Corrigendum for "Dynamic prediction pools: An investigation of financial frictions and forecasting performance"*

Michael Cai William Chen Marco Del Negro Aidan Gleich
Frank Schorfheide

July 19, 2021

1 Introduction

We realized the existence of a coding mistake in Del Negro et al. (2016). The mistake was not in the code for the "dynamic pools", but rather in one of the primitives: the predictive densities for the so-called SWπ model. While the model was estimated correctly, two of the observables were swapped in the computation of the predictive densities. This mistake penalized the predictive ability of the SWπ model relative to its alternative, the SWFF model, and has implications for many of the results shown in the paper. In particular, as discussed in the next section, the dynamic pools weight posterior distribution is substantially more skewed toward the SWπ model before the Great Recession.

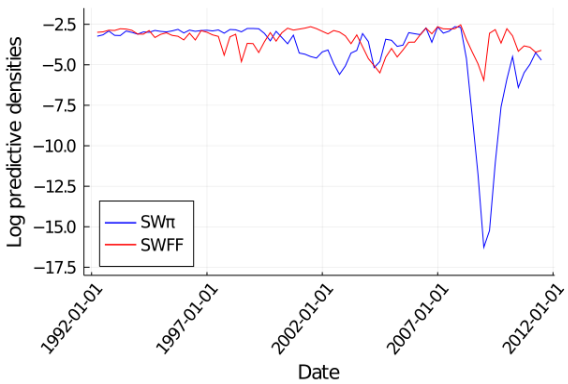
Nonetheless, the broad message of Del Negro et al. (2016) holds: dynamic pools is an attractive alternative for combining density forecasts relative to Bayesian Model Averaging and Static Pools (Geweke and Amisano (2011), Hall and Mitchell (2007)).

2 Replication Figures

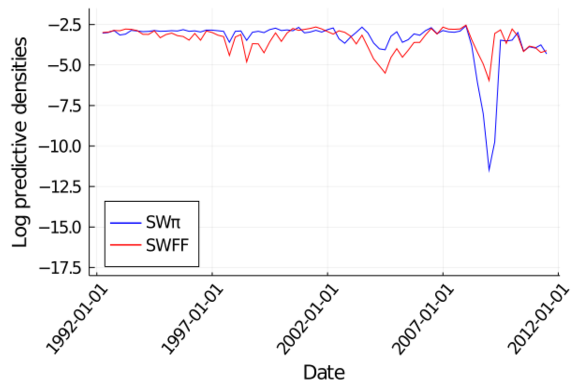
Below we plot the main figures in Del Negro et al. (2016). For each pair of plots, the left column replicates the paper's results, which are obtained using the wrong predictive densities, while the right column shows the results obtained from the correct predictive densities. Refer to the original paper for definitions of the notation and objects.

*Contact Information: M. Cai: Northwestern University, michaeldcai@gmail.com. William Chen: FRB New York, william.chen@ny.frb.org. Marco Del Negro: FRB New York, marco.delnegro@ny.frb.org. A. Gleich: FRB New York, aidan.gleich@ny.frb.org. F. Schorfheide: University of Pennsylvania, schorf@ssc.upenn.edu. The views expressed herein are solely those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System.

Figure 1 depicts the log scores $p(\overline{y}_{t+h,h}|\mathscr{I}^m_{t+}, \mathscr{M}_m)$ (semi-conditional forecasts) for the predictions of four-quarter-ahead ($h = 4$) average output growth and inflation for the SWπ (blue) and the SWFF (red) model. The forecast origins range from 1992:Q1 to 2011:Q2. These scores are subsequently used as the inputs for the density combination.



