Figures/Tables Replication

Michael Cai, Marco Del Negro, Edward Herbst,
Ethan Matlin, Rebecca Sarfati, Frank Schorfheide*

Northwestern University, FRB New York, Federal Reserve Board,

FRB New York, University of Pennsylvania

February 7, 2020

1 SMC Estimation at Work

Table 1: AS Model: Fixed and Adaptive Tempering Schedules

	Fixed	α =0.90	α =0.95	α =0.97	α =0.98
Mean log(MDD)	-1032.60	-1034.21	-1032.48	-1032.07	-1031.92
StdD log(MDD)	0.76	1.48	0.61	0.32	0.22
Schedule Length	200.00	112.17	218.80	350.06	505.46
Resamples	14.63	15.37	15.03	14.99	14.00
Runtime [Min]	1.29	0.88	1.53	2.21	3.13

Notes: Results are based on $N_{run} = 400$ runs of the SMC algorithm. We report averages across runs for the runtime, schedule length, and number of resampling steps.

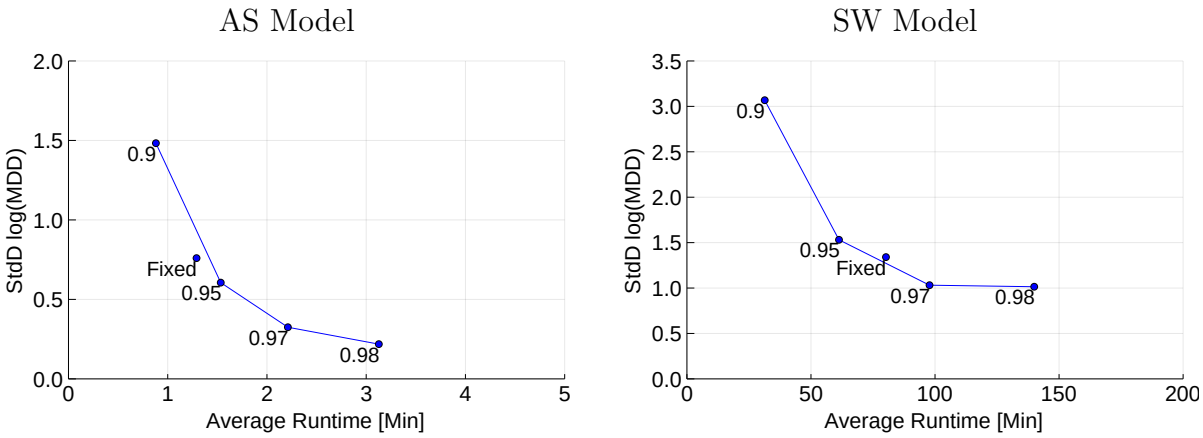
*Correspondence: Michael Cai michaelcai@u.northwestern.edu, Marco Del Negro marco.delnegro@ny.frb.org, Edward Herbst edward.p.herbst@frb.gov, Ethan Matlin ethan.matlin@ny.frb.org, Rebecca Sarfati rebecca.sarfati@ny.frb.org, Frank Schorfheide schorf@ssc.upenn.edu. We thank participants at various conferences and seminars for helpful comments. Schorfheide acknowledges financial support from the National Science Foundation under Grant SES 1851634. The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York, the Federal Reserve Board or the Federal Reserve System.

Table 2: SW Model: Fixed and Adaptive Tempering Schedules

	Fixed	α =0.90	α =0.95	α =0.97	α =0.98
Mean log(MDD)	-1178.43	-1186.23	-1180.04	-1178.31	-1177.72
StdD log(MDD)	1.34	3.07	1.53	1.03	1.01
Schedule Length	500.00	200.53	389.75	618.53	887.42
Resamples	26.75	28.09	27.25	26.33	25.09
Runtime [Min]	80.23	31.36	61.29	97.79	139.99

