

Dynamic Prediction Pools: An Investigation of Financial Frictions and Forecasting Performance

Marco Del Negro Raiden B. Hasegawa Frank Schorfheide*
Federal Reserve Bank Federal Reserve Bank Univ. of Pennsylvania
of New York of New York CEPR and NBER

Preliminary Version: November 22, 2013

Abstract

We provide a methodology for estimating time-varying weights in optimal prediction pools, building on the work by Geweke and Amisano (2011, 2012). We use this methodology to investigate the relative forecasting performance of DSGE models with and without financial frictions from 1992 to 2011 at various forecast horizons. Our results indicate that models with financial frictions produce superior forecasts in periods of financial distress, which may partly explain why macroeconomists may have neglected models with financial frictions prior to the Great Recession. At the same time we show that even in ‘tranquil’ periods the weight on the financial friction model is far from zero, and that using real time estimates this weight increases before the Lehman crisis.

JEL CLASSIFICATION: C52, E31, E32, E37.

KEY WORDS: Optimal pools, Forecasting, Great recession, DSGE models, Bayesian estimation.

*Correspondence: Marco Del Negro (marco.delnegro@ny.frb.org), Raiden B. Hasegawa (Raiden.Hasegawa@ny.frb.org): Research Department, Federal Reserve Bank of New York, 33 Liberty Street, New York NY 10045. Frank Schorfheide (schorf@ssc.upenn.edu): Department of Economics, 3718 Locust Walk, University of Pennsylvania, Philadelphia, PA 19104-6297. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant SES 1061725. The views expressed in this paper do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

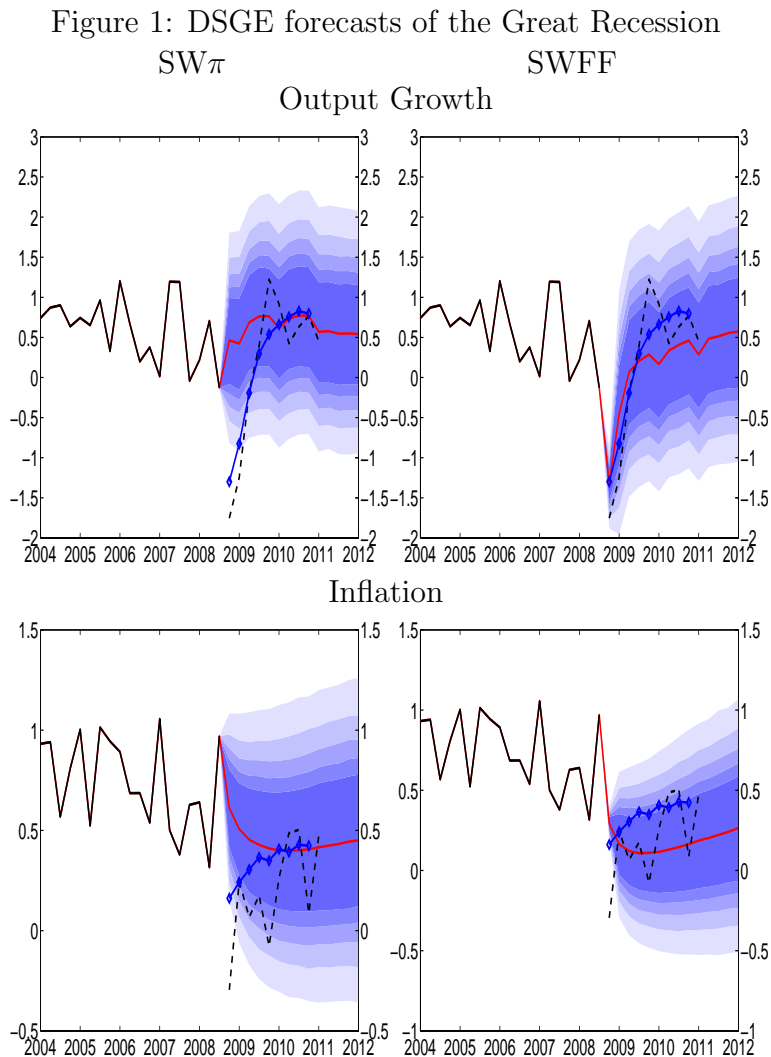
1 Introduction

Many macroeconomists paid scant attention to financial frictions models before the recent crisis. The workhorse DSGE models at several central banks was the model developed by [Smets and Wouters \(2007\)](#) (henceforth, SW), which has no endogenous credit frictions and uses no financial market data other than the fed funds rate as an observable. Yet DSGE models with frictions were available before the crisis – notable examples are [Bernanke, Gertler, and Gilchrist \(1999\)](#), [Kiyotaki and Moore \(1997\)](#), and [Christiano, Motto, and Rostagno \(2003\)](#). [Del Negro and Schorfheide \(2013\)](#) show that a variant of the SW model incorporating financial frictions as in [Bernanke, Gertler, and Gilchrist \(1999\)](#) and [Christiano, Motto, and Rostagno \(2003\)](#) would have forecasted reasonably well the behavior of output growth and inflation in aftermath of the Lehman crisis.

The left panels of [Figure 1](#) show forecasts of the output growth and inflation obtained in the aftermath of the Lehman crisis using a model without financial frictions.¹ The figure shows that this model would have been clueless about what was to come. The right panels show forecasts obtained using the variant with financial frictions, which can take advantage of contemporaneous information coming from financial market spreads. The lesson appears to be that while the available batch of DSGE models would have certainly not forecasted the occurrence of the Lehman crisis, those DSGE models with financial frictions might have provided policymakers with a reasonable outlook for the economy in the aftermath of the crisis. To our knowledge, however, they were not used. With hindsight, one question is: Why not? Did applied macroeconomists in the pre-Lehman period have any reason not to use these models for forecasting? For the future, when there are many models on the table, how should forecasters combine them? If another Lehman is to occur, would we be able to rely on the right model, or on the right combination of models?

We develop a novel methodology to try to address both questions. Our approach builds on the work on forecast combinations by [Geweke and Amisano \(2011\)](#). These authors use the time series of predictive densities (which measure the likelihood of ex post outcomes from the perspective of a model’s ex ante forecast distribution) to construct “optimal pools”, that is, convex combinations of forecasts where the weights are chosen to maximize past forecast performance. If one thinks of models as stocks, and of predictive densities as stock returns,

¹The models and the data on which they are estimated are described in detail in [Section 3](#).



Notes: The figure is taken from Del Negro and Schorfheide (2013). The panels show for each model/vintage the available real GDP growth (upper panel) and inflation (GDP deflator, lower panel) data (black line), the DSGE model’s multi-step (fixed origin) mean forecasts (red line) and bands of its forecast distribution (shaded blue areas; these are the 50, 60, 70, 80, and 90 percent bands, in decreasing shade), the Blue Chip forecasts (blue diamonds), and finally the actual realizations according to the May 2011 vintage (black dashed line) . All the data are in percent, Q-o-Q.shows the filtered mean of λ_t (solid black line) and the 50% , 68% and 90% bands in shades of blue.

the Geweke and Amisano approach can be seen as choosing the weights so to optimize the portfolio’s historical performance. Geweke and Amisano show that because of the benefits from diversification, these pools fare much better in a pseudo-out-of-sample forecasting exercise than “putting all your eggs in one basket” – that is, using only one model to forecast – as well as forecasts combinations based on Bayesian Model Averaging (BMA).

We take the Geweke and Amisano approach and make it time-varying. That is, we postulate a law of motion for the weights respecting the constraint that they remain between zero and one (no model can be short-sold) and use the period by period predictive density of the resulting pool as an observation equation. We obtain a non-linear state space model and conduct inference about the weights using a particle filter. We apply our approach to a two model setting, although it can be extended to the case with multiple models (easily from a conceptual point of view, more challengingly from a computational point of view). From the estimated distributions of the weights we learn whether one model was forecasting better than the other in specific sub periods, how persistent forecast differences are, and whether the weights change rapidly when the economic environment evolves.

Our application focuses on forecasts of four quarters ahead average output growth and inflation obtained from two variants of the SW model, one without (which we call $SW\pi$ to distinguish it from the original SW) and one with financial frictions (which we call SWFF). We find that the relative forecasting performance of the two DSGE models varies considerably over time. During ‘tranquil’ periods, namely periods without significant financial distress such as the mid and late 1990s, the model without financial frictions forecasts better than the alternative. In these period the logarithm of the predictive density, also known as the log score, is uniformly higher for $SW\pi$ and, as a consequence, the distribution of the weights in our dynamic pool is shifted toward this model. However, this distribution is still fairly wide, and puts a non-negligible weight on the SWFF model.

During periods of financial turmoil, the early 2000s and the more recent crisis, the ranking is not surprisingly the opposite. The difference is that the loss in log score from using the $SW\pi$ model in periods of turmoil is much larger than the reverse loss. We find that when the economic environment changes the filtered distribution of the weight in the dynamic pools shifts quite rapidly from one model to the other. The shift seems to be timely as well: by the time the Lehman crisis struck, the filtered distribution had shifted up relative to the pre-crisis period (2006) and was putting considerable mass on SWFF. We conclude that macroeconomists had some reason for not relying exclusively on the model with financial frictions before the crisis. They (and we) had no reason for not relying on it at all, however, given that this implies forfeiting very large benefits from diversification. For the future, the dynamic pool still places most of the weight on the SWFF model, implicitly indicating that the US economy is still undergoing a period of financial stress. Our (pseudo) real time

forecasting results with the dynamic pools suggest that this method may deserve further study.

Our paper is related to several strands of the literature. There is a large body of work on forecast combination dating back to a seminal paper by [Bates and Granger \(1969\)](#). Much of that literature focuses on the combination of point forecasts, whereas we are combining predictive densities using linear pools – building on recent research by [Geweke and Amisano \(2011\)](#). More general methods of combining predictive distributions that include beta-transformations of cumulative densities of linear prediction pools are discussed in [Gneiting and Ranjan \(2013\)](#).

Our paper emphasizes the use of time-varying combination weights. Time-varying weights have been used in the forecast combination literature, e.g., [Terui and van Dijk \(2002\)](#) and [Guidolin and Timmermann \(2009\)](#). With respect to the combination of predictive densities, [Waggoner and Zha \(2012\)](#) extend the Geweke and Amisano approach to a setting in which the combination weights follow a Markov switching process. There are two important differences between our dynamic linear pools and the work by Waggoner and Zha. First, instead of using a Markov-switching process for the combination weights we are using a smoother autoregressive process. Second, we do not attempt to estimate the model parameters and the combination weights simultaneously. Our empirical application focuses on the combination of DSGE models. In practical settings, e.g., forecasting at policy institutions, we think that it is unrealistic (and to some extent undesirable) that the models are being re-estimated when being pooled.²

[Billio, Casarin, Ravazzolo, and van Dijk \(2012\)](#) propose create a predictive density by combining (point) predictors derived from different models using time-varying weights. The evolution of weights may depend on past forecasting performance. Our approach relies on a combination of model-implied predictive density and the period-by-period updating of the combination weights is based on Bayes Theorem. We use a fairly standard particle filter to extract the time-varying combination weights. Recent survey of particle-filtering methods for nonlinear state-space models are provided by [Giordani, Pitt, and Kohn \(2011\)](#) and [Creal \(2012\)](#).

The remainder of this paper is organized as follows. Section 2 presents the methodology. Section 3 describes the models and the data on which they are estimated. Section 4 presents

²In fact, many central banks do not even re-estimate their DSGE models every quarter.

the results and Section 5 concludes.

