FEDERAL RESERVE BANK
PHILADELPHIA

Working Papers
RESEARCH DEPARTMENT

A Structural Approach to Combining External and DSGE Model Forecasts

Thorsten Drautzburg
Federal Reserve Bank of Philadelphia Research Department

WP 23-10

PUBLISHED
June 2023

ISSN: 1962-5361

DOI: https://doi.org/10.21799/frbp.wp.2023.10

A Structural Approach to Combining External and DSGE Model Forecasts

Thorsten Drautzburg [*]

December 2022

Abstract

This note shows that combining external forecasts such as the *Survey of Professional Forecasters* can significantly increase DSGE forecast accuracy while preserving the interpretability in terms of structural shocks. Applied to pseudo real-time from 1997q2 onward, the canonical Smets and Wouters (2007) model has significantly smaller forecast errors when giving a high weight to the SPF forecasts. Incorporating the SPF forecast gives a larger role to risk premium shocks during the global financial crisis. A model with financial frictions favors a larger weight on the DSGE model forecast.

Keywords: Forecasting; model averaging; DSGE model; judgmental forecasts.

JEL codes: C32, C53

[*] Federal Reserve Bank of Philadelphia, email: tdrautzburg[at]gmail.com. This draft has benefited from comments from Elmar Mertens and Mikkel Plagborg-Möller.

1 Introduction

Averaging forecasts from different models is a well established tool for improving forecast accuracy; see, for example, Stock and Watson (2004), Wright (2009), Clark and McCracken (2010), and Geweke and Amisano (2012). Combining forecasts from different models ex post, however, gives up on the benefit of providing a consistent narrative that using a single structural model affords. Full-information estimation of Dynamic Stochastic General Equilibrium (DSGE) models, such as first-generation models like Smets and Wouters (2007, henceforth SW) or models with financial frictions such as Del Negro et al. (2015, henceforth DNGS) as well as structural VARs, in contrast, allow for decomposing forecasts into structural shocks. This note proposes a method for combining external forecasts with structural models to improve forecast accuracy while maintaining their interpretability.

In applications with two quantitative medium-scale DSGE models for the U.S., I find that model-based forecasts can be greatly improved: Mean absolute forecasts errors (MAEs) of output growth can be reduced by more than 50% at horizons up to three years when incorporating professional forecasts. MAE reductions for inflation are up to 30% at some horizons. However, DSGE models can contain valuable information beyond what is contained in the *Survey of Professional Forecasters* (SPF): The pure DNGS does no worse than the SPF-augmented model forecast for inflation forecasts at various horizons. This contrasts with the earlier SW model that always favors a high weight on the SPF forecasts. This corroborates the argument in DNGS that financial frictions help to explain inflationary dynamics in the aftermath of the Global Financial Crisis.

The approach here dates back to at least Del Negro and Schorfheide (2013). Del Negro and Schorfheide (2013) distinguishes two approaches for incorporating forecasts into DSGE models: (1) In the news approach, the outside forecast is an observation of the model forecast. It then improves forecasts insofar as it reveals additional information about the economy. This approach has originally been proposed by Monti (2010). (2) In the noise approach, the outside forecast is modeled as a noisy realization of the truth. The latter approach can improve forecasts even when the model forecast is perfectly pinned down by other observables – and it is the approach taken here. Calibrating the amount of noise controls how far the structural forecasts are tilted towards the model forecast. Both the news-based and the noise-based approaches allow for implicit forecast averaging that allows a structural interpretation of forecasts. If the structural interpretation is not needed, statistical approaches such as Geweke and Amisano (2012) are more natural.

In applications, the literature such as Monti (2010), Del Negro and Schorfheide (2013), and Smets et al. (2014) has focused on nowcasts, or on the long-run (Del Negro and Schorfheide, 2013). The emphasis on nowcasts may be because when forecasts are treated as noisy observations of agents' expectations in models in which the invertibility condition of Fernández-Villaverde et al. (2007, henceforth ABCD in light of their key model and result) holds, observing forecasts only helps to reveal the current state of the economy ahead of the release of official statistics. Here, I show that the gains are also sizable at intermediate horizons.

