde{W}_t(l)(\Pi_{l=1}^k \pi_{t+l-1}^{\iota_w} \pi_*^{1-\iota_w}), \tag{26}$$

The markup above the marginal disutility is distributed to the households in the form of a union dividend.

B.5 Government Policies

The government's nominal budget constraint is given by

$$P_t G_t + B_{t-1} = T_t + \frac{B_t}{R_t}, \tag{27}$$

where G_t is exogenous government spending. Government spending expressed relative to the steady-state-output path follows an $AR(1)$ process,

$$\ln g_t = \rho_g \ln g_{t-1} + \epsilon_t^g, \ \epsilon_t^g \sim N(0, \sigma_g). \tag{28}$$

The central bank follows a nominal interest rate rule,

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*} \right)^{\phi_R} \left[\left(\frac{\pi_t}{\pi_*} \right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*} \right)^{\phi_y} \right]^{1-\phi_R} \left(\frac{Y_t/Y_{t-1}}{Y_t^*/Y_{t-1}^*} \right)^{\phi_{dy}} m_t, \tag{29}$$

where R^* is the steady-state gross nominal interest rate and Y_t^* is the natural output. The parameter ϕ_R determines the degree of interest rate smoothing, and $\phi_\pi, \phi_y, \phi_{dy}$ are feedback coefficients on inflation, the output gap, and output growth, respectively. The monetary-policy shock m_t evolves exogenously according to

$$\ln m_t = \rho_m \ln m_{t-1} + \epsilon_{m,t}. \tag{30}$$

B.6 Priors for the SW model

Table 7 displays our prior distribution. Priors on consumer-preference parameters are centered on standard values and are relatively tight. The intertemporal subsitution elasticity has a mean of unity, and the mean degree of habit persistence is 0.7. The labor-supply elasticity is centered on 2, and the discount rate is calibrated at 0.9925.

Technology parameters are also centered on standard values. The capital share in intermediate-good production and the depreciation rate are calibrated at 0.36 and 0.025, respectively. Prior means for parameters governing the elasticities of capital utilisation and the investment-adjustment cost are the same as in Smets and Wouters (2007). The mode of the share of fixed costs in production is approximately 0.3.

For Calvo-pricing parameters, the probability of re-optimizing prices and wages is normal distributed with a prior mean of 0.75 and a prior standard deviation of 0.1. For the degree of price and wage indexation, we pick a diffuse distribution centred on 0.5.

With respect to shocks, priors for the persistence parameters reflect our belief that government-spending and productivity shocks are persistent while monetary-policy and cost-push shocks decay quickly. Persistence parameters for TFP and government-spending shocks have a mean of 0.7, while those for cost-push and monetary-policy shocks have a mean of 0.3. Priors for the standard deviations are standard. Finally, the prior for policy-rule parameters is the same as in the Rudebusch-Svensson (1999) model.

Table A7: Priors for the Smets-Wouters model

Parameter	Distribution	Mean	Standard Deviation
σ_a^2	Inverse Gamma	0.25	0.2
σ_g^2	Inverse Gamma	0.25	0.2
σ_m^2	Inverse Gamma	0.25	0.2
σ_p^2	Inverse Gamma	0.25	0.2
ρ_a	Beta	0.8	0.1
ρ_g	Beta	0.8	0.1
ρ_m	Beta	0.3	0.1
ρ_p	Beta	0.3	0.1
θ_p	Normal	0.1	0.05
φ	Normal	4.0	0.5
σ_c	Normal	1.0	0.1
λ	Beta	0.7	0.1
σ_l	Normal	2.0	0.2
ζ_p	Beta	0.75	0.1
ζ_w	Beta	0.75	0.1
ι_w	Beta	0.5	0.2
ι_p	Beta	0.5	0.2
θ_u	Beta	0.5	0.2
Φ	Gamma	0.4	0.2
$\phi_\pi - 1$	Gamma	1.0	0.2
ϕ_R	Beta	0.7	0.1
ϕ_y	Normal	0.125	0.05
ϕ_{dy}	Normal	0.125	0.05
λ_w	Calibrated	10	
λ_p	Calibrated	10	
β	Calibrated	0.995	
α	Calibrated	0.36	
δ	Calibrated	0.025	

B.7 Posterior for the SW model

Table 8 summarises the model's posterior distribution. Several features are immediately apparent. First of all, for many parameters, posteriors are not far from the priors. This shows that the large size of the SW model combined with our limited number of data series and short sample pose a number of identification problems. Thus, along several dimensions, parameter values are effectively set via the priors.