2 Linear Pools and Time Variation

We review the idea of optimal linear prediction pools as presented in [Geweke and Amisano \(2011\)](#) in Section 2.1 and provide a comparison to Bayesian Model Averaging in Section 2.2. In Section 2.3 we extend the static prediction pools to dynamic prediction pools by introducing time-varying combination weights for predictive densities. Multi-step forecasting is discussed in Section 2.4.

2.1 The Static Case

Starting point for our analysis are two models ($\mathcal{M}_1, \mathcal{M}_2$) with potentially different time- $t-1$ information sets ($\mathcal{I}_{t-1}^{\mathcal{M}_m}$) and one-step ahead predictive densities $p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$, where $y_t, t = 1, \dots, T$ are the variables of interest. Let $y_{1:t}$ denote the sequence $\{y_1, \dots, y_t\}$ and note that for each of the two models the information set $\mathcal{I}_{t-1}^{\mathcal{M}_m}$ may in general be larger than $y_{1:t-1}$. For instance, the (model-specific) information sets may include additional variables $z_{1:t-1}^m$ such that $\mathcal{I}_{t-1}^{\mathcal{M}_m} = \{y_{1:t-1}, z_{1:t-1}^m\}$ and

$$p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m) = p(y_t|y_{1:t-1}, z_{1:t-1}^m, \mathcal{M}_m).$$

In our applications, y_t includes output growth and inflation, z_t includes consumption, investment, hours per capita, real wage growth, the federal funds rate, long-run inflation expectations, and, for one of the models, also spreads.

We now combine models \mathcal{M}_1 and \mathcal{M}_2 by creating a convex combination of the one-step-ahead predictive densities. Such a combination is called a linear prediction pool:

$$p(y_t|\mathcal{I}_{t-1}, \lambda) = \lambda p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_1}, \mathcal{M}_1) + (1 - \lambda) p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_2}, \mathcal{M}_2), \quad \lambda \in [0, 1]. \quad (1)$$

The expression

$$p_1(y_{1:T}|\lambda) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-1}, \lambda) \quad (2)$$

can be interpreted as the likelihood of the pool because it represents the joint density of $y_{1:T}$ conditional on the combination weight λ . The “1” subscript indicates that it is a combination

of one-step-ahead predictive densities. If the combination weight λ is time-invariant then the prediction pool is static. Following Geweke and Amisano (2011, 2012), we define the optimal (static) pool computed given information at time T as:³

$$\lambda_T^{SP} = \operatorname{argmax}_{\lambda \in [0,1]} p_1(y_{1:T}|\lambda). \quad (3)$$

2.2 Linear Pools versus Bayesian Model Averaging

It is important to note that the (static) pooling of models is very different from Bayesian model averaging (BMA). To fix ideas, consider the case in which $\mathcal{I}_{t-1}^{\mathcal{M}_m} = y_{1:t-1}$ for $m = 1, 2$. In this case, the product of the one-step ahead predictive densities equals the marginal likelihood (see Geweke (2005) and Geweke (2007)):

$$p_1(y_{1:T}|\mathcal{M}_m) = \prod_{t=1}^T p(y_t|y_{1:t-1}, \mathcal{M}_m), \quad (4)$$

where

$$p(y_t|y_{1:t-1}, \mathcal{M}_m) = \int p(y_t|y_{1:t-1}, \theta, \mathcal{M}_m) p(\theta|y_{1:t-1}, \mathcal{M}_m) d\theta,$$

and $p(\theta|y_{1:t-1}, \mathcal{M}_m)$ is the posterior for model \mathcal{M}_m 's parameter vector θ based on the information $y_{1:t-1}$. BMA involves assigning prior probabilities to \mathcal{M}_1 and \mathcal{M}_2 , e.g., let λ_0 be the prior probability associated with \mathcal{M}_1 , and weighting the models by their respective posterior probabilities. The posterior probability of \mathcal{M}_1 is a function of the marginal likelihoods $p_1(y_{1:T}|\mathcal{M}_m)$:

$$\lambda_T^{BMA} = \frac{\lambda_0 p_1(y_{1:T}|\mathcal{M}_1)}{\lambda_0 p_1(y_{1:T}|\mathcal{M}_1) + (1 - \lambda_0) p_1(y_{1:T}|\mathcal{M}_2)} \quad (5)$$

A comparison of (5) and (3) highlights that the weights of the optimal static pool and BMA are different. As discussed in Geweke and Amisano (2011) unless the two models forecast equally well in the sense that the difference in log predictive scores

$$|\ln p_1(y_{1:T}|\mathcal{M}_1) - \ln p_1(y_{1:T}|\mathcal{M}_2)| = O_p(1)$$

is stochastically bounded, the BMA model weight λ_T^{BMA} will either converge to zero or one almost surely, that is, asymptotically there is no model averaging. Under the pooling

³See Amisano and Geweke (2013) for an interesting application to pools consisting of different macroeconomic models, such as VARs, DSGEs, and factor models.

approach, we are creating a convex combination of predictive densities, which may forecast better than any of its components. Thus, unless the data are actually generated from either \mathcal{M}_1 and \mathcal{M}_2 , λ_T^{SP} typically does not converge to either zero or one asymptotically.

In the subsequent empirical application we compare the forecast performance of model pools to the performance of model averages. Because the DSGE models used in the application are estimated based on different information sets which contain model-specific variables z_t^m , we replace the marginal likelihoods (4) by the product of one-step-ahead predictive scores:

$$p_1(y_{1:T}|\mathcal{M}_m) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m) \quad (6)$$

when computing the BMA weights in (5). In the remainder of this section we extend the pooling approach to the dynamic case of time-varying weights and we consider multi-step-ahead forecasts.

2.3 The Dynamic Case

In view of structural changes in the macro economy over the past six decades it is conceivable that the optimal pooling weights also evolve over time. Thus, we proceed by replacing λ with the sequence $\lambda_{1:T} = \{\lambda_1, \dots, \lambda_T\}$. Using the same definition of the $p(y_t|\mathcal{I}_{t-1}, \lambda)$ as in (1), we write the likelihood associated with the dynamic pool as

$$p_1(y_{1:T}|\lambda_{1:T}) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-1}, \lambda_t). \quad (7)$$

Note that a straight maximization with respect to $\lambda_{1:T}$ would yield the uninteresting corner solutions

$$\hat{\lambda}_t = \begin{cases} 1 & \text{if } p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_1}, \mathcal{M}_1) > p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_2}, \mathcal{M}_2) \\ 0 & \text{if } p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_1}, \mathcal{M}_1) < p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_2}, \mathcal{M}_2) \end{cases}, \quad t = 1, \dots, T.$$

In general, we would expect the optimal combination weights to shift slowly over time and exhibit a substantial amount of persistence. We follow the literature on Bayesian non-parametric function estimation (see, e.g., the survey by [Griffin, Quintana, and Steel \(2011\)](#)) and impose a stochastic-process prior on $\lambda_{1:T}$ that implies a “smooth” evolution over time.

We let x_t be a Gaussian autoregressive process with autocorrelation ρ and time-invariant marginal distributions $x_t \sim N(0, 1)$. Moreover, we use a probability-integral transformation

to map x_t onto the unit interval and ensure that the marginal prior distribution of λ_t is uniform. This leads to

$$\begin{aligned} x_t &= \rho x_{t-1} + \sqrt{1 - \rho^2} \varepsilon_t, \quad \varepsilon_t \sim iid N(0, 1), \quad x_0 \sim N(0, 1), \\ \lambda_t &= \Phi(x_t) \end{aligned} \tag{8}$$

where $\Phi(\cdot)$ is the cumulative Gaussian distribution. For $\rho = 1$ this specification nests the one in Geweke and Amisano (2011), given that it imposes the restriction $\lambda_t = \lambda$, all t . For $\rho = 0$ the λ_t are independent from one another.

In the empirical application we will condition on a value of ρ and conduct posterior inference with respect to $\lambda_{1:T}$. Our dynamic model pool can be viewed as a nonlinear state-space model: expression (8) describes the state transition equation and the convex combination of predictive densities in (1) with λ replaced by λ_t is the measurement equation. We use a particle filter to approximate the sequence of densities $p(\lambda_t|y_{1:t})$ with the understanding that $\mathcal{I}_{t-1}^{\mathcal{M}_m}$ affects the inference with respect to λ_t only indirectly, through the model-based predictive densities $p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$. The particle filter and the evaluation of $p(y_t|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$ is described in detail in the appendix.

2.4 Multi-Step Forecasting

In our application we are also interested in multi-step-ahead forecasts. Since the optimal combination weights may vary across horizons, we simply replace the one-step-ahead densities in (1) by h -step-ahead densities:

$$p(y_t|\mathcal{I}_{t-h}, \lambda) = \lambda p(y_t|\mathcal{I}_{t-h}^{\mathcal{M}_1}, \mathcal{M}_1) + (1 - \lambda) p(y_t|\mathcal{I}_{t-h}^{\mathcal{M}_2}, \mathcal{M}_2).$$

Moreover, we replace $p_1(y_{1:T}|\lambda)$ and $p_1(y_{1:T}|\mathcal{M}_m)$ that appear in (2) and (5) by

$$p_h(y_{1:T}|\lambda) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-h}, \lambda), \quad p_h(y_{1:T}|\mathcal{M}_m) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-h}^{\mathcal{M}_m}, \mathcal{M}_m).$$

Finally, for the dynamic pool we define

$$p_1(y_{1:T}|\lambda_{1:T}) = \prod_{t=1}^T p(y_t|\mathcal{I}_{t-1}, \lambda_t).$$

In the empirical analysis y_t is composed of growth rates of output and the price level. The object of interest is typically not the growth rate h periods from now but the average growth

rate over the next h periods. Thus, for multi-step forecasting, we change the argument of the predictive density from y_t to $\bar{y}_{t,h} = \frac{1}{h} \sum_{s=0}^{h-1} y_{t-s}$.

3 The DSGE Models

The model considered is the one used in [Smets and Wouters \(2007\)](#), which is based on earlier work by [Christiano, Eichenbaum, and Evans \(2005\)](#) and [Smets and Wouters \(2003\)](#). It is a medium-scale DSGE model, which augments the standard neoclassical stochastic growth model with nominal price and wage rigidities as well as habit formation in consumption and investment adjustment costs. As discussed before, the model is augmented with financial frictions, as in [Bernanke, Gertler, and Gilchrist \(1999\)](#), [Christiano, Motto, and Rostagno \(2003\)](#), and [Christiano, Motto, and Rostagno \(forthcoming\)](#). All ingredients of the model were however publicly available prior to 2008. As such, the model does not include some of the features that may have been found to be relevant following the crisis.