One challenge of working with external forecasts can be their irregular structure, see Mertens et al. (2022). The forecast consists of a mix of fixed horizon forecasts and fixed event forecasts, i.e., calendar-year forecasts. The forecasts also differ in whether they report quarterly growth rates or growth of annual averages, which corresponds to a weighted moving average of quarterly growth rates. Similar to Mertens et al. (2022), the approach here takes these changes into account by adjusting the timing of the observation equations and their calibrated coefficients by the calendar quarter.[1]

The focus is on utilizing the external forecasts to improve short to medium-run forecast performance. Del Negro and Schorfheide (2013) discuss incorporating SPF expectations of long-run output growth and long-run inflation expectations. These approaches are complementary and can be combined, as the application to the Del Negro et al. (2015) model, which incorporates long-run inflation forecasts, shows.

The code accompanying this note can readily be adopted to other Dynare-based (Adjemian et al., 2011) DSGE models to combine them with real-time SPF forecasts of GDP growth, inflation, interest rates, and the unemployment rate, or a subset thereof. All it takes is to adequately (re-)label the observation equations and to include lines of code that add the extra observation equations for the external forecasts, and the corresponding declarations and observations.

In what follows, this note first motivates the modeling of external forecasts as noisy measures of future realizations and derives its correlation structure across horizons in the context of a standard class of linear DSGE model. It then provides a step-by-step guide for how to incorporate them in DSGE models. Last, it undertakes a pseudo real-time forecasting exercise in the SW and DNGS models, using historical vintages from 1997q2 to 2018q4. The DSGE model forecasts are combined with forecasts from the *Survey of Professional Forecasters* (Federal Reserve Bank of Philadelphia, 2022). An appendix describes the data construction, DSGE model implementation, additional empirical results, and a Monte Carlo study of the proposed approach based on the 3-equation New Keynesian model in Galí (2008).

Notation: In what follows, lower case letters such as x denote scalars. Bold lower case letters denote vectors, e.g., $\mathbf{x} = [x_i]_i$, and bold upper case letters denote matrices, e.g., $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_J]$. $\mathbb{E}[\circ]$ denote the expectation operator, and subscripts, as in $\mathbb{E}_t[\circ]$, denote the conditional information, equal to $\mathbb{E}[\circ|\mathcal{F}_t]$, where $\mathcal{F}_t$ is taken to be the agents' information set. External forecasts are written as $\mathbb{E}_t^x[\circ] = \mathbb{E}[\circ|\mathcal{F}_t^x]$, where $\mathcal{F}_t^x \supseteq \mathcal{F}_t$. The econometrician's information set is often smaller than the agent's information set, although it is assumed to contain the same variables. To highlight this, it is denoted $\mathcal{F}_{t-} \subseteq \mathcal{F}_t$. A key calibrated parameter is the common scale parameter κ of the measurement errors of the external forecasts in the DSGE model.

[1] Mertens et al. (2022) use a statistical model to interpolate SPF forecasts along with their forecast uncertainty. The approach taken here provides an interpolation of the SPF point forecasts using a structural model, although this is not the focus.

2 Incorporating external forecasts in DSGE models

To motivate incorporating external forecasts as noisy future data, start from the canonical representation of the linearized DSGE model in ABCD:

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{B}\boldsymbol{\varepsilon}_t \tag{2.1a}$$

$$\mathbf{y}_t = \mathbf{C}\mathbf{x}_{t-1} + \mathbf{D}\boldsymbol{\varepsilon}_t \tag{2.1b}$$

$\mathbf{x}_t$ contains endogenous and exogenous state variables. $\mathbf{y}_t$ contains the observables of the model, as well as internal, static or purely forward-looking variables.