(a) Replication



(b) Corrected

Figure 1: Log scores comparisons of SWFF with SWπ

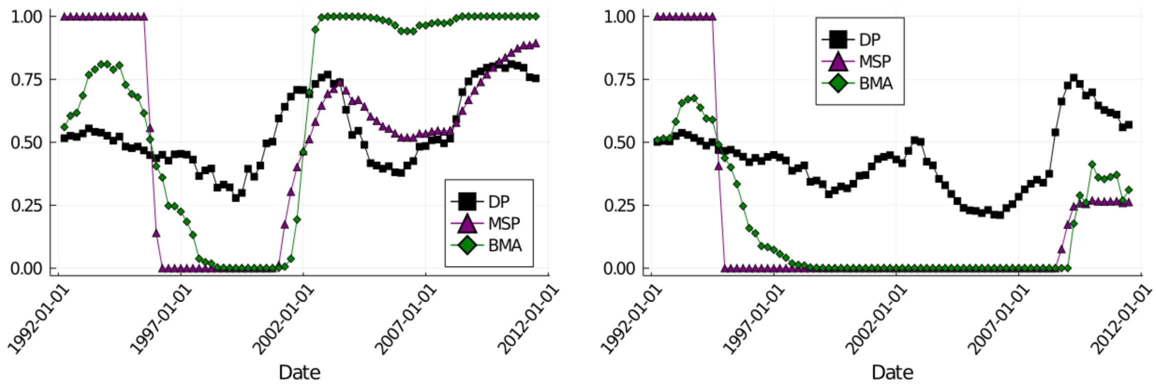
Comparing the replication with the correction, the salient difference is the improved log scores of the SWπ model. While the relative performance of SWFF model still improves during periods of financial turmoil, the SWπ model does not do as poorly, particularly during the early-200-s dot-com bust. As we shall see, this difference will change the real-time behavior of the dynamic prediction pool.

Figure 2 shows the weight on the SWFF model in forecast pools over the period 1992:Q1-2011:Q2 obtained from the proposed dynamic pooling technique ($\widehat{\lambda}^{DP}_{t|t}$–black) as well as the static pool ($\widehat{\lambda}^{MSP}_t$–purple) with weights estimated by maximum likelihood, and BMA ($\widehat{\lambda}^{BMA}_t$-green). All of the weights and the hyperparameter estimates for the DP are computed in real time, based on information that would have been available to the policymaker at the time of the combination of the model forecasts.

Because the SWπ model performs better in the corrected densities, the static pool and BMA weights no longer rise toward 1 starting around 2001:Q3. Instead, they remain close to zero even several quarters after the financial crisis. In contrast, while the dynamic pool does place greater weight on the SWπ model in the early 2000s compared to the replication results, the dynamic pool weights quickly rise to similar magnitudes as the replication when the Great Recession occurs.

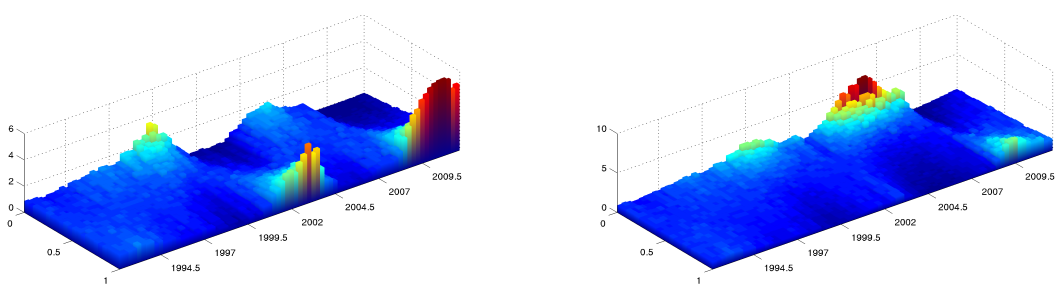
Figures 3 and 4 show the posterior distributions $p^{(h)}_{DP}(\lambda_t|\mathscr{I}^{\mathscr{P}}_t, \mathscr{P})$ and $p^{(h)}_{BSP}(\lambda|\mathscr{I}^{\mathscr{P}}_t, \mathscr{P})$, respectively. Since the SWπ model performs better on the corrected density, the mode of the dynamic pool posterior distribution $p^{(h)}_{DP}(\lambda_t|\mathscr{I}^{\mathscr{P}}_t, \mathscr{P})$ stays much closer to $\lambda = 0$ until the end of the data sample when the Great Recession occurs.