3.1 The Smets-Wouters Model

We begin by briefly describing the log-linearized equilibrium conditions of the [Smets and Wouters \(2007\)](#) model. We follow [Del Negro and Schorfheide \(2013\)](#) and detrend the non-stationary model variables by a stochastic rather than a deterministic trend.⁴ Let \tilde{z}_t be the linearly detrended log productivity process which follows the autoregressive law of motion

$$\tilde{z}_t = \rho_z \tilde{z}_{t-1} + \sigma_z \varepsilon_{z,t}. \quad (9)$$

We detrend all non stationary variables by $Z_t = e^{\gamma t + \frac{1}{1-\alpha} \tilde{z}_t}$, where γ is the steady state growth rate of the economy. The growth rate of Z_t in deviations from γ , denoted by z_t , follows the process:

$$z_t = \ln(Z_t/Z_{t-1}) - \gamma = \frac{1}{1-\alpha} (\rho_z - 1) \tilde{z}_{t-1} + \frac{1}{1-\alpha} \sigma_z \varepsilon_{z,t}. \quad (10)$$

All variables in the following equations are expressed in log deviations from their non-stochastic steady state. Steady state values are denoted by *-subscripts and steady state

⁴This approach makes it possible to express almost all equilibrium conditions in a way that encompasses both the trend-stationary total factor productivity process in [Smets and Wouters \(2007\)](#), as well as the case where technology follows a unit root process.

formulas are provided in the technical appendix of [Del Negro and Schorfheide \(2013\)](#).⁵ The consumption Euler equation is given by:

$$c_t = -\frac{(1 - he^{-\gamma})}{\sigma_c(1 + he^{-\gamma})} (R_t - \mathbb{E}_t[\pi_{t+1}] + b_t) + \frac{he^{-\gamma}}{(1 + he^{-\gamma})} (c_{t-1} - z_t) \\ + \frac{1}{(1 + he^{-\gamma})} \mathbb{E}_t [c_{t+1} + z_{t+1}] + \frac{(\sigma_c - 1)}{\sigma_c(1 + he^{-\gamma})} \frac{w_* L_*}{c_*} (L_t - \mathbb{E}_t[L_{t+1}]), \quad (11)$$

where c_t is consumption, L_t is labor supply, R_t is the nominal interest rate, and π_t is inflation. The exogenous process b_t drives a wedge between the intertemporal ratio of the marginal utility of consumption and the riskless real return $R_t - \mathbb{E}_t[\pi_{t+1}]$, and follows an AR(1) process with parameters ρ_b and σ_b . The parameters σ_c and h capture the degree of relative risk aversion and the degree of habit persistence in the utility function, respectively. The following condition expresses the relationship between the value of capital in terms of consumption q_t^k and the level of investment i_t measured in terms of consumption goods:

$$q_t^k = S'' e^{2\gamma} (1 + \bar{\beta}) \left(i_t - \frac{1}{1 + \bar{\beta}} (i_{t-1} - z_t) - \frac{\bar{\beta}}{1 + \bar{\beta}} \mathbb{E}_t [i_{t+1} + z_{t+1}] - \mu_t \right), \quad (12)$$

which is affected by both investment adjustment cost (S'' is the second derivative of the adjustment cost function) and by μ_t , an exogenous process called the “marginal efficiency of investment” that affects the rate of transformation between consumption and installed capital (see [Greenwood, Hercowitz, and Krusell \(1998\)](#)). The exogenous process μ_t follows an AR(1) process with parameters ρ_μ and σ_μ . The parameter $\bar{\beta} = \beta e^{(1-\sigma_c)\gamma}$ depends on the intertemporal discount rate in the utility function of the households β , the degree of relative risk aversion σ_c , and the steady-state growth rate γ .

The capital stock, \bar{k}_t , evolves as

$$\bar{k}_t = \left(1 - \frac{i_*}{\bar{k}_*} \right) (\bar{k}_{t-1} - z_t) + \frac{i_*}{\bar{k}_*} i_t + \frac{i_*}{\bar{k}_*} S'' e^{2\gamma} (1 + \bar{\beta}) \mu_t, \quad (13)$$

where i_*/\bar{k}_* is the steady state ratio of investment to capital. The arbitrage condition between the return to capital and the riskless rate is:

$$\frac{r_*^k}{r_*^k + (1 - \delta)} \mathbb{E}_t [r_{t+1}^k] + \frac{1 - \delta}{r_*^k + (1 - \delta)} \mathbb{E}_t [q_{t+1}^k] - q_t^k = R_t + b_t - \mathbb{E}_t [\pi_{t+1}], \quad (14)$$

⁵Available at <http://economics.sas.upenn.edu/~schorf/research.htm>.

where r_t^k is the rental rate of capital, r_*^k its steady state value, and δ the depreciation rate. Given that capital is subject to variable capacity utilization u_t , the relationship between \bar{k}_t and the amount of capital effectively rented out to firms k_t is

$$k_t = u_t - z_t + \bar{k}_{t-1}. \quad (15)$$

The optimality condition determining the rate of utilization is given by

$$\frac{1 - \psi}{\psi} r_t^k = u_t, \quad (16)$$

where ψ captures the utilization costs in terms of foregone consumption. Real marginal costs for firms are given by

$$mc_t = w_t + \alpha L_t - \alpha k_t, \quad (17)$$

where w_t is the real wage and α is the income share of capital (after paying markups and fixed costs) in the production function. From the optimality conditions of goods producers it follows that all firms have the same capital-labor ratio:

$$k_t = w_t - r_t^k + L_t. \quad (18)$$

The production function is:

$$y_t = \Phi_p (\alpha k_t + (1 - \alpha)L_t) + \mathcal{I}\{\rho_z < 1\} (\Phi_p - 1) \frac{1}{1 - \alpha} \tilde{z}_t, \quad (19)$$

if the log productivity is trend stationary. The last term $(\Phi_p - 1) \frac{1}{1 - \alpha} \tilde{z}_t$ drops out if technology has a stochastic trend, because in this case one has to assume that the fixed costs are proportional to the trend. Similarly, the resource constraint is:

$$y_t = g_t + \frac{c_*}{y_*} c_t + \frac{i_*}{y_*} i_t + \frac{r_*^k k_*}{y_*} u_t - \mathcal{I}\{\rho_z < 1\} \frac{1}{1 - \alpha} \tilde{z}_t, \quad (20)$$

where again the term $-\frac{1}{1 - \alpha} \tilde{z}_t$ disappears if technology follows a unit root process. Government spending g_t is assumed to follow the exogenous process:

$$g_t = \rho_g g_{t-1} + \sigma_g \varepsilon_{g,t} + \eta_{gz} \sigma_z \varepsilon_{z,t}.$$

Finally, the price and wage Phillips curves are, respectively:

$$\pi_t = \frac{(1 - \zeta_p \bar{\beta})(1 - \zeta_p)}{(1 + \iota_p \bar{\beta}) \zeta_p ((\Phi_p - 1) \epsilon_p + 1)} mc_t + \frac{\iota_p}{1 + \iota_p \bar{\beta}} \pi_{t-1} + \frac{\bar{\beta}}{1 + \iota_p \bar{\beta}} \mathbb{E}_t[\pi_{t+1}] + \lambda_{f,t}, \quad (21)$$

and

$$w_t = \frac{(1 - \zeta_w \bar{\beta})(1 - \zeta_w)}{(1 + \bar{\beta})\zeta_w((\lambda_w - 1)\epsilon_w + 1)} (w_t^h - w_t) - \frac{1 + \iota_w \bar{\beta}}{1 + \bar{\beta}} \pi_t + \frac{1}{1 + \bar{\beta}} (w_{t-1} - z_t - \iota_w \pi_{t-1}) + \frac{\bar{\beta}}{1 + \bar{\beta}} \mathbb{E}_t [w_{t+1} + z_{t+1} + \pi_{t+1}] + \lambda_{w,t}, \quad (22)$$

where ζ_p , ι_p , and ϵ_p are the Calvo parameter, the degree of indexation, and the curvature parameter in the Kimball aggregator for prices, and ζ_w , ι_w , and ϵ_w are the corresponding parameters for wages. w_t^h measures the household's marginal rate of substitution between consumption and labor, and is given by:

$$w_t^h = \frac{1}{1 - h e^{-\gamma}} (c_t - h e^{-\gamma} c_{t-1} + h e^{-\gamma} z_t) + \nu_l L_t, \quad (23)$$

where ν_l characterizes the curvature of the disutility of labor (and would equal the inverse of the Frisch elasticity in absence of wage rigidities). The mark-ups $\lambda_{f,t}$ and $\lambda_{w,t}$ follow exogenous ARMA(1,1) processes

$$\lambda_{f,t} = \rho_{\lambda_f} \lambda_{f,t-1} + \sigma_{\lambda_f} \varepsilon_{\lambda_f,t} + \eta_{\lambda_f} \sigma_{\lambda_f} \varepsilon_{\lambda_f,t-1}, \text{ and}$$

$$\lambda_{w,t} = \rho_{\lambda_w} \lambda_{w,t-1} + \sigma_{\lambda_w} \varepsilon_{\lambda_w,t} + \eta_{\lambda_w} \sigma_{\lambda_w} \varepsilon_{\lambda_w,t-1},$$

respectively. Finally, the monetary authority follows a generalized feedback rule:

$$R_t = \rho_R R_{t-1} + (1 - \rho_R) \left(\psi_1 \pi_t + \psi_2 (y_t - y_t^f) \right) + \psi_3 \left((y_t - y_t^f) - (y_{t-1} - y_{t-1}^f) \right) + r_t^m, \quad (24)$$

where the flexible price/wage output y_t^f is obtained from solving the version of the model without nominal rigidities (that is, Equations (11) through (20) and (23)), and the residual r_t^m follows an AR(1) process with parameters ρ_{r^m} and σ_{r^m} .

3.2 Adding Observed Long Run Inflation Expectations (Model SW π)

In order to capture the rise and fall of inflation and interest rates in the estimation sample, we replace the constant target inflation rate by a time-varying target inflation. While time-varying target rates have been frequently used for the specification of monetary policy rules

in DSGE model (e.g., [Erceg and Levin \(2003\)](#) and [Smets and Wouters \(2003\)](#), among others), we follow the approach of [Aruoba and Schorfheide \(2008\)](#) and [Del Negro and Eusepi \(2011\)](#) and include data on long-run inflation expectations as an observable into the estimation of the DSGE model. At each point in time, the long-run inflation expectations essentially determine the level of the target inflation rate. To the extent that long-run inflation expectations at the forecast origin contain information about the central bank's objective function, e.g. the desire to stabilize inflation at 2%, this information is automatically included in the forecast. [Clark \(2011\)](#) constructs a Bayesian VAR in which variables are expressed in deviations from long-run trends. For inflation and interest rates these long-run trends are given by long-horizon Blue Chip forecasts and the VAR includes equations that capture the evolution of these forecasts. Our treatment of inflation in the DSGE model bears similarities to [Clark \(2011\)](#)'s VAR.

More specifically, for the SW model the interest-rate feedback rule of the central bank (24) is modified as follows:⁶

$$R_t = \rho_R R_{t-1} + (1 - \rho_R) \left(\psi_1 (\pi_t - \pi_t^*) + \psi_2 (y_t - y_t^f) \right) + \psi_3 \left((y_t - y_t^f) - (y_{t-1} - y_{t-1}^f) \right) + r_t^m. \quad (25)$$

The time-varying inflation target evolves according to:

$$\pi_t^* = \rho_{\pi^*} \pi_{t-1}^* + \sigma_{\pi^*} \epsilon_{\pi^*,t}, \quad (26)$$

where $0 < \rho_{\pi^*} < 1$ and $\epsilon_{\pi^*,t}$ is an iid shock. We follow [Erceg and Levin \(2003\)](#) and model π_t^* as following a stationary process, although our prior for ρ_{π^*} will force this process to be highly persistent (see Panel II of Table A-1). The assumption that the changes in the target inflation rate are exogenous is, to some extent, a short-cut. For instance, the learning models of [Sargent \(1999\)](#) or [Primiceri \(2006\)](#) would suggest that the rise in the target inflation rate in the 1970's and the subsequent drop is due to policy makers learning about the output-inflation trade-off and trying to set inflation optimally. We are abstracting from such a mechanism in our specification. We refer to this model as SW π .

⁶We follow the specification in [Del Negro and Eusepi \(2011\)](#), while [Aruoba and Schorfheide \(2008\)](#) assume that the inflation target also affects the intercept in the feedback rule.