Table A8: Posterior for the Smets-Wouters model

Parameter	Prior mean	Post. Mode	Post. Mean	5th %ile	95th %ile
σ_a^2	0.25	0.2400	0.2793	0.1806	0.4215
σ_g^2	0.25	0.1773	0.1869	0.1319	0.2477
σ_m^2	0.25	0.0797	0.0845	0.0657	0.1081
σ_p^2	0.25	0.3655	0.3694	0.2995	0.4525
ρ_a	0.8	0.8478	0.8484	0.7709	0.9165
ρ_g	0.8	0.8377	0.8197	0.6544	0.9436
ρ_m	0.3	0.3135	0.3125	0.1691	0.4657
ρ_p	0.3	0.2228	0.2367	0.1164	0.3770
θ_p	0.1	0.1068	0.1170	0.0385	0.1971
φ	4.0	4.1818	4.1406	3.3510	4.9424
σ_c	1.0	1.1515	1.1715	0.9335	1.4361
λ	0.7	0.8023	0.7760	0.6407	0.8872
σ_l	2.0	1.9697	1.9806	1.6506	2.3063
ζ_p	0.75	0.7191	0.7040	0.5316	0.8491
ζ_w	0.75	0.7923	0.7714	0.6299	0.8900
ι_w	0.5	0.4345	0.4473	0.1436	0.7808
ι_p	0.5	0.1623	0.2291	0.0674	0.4431
θ_u	0.5	0.6865	0.6296	0.3284	0.8696
Φ	0.4	0.4654	0.5114	0.2169	0.8996
$\phi_\pi - 1$	1.0	0.8157	0.8719	0.5989	1.1890
ϕ_R	0.7	0.9566	0.9519	0.9271	0.9725
ϕ_y	0.125	0.1037	0.0859	0.0472	0.1766
ϕ_{dy}	0.125	0.2788	0.2906	0.1831	0.4104

For instance, similar to other studies, we find that the parameters governing the degree of nominal rigidity are not well identified. One notable exception is the degree of price indexation, which is estimated to be considerably lower compared to the prior. Investment-adjustment costs are also weakly identified, most likely because we do not include investment amongst our observations.

For several parameters, however, the data are informative. For instance, the estimated degree of price indexation is considerably lower compared to the prior. The variance of the monetary shock is smaller compared to its prior mean while that of the cost-push shock is larger. The inter-temporal elasticity of substitution is somewhat lower than the prior of unity, while the weight of habits is somewhat higher than the prior, with a mode close to 0.8. The data are also informative about the elasticity of capital utilisation costs, which is estimated to be relatively high. This implies that movements in capital utilisation will not be as pronounced as, for example, in models such as Christiano, Eichenbaum and Evans (2005) where this

elasticity is estimated to be extremely low. The mode of the share of fixed costs in production is higher than the prior, standing at just over 0.4.

The data are also informative about the monetary-policy rule. Notably, the degree of interest rate smoothing is very high despite the tight prior. The response to real variables seems to occur mainly in responding to the growth rate and less in responding to the level of output.

C Bernanke, Gertler and Gilchrist (1999)

The model includes five types of agents - households, entrepreneurs, financial intermediaries, final goods retailers, and the central bank.

C.1 The household's decision problem

The representative household maximises:

$$E_t \sum_{s=0}^{\infty} \beta^s \left[\nu \log c_{t+s} + (1 - \nu) \log (1 - h_{t+s}) \right], \tag{31}$$

subject to the flow budget constraint

$$c_t + b_{t+1} = R_{t-1} b_t + w_t h_t + \Gamma_t - T_t. \tag{32}$$

Aggregate consumption is a Dixit-Stiglitz aggregator of differentiated goods consumption $c_t(i)$,

$$c_t = \left[\int_0^1 c_t(i)^{\frac{\lambda_p - 1}{\lambda_p}} \, di \right]^{\frac{\lambda_p}{\lambda_p - 1}}, \tag{33}$$

h_t represents hours worked, b_t is a real bond which pays out R_t units of the composite consumption good in period $t + 1$, r_t is the rental rate of capital, w_t is the real wage rate, Γ_t are the profits of retailers, T_t is a lump-sum tax, m_t is nominal money holdings. The price of the composite consumption good is

$$p_t = \left[\int_0^1 p_t(i)^{1-\lambda_p} \, di \right]^{\frac{1}{1-\lambda_p}}.$$

The parameter β is the subjective discount factor, λ_p is the elasticity of substitution across varieties, and ν is the weight on consumption in the period utility function of the household.