The change in the posterior distribution of the static pool is more pronounced. Because the SWFF and SWπ log scores from using the incorrect predictive densities alternated in relative performance much more frequently, the static pool's mode switches between $\lambda = 0$



(a) Replication (b) Corrected

Figure 2: Weight on SWFF in real time: Dynamic pool (DP, Prior 1), BMA, and maximum likelihood static pool (MSP)



(a) Replication (b) Corrected

Figure 3: $p_{DP}^{(h)}(\lambda_t|I_t^{\mathscr{P}}, \mathscr{P})$ (using Prior 1)

and $\lambda = 1$ in the same way that the dynamic pool did. However, according to the corrected predictive densities, the SWπ log scores are as good or better than the SWFF log scores before the Great Recession. Due to this change, the static pool places almost all its weight at $\lambda = 0$ for the majority of the data sample. It barely reacts to the Great Recession precisely because it estimates λ as static.

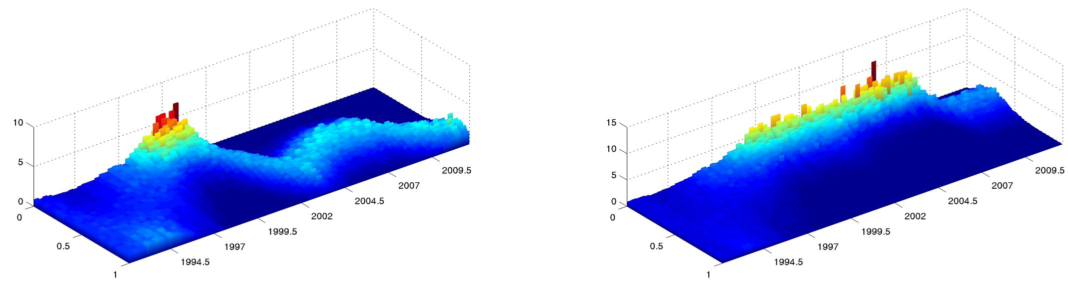
Figures 5, 6, and 7 display the end-of-sample ($t = T$) hyperparameter posterior distributions. Figure 5 shows the posterior $p^{(h)}(\rho|\mathscr{I}_T^{\mathscr{P}}, \mathscr{P})$ (histogram) using the following prior:

$$\text{Hyperparameter Prior 1: } \rho \sim \mathscr{U}[0,1], \quad \mu = 0, \quad \sigma^2 = 1.$$

The corrected posterior has slightly more mass for low ρ, but the overall distribution does not change much.

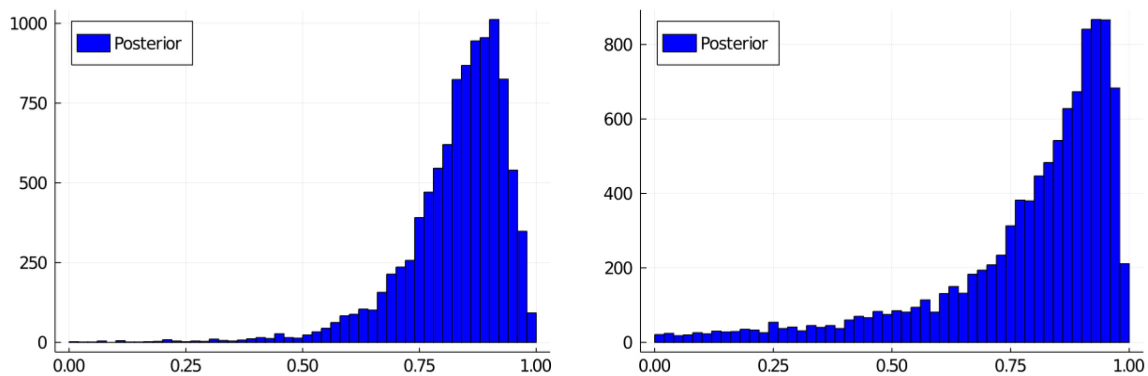
Posterior distributions for σ^2 and $\Phi(\mu)x$ are depicted in Figures 6 and 7, respectively. These results are based on the following prior distribution:

$$\text{Hyperparameter Prior 2: } \rho \sim \mathscr{B}(0.8, 0.1), \quad \mu \sim \mathscr{N}(0, \Phi^{-1}(0.75)), \quad \sigma^2 \sim \mathscr{IG}(1, 4).$$



(a) Replication (b) Corrected

Figure 4: $p_{BSP}^{(h)}(\lambda_t | I_t^{\mathscr{P}}, \mathscr{P})$



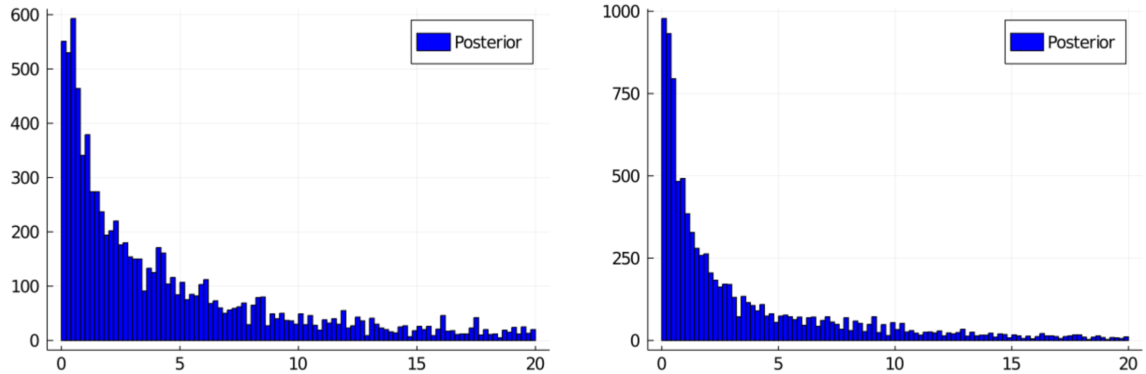
(a) Replication (b) Corrected

Figure 5: Posterior distribution of ρ using Prior 1 ($\rho \sim \mathscr{U}[0,1]$)

Relative to the inverse Gamma prior distribution for σ^2 (not plotted here) the posterior mass is shifted toward the right and is greater than one with very high probability. Recall that for values of $\sigma^2 > 1$ the prior distribution of is U-shaped, which tends to shift the posterior mean of toward one of the endpoints ($\lambda = 0$ or $\lambda = 1$). Thus, the estimation results indicate that the data favor a parameterization in which the posterior mean is more sensitive to the arrival of new information.