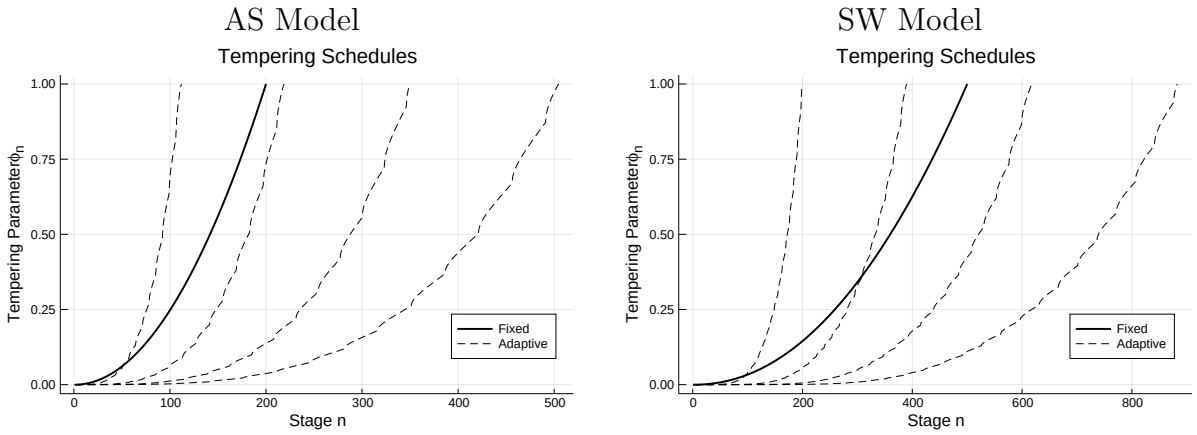
Notes: Results are based on $N_{run} = 200$ runs of the SMC algorithm. We report averages across runs for the runtime, schedule length, and number of resampling steps.

Figure 1: Trade-Off Between Runtime and Accuracy



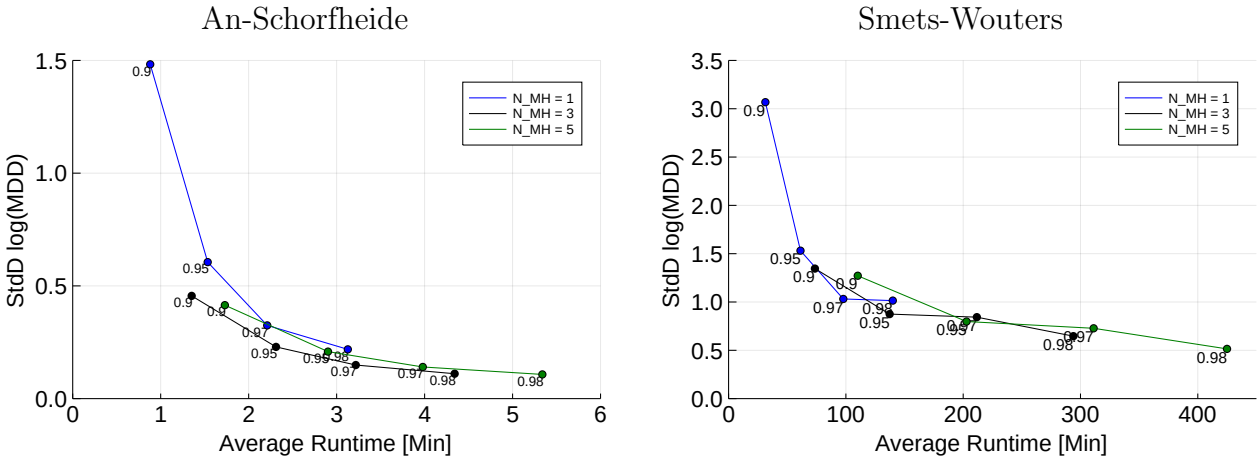
Notes: AS results are based on $N_{run} = 400$ and SW results are based on $N_{run} = 200$ runs of the SMC algorithm.

Figure 2: Tempering Schedules



Notes: The figure depicts (pointwise) median ϕ_n values across $N_{run} = 200$ for AS and $N_{run} = 100$ for SW. The solid lines represent the fixed schedule, parameterized accord ing to Table **??**. The dashed lines represent a range of adaptive schedules: $\alpha = 0.9, 0.95, 0.97, 0.98$.