C.2 The entrepreneurs' problem

Entrepreneurs are risk neutral and have finite lives. They supply labour services to final goods firms inelastically, but their main source of funds are investment

projects. Entrepreneurs are endowed with the technology to make capital goods from consumption goods. They maximise the following objective:

$$E_t \sum_{s=0}^{\infty} (\beta \gamma)^s c_{t+s}^e, \tag{34}$$

subject to the following sequence of budget constraints:

$$c_t^e + q_t k_{t+1} + w_t h_t^h - b_{t+1} = \max \left[\omega_t \left(\frac{y_t}{X_t} + q_t (1-\delta) k_t \right) - R_{t-1} b_t, 0 \right], \tag{35}$$

where c_t^e is entrepreneurial consumption, h_t^h is the employment of household labour, q_t is the price of capital in terms of final goods, and δ is the rate of capital depreciation. Finally, in order to motivate default and agency problems in the model, BGG assume that total revenue is subject to an idiosyncratic i.i.d. shock ω_t which has a mean of unity and a variance of σ_ω^2. Total revenue consists of the value of capital after depreciation $q_t (1-\delta) k_t$ plus the value of entrepreneurs' intermediate goods' output in terms of the final good y_t/X_t, where X_t is the mark-up of retailers over marginal cost.

The entrepreneur has limited liability and can default on his/her debt if the total revenue from his/her project falls short of the value of debt. Furthermore, a fraction γ of entrepreneurs die in every period, which explains the different discount factor of entpreneurs relative to workers.

The technology for the production of intermediate goods is Cobb-Douglas in capital k_t, household labour h_t^h and entrepreneurial labour h_t^e,

$$y_t = \varepsilon_t^a k_t^\alpha \left((h_t^h)^\Omega (h_t^e)^{1-\Omega} \right)^{1-\alpha}. \tag{36}$$

Output is subject to the common productivity shock A_t, α is the share of capital in national income, $\Omega (1-\alpha)$ is the share of household labour, while $(1-\Omega)(1-\alpha)$ is the income share of entrepreneurial labour.

Aggregate capital accumulates with investment net of capital adjustment costs,

$$K_{t+1} = \left[\left(\frac{I_t}{K_t} \right)^{1+\varphi} + 1 - \delta \right] K_t, \tag{37}$$

where φ is the elasticity of capital-adjustment costs with respect to the investment rate. Total factor productivity follows the following process:

$$\ln \varepsilon_t^a = \rho_A \ln \varepsilon_t^a + \eta_t^A. \tag{38}$$

C.3 The problem of the financial intermediary

Under the assumption of no aggregate uncertainty, the risk-neutral perfectly-competitive financial intermediary accepts riskless deposits from households and lends

them to entrepreneurs. Deposits are riskless because the idiosyncratic productivity shock is iid across entrepreneurs and therefore the default loss is perfectly predictable in the aggregate. The intermediary, therefore, expects a return equal to the risk-free rate on each individual contracts it enters into.

Financial contracting takes place in the 'costly state verification' environment described by Townsend (1979). Only the entrepreneur can costlessly find out the revenue from the project, $\omega_t \left(y_t + q_t \left(1 - \delta \right) k_t \right)$. Outsiders such as the financial intermediary can only verify the project output by paying a cost which is a proportion μ of total output. This cost has the interpretation of a bankruptcy cost because, in equilibrium, it is only paid when the entrepreneur declares bankruptcy.

Bernanke, Gertler and Gilchrist (1999) show that the profit-maximisation problem of the financial intermediary can be more conveniently represented as a maximisation of the utility of the entrepeneur subject to a break-even constraint for the intermediary:

$$\max_{b_t, R_t^b} E_t \max \left[\omega_t R_{t+1}^k q_{t+1} \left(n_t + b_t \right) - R_t^b b_t, 0 \right], \tag{39}$$

where $R_{t+1}^k = (r_{t+1} + (1 - \delta) q_{t+1})/q_t$ is the return to holding capital and $k_t = n_t + b_t$ equals total capital purchases by the entrepreneur. The inermediary chooses the debt level b_t and the debt interest rate R_t^b as a function of entrepreneurial net worth n_t in order to maximise the utility of the borrower, which is equal to the expected project revenue net of debt repayments, taking into account the option to default. The break-even constraint is given by:

$$E_t \min \left[R_t^b b_t, \left(1 - \mu \right) \omega_t R_{t+1}^k q_{t+1} \left(n_t + b_t \right) \right] = R_t b_t. \tag{40}$$

C.4 The problem of the retailer

To motivate price stickiness, BGG assume that perfectly-competitive entrepreneurs sell their output to monopolistically-competitive retailers who costlessly differentiate it and sell it to households at a mark up. Retailers set prices according to a Calvo pricing model with backward-looking indexation. With probability $1 - \zeta_p$ a retailer is free to re-optimise his or her price in any given period. With probability ζ_p, it cannot re-optimise but can index its price to a weighted average of last period's inflation rate and the steady state inflation. A retailer who is able to re-optimise its price will choose its new price is $p_t^* (i)$ to maximise

$$\max_{p_t^* (i)} E_t \sum_{s=0}^{\infty} \zeta_p^s \Lambda_{t,t+s} \left[\frac{p_t^* (i) \left(\Pi_{t,t+s} \right)^{\iota_p} \left(\Pi_{t,t+s}^* \right)^{1-\iota_p}}{p_{t+s}} - \frac{1}{X_{t+s}} \right] y_{t+s} (i), \tag{41}$$

where $\Lambda_{t,t+s}$ is the household's stochastic discount factor, $\Pi_{t,t+s}$ is cumulative inflation between t and $t + s$, and ι_p is degree to which retailers who are unable to re-optimise price get to index their price in line with past inflation.