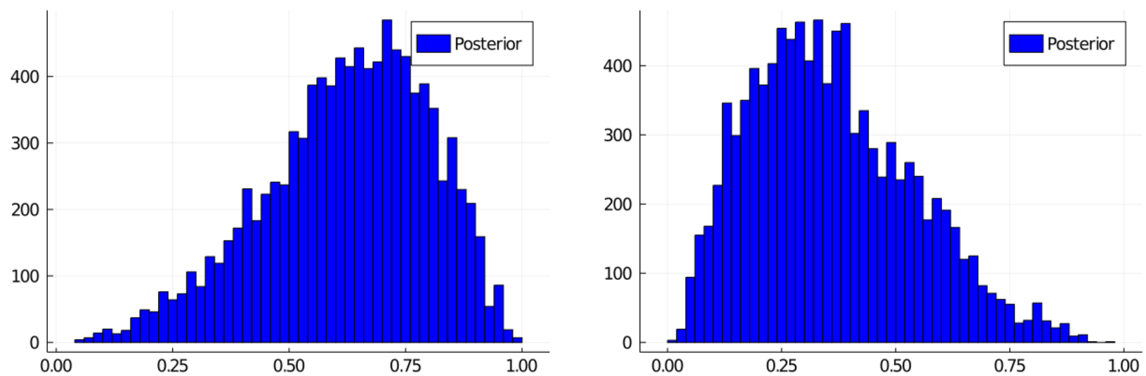
Note that the prior on σ^2 is different from the one reported in the paper ($\mathscr{IG}(2,1)$) because the reported prior is a typo. The true prior used in the code is the prior stated above. The posterior distribution for σ^2 does not change much after using the corrected densities. High values of σ^2 have slightly less mass and a greater concentration around low σ^2. Intuitively, the estimation results indicate that the corrected data favor a parametrization in which the posterior mean is a little more sensitive to the arrival of new information.

The hyperparameter μ determines the location of the seesaw fulcrum, which corresponds to equal weights on both models if $\mu = 0$. Unlike in Figures 5 and 6, the posterior distribution of μ depicted in Figure 7 changes substantially. With the incorrect densities, the distribution skewed left, and the mode occured around $\mu = 0.7$. With the corrected densities, the



(a) Replication (b) Corrected

Figure 6: Posterior distribution of σ^2 using Prior 2 ($\rho \sim \mathscr{B}(0.8, 0.1)$, $\mu \sim \mathscr{N}(0, \Phi^{-1}(0.75))$, $\sigma^2 \sim \mathscr{IG}(1, 4)$)



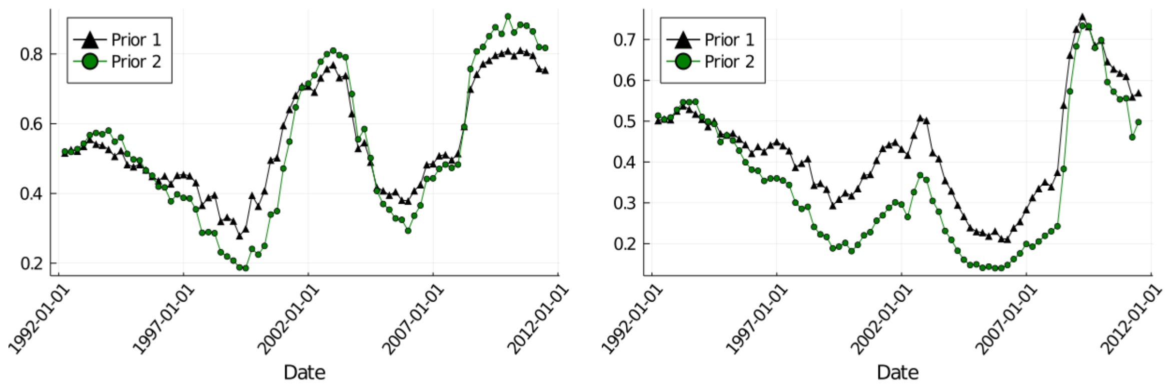
(a) Replication (b) Corrected

Figure 7: Posterior distribution of $\Phi(\mu)$ using Prior 2 ($\rho \sim \mathscr{B}(0.8, 0.1)$, $\mu \sim \mathscr{N}(0, \Phi^{-1}(0.75))$, $\sigma^2 \sim \mathscr{IG}(1, 4)$)

distribution now skews right with a mode around $\mu = 0.3$. Nevertheless, the posterior still assigns substantial probability toward weights greater than 0.5, suggesting that the data remain uninformative about which model is better than the other on average.