Figure 3: Trade-Off Between Runtime and Accuracy – Multiple Metropolis-Hastings Steps



Notes: An-Schorfheide results are based on $N_{run} = 400$ and Smets-Wouters results are based on $N_{run} = 200$ runs of the SMC algorithm.

Table 3: AS Model: Generalized Tempering

	α =0.90	α =0.95	α =0.97	α =0.98
Mean log(MDD)	-1033.95	-1032.54	-1032.06	-1031.93
StdD log(MDD)	1.37	0.61	0.32	0.24
Schedule Length	24.33	47.12	75.74	106.50
Runtime [Min]	0.25	0.48	0.69	0.98

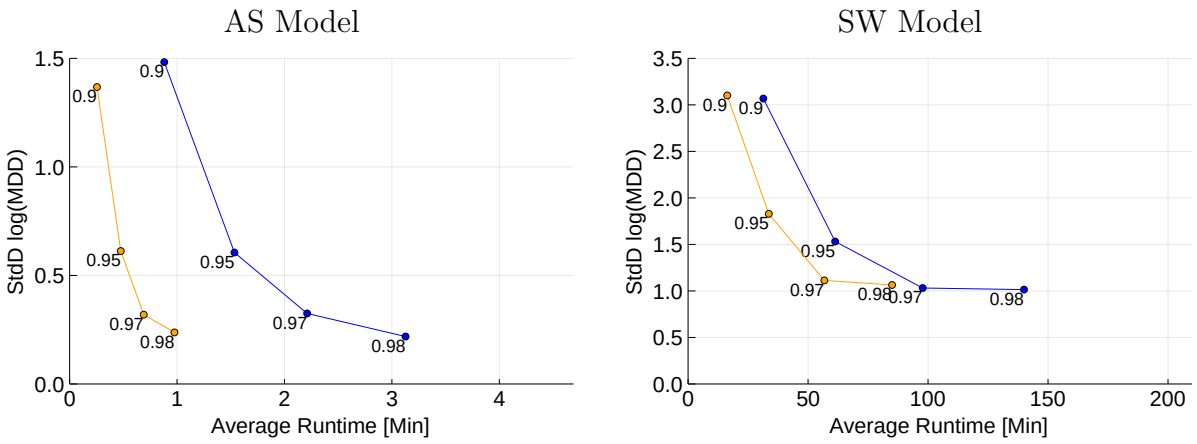
Notes: Results are based on $N_{run} = 400$ runs of the SMC algorithm, starting from particles that represent $p(\theta|Y_{1:T_1})$. We report averages across runs for the runtime, sche dule length, and number of resampling steps.

Table 4: SW Model: Generalized Tempering

	α =0.90	α =0.95	α =0.97	α =0.98
Mean log(MDD)	-1188.93	-1182.08	-1180.05	-1178.90
StdD log(MDD)	3.10	1.83	1.11	1.06
Schedule Length	56.60	115.73	194.01	290.74
Runtime [Min]	16.32	33.64	56.79	85.01

Notes: Results are based on $N_{run} = 200$ runs of the SMC algorithm, starting from particles that represent $p(\theta|Y_{1:T_1})$. We report averages across runs for the runtime, sche dule length, and number of resampling steps.

Figure 4: Trade-Off Between Runtime and Accuracy



Notes: AS results are based on $N_{run} = 400$ and SW results are based on $N_{run} = 200$ runs of the SMC algorithm. Yellow squares correspond to generalized tempering and blue circles correspond to full sample estimation.

