Journal of Economic Dynamics & Control 176 (2025) 105096

Contents lists available at ScienceDirect

Journal of Economic Dynamics and Control

journal homepage: www.elsevier.com/locate/jedc

All models are wrong but all can be useful: Robust policy design using prediction pools ☆

Szabolcs Deák [a,*], Paul Levine [b], Afrasiab Mirza [c], Joseph Pearlman [d]

[a] *University of Exeter, UK*
[b] *University of Surrey, UK*
[c] *University of Birmingham, UK*
[d] *City University London, UK*

ARTICLE INFO

JEL classification:
D52
D53
E44
G18
G23

Keywords:
Bayesian estimation
DSGE models
Financial frictions
Forecasting
Prediction pools
Optimal simple rules

ABSTRACT

We study the design of monetary policy rules robust to model uncertainty using a novel methodology. In our application, policymakers choose the optimal rule by attaching weights to a set of well-established DSGE models with varied financial frictions. The novelty of our methodology is to compute each model's weight based on their relative forecasting performance. Our results highlight the superiority of predictive pools over Bayesian model averaging and the need to combine models when none can be deemed as the true data generating process. In addition, we find that the optimal across-model robust policy rule exhibits attenuation, and nests a price level rule which has good robustness properties. Therefore, the application of our methodology offers a new rationale for price-level rules, namely the presence of uncertainty over the nature of financial frictions.

1. Introduction

"All models are wrong, but some are useful."
— George Box

We propose a general methodology to address the problem of designing simple policy rules when all models are wrong yet, in contrast to this famous quote, *all models can be useful*. We consider an environment with three forms of uncertainty. The first is standard and derives from uncertain future shocks; the second is parameter uncertainty within each competing model, which we refer to as "within-model uncertainty"; the third source of uncertainty is the existence of multiple competing models, referred to as "across-model uncertainty."

The novelty of our methodology lies in the way we handle this third form of uncertainty in the design of simple policy rules. Specifically, we follow the procedure of Geweke and Amisano (2012) to form prediction pools where weights are assigned to models

☆ We acknowledge financial support from the ESRC, grant reference ES/K005154/1 and from the EU Framework Programme 7, grant agreement number 612796.
* Corresponding author at: Department of Economics, University of Exeter, Exeter, UK.
E-mail addresses: s.deak@exeter.ac.uk (Sz. Deák), p.levine@surrey.ac.uk (P. Levine), a.mirza@bham.ac.uk (A. Mirza), Joseph.Pearlman.1@city.ac.uk (J. Pearlman).

https://doi.org/10.1016/j.jedc.2025.105096
Received 22 November 2023; Received in revised form 27 March 2025; Accepted 1 April 2025

Available online 3 April 2025

on the basis of their forecasting accuracy, rather than in-sample data fit as in the common alternative, Bayesian Model Averaging (BMA). These weights are then used to solve for the simple policy rule that is robust to all three forms of uncertainty. Unlike BMA which assumes that one of the models is the true data generating process, prediction pools allow us to consider that all models among a comparative set are misspecified, but they all may be useful at different periods of time.

We apply our methodology to the design of monetary policy rules that are robust to different beliefs about the nature or the importance of financial frictions in the economy. These beliefs are captured by three very different medium-scale new Keynesian DSGE models: a benchmark model without financial frictions and two different financial frictions models. The Smets and Wouters (2007) model (henceforth, SW) is the workhorse model in policy-making institutions for forecasting and policy analysis. The other two models build financial frictions into the SW model along the lines of Gertler et al. (2012) (henceforth, GKQ) and Bernanke et al. (1999) (henceforth, BGG), respectively. These two models represent leading theories in modelling financial frictions in the macroeconomic literature. Hence, our model pool can be motivated by considering a policymaker who is uncertain how to incorporate financial frictions into a DSGE model, or if they should be incorporated at all. Importantly, while potentially useful, none of these models are believed by the policymaker to be the true data generating process.

A key contribution of our paper is that we compare the relative predictive performance of these three DSGE models.[1] While we find that the models that emphasizes frictions between households and financial intermediaries (GKQ) outperform the alternative where frictions are between banks and borrowing firms (BGG) on average, BGG outperforms GKQ during financial crises. Moreover, the SW model without financial friction (SW) performs well relative to these two models on average over the entire horizon even though the latter deliver significantly better predictions during recessions. These results highlight the need to combine models when none can be deemed as the true data generating process as each may have particular strengths at different points in time. As such, current methods like Bayesian Model Averaging (BMA) are generally inadequate for this task. BMA tends to assign all weight to a single model as implicitly it assumes that one of the models under consideration is in fact the true data generating process.

In our novel methodology, we use past predictive performance to determine the weights a policy maker should attach to the different models when designing robust monetary policy rules. An important finding is that using the optimal prediction pool weights as in Geweke and Amisano (2012) leads to a more persistent policy rule that also responds more aggressively to changes in inflation and output growth compared to the policy rules obtained using BMA. In addition, we find that both within-model and across-model robustness results in more inertial and more aggressive rules than the non-robust single model counterparts. Finally, we show that all of the above results continue to hold when the policy maker only employs predictions of inflation and output growth to construct the optimal pooling weights.

We restrict our attention to optimized simple rules, that is to find the welfare-optimal parameter values in a Taylor-type monetary policy rule. There are several reasons for this. First, simple rules are transparent and easy to implement and to communicate. Second, a large literature has arrived at a consensus that they mimic the Ramsey-optimal policy (see, for example, Schmitt-Grohe and Uribe (2007)). Finally, the same literature suggests that they already contain good robustness properties compared with Ramsey-optimal policy.

The final yet crucial contribution of the paper is to show that the optimal robust policy rule is very nearly a price-level rule. Thus we show that indeed such a rule has excellent robustness properties and therefore should be part of a policymakers toolkit when there is uncertainty of the nature of financial frictions in the economy.

Overall, our methodology is suited to a policymaker who does not believe any one model to be correct. She would then be advised to use a forward-looking procedure such as prediction pools to weight models rather than a backward-looking one such as Bayesian model averaging (henceforth BMA) which is based solely on in-sample model fit. Since the former is less likely to see one model completely dominate, she is then more cautious in rejecting models that appear to be inferior. The intuition is that such models, while not particularly accurate most of the time or at all horizons or for all variables, may well be very useful some of the time or at some horizons or for predicting key policy variables. Optimal policy should take this into consideration rather than rejecting the use of models that perform poorly simply from an in-sample likelihood perspective. Moreover, we find that prediction pooling is a far more robust method for attaching weight to models as these are less sensitive to outliers and therefore evolve more smoothly over time than corresponding weights from BMA.

The rest of the paper is organized as follows. Section 2 locates our contribution in the extensive literature on model selection and robust policy design. Section 3 outlines the weighting strategy based on model forecasting performance and how the model weights and the posterior distribution are combined to design policy that is robust with respect to both within-model and across-model uncertainty. Section 4 describes the models in our application and summarizes the characteristics of the posterior distribution that are most important for optimal policy design. Section 5 discusses the optimal weights in our prediction pool while Section 6 reviews the optimal policy rules that are robust within- and across-models and picks out optimal price-level targeting for special attention. Section 8 provides concluding remarks. Full details of the models and their estimation are provided in Online Appendices along with additional results.

[1] Brzoza-Brzezina and Kolasa (2013) also estimate three DSGE models similar to ours using Bayesian methods and evaluate their power to explain the data. However, their investigation focuses on in-sample fit only.

2. Related literature

This paper is related to four strands of methodological literature on Bayesian predictive methods and the use of competing models for forecasting, robust policy, welfare-optimized simple rules and the benefits of inertial rules. We discuss these four strands and how they relate to our paper in turn.

Bayesian Predictive Methods and Competing Models Used for Forecasts. There is an extensive statistical literature on Bayesian predictive methods for assessing, comparing and selecting models (see Vehtari and Ojanen, 2012, for a survey). Within this literature model selection (including more than one model) proceeds via maximization of an expected utility function using the predictive distribution. One particular criterion used in the literature is a scoring rule that measures forecast accuracy where selection then amounts to combining density forecast estimates as a means of improving forecasting accuracy as measured by a scoring rule. We follow Geweke and Amisano (2011) where the utility/loss function is a scoring rule that maximizes forecast accuracy, and BMA is compared with linear combinations of predictive densities. Weights on the component density forecasts are optimized to maximize the score (typically the logarithmic score, of the density combination. Our paper is closely related to Del Negro et al. (2016) who use the same competing models methodology to combine forecasts across two DSGE models, one with and one without financial frictions. Our paper departs from this study in that we seek to use the forecasts to design robust optimal policy across models rather than focus on a counter factual exercise. Ho et al. (2024) adopt an analogous similar pooling exercise to efficiently combine impulse response estimators for the effects of the same economic shock from a class of possible models. They apply optimal pooling to estimating impulse responses when faced with competing VAR and local projection statistical models with different variables. However, neither of these two papers can speak to what would happen in a policy counter-factual and the optimal design of policy that lies at the centre of our paper.

Robust Policy. There is also to a large related literature on robust policy. The spirit of our paper is captured by contributions by Sims (2002, 2007, 2008) who has argued that policymakers are still very far from exploiting the full richness of the Bayesian (or "probability models") approach. A recent literature draws on Hansen and Sargent (2007) in assuming uncertainty is unstructured, with malign Nature 'choosing' exogenous disturbances to minimize the policymaker's welfare criterion ("robust control").[2] Our methodological approach differs from this literature in several important respects. Robust control may be appropriate if little information is available on the uncertainty facing the policymaker. Our approach assumes policymakers are not in such a "Knightian" world and devote considerable resources to assessing the forecasting properties of rival models. It also differs from studies that design robust rules across competing models, but attach probabilities to models under the assumption that one of the models is the true data generating process. For instance, the 'rival models' approach (e.g. Côté et al., 2004; Levin et al., 2003; Adalid et al., 2005) arbitrarily calibrate the relative probabilities of alternative models being true. They define a robust rule as one that "works well" across several (though not necessarily all) models. However, without accounting for how well different models fit the data, it is difficult to assess the value of implementing a rule which performs well in $M-1$ models but poorly in the M^{th} most data-compatible one. Robustness through Bayesian model averaging (e.g. Brock et al., 2007a; Cogley and Sargent, 2005; Levin et al., 2006; Reiss, 2009; Levine and Pearlman, 2010; Levine et al., 2012; Binder et al., 2017, 2018) promotes models with good in-sample fit over models with good forecasting performance by using estimated model probabilities. However, modern monetary policy practices among the inflation-targeting countries are forward-looking and rely heavily on forecasts. This is reflected in our approach which uses a forecasting accuracy criterion to pool models.

Welfare-Optimized Simple Rules. A related literature compares optimized constrained simple rules with their optimal unconstrained counterparts (see, for example, Levine and Currie (1987), Schmitt-Grohe and Uribe 2007; Brock et al. 2007b; Orphanides and Williams 2008; a review is provided by Taylor and Williams 2010). A common finding in this literature is that optimized simple rules can closely mimic optimal policies and perform well in a wide variety of models. By contrast optimal policy can perform very poorly if the policymaker's reference model is mis-specified the reason being that they can be overly fine-tuned to the particular assumptions of the reference model.