3.3 Adding Financial Frictions (Model SWFF)

We now add financial frictions to the SW model building on the work of Bernanke, Gertler, and Gilchrist (1999), Christiano, Motto, and Rostagno (2003), De Graeve (2008), and Christiano, Motto, and Rostagno (forthcoming). In this extension, banks collect deposits from households and lend to entrepreneurs who use these funds as well as their own wealth to acquire physical capital, which is rented to intermediate goods producers. Entrepreneurs are subject to idiosyncratic disturbances that affect their ability to manage capital. Their revenue may thus be too low to pay back the bank loans. Banks protect themselves against default risk by pooling all loans and charging a spread over the deposit rate. This spread may vary as a function of the entrepreneurs' leverage and their riskiness. Adding these frictions to the SW model amounts to replacing equation (14) with the following conditions:

$$E_t \left[\tilde{R}_{t+1}^k - R_t \right] = b_t + \zeta_{sp,b} (q_t^k + \bar{k}_t - n_t) + \tilde{\sigma}_{\omega,t} \quad (27)$$

and

$$\tilde{R}_t^k - \pi_t = \frac{r_*^k}{r_*^k + (1 - \delta)} r_t^k + \frac{(1 - \delta)}{r_*^k + (1 - \delta)} q_t^k - q_{t-1}^k, \quad (28)$$

where \tilde{R}_t^k is the gross nominal return on capital for entrepreneurs, n_t is entrepreneurial equity, and $\tilde{\sigma}_{\omega,t}$ captures mean-preserving changes in the cross-sectional dispersion of ability across entrepreneurs (see Christiano, Motto, and Rostagno (forthcoming)) and follows an AR(1) process with parameters ρ_{σ_ω} and σ_{σ_ω} . The second condition defines the return on capital, while the first one determines the spread between the expected return on capital and the riskless rate.⁷ The following condition describes the evolution of entrepreneurial net worth:

$$\begin{aligned} n_t = & \zeta_{n,\tilde{R}^k} \left(\tilde{R}_t^k - \pi_t \right) - \zeta_{n,R} (R_{t-1} - \pi_t) + \zeta_{n,qK} (q_{t-1}^k + \bar{k}_{t-1}) + \zeta_{n,n} n_{t-1} \\ & - \frac{\zeta_{n,\sigma_\omega}}{\zeta_{sp,\sigma_\omega}} \tilde{\sigma}_{\omega,t-1}. \end{aligned} \quad (29)$$

We refer to this model as SWFF.

⁷Note that if $\zeta_{sp,b} = 0$ and the financial friction shocks $\tilde{\sigma}_{\omega,t}$ are zero, (27) and (28) coincide with (14).

3.4 State-Space Representation, Estimation, and Definition of the Information Set

We use the method in Sims (2002) to solve the log-linear approximation of the DSGE model. We collect all the DSGE model parameters in the vector θ , stack the structural shocks in the vector ϵ_t , and derive a state-space representation for our vector of observables y_t . The state-space representation is comprised of the transition equation:

$$s_t = \mathcal{T}(\theta)s_{t-1} + \mathcal{R}(\theta)\epsilon_t, \quad (30)$$

which summarizes the evolution of the states s_t , and the measurement equation:

$$y_t = \mathcal{Z}(\theta)s_t + \mathcal{D}(\theta), \quad (31)$$

which maps the states onto the vector of observables y_t , where $\mathcal{D}(\theta)$ represents the vector of steady state values for these observables. The measurement equations for real output, consumption, investment, and real wage growth, hours, inflation, and interest rates are given by:

$$\begin{aligned} \text{Output growth} &= \gamma + 100(y_t - y_{t-1} + z_t) \\ \text{Consumption growth} &= \gamma + 100(c_t - c_{t-1} + z_t) \\ \text{Investment growth} &= \gamma + 100(i_t - i_{t-1} + z_t) \\ \text{Real Wage growth} &= \gamma + 100(w_t - w_{t-1} + z_t), \\ \text{Hours} &= \bar{l} + 100l_t \\ \text{Inflation} &= \pi_* + 100\pi_t \\ \text{FFR} &= R_* + 100R_t \end{aligned} \quad (32)$$

where all variables are measured in percent, where π_* and R_* measure the steady state level of net inflation and short term nominal interest rates, respectively and where \bar{l} captures the mean of hours (this variable is measured as an index).

To incorporate information about low-frequency movements of inflation the set of measurement equations (32) is augmented by

$$\begin{aligned} \pi_t^{O,40} &= \pi_* + 100\mathbb{E}_t \left[\frac{1}{40} \sum_{k=1}^{40} \pi_{t+k} \right] \\ &= \pi_* + \frac{100}{40} \mathcal{Z}(\theta)_{(\pi,.)} (I - \mathcal{T}(\theta))^{-1} (I - [\mathcal{T}(\theta)]^{40}) \mathcal{T}(\theta)s_t, \end{aligned} \quad (33)$$

where $\pi_t^{O,40}$ represents observed long run inflation expectations obtained from surveys (in percent per quarter), and the right-hand-side of (33) corresponds to expectations obtained from the DSGE model (in deviation from the mean π_*). The second line shows how to compute these expectations using the transition equation (30) and the measurement equation for inflation. $\mathcal{Z}(\theta)_{(\pi,.)}$ is the row of $\mathcal{Z}(\theta)$ in (31) that corresponds to inflation. The SW π model is estimated using the observables in expressions (32) and (33).

Model SWFF uses in addition spreads as observables. The corresponding measurement equation is

$$\text{Spread} = SP_* + 100E_t \left[\tilde{R}_{t+1}^k - R_t \right], \quad (34)$$

where the parameter SP_* measures the steady state spread. Both models are estimated using quarterly data. The construction of the data set is summarized in Appendix A.

We use Bayesian techniques in the subsequent empirical analysis, which require the specification of a prior distribution for the model parameters. For most of the parameters we use the same marginal prior distributions as Smets and Wouters (2007). There are two important exceptions. First, the original prior for the quarterly steady state inflation rate π_* used by Smets and Wouters (2007) is tightly centered around 0.62% (which is about 2.5% annualized) with a standard deviation of 0.1%. We favor a looser prior, one that has less influence on the model's forecasting performance, that is centered at 0.75% and has a standard deviation of 0.4%. Second, for the financial frictions mechanism we specify priors for the parameters SP_* , $\zeta_{sp,b}$, ρ_{σ_ω} , and σ_{σ_ω} . We fix the parameters corresponding to the steady state default probability and the survival rate of entrepreneurs, respectively. In turn, these parameters imply values for the parameters of (29). A summary of the priors is provided in Table A-1 in Appendix C. Section B.2 in the appendix provides some details on the computation of the predictive density for k observations ahead $p(y_{t:t+k} | \mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$ for DSGE models.

In our empirical analysis we use real time data, so the defining the information set $\mathcal{I}_t^{\mathcal{M}_m}$ requires some care. We use four forecasts per year, corresponding to the information set available to the econometrician on January 1st, April 1st, July 1st, and October 1st of each year. This information set includes the so-called ‘‘first final’’ release of the NIPA data for the quarter ending four months before the forecast date, namely Q3 and Q4 Apr of the previous year for forecasts made in January and April, respectively, and Q1 and Q2 of the current year for forecasts made in July and October, respectively. When plotting results over time we use the convention that t equals the quarter corresponding to the latest NIPA data, so for

instance t would be 2008Q3 for forecasts made on January 1st 2009. Note that in addition to the NIPA data the econometrician can in principle also observe interest rates and spreads for the quarter that just ended (2008Q4 for January 1st 2009). We denote with $\mathcal{I}_t^{\mathcal{M}_m}$ and $\mathcal{I}_{t+}^{\mathcal{M}_m}$ the information sets that exclude and include, respectively, this additional data points.

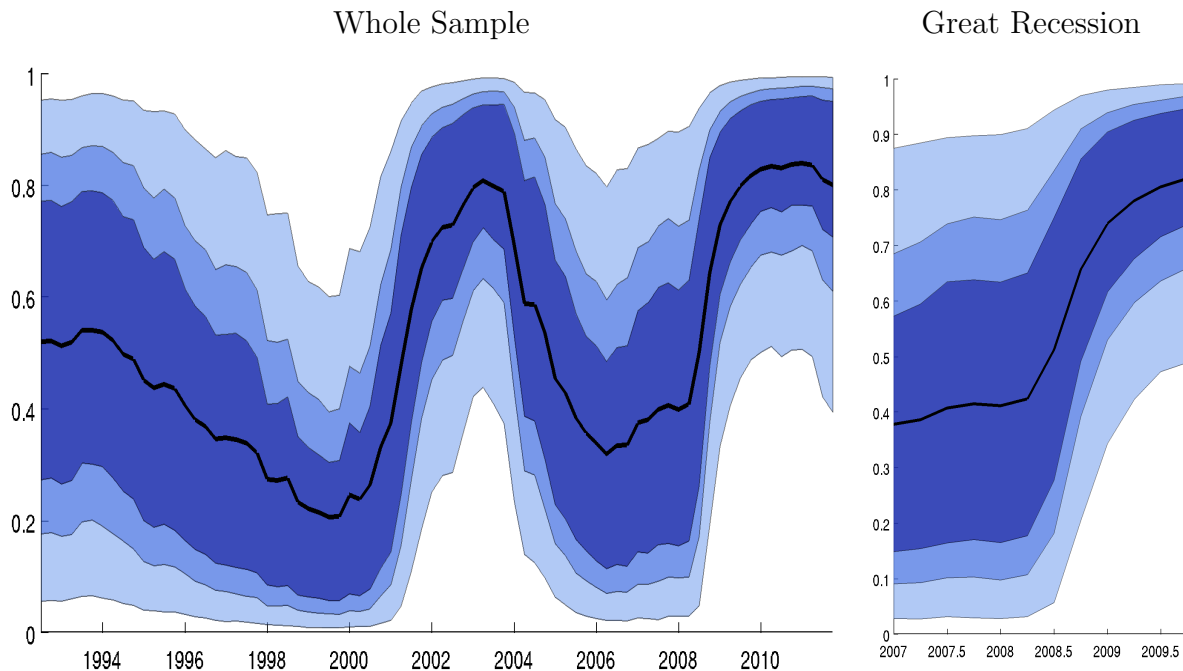
4 Results

We apply the dynamic pools methodology discussed in section 2 to the forecasts of average output growth and inflation one year ahead (hence y_t in the predictive densities formulas is a 2×1 vector containing these two variables) obtained with the two versions of the SW model described earlier, the version with (SWFF) and without (SW π) financial frictions and spreads as observables. We ask the following questions: Is there significant time variation in the relative forecasting performance of the two models, as captured by the estimated distribution of λ_t ? Is it the case that one of the two models always performs better than the other, so that the distribution of λ_t is on either side of .5 for all (or most) t ? Does λ_t change rapidly enough when estimated in real time to offer useful guidance to policy makers or forecasters? Do the dynamic pools perform better in real time than forecasting with static pools, or BMA weights, and by how much?

First, we present results on the evolution of the filtered estimates of λ_t obtained from the state-space model consisting of expressions (8) and (7). We use the convention that $\mathcal{M}_1 = \text{SWFF}$ while $\mathcal{M}_2 = \text{SW}\pi$, therefore λ_t is the weight on the model with financial frictions. For the time being our results are obtained fixing the value of $\rho = .9$ in the law of motion (8). The appendix shows results for $\rho = .75$, which are quite similar.