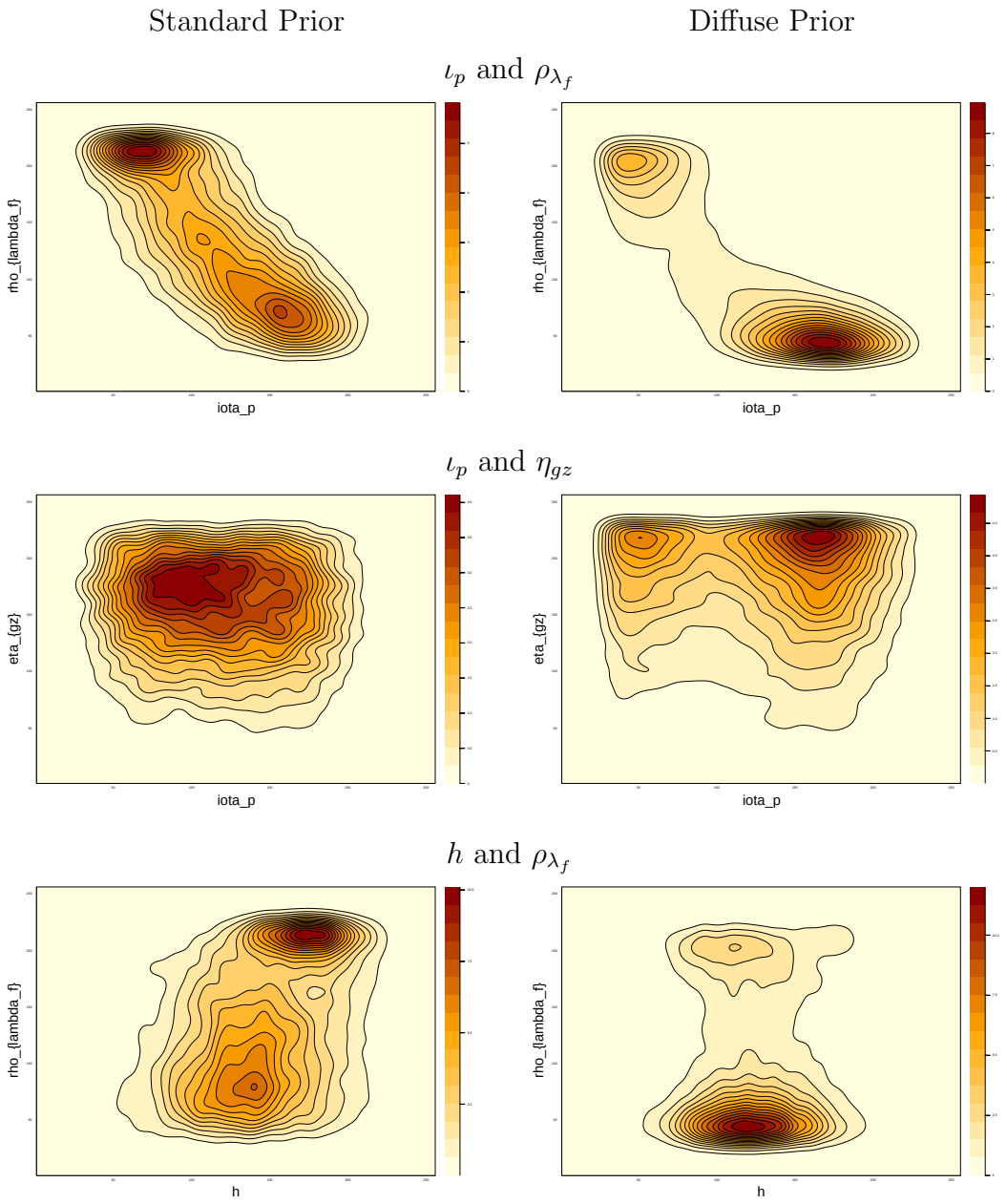
Figure 5: AS Model: Log MDD Increments $\log(p(y_{T_{s-1}+1:T_s}|y_{1:T_{s-1}}))$

Notes: Results are based on $N_{run} = 200$ runs of the SMC algorithm with $\alpha = 0.95$.

Figure 6: AS Model: Evolution of Posterior Means and Coverage Bands

Notes: Sequence of posterior means (red line) and 90% coverage bands black lines. The dashed line indicates the temporal average of the posterior means. We use $\alpha = 0.95$ for the SMC algorithm.