Inertial Taylor-Type Rules. The final strand of literature relates to the benefits of inertial Taylor-type rules of which of price-level targeting is a special case; (see, for example, Svensson, 1999; Schmitt-Grohe and Uribe, 2000; Vestin, 2006; Reiss, 2009; Gaspar et al., 2010; Giannoni, 2014). These papers examine the good determinacy and stability properties of price-level targeting. Holden (2016) shows these benefits extend to a ZLB setting. A very recent literature describes these benefits in terms of "make-up" strategies for central banks and in particular the Federal Reserve (see Powell, 2020; Svensson, 2020). Under such strategies policymakers seek to redress past deviations of inflation from its target. Assuming a make-up rule enjoys credibility, undershooting (overshooting) the target will raise (lower) inflation expectations, lower (raise) the real interest rate and help to stabilize the economy. Inertial Taylor rules have by design the make-up feature as they commit to a response of the nominal interest rate to a weighted average of past inflation with the weights increasing with the degree of inertia. Hebden et al. (2020) provide details of these different makeup

[2] See, for example, Dennis et al. (2009) and Ellison and Sargent (2012). Variants of the Hansen-Sargent approach are developed in Adam and Woodford (2012, 2020).

strategies and analyze their effectiveness using the Federal Reserve US macroeconomic model. Bodenstein et al. (2019) underline the benefits of a price-level rule in a non-rational expectations learning environment provided the rule is permanent (as in our set-up) as opposed to the temporary strategy proposed by Bernanke (2017). In our paper we study optimized inertial Taylor rules that are parameterized so as to encompass a simple form of price-level targeting.

3. Methodology: designing robust rules

We restrict our attention to optimized simple rules. The goal of the policymaker is to choose the parameters of a Taylor-type monetary policy rule to maximize welfare that are robust to both within- and across-model uncertainty. Suppose the parameters of the policy rule are collected in the vector ρ. We use the expected lifetime utility of households

$$\Omega_i(\rho,\psi,\Pi) = \mathbb{E}_0\left[\sum_{t=0}^{\infty}\beta^t U_t(\rho,\psi,\Pi)\right] \qquad \psi\in\Psi_i \tag{1}$$

in model $\mathcal{M}_i$ as our welfare measure, where β is the discount factor, Ψ_i is the parameter space for $\mathcal{M}_i$ and $U_t(\rho,\psi,\Pi)$ denotes utility in period t given policy rule ρ, the vector of parameters $\psi\in\Psi_i$, and the trend gross inflation rate target Π. Moreover, we allow the parameter space Ψ_i to differ, but require the policy rule ρ and the trend inflation rate target Π to be the same across models.

We use the estimated posterior distribution from the Bayesian estimation of the model to account for within-model uncertainty. For a policy rule ρ, welfare in model $\mathcal{M}_i$ is given by

$$\Omega_i(\rho,\Pi) = \int_{\Psi_i}\Omega(\psi,\rho,\Pi)p(\psi|y_{i,1:T},\mathcal{M}_i)d\psi \tag{2}$$

where $p(\Psi_i|y_{i,1:T},\mathcal{M}_i)$ is the joint posterior probability distribution of the model parameters estimated for model $\mathcal{M}_i$ given the history $y_{i,1:T}=\{y_{i,1},\dots,y_{i,T}\}$ for the vector of observed model variables $y_{i,t}$. Notice that, unlike BMA, prediction pools do not require the models to have the same vector of observed variables.

We attach weights to each model to account for across-model uncertainty. Given weights $w=\{w_i\}_{i=1}^m$, the policymaker seeks a common rule ρ^* across every model

$$\rho^* = \arg\max_{\rho}\bar{\Omega}(\rho,w,\Pi) \tag{3}$$

where

$$\bar{\Omega}(\rho,w,\Pi) = \sum_{i=1}^{m}w_i\Omega_i(\rho,\Pi), \tag{4}$$

is a welfare measure that incorporates both within- and across-model uncertainty. Given the weights the chosen simple rule ρ^* is our *optimal robust rule*.

The novelty of our paper lies in the way the weights are constructed for the above policy problem. We use forecasting performance as a criterion for assessing the value of different models. Specifically, we follow the procedure of Geweke and Amisano (2012) to form prediction pools where weights are assigned to models on the basis of the accuracy of their k-period ahead forecasts. Unlike the case of Bayesian model averaging which assumes that one of the models is the true data generating process, prediction pools allow us to consider that all models among a comparative set are misspecified, but they all may be useful at different periods of time.

Let the k-period ahead predictive density of model $\mathcal{M}_i$ for a vector of model variables be:

$$p(Y^f_{T+k}|y_{i,1:T},\mathcal{M}_i) = \int_{\Psi_i}p(Y^f_{T+k}|y_{i,1:T},\psi,\mathcal{M}_i)p(\psi|y_{i,1:T},\mathcal{M}_i)d\psi, \tag{5}$$

where $p(Y^f_{T+k}|y_{i,1:T},\psi,\mathcal{M}_i)$ is the density of k-period ahead predictions of the model given a parameter vector $\psi\in\Psi_i$. Up to time T the vector of forecast variables y^f_t is observed, but after time T this is a random vector Y^f_t. Notice that we require all models to share the same vector of forecast variables, but not the observables used for estimation. The predictive density characterizes out of sample observations that have not been used to estimate the posterior density of the parameter vector ψ. Furthermore, the predictive density is independent of the parameter vector ψ which we have integrated over using the estimated posterior distribution. As such this provides predictions about future observations that fully incorporate the information regarding within-model uncertainty in the data.

We assess each model using the log predictive score function. Given a sample $y^f_{1:T}=\{y^f_1,\dots,y^f_T\}$ of forecast variables, the log predictive score of model $\mathcal{M}_i$ is given by

$$LS(y^f_{1:T},\mathcal{M}_i) = \sum_{t=h}^{T-K}\sum_{k=1}^{K}\log p(y^f_{t+k}|y_{i,1:t},\mathcal{M}_i) \tag{6}$$

where $1\le h\le T$ ensures that there are enough observations in the first subsample to estimate the model. The log predictive score function measures a model's predictive performance by evaluating its predictive density at the available out-of-sample observations.

We use linear prediction pools to assess the predictive performance of a combination of models.[3] Given a sample $y^f_{1:T}$ and a model pool $\mathcal{M} = \{\mathcal{M}_1, \dots, \mathcal{M}_m\}$, the log predictive score of the pool is given by

$$LS(y^f_{1:T}, \mathcal{M}) = \sum_{t=h}^{T-K} \sum_{k=1}^{K} \log \left[\sum_{i=1}^{m} w_i p(y^f_{t+k} | y_{i,1:t}, \mathcal{M}_i) \right]; \qquad \sum_{i=1}^{m} w_i = 1; \qquad w_i \geq 0. \tag{7}$$

The log predictive score function measures the predictive performance of a convex linear combination of the models in the pool. The optimal prediction pool has weights chosen such that the log predictive score of the pool is maximized[4]

$$w_i^* = \arg\max_{w_i} LS(y^f_{1:T}, \mathcal{M}) \tag{8}$$

4. Models

To illustrate our method, we investigate the welfare consequences of alternative monetary policy rules using three medium-scale new Keynesian DSGE models. The first is the Smets and Wouters (2007) model which is the workhorse model in policy-making institutions for forecasting and policy analysis. The other two models build financial frictions into the SW model along the lines of Gertler et al. (2012) and Bernanke et al. (1999), respectively. These two models represent the leading theories in modelling financial frictions in the macroeconomic literature. [5] Hence, our model pool can be motivated by considering a policymaker who is uncertain how to incorporate financial frictions into a DSGE model or if they should be incorporated at all.

We estimate the models with Bayesian methods. For all three models we use the same seven time series as observable variables as in Smets and Wouters (2007): the log difference of real GDP, real consumption, real investment and the real wage, log hours worked, the log difference of the GDP deflator, and the federal funds rate. We use US quarterly data over the sample period 1966:1-2017:4.[6] The parameter prior and posterior distributions are reported in Table D1 in the Appendix, while further details on the estimation procedure are given in Appendix D.

We calibrate several parameters in the estimation procedure that are hard to identify in the models. We estimate the posterior distribution of the remaining 26 parameters, which are common in all three models. The reason that all three models have the same parameter vector estimated is to accommodate BMA in our exercise. BMA requires the models to share the same set of observable variables, but the SW model has no empirical implications for financial variables. Since parameters specific to the banking sector in the GKQ or the BGG model are hard to identify without including financial variables among the observable variables, we decided to fix those parameters and estimate only parameters common in all three models.[7]

4.1. The workhorse new Keynesian model

Our first model follows closely Smets and Wouters (2007). It is a stochastic neoclassical growth model augmented with price and nominal wage stickiness, price and nominal wage indexation, internal habit persistence and investment adjustment costs. The welfare of household j in the model is defined by their expected lifetime utility

$$\Omega(j) = \mathbb{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{\left[C_t(j) - \chi C_{t-1}(j)\right]^{1-\sigma_c}}{1-\sigma_c} \exp\left[(\sigma_c - 1) \frac{H_t(j)^{1+\psi}}{1+\psi} \right] \right] \tag{9}$$

where $C_t(j)$ is real consumption, $H_t(j)$ is hours supplied, β is the discount factor, χ controls habit formation, σ_c is the inverse of the elasticity of intertemporal substitution (for constant labour), and ψ is the inverse of the Frisch labour supply elasticity. The monetary policy rule for the nominal interest rate $R_{n,t}$ in the model is given by the Taylor-type rule

$$\begin{aligned} \log\left(\frac{R_{n,t}}{R_n}\right) &= \rho_r \log\left(\frac{R_{n,t-1}}{R_n}\right) \\ &\quad + (1-\rho_r)\left(\theta_\pi \log\left(\frac{\Pi_t}{\Pi}\right) + \theta_y \log\left(\frac{Y_t}{Y}\right) + \theta_{dy} \log\left(\frac{Y_t}{Y_{t-1}}\right)\right) + MPS_t, \end{aligned} \tag{10}$$

[3] Del Negro et al. (2016) use the terminology static pools to reflect the fact that weights are time invariant.

[4] Logs are used in general since they make the densities globally concave, making the maximization easier.

[5] Our choice is motivated by the fact that central banks most commonly use these models for policy analysis. However, other theories have been proposed in the academic literature. In Section 8 we discuss the possibility of including these theories in our model pool in future research.

[6] As noted, we estimate the models using US data on the standard seven key macroeconomic variables over the sample period 1966:1-2017:4. We have also estimated our models with a sample that stops just before the financial crisis, and a version with the shadow federal funds rate (Wu and Xia, 2016) rather the nominal interest rate. All our results are robust to these differences.

[7] One could estimate the financial frictions models with financial variables included in the set of observables and compute conditional marginal likelihoods from those models instead (Brzoza-Brzezina et al., 2013). This procedure would isolate the effect of financial variables on the marginal likelihoods and would allow us to compare the fit of all three models to the same set of variables. However, Bayesian odds computed this way would be biased towards models estimated on a wider set of data. Alternatively, we could follow Del Negro et al. (2016) and conduct inference based on a pseudo-likelihood function and a pseudo-posterior distribution that leave out the variables that the hypothetical policymakers are not interested in. However, this would make our paper computationally even more challenging.

where Π_t is inflation, Y_t is real output, and MPS_t is a monetary policy shock that follows an AR(1) process. The rest of the model is described in detail in Appendix A.

We review here the estimates for the parameters that are important for our policy problem; the estimated posterior distribution for the rest of the parameters can be found in Table D1 in the Appendix. The estimated Calvo parameter for price setting implies that prices are updated on average every 2.07 quarters, but price indexation is rather weak given that the indexation parameter is estimated to be lower than the prior mean. Nominal wages are updated more frequently, on average every 1.67 quarters, but they are also indexed more strongly to inflation. The Taylor rule parameters are very close to the original estimates of Smets and Wouters (2007) and correspond to the typical estimates in the DSGE literature.

4.2. The banking model with outside equity

Our second model extends the SW model with a banking sector along the lines of Gertler et al. (2012) and de Groot (2014). In the model banks raise deposits and issue outside equity to finance loans to firms. To motivate an endogenous constraint on the bank's ability to raise funds, Gertler et al. (2012) assume that the banker managing the bank may transfer a Θ fraction of assets to their family. Hence, only $1-\Theta$ fraction of assets can be pledged as collateral. Recognizing this possibility, risk averse households limit the funds they lend to banks. Our setup for the banking sector follows closely Gertler et al. (2012), and embeds in our SW model in a similar fashion to Gertler and Karadi (2011). For a detailed description of the model see Appendix B.