2.1 Limitations of the news-approach to modeling external forecasts

The news-approach in Monti (2010) and Del Negro and Schorfheide (2013) treats forecasts as measures of agents' expectations, possibly contaminated by measurement error. This approach is useful only when the underlying model departs from the ABCD benchmark that admits an invertible VAR(∞) representation.

Remark 1. *If the ABCD condition holds for observables $\tilde{\mathbf{y}}_t \subset \mathbf{y}_t$, observed expectations are redundant.*

To see this, let $\tilde{\mathbf{y}}_t = \tilde{\mathbf{C}}\mathbf{x}_{t-1} + \tilde{\mathbf{D}}\boldsymbol{\varepsilon}_t$. The model-implied forecasts are then $\mathbb{E}_t[\tilde{\mathbf{y}}_{t+h}] = \tilde{\mathbf{C}}\mathbf{A}^{h-1}\mathbb{E}_t[\mathbf{x}_t]$ for $h \geq 1$, where $\mathbb{E}_t[\mathbf{x}_t]$ is the best guess of the state of the economy at time t, given all observables $\{\tilde{\mathbf{y}}_{t-s}\}_s$. Under the ABCD condition, which requires $\tilde{\mathbf{D}}$ to be invertible (and thus square) and $\mathbf{A} - \mathbf{B}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{C}}$ to be stable, $\mathbf{x}_t = \sum_{s=0}^{\infty}(\mathbf{A} - \mathbf{B}\tilde{\mathbf{D}}^{-1}\tilde{\mathbf{C}})^s\mathbf{B}\tilde{\mathbf{D}}^{-1}\mathbf{y}_{t-s}$. It is thus revealed by the history of observables $\tilde{\mathbf{y}}_t$.

The news approach in Monti (2010) works in general because she uses external nowcasts. These forecasts of the current period are available ahead of the actual data and thus reveal information. However, the possible gain in information is limited to the information in the nowcast when the ABCD condition holds and the combined forecast necessarily follows the structure of (2.1). In contrast, modeling external forecasts as noisy measures of the actual realizations allows the combined forecast more flexibility, as not only can current states adjust, but the model generally can rationalize the external forecasts also using future shocks.

2.2 Properties of the noise-approach to modeling external forecasts

That leaves us with the second route of treating external forecasts as noisy measures of future data, rather than model expectations. If an external forecast is efficient, then $u^y_{t+h} \equiv y_{t+h} - \tilde{\mathbb{E}}^x_t[y_{t+h}]$ is, as of time t, unforecastable. Here, $\tilde{\mathbb{E}}^x$ denotes the external forecaster's expectations, as opposed to the model agents' expectations $\mathbb{E}$. Re-writing the definition of u^y_{t+h} leads to the observation equation:

$$\tilde{E}^x_t[y_{t+h}] = y_{t+h} + u^y_{t+h}. \tag{2.2}$$

Even though efficient forecast errors cannot be forecast with past information, forecast errors are generally autocorrelated across horizons h.

Remark 2. *Let the data-generating process be given by the canonical state-space model in* (2.1)*. Let the information available to forecasters at time t be denoted* $\mathcal{F}_t^x$*. If* $\mathbf{D}^{-1}$ *exists and* $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=\boldsymbol{0}$*, then the forecast errors* $\mathbf{u}_t^h \equiv \mathbf{y}_{t+h} - \mathbb{E}[\mathbf{y}_{t+h}|\mathcal{F}]$ *follow a VAR(1) structure across horizons h. A sufficient condition for* $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=\boldsymbol{0}$ *is that* $\mathbf{y}_t \subset \mathbf{x}_t$ *and that* $\mathbf{D}^{-1}$ *exists.*

Appendix A.1 provides a proof and a parametric example when the additional forecaster information consists of noisy information about current and future structural shocks.

In what follows, I use real-time data on SPF forecast errors to calibrate the autocorrelation of forecasts errors across horizons. Given data limitations, I consider only first order autocorrelation, which is a good approximation if the conditions of the remark hold approximately.