The left panel of Figure 2 shows that the filtered distribution of λ_t changes substantially over the sample, indicating that there is time-variation in the two models' relative forecasting performance. Consistently with the results in chapter 7 of [Del Negro and Schorfheide \(2013\)](#), the filtered distribution of λ_t shifts toward values above .5 during periods of financial turmoil – the aftermath of the “dot com” bust and the recent financial crisis, while it tends to be below .5 in more “tranquil” periods – the late 90s and the mid-2000s. Both the 50 and 68 percent bands are fairly tightly estimated, especially during periods of financial turmoil.

The results also indicate that the changes in the distribution of λ_t can be quite sudden. For instance, the filtered mean of λ_t increases from .4 to almost .8 in the span of about

Figure 2: Filtered Posterior of λ_t (pool weight on SWFF)

Notes: The figure shows the filtered mean of λ_t (solid black line) and the 50% , 68% and 90% bands in shades of blue.

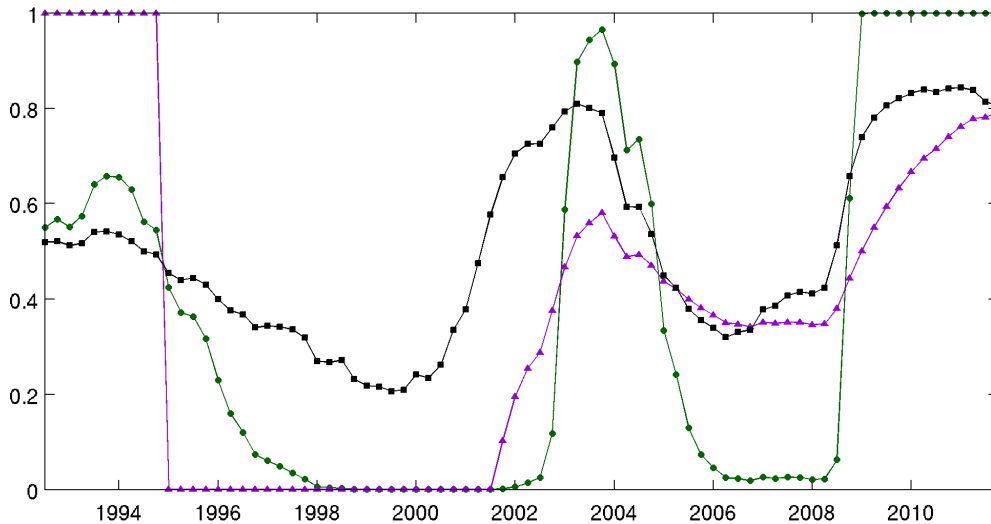
one year, from early 2008 to early 2009. Recall that we are using $h = 4$ quarters ahead data in evaluating the predictive densities. Since the financial turmoil begun in 2007Q3, the econometrician would have to wait at least four quarters (until 2008Q3) to acquire information about the relative performance of the SWFF and SW π model during this period. Yet the *filtered* distribution of λ_t starts shifting upward already in early 2008.⁸

While much of the mass of the distribution of λ_t is fairly concentrated, the 90 percent bands are quite wide however, and cover .5 for almost all periods. A standard results in the model combination literature is that equally weighted predictions are hard to beat. The finding that the equal weight value of .5 is almost always included in one of the tails of the distribution suggests that to some extent this results holds in this application as well: the posterior distribution does not completely rule out weights in the neighborhood of .5. As

⁸These results suggest that even if one is interested in 4-quarters ahead forecasts, it may make sense to modify the measurement equation (7) to include $h = 1$ predictive densities, as these would give a more timely signal about which model forecasts better.

we will see later, the gains in terms of forecasting accuracy of using the time varying pools relative to the equally weighted pools are not dramatic.

Figure 3: Dynamic (*black*), BMA (*green*) and Static Pool (*purple*) Weights in Real Time



Notes: The figure shows the time series of the BMA weight λ_t^{BMA} (green) computed as in equation (5), the static optimal pool weights λ_t^{SP} (purple) computed as in equation (3), and the dynamic pool weights $\lambda_{t|t}$ (black), computed as the posterior mean of the filtered distribution obtained using information available at time t . The weight is the weight on the SWFF model in forecast pools.

In order to gain insights on the properties of our approach, we next compare the evolution of the dynamic pool weights with two alternatives, BMA and static optimal weights. Figure 3 shows the time series of the BMA weight λ_t^{BMA} (green) computed as in (5), the static optimal pool weights λ_t^{SP} (purple) computed as in equation (3), and the dynamic pool weights $\lambda_{t|t}$ (black), computed as the posterior mean of the filtered distribution obtained using information available at time t . The dynamics of the BMA weights present a not too surprising bang-bang pattern. As soon as enough information is available the BMA weight converge to one extreme. However, it switches rapidly toward the opposite extreme as new information favoring the other model becomes available. The static pool weights also tend to favor extremes early in the sample, but later converge toward the middle of the distribution. Dynamic weights rarely if ever assume extreme values (almost all of the realizations of $\lambda_{t|t}$ are between .2 and .8). Most importantly, $\lambda_{t|t}$ appears to be a leading indicator relative to both BMA and static weights.

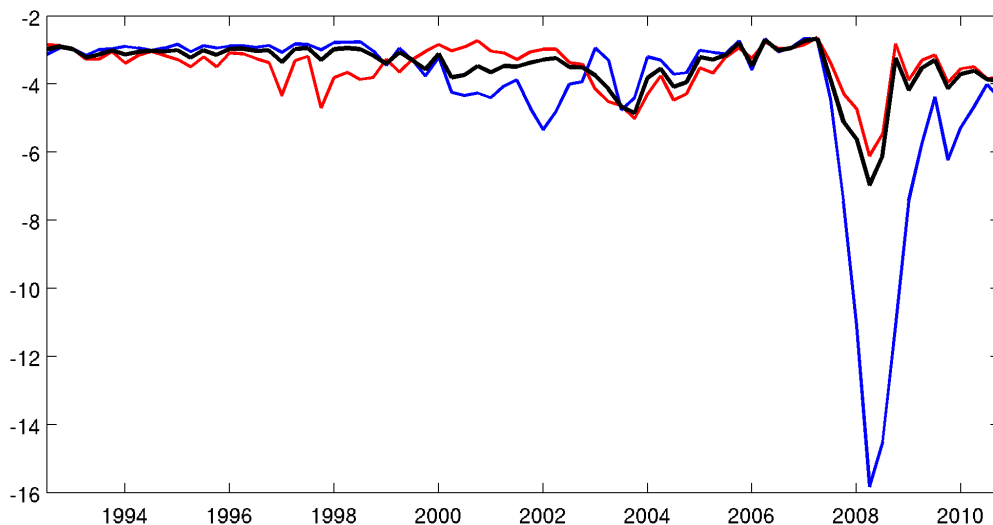
We now ask whether there are any real time gains in forecasting using the time varying

weights relative to: i) each of the models in isolation, ii) a model combination obtained using the (real-time) static weights, iii) the BMA weights, and iv) equal weights. Define the real time predictive density obtained with the dynamic pool as

$$p^{DP}(y_t|\lambda_{t-h|t-h}) = \lambda_{t-h|t-h}p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF) + (1 - \lambda_{t-h|t-h})p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi) \quad (35)$$

where $\lambda_{t-h|t-h}$ is the posterior mean of the filtered distribution obtained using information available at time $t - h$.⁹

Figure 4: Log Scores Comparison: SWFF (red) vs SW π (blue) vs Dynamic Pools (black)

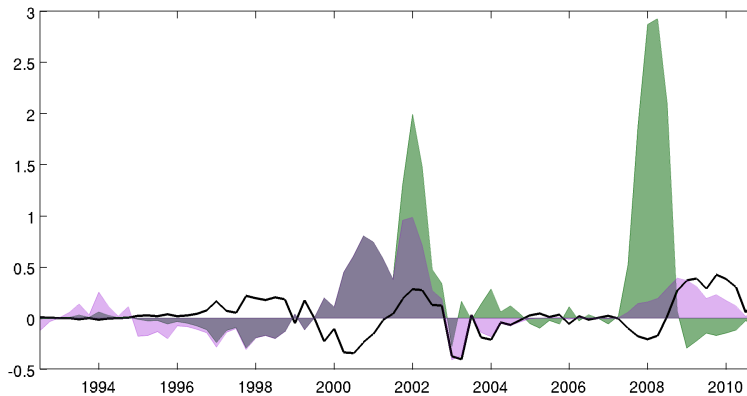


Notes: The figure shows $\log p^{DP}(y_t|\lambda_{t-h|t-h})$ (black), $\log p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF)$ (red), and $\log p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi)$ (blue) over time.

Figure 4 shows the natural logarithms of $p^{DP}(y_t|\lambda_{t-h|t-h})$ (black), $p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF)$ (red), and $p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi)$ (blue) as a function of $t - h$. The lower is the log score, the worse is the model performance as measured by the distance between ex posts outcomes and ex ante forecasts. The pattern of the relative performance of SWFF versus SW π tracks closely the evolution of the distribution of λ_t in figure 2, which is not surprising given that this is precisely the information entering the measurement equation (7). SWFF performs worse than SW π in normal times, and viceversa during periods of financial distress such as the

⁹We are aware it is likely not optimal to use only the posterior mean as opposed to the entire posterior distribution (see Amisano and Geweke (2013)). We will pursue this route in future drafts. We are also aware that in place of $\lambda_{t-h|t-h}$ we could be using $\lambda_{t|t-h}$ by taking advantage of the law of motion for λ_t (8).

Figure 5: Log Score Comparison: Dynamic Pool vs BMA (*green*), Optimal Static Pool (*purple*) and Equally Weighted Pool (*black*)



Notes: The figure shows the difference in the log scores of 1) dynamic pools versus BMA ($\log p^{DP}(y_t|\lambda_{t-h}|t-h) - \log p^{BMA}(y_t|\lambda_{t-h}^{BMA})$, green shaded areas), 2) dynamic pools versus static pools ($\log p^{DP}(y_t|\lambda_{t-h}|t-h) - \log p^{SP}(y_t|\lambda_{t-h}^{SP})$, purple shaded areas), and 3) dynamic pools versus equal weight pools ($\log p^{DP}(y_t|\lambda_{t-h}|t-h) - \log p^{EW}(y_t)$, black lines).

early 2000s or the recent financial crisis. The drop in log score of the model without financial frictions during the Great Recession is remarkable, although also not surprising in light of the results in [Del Negro and Schorfheide \(2013\)](#). The log score obtained from the dynamic pools is never too far below that of the best model at any point in time. While each of the model has periods of not so great forecasting performance (especially $SW\pi$ of course, but also $SWFF$ in the late 1990s) by virtue of diversification the dynamic pool is not severely affected by such episodes.

Table 1: Cumulative Log Scores

Component Models		Model Pooling	
Model	Log score	Method	Log score
$SW\pi$	-306.34	BMA	-275.57
$SWFF$	-259.58	SP	-264.67
		EW	-260.42
		DP	-258.53

The integral of each line is reported in Table 1. The table shows that dynamic pools fare

quite a lot better than using $SW\pi$ only, which is not surprising, but fare only slightly better than using $SWFF$ all the times. We believe that the latter result stems from the fact that our relatively short sample is dominated by the Great Recession episode. The $SW\pi$ model fares so poorly in this episode that any pool that puts positive weight on this model will suffer non-minor losses in log score relative to $SWFF$ (the gap between black and red line in the Great Recession). In the case of dynamic pools, these losses almost fully compensate for the gains relative to $SWFF$ throughout the 1990s.