The posterior distributions of the estimated parameters are close to the estimates in the SW model. The only exceptions that are important for our policy problem are the parameters characterizing intermediate firms' price setting behaviour. Prices are updated less frequently, on average every 2.91 quarters, but they are indexed to inflation more weakly than in the SW model.

4.3. The financial accelerator model

Our third model extends the SW model with a banking sector along the lines of Bernanke et al. (1999). In the model banks collect deposits and lend to entrepreneurs who are subject to idiosyncratic shocks that affect their ability to repay their loans. The financial market friction in this model, which is between the entrepreneur and the bank, is driven by private information. Banks pool loans to protect themselves against credit risk and charge a spread over the deposit rate. The household welfare function and the Taylor-rule is the same as in the SW model. For a detailed description of the model see Appendix C.

The posterior distributions of the estimated parameters resemble the estimates of the SW model, but there a few notable exceptions that are important for our policy problem. Prices are even less flexible than in the GK model and updated on average every 3.34 quarters. Nominal wages are more flexible and are updated on average every 1.76 quarters.

4.4. Measuring the severity of the financial friction across models

Recall that across all models, the ex-post gross real interest rate for the bond issued in period $t-1$ and the gross return on capital can be expressed as follows:

$$R_t = \frac{R_{n,t-1}}{\Pi_t}, \tag{11}$$

$$R_t^K = \frac{r_t^K + (1-\delta)Q_t}{Q_{t-1}}, \tag{12}$$

where Q_t is the price of capital, and r_t^K is the rental rate of capital. Then, we can use the expected discounted spread between the return on capital and the risk-free bond as a measure of the severity of financial frictions present in a model.

For the SW model, as financial frictions are absent, this spread is zero:

$$\mu_t^{SW} = \mathbb{E}_t[\Lambda_{t,t+1}(R_{t+1}^K - R_{t+1}) = 0, \tag{13}$$

where $\Lambda_{t,t+1}$ is the household's stochastic discount factor between periods t and $t+1$. l product of capital in the wholesale sector. From Appendix B the counterparts for the GKQ and BGG models are respectively

$$\mu_t^{GKQ} = \mathbb{E}_t[\Lambda_{t,t+1}\Omega_{t+1}(R_{t+1}^K - R_{t+1})] > 0, \tag{14}$$

$$\mu_t^{BGG} = \mathbb{E}_t[R_{t+1}^K - R_{t+1}] = \mathbb{E}_t[(\rho_{t+1}(\bar{\psi}_{t+1}) - 1)R_{t+1}] > 0, \tag{15}$$

where for the GKQ model a wedge Ω_{t+1} arises from the incentive constraint of the bankers, and for the BGG model $\rho_t(\bar{\psi}_t) > 1$ is the premium on external finance where $\bar{\psi}_t$ is the productivity threshold at which a firm defaults on its loan from the bank.

As the household-bank GKQ friction disappears, $\Omega_{t+1} \to 1$, and therefore $\mu_t^{GKQ} \to \mu_t^{NK}$. In the BGG model, the external finance premium approaches 1 as we reduce the monitoring costs to zero and therefore $\mu_t^{BGG} \to \mu_t^{NK}$ in that case.

5. Forecasting performance

We use standard Bayesian methods to estimate our models repeatedly with an increasing window of data, and compute log predictive scores (6) and (7) for predictions made by our estimated models. Each estimation sample starts at 1966:1. The first sample

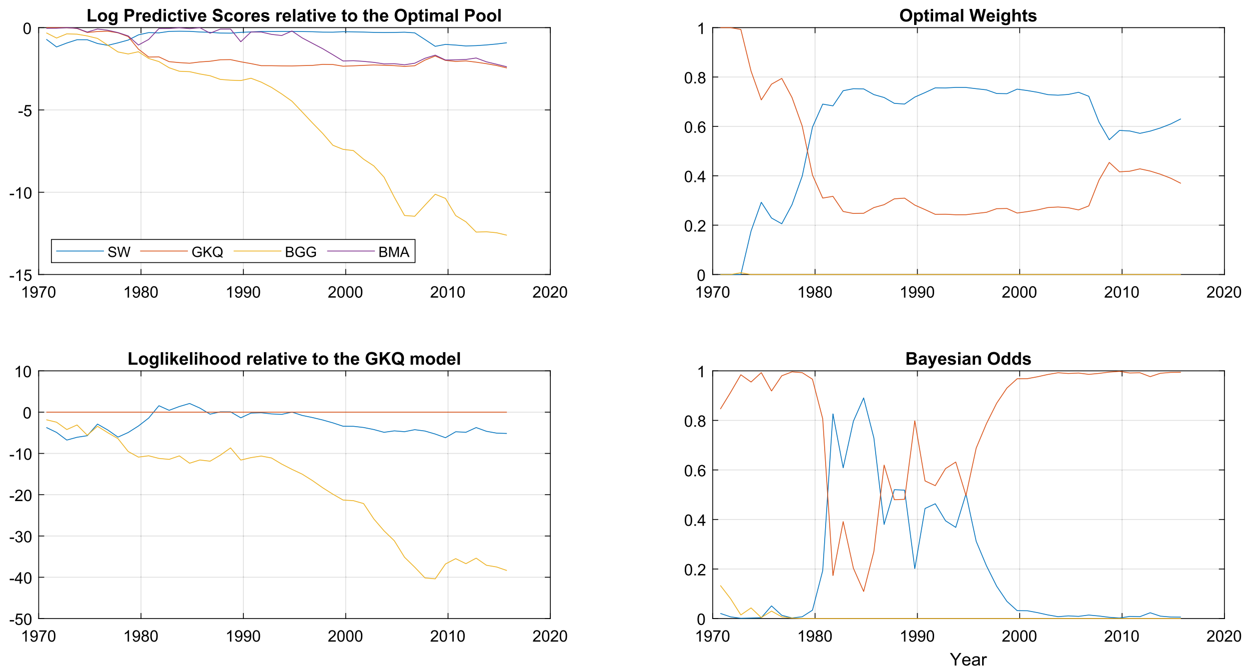

Fig. 1. Model weights based on prediction pools and Bayesian model averaging.

ends at 1970:4 ($h = 20$). We assess our models based on how well they predict all seven observable variables jointly up to eight quarters ahead ($K = 8$). This corresponds to the 2-year forecasting and inflation-targeting of most central banks.[8] We increase the sample size by four quarters each time and repeat the same steps.[9]

The top panel of Fig. 1 shows the log predictive score function for each individual model, and for the pool using the Bayesian odds as weights, relative to the optimal pool's. Note that by construction the optimal pool is the linear convex combination, so it provides the best forecast. We employ prediction pooling to aggregate these relative predictive performance differences over time and the second panel of Fig. 1 shows the resulting optimal prediction pool weights over the sample period 1970:4-2015:4. To obtain these weights we solve the optimization problem (8) recursively. At each point in time we use the log predictive scores up to that point to determine the weights as if our full sample ended there.

It is clear from both panels that the SW and GKQ models provide better predictions than BGG overall. BGG falls quickly behind the other two models and the predictive score difference accelerates further in the second half of the sample. However, for the 2007-2009 crisis we see a marked improvement in the performance of BGG, but the weight associated with the model remains zero. This is due to the fact that the model accumulates such a log predictive score deficit by this point in the sample that this improvement is not large enough to gain the model a positive weight. Weights of approximately 63%, 37% and 0% are assigned to the SW, GKQ and BGG models, respectively, by the end of the sample period.

The bottom panels of Fig. 1 show an interesting and important contrast to the top two. The third panel shows how the difference between our models' log-likelihood evolve over our sample. Since for most of our sample period the GKQ model provides the best fit, we plot each model's log-likelihood relative to GKQ's. The last panel shows the evolution of the Bayesian odds over our sample, given a uniform prior belief of the policymaker over the competing models. While for most of the sample period the GKQ model provides the best fit and gets the highest odds, it receives much lower weight in the optimal prediction pool. Had the policymaker used BMA to attach weights to the models, she would have put most of her faith in the GKQ model while ignoring the other two models entirely. In fact, with the exception of the few years in the middle of our sample, BMA have the tendency to assign almost all the weight to one model only in our model pool. Moreover, the optimal prediction pool weights change slowly over time while large changes in Bayesian odds can be bought by adding only a handful of observations to the sample, especially in the middle of the sample.

The financial crisis then demonstrates the differences between the two weighting schemes very well. Del Negro et al. (2016) find that during periods of financial turmoil models that incorporate financial frictions as in the financial frictions model used in that paper (BGG) provide better forecasts, while the opposite occurs during more tranquil periods. We extend their results to models with

8 But note that the inflation targeting rule is only one-period-ahead. This is motivated by Batini et al. (2006) that examines forward-looking rules up to 8 periods ahead and finds that they are increasingly prone to the problem of indeterminacy as the forward horizon increases. As a consequence the stabilization performance of optimized rules of this type worsens too.

9 We modify Dynare's forecasting routine to calculate the predictive density (5) using each model's estimated posterior distribution. We increase the sample size by four quarters each time and repeat the same steps. We re-estimate the models only every four quarters to reduce the computational complexity of the task. This way we need to estimate each model only 46 times, and our forecasting periods do not overlap each other.

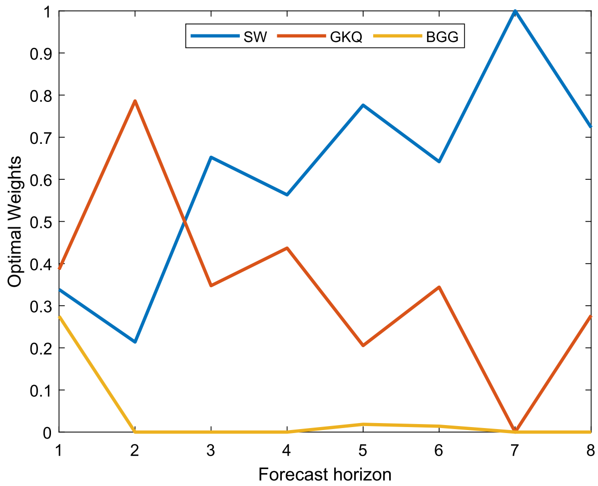

Fig. 2. Optimal prediction pool weights at the end of our sample period using different forecast horizons.

GKQ style financial frictions. We confirm that during tranquil periods the importance of both financial frictions are reduced, but less so for the financial friction between households and banks (GKQ). We also find a marked relative improvement in the forecasting performance of both models with financial frictions around 2007 with a large log predictive score differential over a few quarters (see upper panel). This produces a noticeable jump in the optimal prediction pool weight assigned to the GKQ model, but not in the weight assigned to the BGG model for reasons explained above. At the same time the BMA weights are unchanged during this period since the relative fit of the different models in the pool are not affected by the financial crisis.

The difference in rankings across models between prediction pooling and BMA is due to the standard trade-off between in-sample fit and out-of-sample predictive performance. The financial frictions models fit the in-sample data quite well relative to the SW model (and therefore have higher marginal log-likelihoods) yet they tend to predict poorly (especially in non-crises periods) in comparison to the SW model as they tend to over-fit the data. As a result, the prediction pools assign a significant weight to the SW model. Nevertheless, the situation is reversed at particular times (e.g. crisis times) when model complexity may be beneficial for forecasting and therefore prediction pools should also attach significant weight to the financial frictions models.