3 Augmenting DSGE models with SPF forecasts

As Del Negro and Schorfheide (2013) note, the likelihood function is approximately unaffected by the final observation. I thus treat as given the Bayesian estimate of the DSGE model parameters $\boldsymbol{\theta}$, which determine the coefficient matrices $\mathbf{A},\mathbf{B},\mathbf{C},\mathbf{D}$ in (2.1). Incorporating external forecasts in the DSGE model then requires the following steps (in addition to the standard model estimation):

1. Introduce extra observation equations.
2. Calibrate the parameters of the extra observation equations.
3. Apply the Kalman filter to the expanded information set.

Each of these steps is now detailed, followed by a discussion of the data and sample period. Last, I briefly discuss the two DSGE models that I consider here.

3.1 Extra observation equations

The date t external forecasts correspond to observations across different future time periods when mapped into `Dynare` observation equations. To include these observed external forecasts $FC_s(\mathbb{E}_{t-}^x[y_{t+h}])$ from time period t for period $s=t+h$ for model variable y_{t+h} in the data requires extra observation equations of the following form:

$$FC_s(\mathbb{E}_{t-}^x[y_{t+h}]) = \begin{cases} \text{missing} & s \neq t+h \\ y_{t+h} + u_{y,h,t} & s = t+h. \end{cases} \tag{3.1}$$

The Kalman filter easily accommodates the missing data. The within-period correlation across forecast horizons of $u_{y,h,t}$ in (2.2) now translates into autocorrelation across observed time periods when mapped to the DSGE model as a noisy measure of the future realization as in (3.1).

An additional adjustment is needed to account for the per-capita structure of the DSGE models considered here. In contrast, the external forecasts are for real variables that are not in per-capita terms. I thus include the identity that real GDP growth is the sum of real GDP per capita and population growth as an additional observation equation. It is modeled as an exogenous AR(1) process.

The extra equations are listed in Appendix B.

3.2 Calibrating the extra observation equations

I compute forecast errors using the latest data vintage to avoid using different vintages for forecasts of different horizons. Using these forecast errors, I estimate the persistence of forecast errors across horizons as well as possible cross-correlations of forecast errors by ordinary least squares. The estimated root-mean squared error (RMSE) for each forecast and forecast horizon is the basis for the calibration of the measurement equations corresponding to the external forecasts.

$$u_{y,0,t} \equiv y_t - \mathbb{E}^x_{t-}[y_{t+0}] = \sigma_{y0}\epsilon_{y,0,t} \tag{3.2a}$$

$$u_{y,1,t} \equiv y_{t+1} - \mathbb{E}^x_{t-}[y_{t+1}] = \rho_{y0}u_{y,0,t} + \sigma_{y1}\epsilon_{y,1,t} \tag{3.2b}$$

$$u_{y,4,t} \equiv y_{t+4} - \mathbb{E}^x_{t-}[y_{t+4}] = \rho_{y,4}u_{y1,t} + \sigma_{y,4}\epsilon_{y,4,t} \tag{3.2c}$$

$$u_{y,cy(t)+1,t} \equiv y_{cy(t)+1} - \mathbb{E}^x_{t-}[y_{cy(t)+1}] = \rho_{y,cy(t)+1}u_{y,4,t} + \sigma_{y,cy(t)}\epsilon_{y,cy(t)+1,t}, \tag{3.2d}$$

$$u_{y,cy(t)+s,t} \equiv y_{cy(t)+s} - \mathbb{E}^x_{t-}[y_{cy(t)+s}] = \rho_{y,cy(t)+s}u_{y,cy(t)+s-1,t} + \sigma_{y,cy(t)}\epsilon_{y,cy(t)+s,t}, \qquad s \in \{2,3\}, \tag{3.2e}$$

where y is a stand-in for growth in the GDP deflator or in real GDP. $cy(t)$ denotes the calendar-year at time t, so that $cy(t) + s$ is s calendar years in the future. The observation equations for the TBill rate forecast are similar, but have additional covariance terms with the contemporaneous GDP deflator or real GDP forecast error.