Figure 7: SW Model: Posterior Contours for Selected Parameter Pairs

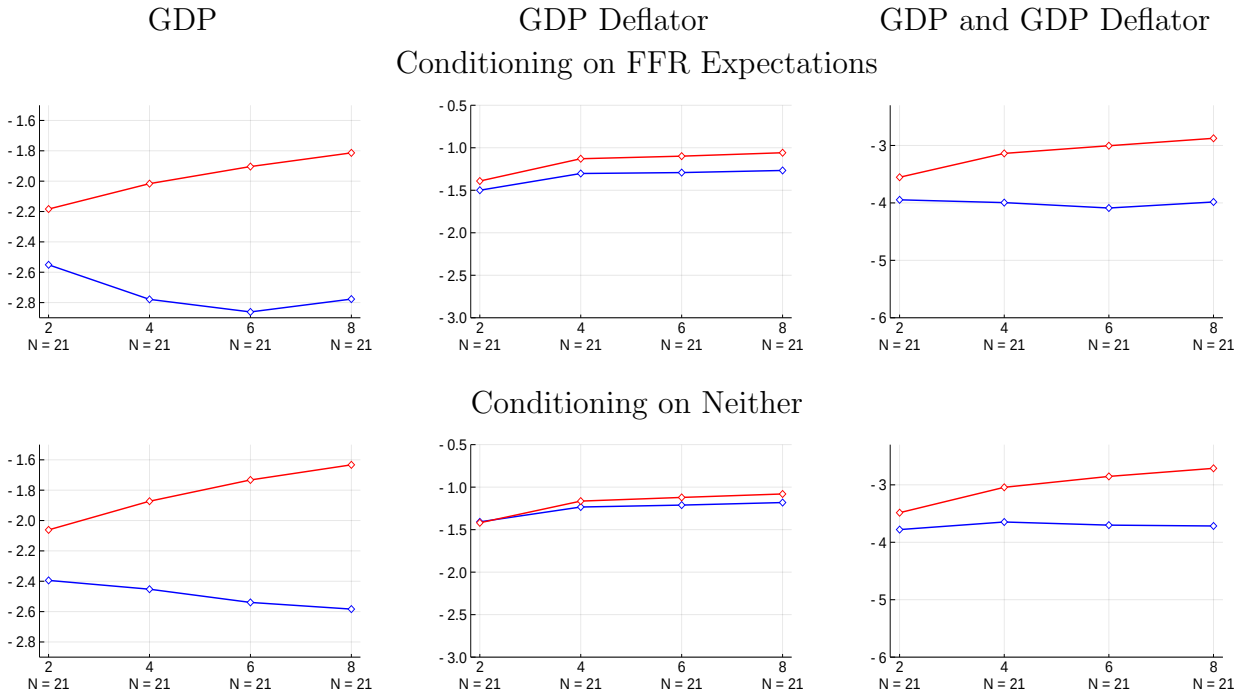


Notes: Estimation sample is 1960:Q1 to 1991:Q3. We use $\alpha = 0.98$ for the SMC algorithm. Plots show a two-dimensional visualization of the full-dimension joint posterior.