Define the real time log scores obtained using BMA, static pool weights, and equal weights as

$$\begin{aligned} p^{BMA}(y_t|\lambda_{t-h}^{BMA}) &= \lambda_{t-h}^{BMA} p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF) + (1 - \lambda_{t-h}^{BMA}) p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi), \\ p^{SP}(y_t|\lambda_{t-h}^{SP}) &= \lambda_{t-h}^{SP} p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF) + (1 - \lambda_{t-h}^{SP}) p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi), \text{ and} \\ p^{EW}(y_t) &= .5p(y_t|\mathcal{I}_{t-h}^{SWFF}, SWFF) + .5p(y_t|\mathcal{I}_{t-h}^{SW\pi}, SW\pi), \end{aligned} \quad (36)$$

respectively. Figure 5 shows the difference in the log scores of 1) dynamic pools versus BMA ($\log p^{DP}(y_t|\lambda_{t-h|t-h}) - \log p^{BMA}(y_t|\lambda_{t-h}^{BMA})$, green shaded areas), 2) dynamic pools versus static pools ($\log p^{DP}(y_t|\lambda_{t-h|t-h}) - \log p^{SP}(y_t|\lambda_{t-h}^{SP})$, purple shaded areas), and 3) dynamic pools versus equal weight pools ($\log p^{DP}(y_t|\lambda_{t-h|t-h}) - \log p^{EW}(y_t)$, black lines). Positive values show that dynamic pools perform better than the alternative.

The loss in performance resulting from using BMA weights can be very large. As emphasized in Amisano and Geweke (2013), this model combination approach suffers from using weights that are often close to the extremes of the $[0, 1]$ interval, thereby forfeiting the benefits from diversification. Table 1 shows that the cumulative log score from using BMA in real time is significantly worse than that of any other model combination.

Static pools fare better, but are also slower to adjust to the changing economic environment. For instance, in the latest period the weight given to the financial friction model rises at a moderate pace relative to the weight in dynamic pools (see figure 3). There are periods where the static pools outperform the dynamic ones, but these are relatively short lived and the losses are small. As a consequence, the cumulated log score difference between the two models amounts to 6 points, which is non negligible.

Finally, the equal weighted pool performance is comparable to that of dynamic pools. Since this pool is by construction well diversified, it never suffers the large losses associated with the BMA pool or, at least in the early part of the sample, with the static pool.¹⁰ As a

¹⁰Amisano and Geweke (2013) find that equal weighted pools perform better than static pools in real time.

consequence, the improvement in performance of dynamic pools relative to equal weights is only modest.

5 Conclusions

This paper provides a methodology for estimating time-varying weights in optimal prediction pools. In our application we combine predictive densities from two DSGE models, with and without financial frictions. However, the same method could be used to combine other classes of time series models. Extensions to pools of more than two models are conceptually straightforward but may pose computational challenges.

From the substantive point of view we find that the model without financial frictions forecasts better in times without significant financial distress, but that the relative forecasting performance changes dramatically in more turbulent times. This findings begs the question of whether this is a result of using linearized models: A non-linear model with financial friction may look (and forecast) very much like one without friction in tranquil times, but have very different dynamics when the financial constraints become binding (e.g., see [Brunnermeier and Sannikov \(forthcoming\)](#), [Dewachter and Wouters \(2012\)](#), or the estimated DSGE model of [Bocola \(2013\)](#)). One could therefore interpret the findings in this paper as evidence in favor of an encompassing non-linear model. Whether such non-linear model would forecast better than the pool of models considered here is a question for future research.

References

- AMISANO, G., AND J. GEWEKE (2013): “Prediction with Macroeconomic Models,” *ECB Working Paper 1537*.
- ARUOBA, S. B., AND F. SCHORFHEIDE (2008): “Insights from an Estimated Search-Based Monetary Model with Nominal Rigidities,” *Working Paper*.
- BATES, J., AND C. W. GRANGER (1969): “The Combination of Forecasts,” *Operations Research*, 20(4), 451–468.

- BERNANKE, B., M. GERTLER, AND S. GILCHRIST (1999): “The Financial Accelerator in a Quantitative Business Cycle Framework,” in *Handbook of Macroeconomics*, ed. by J. B. Taylor, and M. Woodford, vol. 1C. North Holland, Amsterdam.
- BILLIO, M., R. CASARIN, F. RAVAZZOLO, AND H. K. VAN DIJK (2012): “Time-varying Combinations of Predictive Densities using Nonlinear Filtering,” *Tinbergen Institute Discussion Paper*, 118/III(4).
- BOCOLA, L. (2013): “The Pass-Through of Sovereign Risk,” *Manuscript, University of Pennsylvania*.
- BRUNNERMEIER, M. K., AND Y. SANNIKOV (forthcoming): “A Macroeconomic Model with a Financial Sector,” *American Economic Review*.
- CHRISTIANO, L., R. MOTTO, AND M. ROSTAGNO (2003): “The Great Depression and the Friedman-Schwartz Hypothesis,” *Journal of Money, Credit and Banking*, 35, 1119–1197.
- CHRISTIANO, L. J., M. EICHENBAUM, AND C. L. EVANS (2005): “Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy,” *Journal of Political Economy*, 113, 1–45.
- CHRISTIANO, L. J., R. MOTTO, AND M. ROSTAGNO (forthcoming): “Risk Shocks,” *American Economic Review*.
- CLARK, T. (2011): “Real-Time Density Forecasts From Bayesian Vector Autoregress with Stochastic Volatility,” *Journal of Business & Economic Statistics*, 29(3), 327–341.
- CREAL, D. (2012): “A Survey of Sequential Monte Carlo Methods for Economics and Finance,” *Econometric Reviews*, 31(3), 245–296.
- DE GRAEVE, F. (2008): “The External Finance Premium and the Macroeconomy: US Post-WWII Evidence,” *Journal of Economic Dynamics and Control*, 32(11), 3415 – 3440.
- DEL NEGRO, M., AND S. EUSEPI (2011): “Fitting Observed Inflation Expectations,” *Journal of Economic Dynamics and Control*, 35, 2105–2131.
- DEL NEGRO, M., AND F. SCHORFHEIDE (2013): “DSGE Model-Based Forecasting,” in *Handbook of Economic Forecasting, Volume 2*, ed. by G. Elliott, and A. Timmermann. Elsevier.

- DEWACHTER, H., AND R. WOUTERS (2012): “Endogenous risk in a DSGE model with capital-constrained Financial intermediaries,” *National Bank of Belgium Working Paper*, 235.
- EDGE, R., AND R. GÜRKAYNAK (2010): “How Useful Are Estimated DSGE Model Forecasts for Central Bankers,” *Brookings Papers of Economic Activity*, p. forthcoming.
- ERCEG, C. J., AND A. T. LEVIN (2003): “Imperfect Credibility and Inflation Persistence,” *Journal of Monetary Economics*, 50, 915–944.
- GEWEKE, J. (2005): *Contemporary Bayesian Econometrics and Statistics*. Wiley.
- (2007): “Bayesian Model Comparison and Validation,” *American Economic Review Papers and Proceedings*, 97, 60–64.
- GEWEKE, J., AND G. AMISANO (2011): “Optimal prediction pools,” *Journal of Econometrics*, 164, 130141.
- (2012): “Prediction with Misspecified Models,” *American Economic Review: Papers & Proceedings*, 103(3), 482–486.
- GIORDANI, P., M. K. PITT, AND R. KOHN (2011): “Bayesian Inference for Time Series State Space Models,” in *Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. K. van Dijk. Oxford University Press.
- GNEITING, T., AND R. RANJAN (2013): “Combining Predictive Distributions,” *Electronic Journal of Statistics*, 7, 1747–1782.
- GREENWOOD, J., Z. HERCOVITZ, AND P. KRUSELL (1998): “Long-Run Implications of Investment-Specific Technological Change,” *American Economic Review*, 87(3), 342–36.
- GRIFFIN, J., F. QUINTANA, AND M. STEEL (2011): “Flexible and Nonparametric Modeling,” in *Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. K. van Dijk. Oxford University Press.
- GUIDOLIN, M., AND A. TIMMERMANN (2009): “Forecasts of US short-term interest rates: A flexible forecast combination approach,” *Journal of Econometrics*, 150, 297–311.

- KIYOTAKI, N., AND J. MOORE (1997): “Credit Cycles,” *Journal of Political Economy*, 105(2), 211–248.
- PRIMICERI, G. (2006): “Why Inflation Rose and Fell: Policymakers Beliefs and US Postwar Stabilization Policy,” *Quarterly Journal of Economics*, 121, 867–901.
- SARGENT, T. J. (1999): *The Conquest of American Inflation*. Princeton University Press, Princeton.
- SIMS, C. A. (2002): “Solving Linear Rational Expectations Models,” *Computational Economics*, 20((1-2)), 1–20.
- SMETS, F., AND R. WOUTERS (2003): “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area,” *Journal of the European Economic Association*, 1(5), 1123 – 1175.
- (2007): “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach,” *American Economic Review*, 97(3), 586 – 606.
- TERUI, N., AND H. K. VAN DIJK (2002): “Combined forecasts from linear and nonlinear time series models,” *International Journal of Forecasting*, 18(4), 421438.
- WAGGONER, D., AND T. ZHA (2012): “Confronting model misspecification in macroeconomics,” *Journal of Econometrics*, 171(2), 167184.

Appendix for *Time-varying Prediction Pools*

Marco Del Negro, Raiden B. Hasegawa, and Frank Schorfheide

A Data

Real GDP (GDPC), the GDP price deflator (GDPDEF), nominal personal consumption expenditures (PCEC), and nominal fixed private investment (FPI) are constructed at a quarterly frequency by the Bureau of Economic Analysis (BEA), and are included in the National Income and Product Accounts (NIPA). Average weekly hours of production and nonsupervisory employees for total private industries (AWHNONAG), civilian employment (CE16OV), and civilian noninstitutional population (LNSINDEX) are produced by the Bureau of Labor Statistics (BLS) at the monthly frequency. The first of these series is obtained from the Establishment Survey, and the remaining from the Household Survey. Both surveys are released in the BLS Employment Situation Summary (ESS). Since our models are estimated on quarterly data, we take averages of the monthly data. Compensation per hour for the nonfarm business sector (COMPNFB) is obtained from the Labor Productivity and Costs (LPC) release, and produced by the BLS at the quarterly frequency. All data are transformed following [Smets and Wouters \(2007\)](#). Let Δ denote the temporal difference operator. Then:

$$\begin{aligned}
 \text{Output growth} &= 100 * \Delta \text{LN}((GDPC)/LNSINDEX) \\
 \text{Consumption growth} &= 100 * \Delta \text{LN}((PCEC/GDPDEF)/LNSINDEX) \\
 \text{Investment growth} &= 100 * \Delta \text{LN}((FPI/GDPDEF)/LNSINDEX) \\
 \text{Real Wage growth} &= 100 * \Delta \text{LN}(COMPNFB/GDPDEF) \\
 \text{Hours} &= 100 * \text{LN}((AWHNONAG * CE16OV/100)/LNSINDEX) \\
 \text{Inflation} &= 100 * \Delta \text{LN}(GDPDEF).
 \end{aligned}$$

The federal funds rate is obtained from the Federal Reserve Board's H.15 release at the business day frequency. We take quarterly averages of the annualized daily data and divide by four. In the estimation of the DSGE model with financial frictions we measure *Spread* as the annualized Moody's Seasoned Baa Corporate Bond Yield spread over the 10-Year Treasury Note Yield at Constant Maturity. Both series are available from the Federal

Reserve Board's H.15 release. Like the federal funds rate, the spread data is also averaged over each quarter and measured at the quarterly frequency. This leads to:

$$\begin{aligned} FFR &= (1/4) * FEDERAL FUNDS RATE \\ Spread &= (1/4) * (BaaCorporate - 10yearTreasury) \end{aligned}$$

The long-run inflation forecasts used in the measurement equation (33) are obtained from the Blue Chip Economic Indicators survey and the Survey of Professional Forecasters (SPF) available from the FRB Philadelphia's Real-Time Data Research Center. Long-run inflation expectations (average CPI inflation over the next 10 years) are available from 1991:Q4 onwards. Prior to 1991:Q4, we use the 10-year expectations data from the Blue Chip survey to construct a long time series that begins in 1979:Q4. Since the Blue Chip survey reports long-run inflation expectations only twice a year, we treat these expectations in the remaining quarters as missing observations and adjust the measurement equation of the Kalman filter accordingly. Long-run inflation expectations $\pi_t^{O,40}$ are therefore measured as

$$\pi_t^{O,40} = (10\text{-YEAR AVERAGE CPI INFLATION FORECAST} - 0.50)/4.$$

where .50 is the average difference between CPI and GDP annualized inflation from the beginning of the sample to 1992. We divide by 4 since the data are expressed in quarterly terms.