Fig. 2 shows how the weights assigned to each model at the end of our sample period depend on the forecast horizon used. These weights, unlike the ones depicted in Fig. 1, are computed based on how well models predict for a given horizon only. At shorter horizons all three models are useful in terms of predictability and ought to be employed by a policymaker when designing policy. But the performance of models with financial frictions declines while the performance of the SW model increases with the horizon to the point where the latter dominates the pool. Nonetheless, the weights reported in Fig. 1 are less extreme since prediction pools leverage the strengths of different models at different horizons when assigning a weight to each model when we take into forecasting performance account up to eight quarters ahead.

Overall, our results suggest that if the policymaker does not believe any model to be correct, she must nevertheless be much more cautious in rejecting models that appear to be inferior. The intuition here is that such models while not particularly accurate most of the time or at all horizons or for all variables, may well be very useful some of the time or at some horizons or for predicting key policy variables. Thus the optimal policy should take this into consideration rather than rejecting the use of such models.

Prediction pools leverage the strengths of the various models along these three dimensions and assign model weights accordingly. They take into account the fact that the usefulness of different models vary over the sample period. Consequently, prediction pools assign positive weights to more than one model that also tend to evolve more smoothly over time. BMA on the other hand provides more extreme weights that most of the time reject all but one model. These weights are also much more subject to big changes when there are outliers in the data. Second, conventional wisdom is that monetary policy works over a horizon of between two to eight quarters. Prediction pools can account for the relative strength of the various models at forecasting over different horizons when assigning weights. BMA on the other hand ignores this dimension since the estimate of the state vector is updated continually as new observations become available.

6. Optimal robust rules

This section analyses the optimal simple rules for our model economies. First, we define the optimal simple rule we compute. Next, we examine optimal simple rules designed for individual models. Then, we proceed to study optimal simple rules using different ways to attach weights to models. Finally, we examine the welfare cost of suboptimal policy using two different sets of criteria: i) the cost of implementing the rule designed for model i in the environment of model $k \neq i$ and ii) the cost of implementing a rule with deviations from the optimal parameter values.

We assign weights to the SW, BGG and GKQ models in our policy design exercise based on relative predictive performance over the entire sample period from 1966Q1 to 2017Q4. The results in following sections are conditional on the choice of these weights.

However, we also compute corresponding results using alternative end-dates. We show in the robustness section 7 that the subsequent results are independent of this choice.

6.1. Computation and welfare measures

We seek policy parameters $\rho = \begin{bmatrix} \rho_r & \alpha_\pi & \alpha_y & \alpha_{dy} \end{bmatrix}$ in the inertial Taylor-rule

$$\log\left(\frac{R_{n,t}}{R_n}\right) = \rho_r \log\left(\frac{R_{n,t-1}}{R_n}\right) + \alpha_\pi \log\left(\frac{\Pi_t}{\Pi}\right) + \alpha_y \log\left(\frac{Y_t}{Y}\right) + \alpha_{dy} \log\left(\frac{Y_t}{Y_{t-1}}\right) \tag{16}$$

that maximize the unconditional lifetime utility of households given the policy decision on the trend gross inflation rate target Π.[10] We estimate the model using the same form of the Taylor-rule (10) and the same priors on its parameters as in Smets and Wouters (2007). However, we re-parametrize the feedback coefficients by setting $\alpha_\pi = (1-\rho_r)\theta_\pi$, $\alpha_y = (1-\rho_r)\theta_y$, and $\alpha_{dy} = (1-\rho_r)\theta_{dy}$ to allow for the possibility of a price level rule ($\rho_r = 1$) when computing optimized simple rules.

Optimal policy for a particular model $\mathcal{M}_i$ is given by the rule $\rho_i^* = \arg\max_\rho \Omega_i(\rho, \Pi)$, where $\Omega_i(\rho, \Pi)$ is our chosen welfare measure defined in (2). Optimal policy across a pool of models $\mathcal{M} = \{\mathcal{M}_i\}_{i=1}^m$ with associated weights $w = \{w_i\}_{i=1}^m$ is given by a rule ρ^* common to every model in the pool that solves

$$\max_\rho \bar{\Omega}(\rho, \Pi, w) \tag{17}$$

given a trend inflation rate Π common to every model and

$$\bar{\Omega}(\rho, \Pi, w) = \sum_{i=1}^m w_i \int_{\Psi_i} \Omega_i(\rho, \psi, \Pi) p(\psi | y_{i,1:T}, \mathcal{M}_i) d\psi \approx \sum_{i=1}^m w_i \sum_{j=1}^N \Omega_i(\rho, \psi_j, \Pi), \tag{18}$$

where ψ_j is a draw from parameter space Ψ_i. The optimization problem defined by (17) can incorporate both *within-model* robustness, by averaging of N draws for each model $\mathcal{M}_i$ and *across-model* robustness by using a weighted average of model specific welfare measures. Calculation of (18) is numerically facilitated by MCMC methods that allow the sampling of a given distribution by simulating an appropriately-constructed Markov chain. Thus, the integral can be approximated by a sum. For each model i, we draw $N = 500$ parameter vectors $\psi_j \in \Psi_i, j \in [1, N]$ from the estimated posterior distributions to approximate the integral above.

We compare alternative policies in terms of consumption equivalent welfare changes. Consider two alternative policies ρ_1 and ρ_2. The consumption equivalent welfare cost of adopting ρ_2 in model $\mathcal{M}_i$ is the fraction ω^i of the consumption stream households are willing to give up to be as well off under ρ_2 as under ρ_1. It is implicitly defined by the indifference condition

$$\Omega_i(\rho_2, \psi) = \mathbb{E}_0 \left[\sum_{t=0}^\infty \beta^t U((1-\omega^i) C_t(\rho_1, \psi), H_t(\rho_1, \psi)) \right] \tag{19}$$

This welfare measure allows us to compare the effect of the same policy in different model economies directly or aggregate the effect of the policy across models. To compute the welfare cost of a policy rule common to all models in a pool, we calculate the weighted average of the associated welfare cost from each model using the weights assigned to each model.

$$\omega = \sum_{i=1}^3 w_i \omega^i \tag{20}$$

Before turning to optimal rules we should note the magnitude of the *costs of the business cycle* in our three estimated models. They are, in consumption equivalent percentage terms, 0.49 for the SW model, 0.39 for GKQ and 0.29 for BGG.[11] These are much higher than those reported in the seminal study by Lucas (1987) and updated in Lucas (2003) which are less than 0.01. The reason for this is his choice of utility function which, unlike our SW-type models, excludes hours (and therefore leisure) in these original studies.[12] We view the benefits of stabilization rules in terms of the proportion of these costs that are removed.

6.2. Optimal within-model robust rules

The optimal simple rules are considerably different from the estimated rule for each model (Table 1). The sub-optimality of the estimated rules is well-established in the literature. Our consumption-equivalent measure ω^i indicates that using optimized rules provides a consumption equivalent welfare gain over the estimated rule ranging from around 0.050 percent for the GKQ model to 0.117 percent for the BGG model.

[10] Note that the trend (steady state) gross inflation rate target in the rule which we take to be $\Pi \geq 1$ (ruling out a liquidity trap) uniquely pins down the rest of the steady state including the gross nominal interest rate R_n.

[11] These are found by comparing the estimated stochastic models with their deterministic counterparts. Differences reflect different model features and different estimates of the parameters including those for shock processes and rules.

[12] But as Gali et al. (2007) point out even in the SW framework we may still be seriously understating the welfare costs of recessions because a reduction in hours leads to a increased welfare benefit of leisure. We return to this issue in our concluding remarks.

Table 1
Optimized simple rules.

	SW		GKQ		BGG	
	Estimated	Optimal	Estimated	Optimal	Estimated	Optimal
	Taylor rule parameters					
Interest rate smoothing (ρ_r)	0.667	0.003	0.665	0.707	0.718	0.834
Feedback on inflation (α_π)	0.661	8.374	0.653	1.176	0.566	2.858
Feedback on output (α_y)	0.002	0.042	0.002	0.003	0.002	0.021
Feedback on output growth (α_{dy})	0.061	1.028	0.060	0.188	0.057	0.100
	Consumption equivalent welfare gain over the estimated rule					
Welfare gain, $\omega^i(\%)$	0.000	0.100	0.000	0.050	0.000	0.117
	Unconditional standard deviation of key variables					
Nominal interest rate	0.695	0.993	0.788	0.651	0.898	0.740
Inflation	0.585	0.188	0.620	0.350	0.623	0.148
Real output	0.178	0.178	0.179	0.178	0.143	0.144
Output growth	0.858	0.830	0.835	0.808	0.792	0.850
	Probability of hitting the zero lower bound					
Probability of ZLB (%)	2.802	8.968	4.573	2.100	6.924	3.678

Note: ω_i shows the fraction of the consumption stream households need to give up to have the same welfare as under the estimated rule. Some of the standard deviations are scaled by 100 for ease of presentation. p_{ZLB} is the probability that the gross nominal interest rate is below one in the normal distribution $\mathcal{N}(R_N, \sigma_{R_N})$, where R_N is the deterministic steady state of the nominal interest rate.

Across-model differences of optimal rules are considerable. The optimal simple rule designed for the SW model is close to a pure inflation targeting rule. It has almost zero interest rate smoothing parameter ρ_r but sees a very aggressive immediate response to changes in inflation captured by the feedback parameter α_π.[13] Consequently, the welfare gains are closely related to the large decrease in inflation volatility implied by the optimal rules: the standard deviation of inflation decreases by a factor of three from 0.00585 to 0.00188. By contrast the optimal rules for the GKQ and BGG models have considerable interest rate smoothing but far more muted response to inflation. The welfare gains from decreased inflation volatility, which works through agents' inflation expectations, are much smaller.

The lack of interest rate smoothing in the optimal policy for the SW model results in a higher nominal interest rate volatility than under the estimated rule. An important consequence of the wider interest rate distribution is that the unconditional probability of hitting the zero lower bound p_{ZLB} is much higher than with the estimated policy rule–over 9 percent per quarter or about once every 3 years.[14] This contrasts with about 2 percent per quarter for the optimal rules in the GKQ and BGG models, or about once every 12.5 years. Available data suggest that zero lower bound episodes are rare but long-lived (Dordal-i-Carreras et al., 2016; Ramey and Zubairy, 2018). The U.S. post-WWII experience (seven years at the zero lower bound over seventy years) indicates that unconditional probabilities below 10 percent are empirically plausible. However, a policymaker may want to set an upper bound on this probability as part of the policy mix.[15]

6.2.1. *Comparison of impulse responses*

In this section we investigate the models' impulse response functions to demonstrate that differences in optimal policies across models are due to the different financial frictions in the GKQ and BGG models, and because the SW model has no such friction. Thus, we highlight the model uncertainty problem facing a policymaker.[16]

Specifically, we investigate the across-model differences of optimized rules using impulse responses to the monetary policy shock $\epsilon_{MPS,t}$ in the Taylor rule and to the labour productivity shock $\epsilon_{A,t}$. As is standard, the former is chosen in order to understand the monetary policy transmission mechanism, whilst the latter is the most important shock in the variance decomposition analysis of the estimated model. It should be stressed that although we can consider impulse responses to individual shocks our optimized rules are

[13] We do not impose an upper bound on the policy parameter α_π during optimization. Imposing the upper bound $\alpha_\pi \leq 3$ as in Schmitt-Grohe and Uribe (2007) for example would reduce the welfare gain for the SW model slightly, but would be binding for this optimal simple rule only throughout our paper.

[14] The unconditional probability of hitting the zero lower bound is computed from a normal approximation of the gross nominal interest rate's ergodic distribution. Let R_N and σ_{R_N} denote the deterministic steady state and the unconditional standard deviation of the gross nominal interest rate, respectively, computed from a second-order approximation around a deterministic steady state. Then the probability of hitting the zero lower bound, p_{ZLB} is given by the probability that the gross nominal interest rate is below one in the normal distribution $\mathcal{N}(R_N, \sigma_{R_N})$.