2 Predictive Density Evaluations

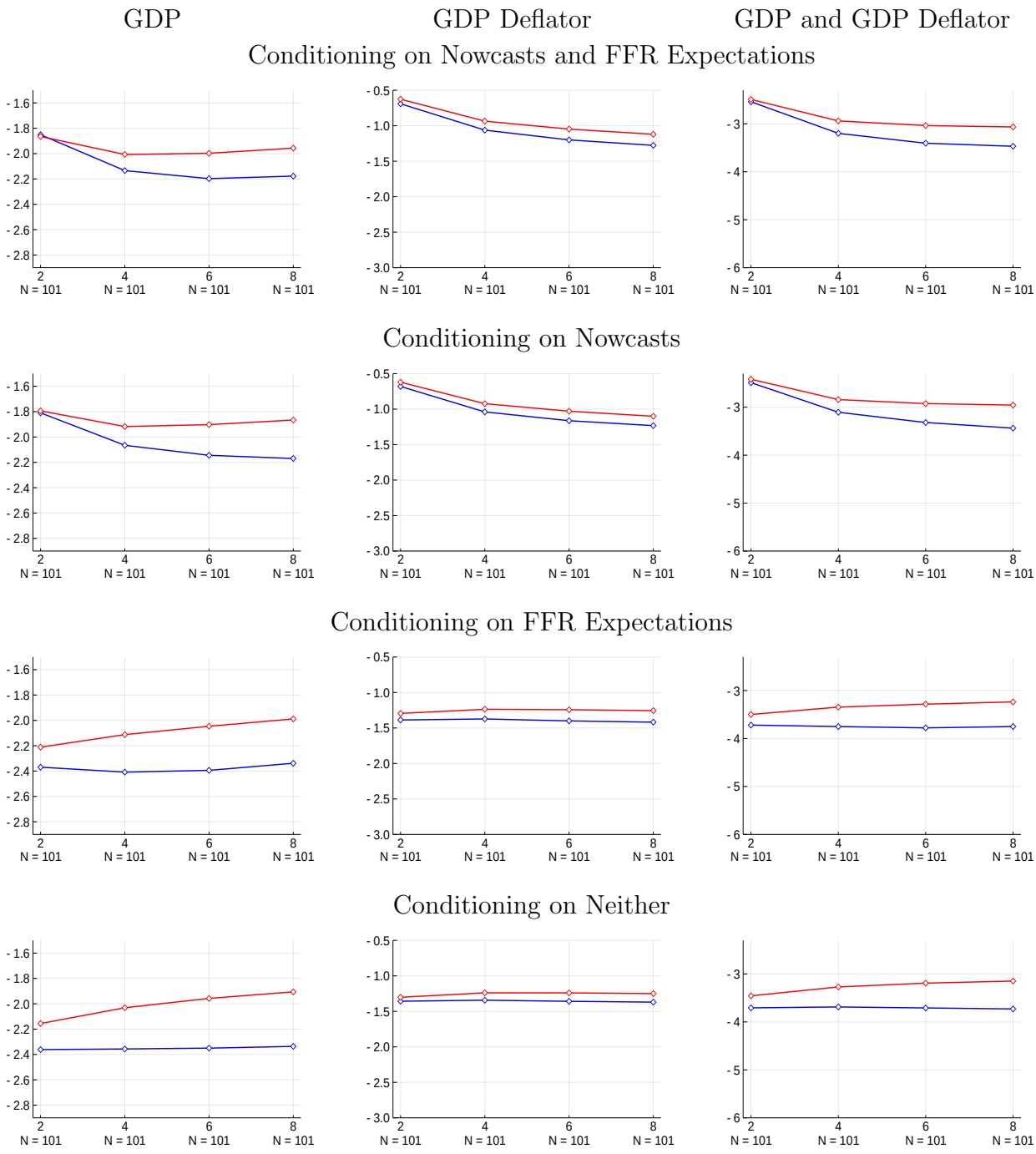
2.1 Log Predictive Density Scores with Standard Prior

Figure 10: Average Log Predictive Density Scores for SW vs SWFF: Post-Recession Sample