43

C.5 Government policies

The monetary authority sets the nominal interest rate according to the following Taylor-type rule

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\phi_R} \left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{\phi_y}\right]^{1-\phi_R} \left(\frac{Y_t/Y_{t-1}}{Y_t^*/Y_{t-1}^*}\right)^{\phi_{dy}} m_t, \tag{42}$$

where

$$\ln m_t = \rho_m \ln m_{t-1} + \epsilon_{m,t}, \epsilon_{m,t} \sim N\left(0, \sigma_m^2\right) \tag{43}$$

is an exogenous monetary-policy shock and Y_t^* denotes the level of output under flexible prices and wages. The fiscal authority runs a balanced budget in every period, using seigniorage and lump sum tax revenues (levied on the household) to fund its expenditure,

$$G_t = \frac{M_t - M_{t-1}}{P_t} + T_t. \tag{44}$$

Government expenditures are exogenous and evolve as

$$\ln G_t = \rho_g \ln G_{t-1} + \epsilon_t^g, \epsilon_{g,t} \sim N\left(0, \sigma_g^2\right) \tag{45}$$

C.6 Market clearing

The goods market clears when

$$Y_t = C_t^h + C_t^e + I_t + DC_t + G_t, \tag{46}$$

where C_t^h and C_t^e are, respectively, the aggregate consumption levels of households and entrepreneurs, I_t is aggregate investment, and DC_t is the total verification cost paid by the financial intermediary to audit bankrupt entrepreneurs.

C.7 Priors

The priors on the BGG model (at least for those parameters that overlap) do not differ much from those of the SW model described in the previous subsection. Those are described in Table 9. Since we do not use data on investment or private interest rates in our estimation procedure, identifying the parameters that govern the financial contracting problems is problematic. Therefore, we calibrate those to the values chosen by BGG. The variance of the idiosyncratic productivity shock is calibrated at 0.28, the costs of bankruptcy (or monitoring costs in the costly-state-verification framework of the paper) are set at 0.12 of firm output. The share of capital in output is set at 0.36 and the depreciation rate at 0.025 per quarter. The BGG model assumes log-utility in both consumption and leisure. The weight on

leisure in period utility is calibrated to ensure that individuals work approximately one third of their total time endowment.

Table A9: Priors for the BGG

Parameter	Distribution	Mean	Standard Deviation
ν	Gamma	0.5	0.2
ζ_p	Beta	0.7	0.1
ι_p	Beta	0.5	0.2
ϕ_R	Beta	0.7	0.1
$\phi_\pi - 1$	Gamma	1.0	0.2
ϕ_y	Normal	0.125	0.05
ϕ_{dy}	Normal	0.125	0.05
ρ_a	Beta	0.8	0.1
ρ_g	Beta	0.8	0.1
ρ_m	Beta	0.3	0.1
ρ_p	Beta	0.3	0.1
σ_a^2	Inverse Gamma	0.25	0.2
σ_g^2	Inverse Gamma	0.25	0.2
σ_m^2	Inverse Gamma	0.25	0.2
σ_p^2	Inverse Gamma	0.25	0.2
λ_p	Calibrated	10	
β	Calibrated	0.995	
α	Calibrated	0.36	
δ	Calibrated	0.025	
μ	Calibrated	0.12	
σ_ω^2	Calibrated	25	
Ω	Calibrated	0.99	

C.8   Posteriors

Estimation results are displayed in Table 10. In many respects, estimates for the BGG model agree with those for the SW model. Monetary shocks are relatively small and cost-push and demand shocks relatively large. The perisistence of TFP and government spending shocks are accurately estimated as very high. The results differ, however, in one important respect. The degree of price stickiness is considerably lower than in the SW model, implying that prices are re-optimized once every 1.5 quarters. Because our estimates imply little nominal rigidity, inflation is volatile but not persistent.