[15] See Deak et al. (2023) which proposes a general mandate framework for delegating monetary policy to a central bank (CB) that emphasizes simplicity in both the objectives entering the welfare criterion and those in the instrument rule. The mandate includes a Taylor-type log-linear nominal interest-rate rule with target variables that match those in the loss function and a zero-lower-bound (ZLB) constraint on the nominal interest rate taking the form of an unconditional probability of ZLB episodes.

[16] The models are estimated separately (essential for computing the pooling weights) and therefore also differ because the estimated parameters differ, including those for the shock process. Thus multiple things are happening at the same time, making the interpretation of the irfs to essentially different estimated shocks processes less straightforward. However, in defence of the comparison, different estimates are an important part of the model uncertainty we are trying to capture for the robust design problem.

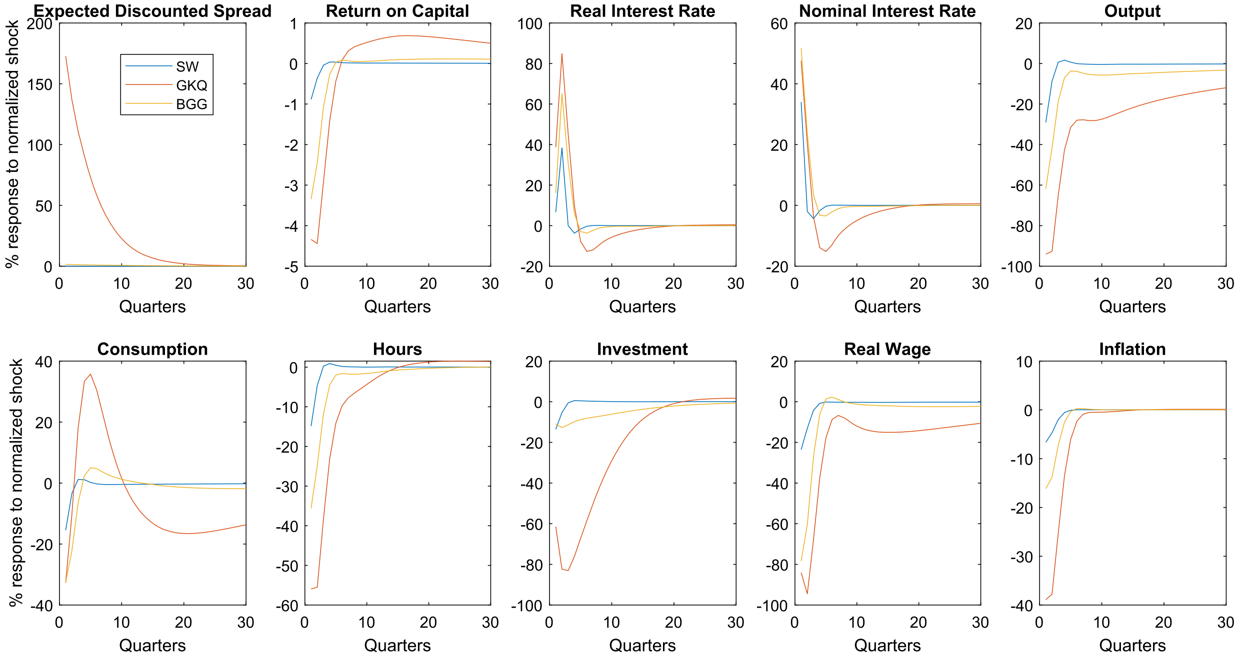

Fig. 3. Impulse Responses to a Monetary Policy Shock Compared across Models.

only welfare optimal for all shocks in the deterministic steady state. Optimized rules are found by solving the optimization problem (17) given the estimated joint distribution of all parameters which include those in the assumed AR(1) processes of shocks.

Fig. 3 plots the impulse responses to a positive monetary policy shock (a tightening of monetary policy). For all models we see the familiar result of a monetary tightening: the gross return on capital (R_t^K in (11)) immediately falls causing a drop in investment. The real interest rate rises causing an immediate increase in household savings and a decrease in consumption. Output, hours and the real wage fall balancing supply and demand in output and labour markets. The nominal interest rises despite the fall in inflation and output owing to the shift in the Taylor rule brought about by the positive monetary shock.

Fig. 4 plots the corresponding impulse responses to a positive technology shock. Now all models we see that the gross return on capital (R_t^K) immediately *rises* causing a *rise* in investment. The real interest rate still rises owing to the 'Taylor principle' in the interest rate rule but consumption increases as a result of an increase in the real wage. Output rises but households enjoy more leisure and hours fall.

Figs. 3 and 4 show both the familiar *accelerator* effects (for the monetary shock) but also *attenuation* effects (for the productivity shock) of financial frictions when accompanied by an inertial Taylor rule (see, for example, Wieland and Yang (2020)) This arises from changes in the expected spread and its effect on investment, output and inflation. In both cases make-up strategies in the form of an inertial monetary rule or at its extreme, a price-level rule respond to a prolonged mismatch between outcomes and targets of inflation and output specified in the nominal interest rate rule.

6.3. *Optimal across-model robust rules*

We compare two sets of policy rules computed using two different set of pooling weights. The first uses BMA and the second which we call 'Prediction Pool' uses optimal prediction pool weights.

The choice of pooling weights matters for the optimized rule (Table 2). The optimal simple rule obtained using the optimal prediction pool weights implies a higher welfare gain over the estimated rule compared to the policy rules obtained using the BMA weights, 0.068 against 0.050 consumption equivalent percent. This higher welfare gain is explained by a combination of two factors: differences in i) the weights and ii) in the rules. The first is a composition effect since the weights used in (20) to aggregate the welfare gains from the individual models are different. If we aggregated the welfare gains of the prediction pool optimal rule from the individual models using the BMA weights, then the aggregate welfare gain would be only 0.039. Hence, differences in the optimal rules would imply a higher difference in welfare gains were the weights the same.

The second factor is the differences in the rules themselves. Using the optimal prediction pool weights implies a more persistent policy rule that also responds more aggressively to changes in inflation and output growth compared to the policy rules obtained using the BMA weights. The welfare gains from the individual models in the ω^i row reveal that the prediction pool rule trades off welfare gains from the GKQ model for welfare gains from the SW model given that it attaches more weight to the latter. Interestingly, inflation volatility is reduced in all three models the most by the prediction pool optimal rule, but the likely welfare gains from this reduced volatility are partly negated by the interest rate smoothing parameter being above its optimal level in all three models.

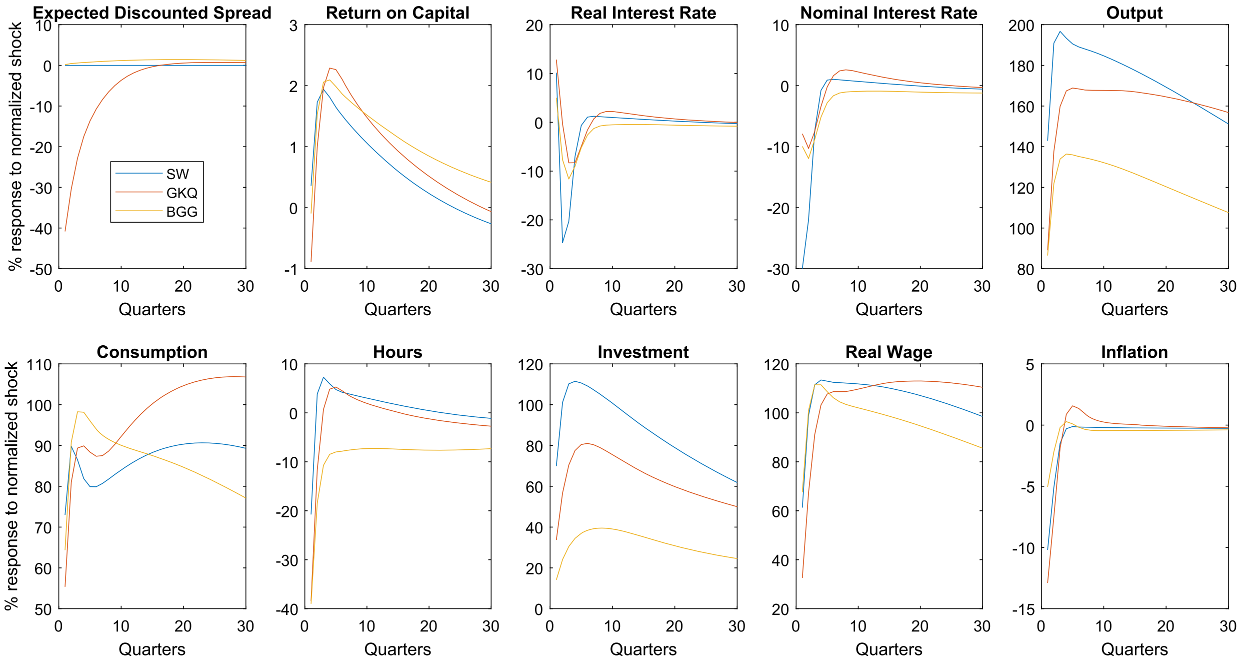

Fig. 4. Impulse Responses to a Technology Shock Compared across Models. *Note:* These graphs compare the impulse responses in deviation form relative to the deterministic steady state. The impulse responses are responses to a unit shock of size 1, multiplied by 100. The rule is the optimized rule in Table 1 computed for the SW, GKQ and the BGG models.

Table 2
Optimized robust simple rules.

	BMA			Prediction pool		
	SW	GKQ	BGG	SW	GKQ	BGG
Model weights (%)	0.57	99.43	0.00	63.06	36.94	0.00
Taylor rule parameters						
Interest rate smoothing (ρ_r)		0.707			0.769	
Feedback on inflation (α_π)		1.179			1.936	
Feedback on output (α_y)		0.003			0.006	
Feedback on output growth (α_{dy})		0.188			0.326	
Consumption equivalent gain over estimated rule						
Welfare gain in each model, ω^i(%)	0.066	0.050	0.096	0.085	0.039	0.110
Weighted average gain		0.050			0.068	
Probability of hitting the zero lower bound						
Probability of ZLB (%)	1.564	2.098	3.994	1.645	2.190	3.048

Note: ω_i shows the fraction of the consumption stream households need to give up to have the same welfare as under the estimated rule in each model. The average gain is the weighted average of these gains using the weights assigned to each model. p_{ZLB} is the probability that the gross nominal interest rate is below one in the normal distribution $\mathcal{N}(R_N, \sigma_{R_N})$, where R_N is the deterministic steady state of the nominal interest rate.

Finally, across-model robustness removes the high probability of hitting the nominal interest rate lower bound seen in Table 1 for the SW model with only within-model robustness. Since the optimal values for the interest rate smoothing parameter are well above the estimated values regardless of the weights we use, interest rate volatility decreases for every optimal policy and so does the probability of hitting the zero lower bound.

Fig. 5 confirms that the choice of optimal pooling versus BMA weighting for the transmission of monetary policy for our three models does matter. Taking into account both the differences in the persistence parameter ρ_r and the direct responses, the optimal rules computed in Table 2 using these two approaches are quite different from each other. To see this compute the long-run responses $\alpha_\pi/(1-\rho_r)$ and $\alpha_{dy}/(1-\rho_r)$: these are 4.25 and 0.704 for the BMA rule and 8.823 and 1.241 for the optimal prediction rule; that is almost *twice* as strong. A comparison of the impulse responses for each model for the monetary shock in Fig. 5 confirms a significance difference. In a stochastic steady state, differences across all 7 shocks hitting our model economies then leads to significant welfare effects to which we now turn.