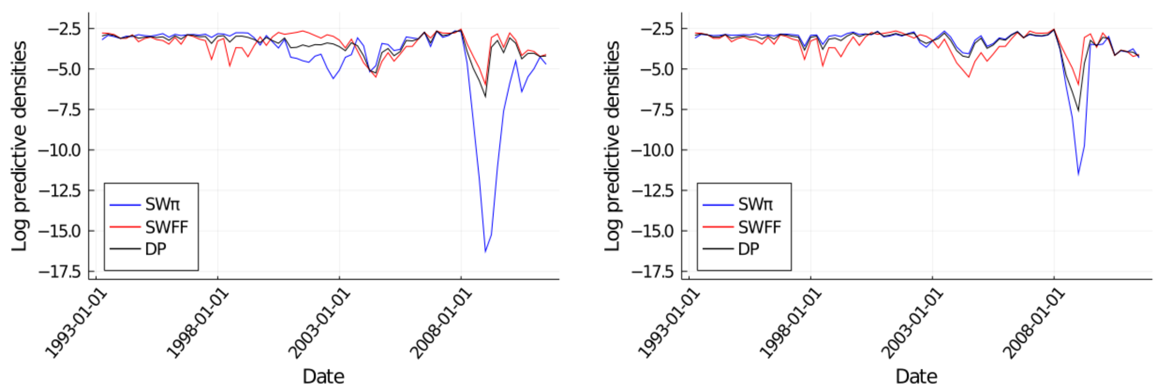
Figure 8 shows the effect of estimating the hyperparameters μ and σ on the evolution of the posterior mean $\widehat{\lambda}_{t|t}^{DP}$. The figure compares the weight on the SWFF under Priors 1 and 2. The effect of estimating μ and σ remain similar to the original results: the swings of $\widehat{\lambda}_{t|t}^{DP}$ are more prounounced under Prior 2 than under Prior 1.

Figures 9 and 10 evaluate the performance of Prior 2. Figure 9 compares the log score of the dynamic pool (black line) over time to that of its two components: the SWFF model (red) and SW model (blue). Recall that the forecasts are generated based on an information set that includes current interest rates and spreads (denoted by $\mathscr{I}_{t^+}^{\mathscr{P}}$). The predictive density associated with the dynamic pool, $p_{DP}^{(h)}(\overline{y}_{t+h,h}|\mathscr{I}_{t^+}^{\mathscr{P}}, \mathscr{P})$ is a linear combination of the predictive densities of the two DSGE models, $p^{(h)}(\overline{y}_{t+h,h}|\mathscr{I}_{t^+}^{m}, \mathscr{M}_m)$, and, for every $\overline{y}_{t+h,h}$, has



(a) Replication (b) Corrected

Figure 8: $\hat{\lambda}_{t|t}^{DP}$: Fixed vs. Estimated μ and σ (Prior 1 vs. Prior 2)



(a) Replication (b) Corrected

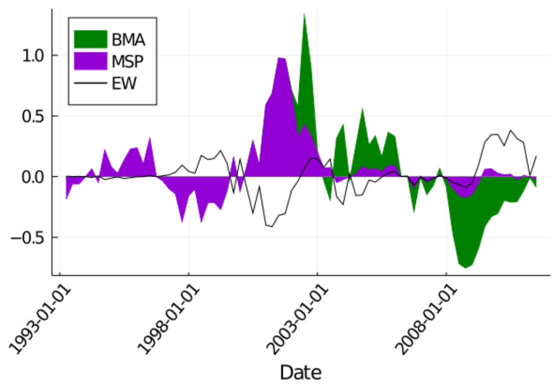
Figure 9: Log scores comparisons over time (using Prior 2)

to lie between the two model-specific density values.