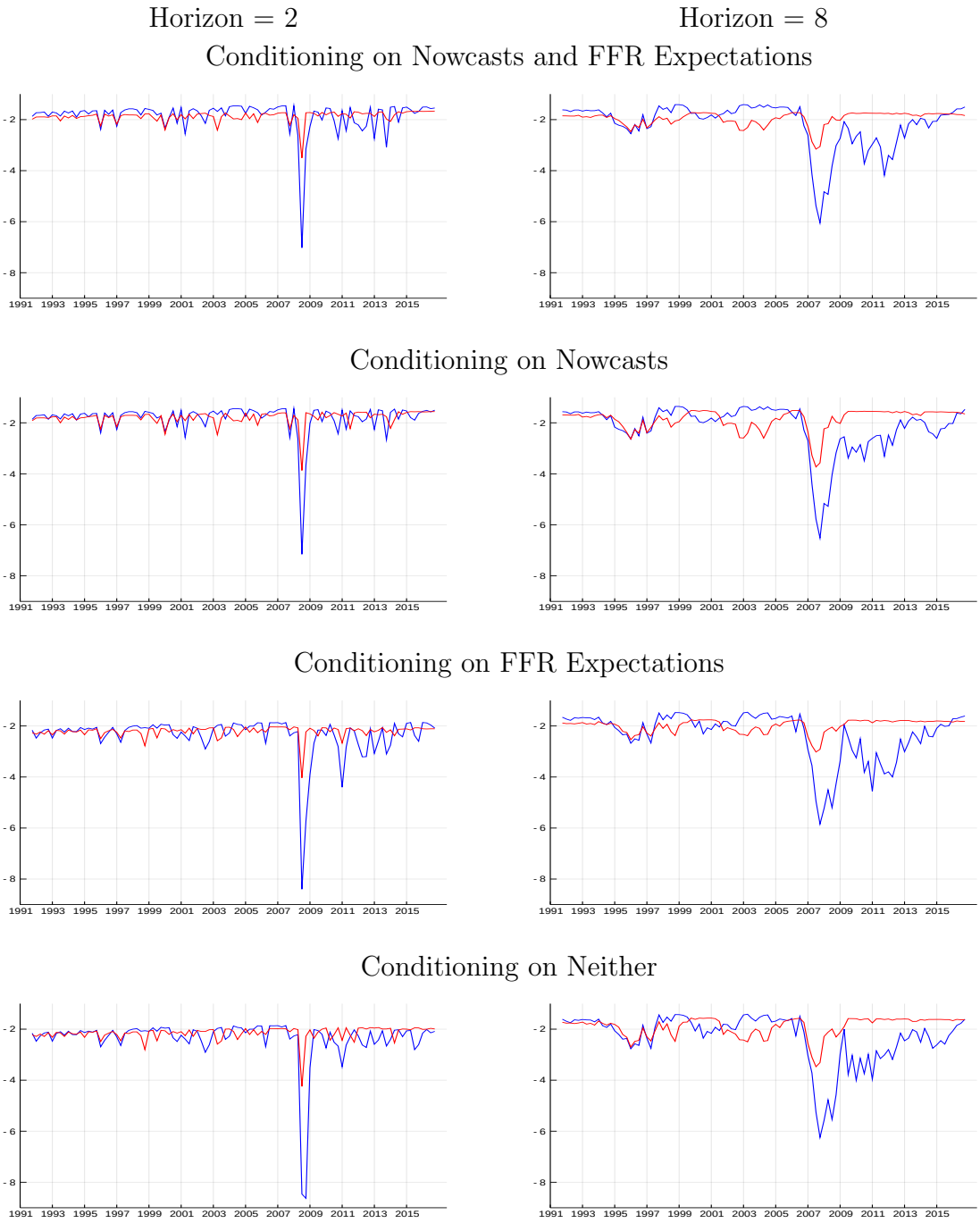
Note: These panels compare the log predictive densities from the SW DSGE model (blue diamonds) and with the SWFF DSGE model (red diamonds) averaged over two, four, six, and eight quarter horizons for output growth and inflation individually, and for both together. Forecast origins from April 2011 to April 2016 only are included in these calculations.

Figure 8: Average Log Predictive Scores for SW vs SWFF



Note: These panels compare the log predictive densities from the SW DSGE Model (blue diamonds) with the SWFF DSGE model (red diamonds) averaged over two, four, six, and eight quarter horizons for output growth and inflation individually, and for both together. Forecast origins from January 1992 to January 2017 only are included in these calculations.