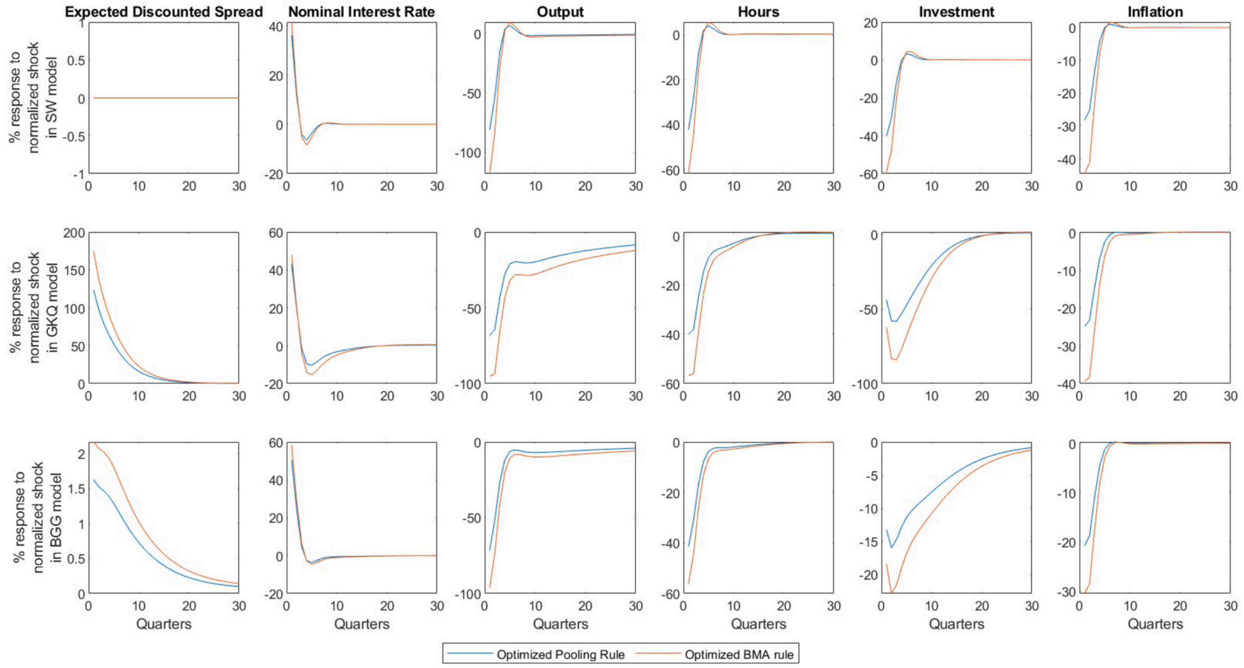

Fig. 5. Impulse Responses to a Monetary Policy: Pooling versus BMA Optimized Rules. *Note:* These graphs compare the impulse responses in deviation form relative to the deterministic steady state. The expected discounted spreads are defined in (13)-(15). The impulse responses are responses to a unit shock of size 1, multiplied by 100. The rules are the optimized rules in Table 2 computed for the prediction pool and BMA.

Table 3
Welfare cost of robust optimal policy i in model $j \neq i$.

		True model		
		SW	GKQ	BGG
Policy rule	SW	0.000	0.137	0.005
	GKQ	0.034	0.000	0.021
	BGG	0.008	0.051	0.000
	$Pool$	0.014	0.011	0.007

Note: The table shows what happens when an optimal simple rules optimized for model i is used in model $j \neq i$. The first column shows the consumption equivalent welfare loss in the SW model relative to the welfare attained using the robust simple rule optimized for the SW model if, for example, we use the robust simple rules optimized for the SW, GKQ, BGG models, respectively. The last row shows the welfare cost incurred in model i when instead of using the robust simple rule optimized for model i we use the robust optimal simple rule obtained with the optimal prediction pool weights.

6.4. *The welfare cost of suboptimal policy*

In this section we examine the welfare cost of suboptimal policy using two different sets of criteria. First, we quantify the cost of implementing the rule designed for model i in the environment of model $j \neq i$. Then, we analyse the cost of implementing a rule with deviations from the optimal parameter values.

Table 3 shows the welfare cost of using a rule optimized for a specific model in another model. This is a counterfactual exercise that shows the cost of incorrectly identifying the data generating process. For example, the first row shows that if we use the robust simple rule optimized for the SW model in the GKQ and BGG models, then the welfare loss is 0.137 and 0.005 percent of consumption respectively relative to that from the robust simple rules optimized for the latter two models themselves. The results show that incorrectly identifying the GKQ model as the data generating process implies the largest welfare costs. Using the rule optimized for the SW model in the BGG model, or vice versa, imply rather small welfare losses. Using the robust simple rule optimized for the GKQ model in the other two models, however, implies much larger welfare losses than the other way around. The final row shows the welfare cost of using the robust rule optimized for the prediction pool in Table 2 relative to the model specific robust optimal rules reported in Table 1. These now avoid the large costs of the single model optimized rules and the costs are generally small relative to the gains from using optimal rules.

Although the welfare gains from using the optimal robust policy across models as opposed to rules designed for each model remain modest, there can be very significant gains if there are small deviations from the precise optimized rules. Fig. 6 shows how

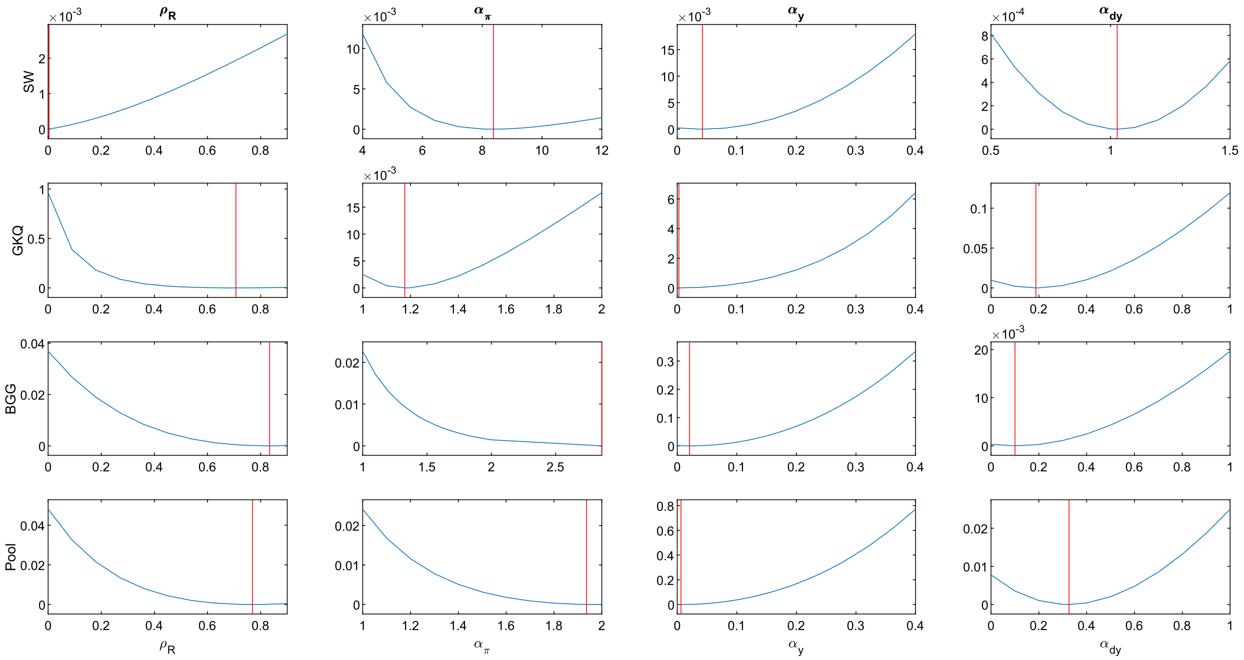

Fig. 6. The cost of bad policy I. *Note:* The graph shows the welfare cost of deviating from the optimal Taylor-rule parameters in the optimal robust rule. Each column shows the welfare cost of changing a single parameter in the Taylor-rule while keeping all other parameters constant. The first three rows show the welfare cost of deviating from the robust rule optimized for that particular model. The last row shows the welfare cost of deviating from the common robust rule in all three models at the same time optimized for the prediction pool. Welfare costs are measured in consumption equivalent welfare changes. The vertical red line shows the optimal value of the parameter that is changed in each column. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

sensitive welfare is to changes in the coefficients in the robust optimal simple rule in each model. Each panel in the graph shows the consumption equivalent variation in welfare when changing a single parameter in the robust optimal simple rule at a time in a given model. Each column shows the welfare consequences of changing a parameter in the Taylor rule, while each row shows the welfare costs of deviating from the optimal parameter values in a given model.

In the SW model welfare is quite flat around the optimal point along every dimension. Deviations have small effects and welfare is most sensitive to changes in the output feedback parameter α_y.[17] In the financial friction models deviations from the optimal value α_y still dominate the policymaker's objective: deviations can have large welfare consequences that can reach as much as a 5% consumption equivalent. However, welfare in those models is also more sensitive to deviations of the interest rate smoothing parameter, ρ_r, from its optimal value. The panels in the last row show the welfare cost of deviating from the optimal parameter values of the common robust rule optimized for the prediction pool in all three models at the same time. These panels are essentially weighted averages of the three panels above them using the weights assigned to each model in the prediction pool. The numbers here avoid the large welfare losses in the GKQ model due to the fact that the SW model receives the largest weight in the prediction pool.

To summarize this sub-section, the welfare gains from using the correct robust rather than non-robust rule in model i when model $j \neq i$ turns out to be the actual DGP are apparently quite low and of the order of a 0.137% consumption equivalent permanent loss at most. However using the robust optimal rule does avoid substantial losses from deviating slightly from the rule. In addition it should be noted that optimized simple rules *in general* have good robustness properties and using non-optimized rules can lead to substantial losses compared with all such we have considered in this paper. One such 'bad policy' is the non-optimized inertial Taylor rule examined in Hebden et al. (2020). In our parametrization it has parameter values: $\rho_r = 0.85$, $\alpha_\pi = (1-0.85)1.5 = 0.2250$, $\alpha_y = (1-0.85)0.5 = 0.0750$ and $\alpha_{dy} = 0$. With DGPs given by the estimated SW, GKQ and BGG models at the mean of estimates we find this rule actually decreases welfare compared with the estimated rule in all cases with substantial consumption equivalent percentage losses of up to 30% for the GKQ model which is close to instability under this rule.

6.5. A price level rule

Our results then show that even small deviations of ρ_r and α_y have serious welfare consequences with losses far greater than any gains relative to the estimated rule. So suppose that the monetary policymaker commits to a rule with $\rho_r = 1$ and $\alpha_y = \alpha_{dy} = 0$. Then integrating (16) and putting $\frac{\Pi_t}{\Pi} = \frac{P_t/P_{t-1}}{\bar{P}_t/\bar{P}_{t-1}}$ where $\bar{P}_t$ is the price trend in the constant inflation rate steady state, we arrive at the rule

[17] Schmitt-Grohe and Uribe (2007) termed this result the 'importance of not responding to output'.

Table 4
Optimized simple robust price level rules.

	SW	GKQ	BGG	BMA			Prediction Pool		
				SW	GKQ	BGG	SW	GKQ	BGG
	Taylor rule parameters								
Interest rate smoothing (ρ_r)	1.000	1.000	1.000		1.000			1.000	
Feedback on inflation (α_π)	4.911	1.218	2.608		1.219			1.695	
Feedback on output (α_y)	0.000	0.000	0.000		0.000			0.000	
Feedback on output growth (α_{dy})	0.000	0.000	0.000		0.000			0.000	
	Consumption equivalent welfare gain over the estimated rule								
Welfare gain in each model, $\omega^i(\%)$	0.090	0.022	0.114	0.079	0.017	0.112	0.080	0.016	0.112
Weighted average gain across models					0.023			0.056	
	Probability of hitting the zero lower bound								
Probability of ZLB (%)	3.469	0.689	2.750	0.734	1.102	1.931	0.788	1.175	1.993

Note: Simple price level rules use the restriction $\rho_r = 1$ and $\alpha_y = \alpha_{dy} = 0$ and only the value of α_π is optimized. ω_i shows the fraction of the consumption stream households need to give up to have the same welfare as under the estimated rule in each model. p_{ZLB} is the probability that the gross nominal interest rate is below one in the normal distribution $\mathcal{N}(R_N, \sigma_{R_N})$, where R_N is the deterministic steady state of the nominal interest rate.