Many macroeconomic time series get revised multiple times by the statistical agencies that publish the series. In many cases the revisions reflect additional information that has been collected by the agencies, in other instances revisions are caused by changes in definitions. For instance, the BEA publishes three releases of quarterly GDP in the first three months following the quarter. Thus, in order to be able to compare DSGE model forecasts to real-time forecasts made by private-sector professional forecasters or the Federal Reserve Board, it is important to construct vintages of real time historical data. We follow the work by [Edge and Gürkaynak \(2010\)](#) and construct data vintages that are aligned with the publication dates of the Blue Chip survey. You can find a more detailed description of how we construct this real-time dataset in (CITE FORECAST HANDBOOK CHAPTER).

B Computational Details

B.1 Particle Filter

Initialization: $x_0^i \sim iidN(0,1)$, $w_0^i = 1$ (note: the weights sum up to N , not 1). For $t = 1, \dots, T$

1. Propagate particles forward

$$\hat{x}_t^i = \rho x_{t-1}^i + \sqrt{1 - \rho^2} \varepsilon_t^i, \quad \varepsilon_t^i \sim N(0, 1)$$

2. compute $\hat{\lambda}_t^i = \Phi(\hat{x}_t^i)$

3. compute unnormalized weights $\tilde{w}_t^i = p(y_t | \hat{\lambda}_t^i) w_{t-1}^i$

4. Normalize weights $\hat{w}_t^i = \frac{\tilde{w}_t^i}{\frac{1}{N} \sum_{i=1}^N \tilde{w}_t^i}$

5. (a) If $ESS = \frac{N^2}{\sum_{i=1}^N \hat{w}_t^i}$ is small ($< N/2$), then resample particles by drawing from

multinomial distribution with support on $\{\hat{x}_t^i\}_{i=1}^N$ and probabilities $\{\frac{\tilde{w}_t^i}{N}\}_{i=1}^N$, leading to $\{x_t^i, w_t^i\}_{i=1}^N$, where $w_t^i = 1$, all i .

- (b) Otherwise (if $ESS = \frac{N^2}{\sum_{i=1}^N \hat{w}_t^i} > N/2$) set $x_t^i = \hat{x}_t^i$, $w_t^i = \hat{w}_t^i$.

We are interested in the evaluation of our state-space model log likelihood for different choices of ρ , the persistence of the particle propagation, in order to find the MLE estimate of ρ . Since evaluating the likelihood is computationally intensive, we refrain from conducting full bayesian inference on the parameter. To calculate our log likelihood, first note that

$$p(\lambda_t | y_{1:t}, \rho) \approx \sum_{i=1}^N \left(\frac{w_t^{(i)}}{N} \right) \delta_{\lambda_t^{(i)}}(\lambda_t)$$

and that, given $\hat{\lambda}_t \sim p(\lambda_t | \lambda_{t-1}, \rho)$, we can write

$$p(y_t | y_{1:t-1}, \rho) \approx \sum_{i=1}^N \left(\frac{w_{t-1}^{(i)}}{N} \right) p(y_t | \hat{\lambda}_t^i).$$

Hence, we can approximate the full log likelihood of the model as

$$\log \mathcal{L}(\rho; y_{1:T}) \approx \sum_{j=1}^T \log p(y_t | y_{1:t-1}, \rho).$$

In figures A-1, A-2 and A-3 we graphically examine the Monte Carlo variance of our estimate of $p(\lambda_t | y_{1:t}, \rho)$, the filtered densities of the model combination weights. Even when resampling in every time period, the Monte Carlo variance of our filtered densities is at an acceptable level.

Figure A-1: Diagnostic: Accuracy of $\widehat{\mathbb{E}}_t[\lambda_t]$ - 100 Groups, 1000 Particles/Group

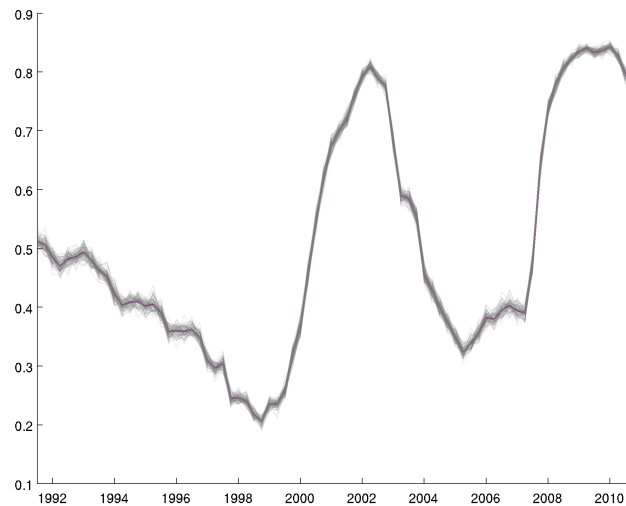


Figure A-2: Diagnostic: Accuracy of λ_t 95th-percentile - 100 Groups, 1000 Particles/Group

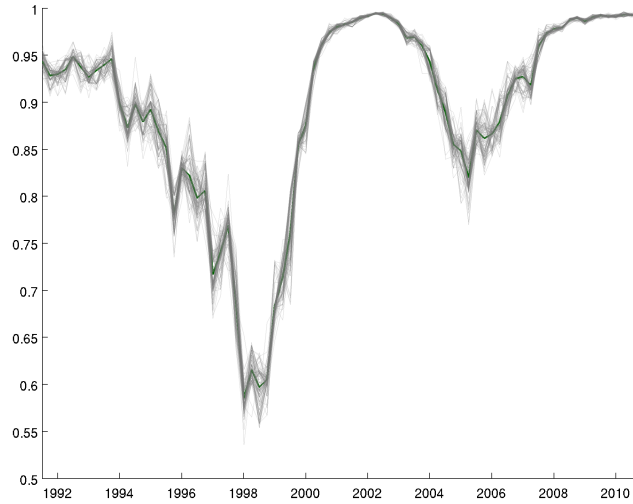
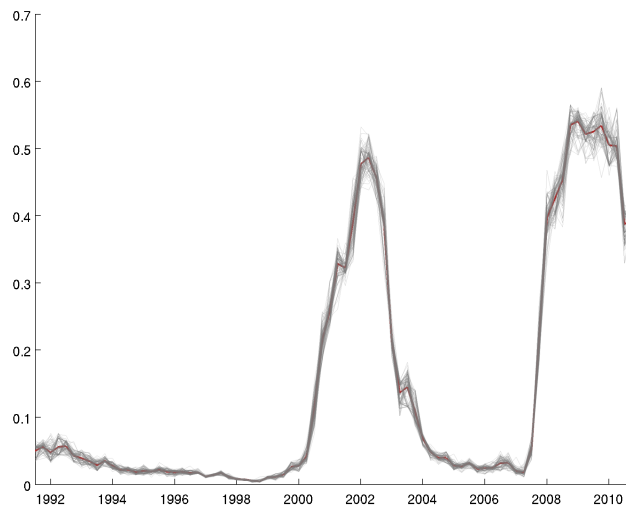


Figure A-3: Diagnostic: Accuracy of λ_t 5th-percentile - 100 Groups, 1000 Particles/Group



B.2 Computing the predictive density

This sections provides some details on the computation of the predictive density for k observations ahead $p(y_{t:t+k}|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$ for DSGE models.

For DSGE models we have a state-space representation which is comprised of the transition equation:

$$s_t = \mathcal{T}(\theta)s_{t-1} + \mathcal{R}(\theta)\epsilon_t, \quad \epsilon_t \sim N(0, \mathcal{Q}) \quad (\text{A-1})$$

which summarizes the evolution of the states s_t , and the measurement equation:

$$y_t = \mathcal{Z}(\theta)s_t + \mathcal{D}(\theta), \quad (\text{A-2})$$

which maps the states onto the vector of observables y_t , where $\mathcal{D}(\theta)$ represents the vector of steady state values for these observables.

1. Generate draws from the posterior $p(\theta|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m)$. For each draw θ^i evaluate $\mathcal{T}(\theta^i), \mathcal{R}(\theta^i), \mathcal{Z}(\theta^i), \mathcal{D}(\theta^i)$.
2. Run Kalman filter to obtain $s_{t-1|t-1}$ and $P_{t-1|t-1}$.
3. Compute $\hat{s}_{t|t-1} = s_{t|I_{t-1}^{\mathcal{M}_1}}$ and $\hat{P}_{t|t-1} = P_{t|I_{t-1}^{\mathcal{M}_1}}$ as
 - (a) Unconditional: $\hat{s}_{t|t-1} = \mathcal{T}s_{t-1|t-1}, \hat{P}_{t|t-1} = \mathcal{T}P_{t-1|t-1}\mathcal{T}' + \mathcal{R}\mathcal{Q}\mathcal{R}'$.
 - (b) Semiconditional (time t spreads, and FFR): run updating step based on these two observables.
4. Build recursively for $j = 1, \dots, k$ the objects $\hat{s}_{t+j|t-1} = \mathcal{T}s_{t+j-1|t-1}, \hat{P}_{t+j|t-1} = \mathcal{T}P_{t+j-1|t-1}\mathcal{T}' + \mathcal{R}\mathcal{Q}\mathcal{R}'$ and construct the matrices

$$\hat{s}_{t:t+k|t-1} = \begin{bmatrix} \hat{s}_{t|t-1} \\ \vdots \\ \hat{s}_{t+k|t-1} \end{bmatrix}$$

and

$$\hat{P}_{t:t+k|t-1} = \begin{bmatrix} \hat{P}_{t|t-1} & \hat{P}_{t|t-1}\mathcal{T}' & \dots & \hat{P}_{t|t-1}\mathcal{T}^{k-1} \\ \mathcal{T}\hat{P}_{t|t-1} & \hat{P}_{t+1|t-1} & \dots & \hat{P}_{t+1|t-1}\mathcal{T}^{k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{T}^k\hat{P}_{t|t-1} & \mathcal{T}^{k-1}\hat{P}_{t+1|t-1} & \dots & \hat{P}_{t+k|t-1} \end{bmatrix}.$$

Thereby obtaining the distribution of $s_{t:t+k}|I_{t-1}^{\mathcal{M}_1}: s_{t:t+k}|I_{t-1}^{\mathcal{M}_1} \sim N(\hat{s}_{t:t+k|t-1}, \hat{P}_{t:t+k|t-1})$.