Table A10: Posterior for the BGG

Parameter	Prior Mean	Post. Mode	Post. Mean	5th %ile	95th %ile
μ	0.5	0.6660	0.7427	0.4266	1.1335
ζ_p	0.75	0.3542	0.3555	0.2723	0.4388
ι_p	0.5	0.3243	0.3396	0.1078	0.6397
ϕ_R	0.7	0.6667	0.6580	0.5288	0.7643
ϕ_π - 1	1	0.9755	1.0274	0.7334	1.3681
ϕ_y	0.125	0.1068	0.1242	0.0552	0.2143
ϕ_{dy}	0.125	0.0925	0.1075	0.0481	0.1856
ρ_a	0.8	0.8932	0.8840	0.8339	0.9264
ρ_g	0.8	0.8932	0.8840	0.8339	0.9264
ρ_m	0.3	0.2741	0.2767	0.2303	0.3316
ρ_p	0.3	0.2938	0.3151	0.1586	0.4952
σ_a^2	0.25	0.2741	0.2767	0.2303	0.3316
σ_g^2	0.25	0.6392	0.6604	0.4886	0.8732
σ_m^2	0.25	0.1465	0.2025	0.1050	0.3748
σ_p^2	0.25	0.3077	0.3242	0.2228	0.4535

D   A small-open-economy model a la Gali and Mona-celli (2005)

In this section we consider a small open economy framework, which follows closely the specification Gali and Monacelli (2005) (GM hereafter) and De Paoli (2009). A small open economy is characterised as a limiting case of a two-country dynamic general equilibrium model,[12] and monopolistic competition and sticky prices are introduced in order to address issues of monetary policy. In particular, the model assumes that home price setting follows a Calvo-type contract and features complete pass-through, as producers set prices set in their own currency. In addition, the law of one price holds, but deviations from purchasing power parity arise because of home bias in consumption. Finally, domestic and foreign agents optimally share risk.

D.1   Preferences

We consider two countries, H (Home) and F (Foreign). The world economy is populated with a continuum of agents of unit mass, where the population in the segment $[0, n)$ belongs to country H and the population in the segment $(n, 1]$ belongs

---

[12]Gali and Monacelli (2005) assume that the world is populated by a continuum of small open economies, but the final equilibrium conditions for the two representations are identical.

to country F. The utility function of a consumer j in country H is given by

$$U_t = E_t \sum_{i=0}^{\infty} \beta^i \left[U(\varepsilon_{t+i}^a, C_{t+i}(a)) - V(\varepsilon_{t+i}^a, y_{t+i}(a)) \right]$$

with $U(C_{t+i}(a)) = \dfrac{(C_{t+i}(a)/\varepsilon_{t+i}^a)^{1-\sigma_c}}{1-\sigma_c}$ and $V(\varepsilon_{t+i}^a, y_{t+i}(a)) = \dfrac{(y_{t+i}(a)/\varepsilon_{t+i}^a)^{1+\sigma_l}}{1+\sigma_l}$ (47)

Households obtain utility from consumption $C_{t+i}(a)$ and disutility from producing a differenciated domestic goods $y_{t+i}(a)$. The parameter $\beta \in (0,1)$ represents their subjective discount factor, σ_c is the inverse of the intertemporal elasticity of substitution, σ_l is the inverse elasticity of labor supply and productivity shocks are denoted by ε_s^a. C is a C.E.S. (constant elasticity of substitution) aggregate of home and foreign goods, defined by

$$C = \left[v^{\frac{1}{\theta}} C_H^{\frac{\theta-1}{\theta}} + (1-v)^{\frac{1}{\theta}} C_F^{\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}. \tag{48}$$

The parameter $\theta > 0$ is the intratemporal elasticity of substitution between home and foreign-produced goods, C_H and C_F. As in Sutherland (2005), the parameter determining home consumers' preferences for foreign goods, $(1-v)$, is a function of the relative size of the foreign economy, $(1-n)$, and of the degree of openness, α; more specifically, $(1-v) = (1-n)\alpha$.

Similar preferences are specified for the rest of the world

$$C^* = \left[v^{*\frac{1}{\theta}} C_H^{*\frac{\theta-1}{\theta}} + (1-v^*)^{\frac{1}{\theta}} C_F^{*\frac{\theta-1}{\theta}} \right]^{\frac{\theta}{\theta-1}}, \tag{49}$$

with $v^* = n\alpha$. That is, foreign consumers' preferences for home goods depend on the relative size of the home economy and the degree of openness. Note that the specification of v and v^* generates a home bias in consumption.

The sub-indices C_H (C_H^*) and C_F (C_F^*) are Home (Foreign) consumption of the differentiated products produced in countries H and F. These are defined as follows

$$C_H = \left[\left(\frac{1}{n}\right)^{\frac{1}{\lambda_p}} \int_0^n c(z)^{\frac{\lambda_p-1}{\lambda_p}} dz \right]^{\frac{\lambda_p}{\lambda_p-1}}, \qquad C_F = \left[\left(\frac{1}{1-n}\right)^{\frac{1}{\lambda_p}} \int_n^1 c(z)^{\frac{\lambda_p-1}{\lambda_p}} dz \right]^{\frac{\lambda_p}{\lambda_p-1}},$$
$$\tag{50}$$