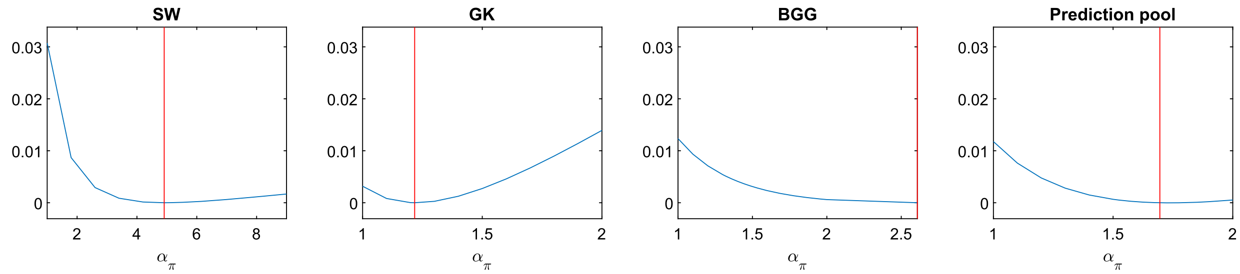

Fig. 7. The cost of bad policy II. *Note:* The graph shows the welfare cost of deviating from the optimal value of α_π in the optimal robust price level rule. Price level rules use the restriction $\rho_r = 1$ and $\alpha_y = \alpha_{dy} = 0$. The first three panel show the welfare cost of deviating from the robust price level rule optimized for that particular model. The last panel shows the welfare cost of deviating from the common robust price level rule in all three models at the same time optimized for the prediction pool. Welfare costs are measured in consumption equivalent welfare changes. The vertical red line shows the optimal value of the parameter that is changed in each column.

$$\log\left(\frac{R_{n,t}}{R_n}\right) = \alpha_\pi \log\left(\frac{P_t}{\bar{P}_t}\right) \qquad \Rightarrow \qquad \frac{R_{n,t}}{R_n} = \left(\frac{P_t}{\bar{P}_t}\right)^{\alpha_\pi} \tag{21}$$

which is a *price-level rule* that adjusts the deviation of the nominal interest rate to changes in the price level relative to its long-run trend.

We have described the literature on the benefits of price-level targeting versus inflation targeting in Section 2 where we describe price-targeting (and indeed any inertial Taylor rule) as possible makeup strategy as follows. It anchors expectations as follows: faced with of an unexpected temporary rise in inflation, price-level stabilization commits the policymaker to bring inflation below the target in subsequent periods. In contrast, with inflation targeting, the drift in the price level is accepted.

Our results in Table 4 and Fig. 7 indicate a further benefit of price-level targeting: when the robust rule is implemented, even with departures from its optimal setting of the single feedback parameter α_π that defines the policy, it remains robust across models with and without financial frictions.

7. Robustness

In this section we demonstrate that our benchmark results are not affected by the particular estimation strategy we have employed. To conserve space we have delegated detailed description of each exercise to the online Appendix. Specifically, we show that the following do not affect our results:

1. Alternative samples. We carry out the optimal weighting and policy exercises using the full sample. However, we can compute the optimal policy at each point in our sample using observations up to that date to show that the choice of sample size does not affect our results. This is essentially akin to using a dynamic weighting scheme as in Del Negro et al. (2016).
2. Unconventional monetary policy. We re-estimate all models using the shadow federal funds rate of Wu and Xia (2016) to obtain estimates that incorporate the effects of unconventional monetary policy near or at the zero lower bound. Again, results are very similar to those presented above.
3. Alternative formulation of the Taylor rule. We also estimate a version of the three models in our pool with an alternative formulation of the Taylor rule. As documented by Ascari et al. (2011) and Coibion and Gorodnichenko (2012), a Taylor rule

with a degree of interest rate smoothing of order 2 fits the data better. In all other respects the three models are identical to the ones used to obtain the benchmark results in the paper. We find that the data rejects this alternative formulation in favour of our benchmark formulation. The loglikelihood values for all three models are smaller and a likelihood ratio test accepts restricting the second order interest rate parameter to be zero with a p-value well below 1%.[18]

4. Alternative set of forecast variables.[19] We assign weights to our models based on how well they predict all seven observable variables jointly. One might argue, however, that central banks are mostly interested in targeting inflation and output growth, the two variables entering the Taylor rule in our model. For this reason we have also computed the prediction pool weights based on the three models' ability to forecast inflation and output growth only. In contrast to the benchmark results in our paper, the BGG model gets the highest weights for most of our sample period (Fig. 10). However, the SW model also gets a non-trivial weight. Hence, our benchmark results in the paper are robust to using a different set of forecast variables in the sense that more than one model gets a non-trivial weight in the prediction pool. We have also computed the optimized simple rule with these weights and we can conclude that our results are not due to the specific set of forecast variables used in our exercise (Table F6). The optimized simple rule is still characterized by very high values for the interest rate smoothing parameter and the feedback on inflation. If anything, this optimized rule is even closer to a price level targeting rule than the one in Table 2.

8. Conclusion and directions for future research

This paper studies the problem of designing robust simple rules when the policymaker has at her disposal a finite set of models, none of which are believed to be the true data generating process. We assign weights to models on the basis of the accuracy of their 4-period ahead forecasts rather than their in-sample fit, consistent with the forward-looking viewpoint of the policymaker. We study the robust optimal policy problem in the form of an optimized Taylor-type inertial nominal interest rate rule under this weighting scheme using three estimated models exemplifying the policymakers' uncertainty about how to incorporate financial frictions into the canonical DSGE model of Smets and Wouters (2007). In comparison with Bayesian model averaging, we find that our prediction pool choice has a significant impact on the robust optimized rule. We find that the price level rule has particularly good robustness properties.

Our approach provides a very general framework for the combination of models in a policy design problem. It only requires models to share the same policy instrument, to provide a k-period ahead predictive density given macro-economic data, and to have a welfare criterion to rank alternative policies. The models in the pool do not need to share the estimated parameter vector, nor even the observables; they can be nested as well as non-nested. Thus, the methodology can be applied to a wide range of macroeconomic models from mainstream DSGE, behavioural to agent-based, and indeed to other non-macroeconomic settings as long as these three requirements are met.

Within the DSGE framework, several open questions offer possible avenues for further research that address the limitations of our paper. The first concerns the *choice of model sets* used to address model uncertainty. Our choice is motivated by the fact that central banks most commonly use these models for policy analysis. However, other theories have been proposed in the academic literature. Following Brunnermeier and Sannikov (2014, 2016), a nascent literature has developed that incorporates a novel set of financial frictions in continuous-time macroeconomic models. This approach offers several key advantages, in particular, it permits a complete characterization of the economic dynamics instead of focusing on approximations around the steady state, as we do for all the models considered in this paper. But retaining our approach, we can extend our choice of models with financial frictions to include for example a housing sector as in Iacoviello (2005), Iacoviello and Neri (2010) and Adam and Woodford (2020). A model that combines firm-bank and bank-household frictions as in Rannenberg (2016) is another possible competing model. Our policy problem can also be extended to accommodate macro-prudential instruments where banks are subjected to a common capital, or equivalently leverage-ratio requirement (see, for example, Karmakar, 2016). An important line of future research would incorporate predictions from these models in our pooling and optimal policy exercises.

The second area concerns the *absence of non-linearities* in the models. In this paper one important aspect is the zero lower bound constraint on the nominal interest rate and the possibility of an occasionally binding constraint (OBC). Our estimated NK, BGG and GKQ models are linearized and the constraint is always binding. Solving non-linear models with an OBC is addressed in Dou et al. (2020) and Holden (2023). The results we have obtained in the paper examine the possibility of a hitting a zero-lower bound for monetary policy, they are not designed to minimize this outcome as in Levine et al. (2012) and Deák et al. (2020). In a recent paper Dordal-i-Carreras et al. (2016) demonstrate that a regime switching representation of risk premium shocks can generate a realistic distribution of such zero-lower bound durations in a standard New Keynesian model. Incorporating zero-lower bound considerations into our policy problem may reveal policy trade-offs that are not present in our results.

The third area concerns a possible *asymmetry of doubts* faced by policymakers and agents within the model. This heightened scope for expectation differences between private and public sectors is largely ignored by the literature. There is a need to admit cases where the private sector may also believe an *incorrect* state of the world defined in terms of a model variant or parameter draw. This case of *model-inconsistent expectations* needs to be factored into truly robust Bayesian rules.[20]

[18] We would like to thank an anonymous referee for suggesting this robustness exercise.

[19] We would like to thank an anonymous referee for suggesting this exercise.

[20] Frankel and Rockett (1988) study model-inconsistent expectations in a quite different sense: in a world of interdependent economies each with their own central banks, the latter may each believe in different models. A first attempt at computing a model-consistent Bayesian rule is to be found in Levine et al. (2008), but using an insufficiently large number of draws from the posterior distribution to combine within- and between model-robustness in a consistently Bayesian fashion.

Fourth, the use of *real-time data* when analysing the out-of-sample forecast performances of competing models to compute weights would be an interesting exercise. An application of density forecasting using real time data, but without model pooling, is provided by McAdam and Warne (2019).

Fifth, the *set of forecast variables* and the *forecast horizon used* to compute the model weights can be investigated further. One might argue that central banks are mostly interested in forecasting variables that enter their loss function; e.g. inflation, output growth, and the nominal interest rate. Moreover, we can compute model weights based on a different forecast horizon in each model in the pool, choosing the forecast horizon based on what is revealed by the policy in each model.

Finally, the section in Related Literature surveyed the recent literature on inertial Taylor-Type rules and "makeup strategies". The aim is to offset, at least in part, past departures of inflation from its long-run target. The stabilization benefits highlighted in our paper stem from their effect on expectations and hinge critically on the RE assumption and the credibility of commitment. Price-level targeting (PLT) is a particular type of make-up strategy. Our paper finds that a PLT rule has strong robustness properties across the three models irrespective of the choice of forecast variables and the forecasting period.[21] But this does assume RE, so a natural question is whether this result remains in a *non-RE environment* such as that studied in Bodenstein et al. (2019) and Arifovic et al. (2023). Deák et al. (2023) uses the methodology of this paper to design a robust Taylor-type monetary rule across a RE "workhorse" NK model and competing non-RE behavioural alternatives. The latter consist of two models with "Euler learning" and a bounded rational one with myopia. Future work could conduct a similar robustness exercise across two dimensions: models both with and without financial frictions and models with different expectations assumptions.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jedc.2025.105096.