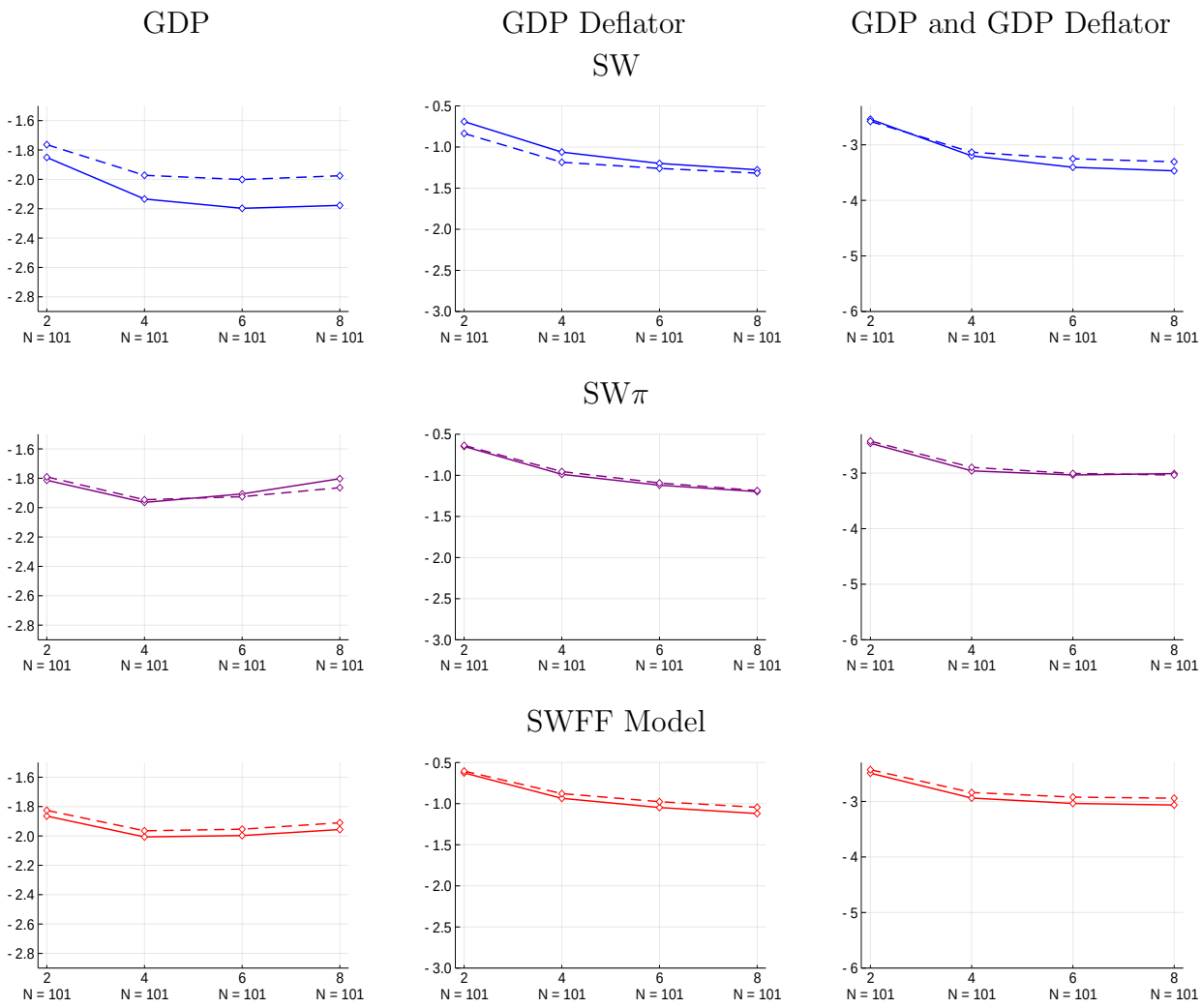
Figure 9: Log Predictive Scores for GDP Growth Over Time — SW vs SWFF



Note: These panels compare $\log p(\bar{y}_{t+h,h}|\mathcal{I}_t^m, \mathcal{M}_m)$ from the SW DSGE model (blue) with the SWFF DSGE model (red) averaged over two, four, six, and eight quarter horizons for output growth. Forecast origins from January 1992 to January 2017 only are included in these calculations. The x-axis shows the quarter in which the forecasts were generated (time $T+1$ in the parlance of section **??**).

2.2 Log Predictive Density Scores with Diffuse Priors

Figure 11: Comparison of Predictive Densities under Standard and Diffuse Priors



Note: These panels compare the predictive densities estimated with the standard (solid line) and the diffuse (dashed line) prior for the SW (blue, top row), SWπ (purple, middle row), and SWFF (red, bottom row) DSGE models. The predictive densities are averaged over two, four, six, and eight quarter horizons for output growth and inflation individually, and for both together. Forecast origins fro m January 1992 to January 2017 only are included in these calculations. The forecasts associated with these predictive densities are generated conditioning on nowcasts and FFR expectations.