$$C_H^* = \left[\left(\frac{1}{n}\right)^{\frac{1}{\lambda_p}} \int_0^n c^*(z)^{\frac{\lambda_p-1}{\lambda_p}} dz \right]^{\frac{\lambda_p}{\lambda_p-1}}, \qquad C_F^* = \left[\left(\frac{1}{1-n}\right)^{\frac{1}{\lambda_p}} \int_n^1 c^*(z)^{\frac{\lambda_p-1}{\lambda_p}} dz \right]^{\frac{\lambda_p}{\lambda_p-1}},$$
$$\tag{51}$$

where $\lambda_p > 1$ is the elasticity of substitution across the differentiated products. The consumption-based price indices that correspond to the above specifications of preferences are given by

$$P = \left[v P_H^{1-\theta} + (1-v)(P_F)^{1-\theta} \right]^{\frac{1}{1-\theta}}, \tag{52}$$

47

and

$$P^* = \left[v^* P_H^{*1-\theta} + (1 - v^*) \left(P_F^* \right)^{1-\theta} \right]^{\frac{1}{1-\theta}}, \tag{53}$$

where P_H (P_H^*) is the price sub-index for home-produced goods expressed in the domestic (foreign) currency and P_F (P_F^*) is the price sub-index for foreign produced goods expressed in the domestic (foreign) currency:

$$P_H = \left[\left(\frac{1}{n} \right) \int_0^n p\,(z)^{1-\lambda_p}\,dz \right]^{\frac{1}{1-\lambda_p}}, P_F = \left[\left(\frac{1}{1-n} \right) \int_n^1 p\,(z)^{1-\lambda_p}\,dz \right]^{\frac{1}{1-\lambda_p}}, \tag{54}$$

$$P_H^* = \left[\left(\frac{1}{n} \right) \int_0^n p^*\,(z)^{1-\lambda_p}\,dz \right]^{\frac{1}{1-\lambda_p}}, P_F^* = \left[\left(\frac{1}{1-n} \right) \int_n^1 p^*\,(z)^{1-\lambda_p}\,dz \right]^{\frac{1}{1-\lambda_p}}. \tag{55}$$

We assume that the law of one price holds, so

$$p(h) = Sp^*(h) \text{ and } p(f) = Sp^*(f), \tag{56}$$

where the nominal exchange rate, S_t, denotes the price of foreign currency in terms of domestic currency. Equations (52) and (53), together with condition (56), imply that $P_H = SP_H^*$ and $P_F = SP_F^*$. However, as Equations (54) and (55) illustrate, the home bias specification leads to deviations from purchasing power parity; that is, $P \neq SP^*$. For this reason, we define the real exchange rate as $Q \equiv \frac{SP^*}{P}$.

From consumers' preferences, we can derive the total demand for a generic good h, produced in country H, and the demand for a good f, produced in country F

$$y_t^d(h) = \left[\frac{p_t(h)}{P_{H,t}} \right]^{-\lambda_p} \left[\frac{P_{H,t}}{P_t} \right]^{-\theta} \left[vC_t + \frac{v^*(1-n)}{n} \left(\frac{1}{Q_t} \right)^{-\theta} C_t^* \right], \tag{57}$$

$$y_t^d(f) = \left[\frac{p_t(f)}{P_{F,t}} \right]^{-\lambda_p} \left[\frac{P_{F,t}}{P_t} \right]^{-\theta} \left[\frac{(1-v)n}{1-n} C_t + (1 - v^*) \left(\frac{1}{Q_t} \right)^{-\theta} C_t^* \right]. \tag{58}$$

Finally, to portray a small open economy, we use the definition of v and v^* and take the limit for $n \to 0$. Consequently, conditions (57) and (58) can be rewritten as

$$y^d(h) = \left[\frac{p_t(h)}{P_{H,t}} \right]^{-\lambda_p} \left[\frac{P_{H,t}}{P_t} \right]^{-\theta} \left[(1 - \alpha)C_t + \alpha \left(\frac{1}{Q_t} \right)^{-\theta} C_t^* \right], \tag{59}$$

$$y^d(f) = \left[\frac{p_t^*(f)}{P_{F,t}^*} \right]^{-\lambda_p} \left[\frac{P_{F,t}^*}{P_t^*} \right]^{-\theta} C_t^*. \tag{60}$$

Equations (59) and (60) show that external changes in consumption affect demand in the small open economy, but the opposite is not true. Moreover, movements in the real exchange rate do not affect the rest of the world's demand.