References

Adalid, R., Coenen, G., McAdam, P., Siviero, S., 2005. The performance and robustness of interest-rate rules in models of the euro area. Int. J. Cent. Bank. 1 (1), 95–132.
Adam, K., Woodford, M., 2012. Robustly optimal monetary policy in a micro-founded new Keynesian model. J. Monet. Econ. 59, 468–487.
Adam, K., Woodford, M., 2020. Robustly Optimal Monetary Policy in a New Keynesian Model with Housing. CESifo Working Paper No. 8127.
Arifovic, J., Salle, I., Troung, H., 2023. History-Dependent Monetary Policy: a Lab Experiment and a HENK Model. Discussion Paper Series TI 2023-028/VI. Tinbergen Institute, Amsterdam.
Ascari, G., Castelnuovo, E., Rossi, L., 2011. Calvo vs. Rotemberg in a trend inflation world: an empirical investigation. J. Econ. Dyn. Control 35 (11), 1852–1867.
Batini, N., Justiniano, A., Levine, P., Pearlman, J., 2006. Robust inflation-forecast-based rules to shield against indeterminacy. J. Econ. Dyn. Control 30 (9–10), 1491–1526.
Bernanke, B., 2017. Monetary policy in a new era. Technical report. Brookings Institution.
Bernanke, B., Gertler, M., Gilchrist, S., 1999. The financial accelerator in quantitative business cycles. In: Woodford, M., Taylor, J.B. (Eds.), Handbook of Macroeconomics, vol. 1C. Elsevier Science, Amsterdam.
Binder, M., Lieberknecht, P., Quintana, J., Wieland, V., 2017. Model Uncertainty in Macroeconomics: on the Implications of Financial Frictions. C.E.P.R. Discussion Papers and Chapter 23 of The Oxford Handbook of the Economics of Central Banking, 2019, Edited by Mayes, D. G. and Siklos, P. L. and Sturm, J-E.
Binder, M., Lieberknecht, P., Quintana, J., Wieland, V., 2018. Robust macroprudential policy rules under model uncertainty. In: Annual Conference 2018, German Economic Association.
Bodenstein, M., Hebden, J., Winkler, F., 2019. Learning and misperception: implications for price-level targeting. In: Finance and Economics Discussion Series 2019-078, Board of Governors of the Federal Reserve System (U.S.).
Brock, W.A., Durlauf, S.N., West, K.D., 2007a. Model uncertainty and policy evaluation: some theory and empirics. J. Econom. 127 (2), 629–664.
Brock, W.A., Durlauf, S.N., Nason, J.M., Rondina, G., 2007b. Simple versus optimal rules as guides to policy. J. Monet. Econ. 54 (5), 1372–1396.
Brunnermeier, M., Sannikov, Y., 2016. The I Theory of Money. CEPR Discussion Papers 11444, C.E.P.R. Discussion Papers.
Brunnermeier, M.K., Sannikov, Y., 2014. A macroeconomic model with a financial sector. Am. Econ. Rev. 104 (2), 379–421.
Brzoza-Brzezina, M., Kolasa, M., 2013. Bayesian evaluation of DSGE models with financial frictions. J. Money Credit Bank. 45 (8), 1451–1476.
Brzoza-Brzezina, M., Kolasa, M., Makarski, K., 2013. The anatomy of standard DSGE models with financial frictions. J. Econ. Dyn. Control 37 (1), 32–51.
Cogley, T., Sargent, T., 2005. The conquest of U.S. inflation: learning and robustness to model uncertainty. Rev. Econ. Dyn. 8 (2), 528–563.
Coibion, O., Gorodnichenko, Y., 2012. Why are target interest rate changes so persistent? Am. Econ. J. Macroecon. 4 (4), 126–162.
Côté, D., Kuszczak, J., Lam, J.-P., Liu, Y., St-Amant, P., 2004. The performance and robustness of simple monetary policy rules in models of the Canadian economy. Can. J. Econ. 37 (4), 978–998.
de Groot, O., 2014. The risk channel of monetary policy. Int. J. Cent. Bank. 10 (2), 115–160.
Deák, S., Levine, P., Pham, S., 2020. Mandates and Monetary Rules in a New Keynesian Framework. School of Economics Discussion Papers DP0l/20. University of Surrey.
Deák, S., Levine, P., Mirza, A., Pham, S., 2023. Negotiating the Wilderness of Bounded Rationality through Robust Policy. School of Economics Discussion Papers 02/23. University of Surrey.
Deak, S., Levine, P., Pham, S.T., 2023. Simple mandates, monetary rules and trend inflation. Macroecon. Dyn. https://doi.org/10.1017/S136510052300024X. Published Online, 7 July, 2023.
Del Negro, M., Hasegawa, R., Schorfheide, F., 2016. Dynamic prediction pools: an investigation of financial frictions and forecasting performance. J. Econom. 192, 391–405.
Dennis, R., Leitemo, K., Soderstrom, U., 2009. Methods for robust control. J. Econ. Dyn. Control 33 (8), 1604–1616.
Dordal-i-Carreras, M., Coibion, O., Gorodnichenko, Y., Wieland, J., 2016. Infrequent but Long-Lived Zero-Bound Episodes and the Optimal Rate of Inflation. NBER Working Papers 22510. National Bureau of Economic Research, Inc.
Dou, W., Lo, A.M., Muley, A., Uhlig, H., 2020. Macroeconomic models for monetary policy: a critical review from a finance perspective. Annu. Rev. Econ. 12, 95–140.
Ellison, M., Sargent, T.J., 2012. A defense of the FOMC. Int. Econ. Rev. 53 (4), 1047–1065.

[21] Earlier versions of the paper used one period ahead forecasting performance to determine weights.

Frankel, J., Rockett, K.E., 1988. International macroeconomic policy coordination when policymakers do not agree on the true model. Am. Econ. Rev. 78 (3), 318–340.
Gali, J., Gertler, M., Lopez-Salido, J.D., 2007. Markups, gaps, and the welfare costs of business fluctuations. Rev. Econ. Stat. 89 (1), 45–59.
Gaspar, V., Smets, F., Vestin, D., 2010. Is time ripe for price level path stability? In: Siklos, P.L., Bohl, M.T., Wohar, M.E. (Eds.), Challenges in Central Banking: the Current Institutional Environment and Forces Affecting Monetary Policy. Cambridge University Press.
Gertler, M., Karadi, P., 2011. A model of unconventional monetary policy. J. Monet. Econ. 58 (1), 17–34.
Gertler, M., Kiyotaki, N., Queralto, A., 2012. Financial crises, bank risk exposure and government financial policy. J. Monet. Econ. 59 (S), 17–34.
Geweke, J., Amisano, G., 2011. Optimal prediction pools. J. Econom. 164 (1), 130–141.
Geweke, J., Amisano, G., 2012. Prediction with misspecified models. Am. Econ. Rev. 102 (3), 482–486.
Giannoni, M., 2014. Optimal interest-rate rules and inflation stabilization versus price-level stabilization. J. Econ. Dyn. Control 41 (C), 110–129.
Hansen, L., Sargent, T.J., 2007. Robustness. Princeton University Press.
Hebden, J., Herbst, E.P., Tang, J., Topa, G., Winkler, F., 2020. How Robust Are Make-up Strategies to Key Alternative Assumptions? Finance and Economics Discussion Series 2020-069. Federal Reserve Board.
Ho, P., Lubik, T., Matthes, C., 2024. Averaging impulse responses using prediction pools. J. Monet. Econ. 146.
Holden, T., 2016. Existence and uniqueness of solutions to dynamic models with occasionally binding constraints. EconStor Preprints 130142, ZBW - German National Library of Economics.
Holden, T., 2023. Existence and uniqueness of solutions to dynamic models with occasionally binding constraints. Rev. Econ. Stat. 105 (6), 1481–1499.
Iacoviello, M., 2005. House prices, borrowing constraints, and monetary policy in the business cycle. Am. Econ. Rev. 95 (3), 739–764.
Iacoviello, M., Neri, S., 2010. Housing market spillovers: evidence from an estimated DSGE model. Am. Econ. J. Macroecon. 2 (2), 125–164.
Karmakar, S., 2016. Macroprudential regulation and macroeconomic activity. J. Financ. Stab. 25 (C), 166–178.
Levin, A., Wieland, V., Williams, J.C., 2003. The performance of inflation forecast-based rules under model uncertainty. Am. Econ. Rev. 93, 622–645.
Levin, A., Onatski, A., Williams, J.C., Williams, N., 2006. Monetary policy under uncertainty in micro-founded macroeconomic models. In: Gerler, M., Rogoff, K. (Eds.), NBER Macroeconomics Annual. 2005, pp. 229–387.
Levine, P., Currie, D.A., 1987. The design of feedback rules in linear stochastic rational expectations models. J. Econ. Dyn. Control 11, 1–28.
Levine, P., Pearlman, J., 2010. Robust monetary rules under unstructured and structured uncertainty. J. Econ. Dyn. Control 34 (3), 456–471.
Levine, P., McAdam, P., Pearlman, J., Pierse, R., 2008. Risk Management in Action: Robust Monetary Rules Under Structured Uncertainty. ECB Working Paper No. 870.
Levine, P., McAdam, P., Pearlman, J., 2012. Probability models and robust policy rules. Eur. Econ. Rev. 56 (2), 246–262.
Lucas, R.E., 1987. Models of Business Cycles. Basil Blackwell, Oxford.
Lucas, R.E., 2003. Macroeconomic priorities. Am. Econ. Rev. 93, 1–14.
McAdam, P., Warne, A., 2019. Euro area real-time density forecasting with financial or labor market frictions. Int. J. Forecast. 35 (2), 580–600. https://doi.org/10.1016/j.ijforecast.2018.10.013.
Orphanides, A., Williams, J.C., 2008. Learning, expectations formation, and the pitfalls of optimal control monetary policy. J. Monet. Econ. 55S, S80–S96.
Powell, J.H., 2020. New Economic Challenges and the Fed's Monetary Policy Review. Speech Given at "Navigating the Decade Ahead: Implications for Monetary Policy," an Economic Policy Symposium Sponsoredthe Federal Reserve Bank of Kansas City. Jackson Hole.
Ramey, V.A., Zubairy, S., 2018. Government spending multipliers in good times and in bad: evidence from US historical data. J. Polit. Econ. 126 (2), 850–901.
Rannenberg, A., 2016. Bank leverage cycles and the external finance premium. J. Money Credit Bank. 48 (8), 1477–1479.
Reiss, R., 2009. Optimal monetary policy rules in an estimated sticky-information model. Am. Econ. J. Macroecon. 1 (2), 1–28.
Schmitt-Grohe, S., Uribe, M., 2000. Price level determinacy and monetary policy under a balanced-budget requirement. J. Monet. Econ. 45, 211–246.
Schmitt-Grohe, S., Uribe, M., 2007. Optimal simple and implementable monetary and fiscal rules. J. Monet. Econ. 54 (6), 1702–1725.
Sims, C.A., 2002. The role of models and probabilities in the monetary policy process. Brookings Pap. Econ. Act. 2, 1–62.
Sims, C.A., 2007. Monetary Policy Rules. CEPS Working Paper No. 155.
Sims, C.A., 2008. Improving monetary policy models. J. Econ. Dyn. Control 32 (8), 2460–2475.
Smets, F., Wouters, R., 2007. Shocks and frictions in US business cycles: a Bayesian DSGE approach. Am. Econ. Rev. 97 (3), 586–606.
Svensson, L.E., 1999. Price level targeting versus inflation targeting. J. Money Credit Bank. 31, 277–295.
Svensson, L.E.O., 2020. Monetary Policy Strategies for the Federal Reserve.
Taylor, J.B., Williams, J.C., 2010. Simple and robust rules for monetary policy. In: Friedman, B.M., Woodford, M. (Eds.), Handbook of Monetary Economics. In: Handbook of Monetary Economics, vol. 3. Elsevier, pp. 829–859. chapter 15.
Vehtari, A., Ojanen, J., 2012. A survey of Bayesian predictive methods for model assessment, selection and comparison. Stat. Surv. 6, 142–228.
Vestin, D., 2006. Price level targeting versus inflation targeting. J. Monet. Econ. 53, 1361–1376.
Wieland, J.F., Yang, M.J., 2020. Financial dampening. J. Money Credit Bank. 52 (1), 79–113.
Wu, J.C., Xia, F.D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. J. Money Credit Bank. 48 (2–3), 253–291.