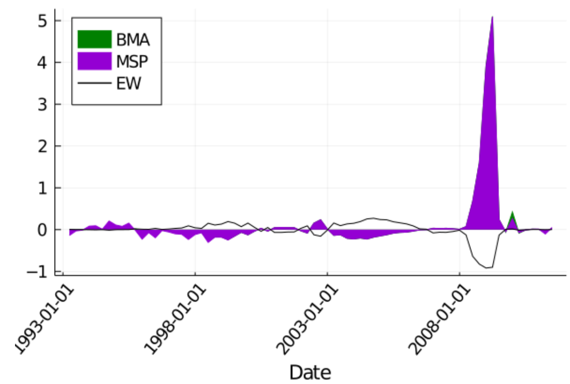
Figure 10 shows the log predictive score differences between the dynamic pool and the following alternatives: BMA (green area), maximum likelihood static pools (purple area), and equal weights (black line). Positive differentials favor the dynamic pool. In the early part of the sample there are no major differences in forecasting performance.

After correcting the predictive densities, the large performance differentials in the early 2000s vanishes. However, around the time of the Great Recession, large performance differentials persist. Again, following the eight-year period in which SWπ has been the dominant model, both BMA and MSP are caught off guard by the change in the relative forecast performance of the two DSGE models. The DP also struggles to adjust, as evidenced by the fact that it is still forecasting worse than the equal-weights combination scheme for part of the Great Recession period.

Table 1 shows the cumulative log scores for three DP specifications, as well as the difference between the cumulative log scores of the dynamic pools and that of equal weights,



(a) Replication



(b) Corrected

Figure 10: Dynamic Pool (prior 2) relative to BMA, SP, and Equal Weights

Table 1: Cumulative log scores / differentials computed over the period 1992:Q1-2011:Q2. This table adds an additional column “Orig. DP” because the paper’s original code for this computation erroneously set $\mu = 0$ even when it was being estimated by Prior 2. Differentials are computed relative to DP (not Orig. DP)

Prior	Density	Log Score	Differentials			
		DP	Orig. DP	EW	BMA	MSP
Prior 1	Replication	-256.91	—	1.34	4.07	4.95
	Corrected	-246.07	—	-0.35	9.13	8.65
Prior 2	Replication	-257.31	-256.43	0.94	3.67	4.56
	Corrected	-246.26	—	-0.55	8.93	8.46
Prior 3	Replication	-255.90	—	2.35	5.07	5.96
	Corrected	-245.10	—	0.61	10.10	9.62

BMA, and the static pool (MSP), respectively. We report the log scores from both the replication and correction. Positive differentials favor the dynamic pool. The table also has an additional column, “Orig. DP”, because there was a small error in the code computing the log scores for Prior 2. The code computing the log score reported in the paper erroneously set $\mu = 0$. Correcting this error leads to a slightly higher cumulative log score.

The three dynamic pool specifications continue to have similar performances, but the comparison with the other pooling methods changes. Equal weights now slightly outperform the dynamic pool, owing to the fact that it takes the dynamic pool longer before it weights the SWFF model more than the SWπ model. On the other hand, the performance of the dynamic pool improves even further relative to BMA and the static pool.

For reference, the priors are

- Prior 1: $\rho \sim \mathscr{U}[0,1]$, $\mu = 0$, $\sigma^2 = 1$
- Prior 2: $\rho \sim \mathscr{B}(0.8, 0.1)$, $\mu \sim \mathscr{N}(0, \Phi^{-1}(0.75))$, $\sigma^2 \sim \mathscr{IG}(1,4)$
  - Note again that the Prior 2 reported in the paper was a typo.
- Prior 3: $\rho \sim \mathscr{B}(0.8, 0.1)$, $\mu = 0$, $\sigma^2 \sim \mathscr{IG}(1,4)$

References

Del Negro, Marco, Raiden B Hasegawa, and Frank Schorfheide, "Dynamic prediction pools: an investigation of financial frictions and forecasting performance," *Journal of Econometrics*, 2016, *192* (2), 391–405.

Geweke, John and Gianni Amisano, "Optimal prediction pools," *Journal of Econometrics*, 2011, *164* (1), 130 – 141. Annals Issue on Forecasting.

Hall, Stephen G. and James Mitchell, "Combining Density Forecasts," *International Journal of Forecasting*, 2007, *23*, 1–13.