D.2 Price-setting Mechanism

Prices follow a Calvo-style partial adjustment rule. Producers of differentiated goods know the form of their individual demand functions (given by Equations (59) and (60)), and maximize profits taking overall market prices and products as given. In each period a fraction, $\zeta_p \in [0,1)$, of randomly chosen producers is not allowed to change the nominal price of the goods they produce. The remaining fraction of firms, given by $(1 - \zeta_p)$, chooses prices optimally by maximizing the expected discounted value of profits. The optimal choice of producers that can set their price $\tilde{p}_t(j)$ at time T is, therefore

$$E_t\left\{\sum(\zeta_p\beta)^{T-t}U_c(C_T)\left(\frac{\tilde{p}_t(j)}{P_{H,T}}\right)^{-\lambda_p}Y_{H,T}\left[\frac{\tilde{p}_t(j)}{P_{H,T}}\frac{P_{H,T}}{P_T} - \frac{\lambda_p V_y\left(\tilde{y}_{t,T}(j),\varepsilon_t^a\right)}{(\lambda_p - 1)U_c(C_T)}\right]\right\} = 0.$$
(61)

Given the Calvo-type setup, the price index evolves according to the following law of motion,

$$(P_{H,t})^{1-\sigma} = \zeta_p P_{H,t-1}^{1-\sigma} + \left(1 - \zeta_p\right)\left(\tilde{p}_t(h)\right)^{1-\sigma}.$$
(62)

The rest of the world has an analogous price setting mechanism.

D.3 Complete Markets

Agents have access to state-contingent claims that allow them optimally to share risk with the rest of the world. Following Chari *et al* (2002), this asset market structure implies the following risk-sharing condition,

$$\frac{U_C\left(C_t^*\right)}{U_C\left(C_t\right)} = \frac{S_t P_t^*}{P_t}.$$
(63)

D.4 Government policies

The monetary authority sets the nominal interest rate according to the following Taylor-type rule

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*}\right)^{\phi_R}\left[\left(\frac{\pi_t}{\pi_*}\right)^{\phi_\pi}\left(\frac{Y_t}{Y_t^*}\right)^{\phi_y}\right]^{1-\phi_R}\left(\frac{Y_t/Y_{t-1}}{Y_t^*/Y_{t-1}^*}\right)^{\phi_{dy}}m_t,$$
(64)

where

$$\ln m_t = \rho_m \ln m_{t-1} + \epsilon_{m,t}, \epsilon_{m,t} \sim N\left(0,\sigma_m^2\right)$$
(65)

is an exogenous monetary-policy shock and Y_t^* denotes the level of output under flexible prices.

D.5 Estimation

Following Lubik and Schorfheide (2007) (LS hereafter), we estimate a simplified version of GM in which $\sigma_l = 0$ and $\theta = 1$. The system of equilibrium conditions is estimated with variables measured in percentage deviations from a balanced growth path, induced by the technology process ε_t^a,

$$\ln \varepsilon_t^a = \ln \varepsilon_{t-1}^a + z_t^a, \ \ \ln z_t^a = \rho_z \ln z_{t-1}^a + \eta_t^a, \ \eta_t^a \sim N(0, \sigma_a). \tag{66}$$

The estimated system can be summarised by a Phillips curve (PC), a forward looking IS equation (IS) and a risk sharing equation (RS) and the policy rule (PR):

$$\pi_t + \alpha \Delta s_t = k/\alpha'(y_t - \bar{y}_t) + \beta(E_t \pi_{t+1} + \alpha E_t \Delta s_{t+1}), \tag{PC}$$

$$y_t = E_t y_{t+1} - \alpha'(R_t - E_t \pi_{t+1} - \alpha E_t \Delta s_{t+1}) + E_t \Delta \bar{y}_{t+1} - E_t z_{t+1}^a, \tag{IS}$$

$$\Delta y_t = \Delta y_t^* + \alpha' \Delta s_t, \tag{RS}$$

$$R_t = \phi_R R_{t-1} + (1 - \phi_R) \left[(1 + \tilde{\phi}_\pi)\pi_t + \phi_y(y_t - \bar{y}_t) \right] + \phi_{dy}(\Delta y_t - \Delta \bar{y}_t) + \varepsilon_t^R. \tag{PR}$$

The variable y denotes domestic output, s represents the terms of trade (note that $(1 - \alpha)s = -q$), $\bar{y} = -\alpha(2 - \alpha)(1 - \tau)/\tau y^*$ is potential output, R is the nominal interest rate, π represents CPI inflation, and output in the rest of the world follows

$$y_t^* = \rho^* y_{t-1}^* + \eta_t^*, \ \eta_t^* \sim N(0, \sigma_{y^*}). \tag{67}$$

Moreover, we define $\phi_\pi = 1 + \tilde{\phi}_\pi$, $\alpha' = \sigma_c^{-1} + \alpha(2 - \alpha)(1 - \sigma_c^{-1})$, and $k = (1 - \zeta_p \beta)(1 - \zeta_p)/\zeta_p$.