5. The distribution of $y_{t:t+k} = \tilde{\mathcal{D}} + \tilde{\mathcal{Z}}s_{t:t+k}$ is

$$y_{t:t+k}|I_{t-1}^{\mathcal{M}_1} \sim N(\tilde{\mathcal{D}} + \tilde{\mathcal{Z}}\hat{s}_{t:t+k|t-1}, \tilde{\mathcal{Z}}\hat{P}_{t:t+k|t-1}\tilde{\mathcal{Z}}')$$

where $\tilde{\mathcal{Z}} = I_{k+1} \otimes \mathcal{Z}$ and $\tilde{\mathcal{D}} = 1_{k+1} \otimes \mathcal{D}$ (note $I_1 = 1_1 = 1$)

6. Compute

$$p(y_{t:t+k}|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m) = \phi(y_{t:t+k}^o; \tilde{\mathcal{D}} + \tilde{\mathcal{Z}}\hat{s}_{t:t+k|t-1}, \tilde{\mathcal{Z}}\hat{P}_{t:t+k|t-1}\tilde{\mathcal{Z}}') \quad (\text{A-3})$$

where $y_{t:t+k}^o$ are the actual observations and ϕ is the multivariate normal probability density.

7. If using linear functions of $y_{t:t+k}$ (e.g., four quarter averages, etc.), then write these functions as $f_{t:t+k} = Fy_{t:t+k}$ and the predictive density becomes

$$p(Fy_{t:t+k}|\mathcal{I}_{t-1}^{\mathcal{M}_m}, \mathcal{M}_m) = \phi(Fy_{t:t+k}^o; F\tilde{\mathcal{D}} + F\tilde{\mathcal{Z}}\hat{s}_{t:t+k|t-1}, F\tilde{\mathcal{Z}}\hat{P}_{t:t+k|t-1}\tilde{\mathcal{Z}}'F') \quad (\text{A-4})$$

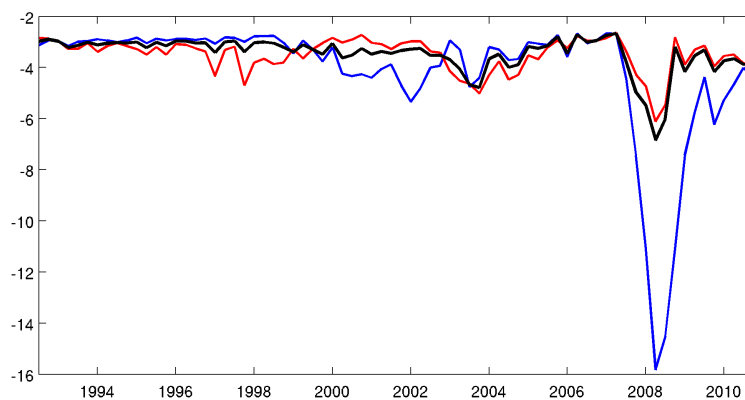
C Additional Tables and Figures

Table A-1 summarizes the prior distribution. Figures A-4 to A-9 are the same as those presented in the main results section but with $\rho = 0.75$.

Table A-1: Priors

	Density	Mean	St. Dev.		Density	Mean	St. Dev.
Panel I: Smets-Wouters Model (SW)							
<i>Policy Parameters</i>							
ψ_1	Normal	1.50	0.25	ρ_R	Beta	0.75	0.10
ψ_2	Normal	0.12	0.05	ρ_{r^m}	Beta	0.50	0.20
ψ_3	Normal	0.12	0.05	σ_{r^m}	InvG	0.10	2.00
<i>Nominal Rigidities Parameters</i>							
ζ_p	Beta	0.50	0.10	ζ_w	Beta	0.50	0.10
<i>Other “Endogenous Propagation and Steady State” Parameters</i>							
α	Normal	0.30	0.05	π^*	Gamma	0.75	0.40
Φ	Normal	1.25	0.12	γ	Normal	0.40	0.10
h	Beta	0.70	0.10	S''	Normal	4.00	1.50
ν_l	Normal	2.00	0.75	σ_c	Normal	1.50	0.37
ι_p	Beta	0.50	0.15	ι_w	Beta	0.50	0.15
r_*	Gamma	0.25	0.10	ψ	Beta	0.50	0.15
(Note $\beta = (1/(1 + r_*/100))$)							
<i>ρs, σs, and ηs</i>							
ρ_z	Beta	0.50	0.20	σ_z	InvG	0.10	2.00
ρ_b	Beta	0.50	0.20	σ_b	InvG	0.10	2.00
ρ_{λ_f}	Beta	0.50	0.20	σ_{λ_f}	InvG	0.10	2.00
ρ_{λ_w}	Beta	0.50	0.20	σ_{λ_w}	InvG	0.10	2.00
ρ_μ	Beta	0.50	0.20	σ_μ	InvG	0.10	2.00
ρ_g	Beta	0.50	0.20	σ_g	InvG	0.10	2.00
η_{λ_f}	Beta	0.50	0.20	η_{λ_w}	Beta	0.50	0.20
η_{gz}	Beta	0.50	0.20				
Panel II: Model with Long Run Inflation Expectations (SWπ)							
ρ_{π^*}	Beta	0.50	0.20	σ_{π^*}	InvG	0.03	6.00
Panel III: Financial Frictions (SWFF)							
SP_*	Gamma	2.00	0.10	$\zeta_{sp,b}$	Beta	0.05	0.005
ρ_{σ_w}	Beta	0.75	0.15	σ_{σ_w}	InvG	0.05	4.00

Notes: Smets and Wouters (2007) original prior is a $Gamma(.62, .10)$. The following parameters are fixed in Smets and Wouters (2007): $\delta = 0.025$, $g_* = 0.18$, $\lambda_w = 1.50$, $\varepsilon_w = 10$, and $\varepsilon_p = 10$. In addition, for the model with financial frictions we fix the entrepreneurs’ steady state default probability $\bar{F}_* = 0.03$ and their survival rate $\gamma_* = 0.99$. The columns “Mean” and “St. Dev.” list the means and the standard deviations for Beta, Gamma, and Normal distributions, and the values s and ν for the Inverse Gamma (InvG) distribution, where $p_{IG}(\sigma|\nu, s) \propto \sigma^{-\nu-1} e^{-\nu s^2/2\sigma^2}$. The effective prior is truncated at the boundary of the determinacy region. The prior for \bar{l} is $\mathcal{N}(-45, 5^2)$.

Figure A-4: $[\rho = 0.75]$ Log Scores: SW+ π FF (*red*), SW+ π (*blue*) and Pooled (*black*)Table A-2: $[\rho = 0.75]$ Cumulative Log Scores

Component Models		Model Pooling	
Model	Log score	Method	Log score
SW π	-306.34	BMA	-275.57
SWFF	-259.58	SP	-264.67
		EW	-260.42
		DP	-256.75

Figure A-5: $[\rho = 0.75]$ Log Score Comparison: Pooled over SW+ π FF (*red*), Pooled over SW+ π (*blue*)

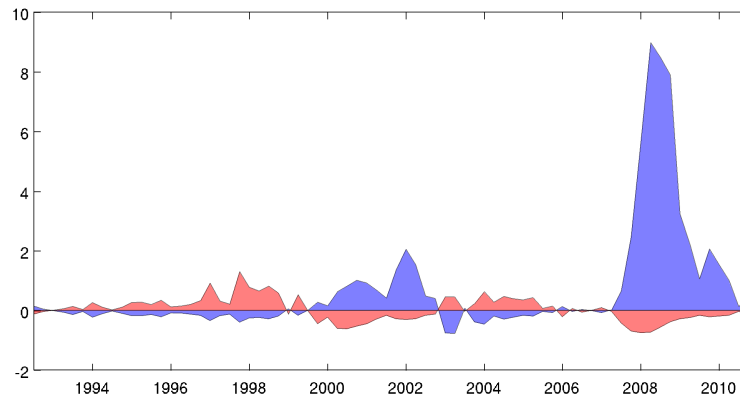


Figure A-6: $[\rho = 0.75]$ Filtered Posterior of λ_t (pool weight on SW+ π FF), with 50%, 68% and 90% bands

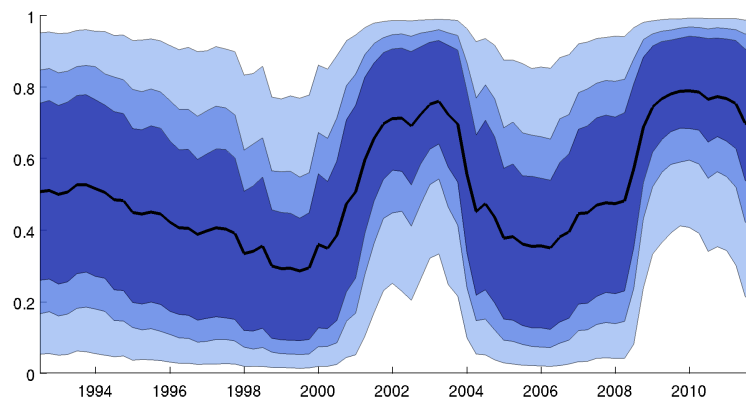


Figure A-7: [$\rho = 0.75$] Filtered Posterior of λ_t (pool weight on SW+ π FF), with 50%, 68% and 90% bands: Great Recession Period

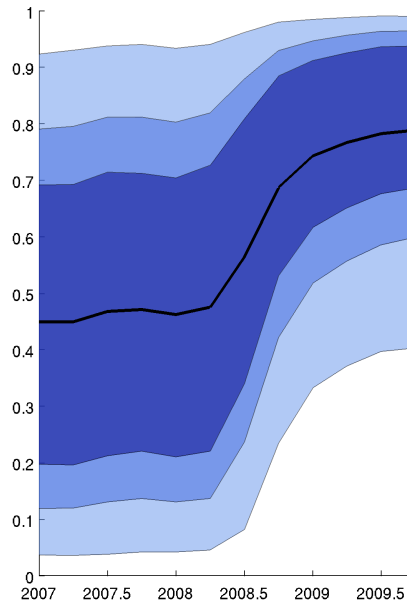


Figure A-8: [$\rho = 0.75$] Dynamic Pool Weight (*black*), BMA (*green*) and Optimal Static Pool (*purple*) [All weights on SW+ π FF]

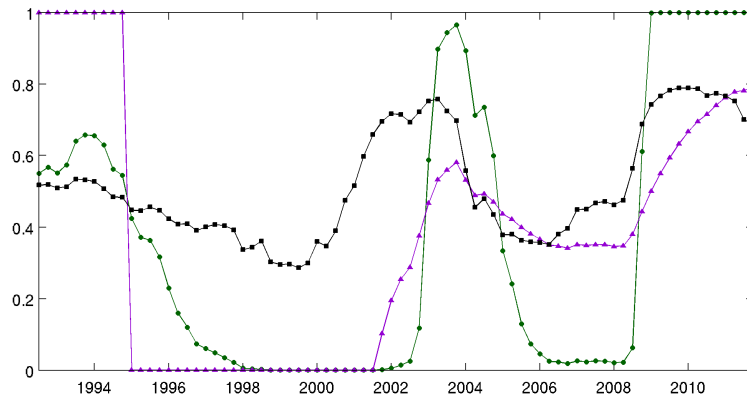


Figure A-9: [$\rho = 0.75$] Log Score Comparison: Dynamic Pool over BMA (*green*) and Dynamic Pool over Optimal Static Pool (*purple*) and Dynamic Pool over Equally Weighted Pool (*black*)