D.5.1 Priors

The choice of prior mean, distribution and standard deviation for the remaining parameters follows LS and are presented in Table 11. The mean of the intertemporal elasticity of substitution (σ_c^{-1}) is set to 0.5 with standard deviation of 0.2. The assumption implies an average coefficient of risk aversion higher than unity. Given that the elasticity of intratemporal substitution is set to unity $(\theta = 1)$, we have that $\sigma_c \theta > 1$. Thus, under the prior mean, domestic and foreign goods are substitutes in utility. The prior for the slope of the Phillips curve (k) is centered at 0.5 and has a standard deviation of 0.25. The policy-rule parameters $\tilde{\phi}_\pi$, ϕ_y tand ϕ_{dy} are centered

at 0.54, 0.25 and 0.25 respectively and are assumed to follow a gamma distribution. In addition, the prior mean of β is set to yield an annual interest rate of 2.51 percent, and the degree of openness is centered at 0.2.

As chosen by LS when estimating the model on UK data, the standard deviations of productivity and external shocks (σ_a, σ_{y^*}) are centered at 1.5 with a standard deviation of 4, but the mean of the standard deviation of monetary policy shocks (σ_m), is set at 0.5. These follow an inverted gamma distribution. The persistence of productivity and foreign shocks (ρ_z, ρ_{y^*}) is centered at 0.2 and 0.9, respectively, while the persistence of the interest rate (ϕ_R) has mean 0.5. Persistence parameters are assumed to follow a beta distribution.

Table A11: Priors for the Gali-Monacelli model

Parameter	Distribution	Mean	Standard Deviation
ϕ_R	Beta	0.5	0.2
$\tilde{\phi}_\pi = \phi_\pi - 1$	Gamma	0.54	0.5
ϕ_y	Gamma	0.25	0.13
ϕ_{dy}	Gamma	0.25	0.13
ρ_z	Beta	0.2	0.1
ρ_{y^*}	Beta	0.9	0.05
σ_a	Inverse Gamma	1.5	4
σ_{y^*}	Inverse Gamma	1.5	4
σ_m^2	Inverse Gamma	0.5	4
k^*	Gamma	0.5	0.25
R	Gamma	2.51	1
α	Beta	0.2	0.05
σ_c^{-1}	Gamma	0.5	0.2
σ_l	Calibrated	0	
θ	Calibrated	1	
Note:	$^*k = (1 - \zeta_p \beta)(1 - \zeta_p)/\zeta_p$		

D.5.2 Posteriors

Estimation results are shown in Table 12. The results present a tight posterior distribution for the persistence in the policy rule (ϕ_R), with the posterior mode at 0.76. This estimate is between the levels found in the SW and BGG models. The posterior mode for the coefficient of inflation in the policy rule (ϕ_π) is 1.07, which is also similar to the one obtained in SW and BGG, though the posterior distribution in GM present a larger positive skew. The coefficient on output gap and output growth in the policy rule $(\phi_y$ and $\phi_{dy})$ are close to their specified priors, with the posterior mode of 0.27 and 0.24 respectively. The posterior distribution for the external shock persistence and standard deviation are tight and the posterior modes do not depart

significantly from the prior mean. On the other hand the posterior mode for the standard deviation for the other shocks is much smaller than the assumed in the prior distribution.

The posterior mode for the slope of the Phillips curve suggest a degree of price stickiness below the one found in the SW model but above the one found in the BGG model. The estimates for the rates of return suggest a subjective discount factor similar to the one calibrated in the previous models (that is, the posterior mode for R is consistent with a β equal to 0.99). Turning to the parameters with a direct international dimension, the posterior distribution for the degree of openness is concentrated near the mode. And the estimated mode is around 0.34, which implies an import share slightly larger than the one usually considered for the UK. Finally the posterior mode for the elasticity of intertemporal substitution is estimated at 0.67, which implies a coefficient of risk aversion of around 1.5, and suggests that UK imports tend to be substitutes to domestically produced goods.

Table A12: Gali-Monacelli model posterior

Parameter	Prior mean	Post. Mode	Post. Mean	5th %ile	95th %ile
ϕ_R	0.5	0.7618	0.7434	0.6368	0.8586
$\tilde{\phi}_\pi = \phi_\pi - 1$	0.54	1.0762	1.3356	0.4728	2.1931
ϕ_y	0.25	0.2662	0.3312	0.0846	0.5723
ϕ_{dy}	0.25	0.2390	0.2902	0.1065	0.4695
ρ_z	0.2	0.1132	0.1059	0.0355	0.1746
ρ_{y^*}	0.9	0.9423	0.9382	0.9152	0.9612
σ_a	1.5	0.4262	0.4333	0.3586	0.5077
σ_{y^*}	1.5	1.1509	1.6718	1.105	2.2376
σ_m	0.5	0.2182	0.2562	0.1719	0.339
k	0.5	0.4324	0.6143	0.1684	1.0398
R	2.51	2.0992	2.5012	0.9385	4.0312
α	0.2	0.3368	0.3376	0.2536	0.4219
σ_c^{-1}	0.5	0.6716	0.7080	0.5953	08230