To parameterize the extent of the forecast tilt towards the SPF forecasts, I then scale all the measurement error standard deviations $\sigma_{y,h}$ down (or up) by a common factor κ. As $\kappa \nearrow \infty$, the observed forecasts are pure measurement error, while the model interprets them as equal to the realizations as $\kappa \searrow 0$.

3.3 DSGE models

The first DSGE model considered is SW: A DSGE model featuring Calvo-sticky prices and nominal wages with indexation, and real frictions estimated using Bayesian methods on seven observables, including (GDP-deflator) inflation, real GDP growth, and the Federal Funds Rate (FFR). The second model is DNGS. This model extends the first model by including financial frictions in the form of a financial accelerator and by using long-run SPF inflation expectations to flexibly model low-frequency inflation movements. In addition, it uses the BAA-10-year Treasury bond spread in the estimation. Appendix D lists the estimated parameters along with the distribution of the posterior modes resulting from the recursive estimation.

3.4 Data and sample period

The baseline observables and variable definitions in the DSGE model closely follow SW, with three small modifications. First, I treat durable consumption as part of investment, as in Drautzburg and Uhlig (2015). Second, I substitute the Wu and Xia (2016) shadow rate for FFR, when the FFR runs below 0.1. Third, I exponentially smooth (log) population levels to avoid artificial jumps in population levels. The final macro dataset is based on data from U.S. Bureau of Economic Analysis (nd); U.S. Bureau of Labor Statistics (nd); Board of Governors of the Federal Reserve System (nd); Moody's (nd); Federal Reserve Bank of Atlanta (nd).

I use median SPF forecasts (Federal Reserve Bank of Philadelphia, 2022) for the TBill Rate, output growth, and inflation. After adjusting for the mean difference between the TBill rate and the FFR by subtracting a 4-quarter moving average of their difference, its forecast serves as a counterpart to the FFR. I do not use TBill forecasts when the one-quarter lagged FFR is less than 0.1pp, since during that time I use the shadow rate as our measure of the short rate. For all three variables I use the nowcast, the one-quarter ahead forecast, the forecast over the next quarter, and the forecast for the next calendar year – in either levels (TBill rate) or growth rates (real GDP and GDP deflator). For real GDP growth, I also use the growth of the annual averages two and three calendar years out once they become available.

With the exception of the shadow rate data and SPF forecasts, which are downloaded from the Atlanta Fed and Philadelphia Fed, the `Stata` code accompanying this note downloads all vintage data from FRED.

The estimation sample runs from 1947q1 with the likelihood evaluation beginning in 1948q1. The estimation sample is expanding along with the pseudo real-time forecast date, which begins in 1997q2.[2] To exclude data from the COVID-19 pandemic, the baseline exercise ends in 2018q4. However, results are robust to extending the sample to 2022q3.

4 Results

4.1 Forecast performance

The hair plots in Figure 1 contrast the forecasts with the realized data with forecasts dating from 1997q2 through 2018q4 for quarterly real GDP growth on the left and GDP-deflator inflation on the right. Both series are in percent [not annualized]. The figures contain four sets of lines. First, a single green line displaying the 2018q4 data vintage. Second, a set of gray lines displaying earlier

[2]The real-time evaluation sample balances the real-time availability of different series. Vintage NIPA data is available early. Hourly compensation data is available beginning in 1997q2, which also marks the beginning of the estimation sample. Vintage Population and labor input data is available only beginning in 2011q1, but is subject to only minor revisions.

vintages that generally tracks the green line, but that can diverge from the latest vintage noticeably during some periods, such as in 2007 for output growth or in the early 2000s for inflation. The third set of lines shows the pure DSGE model forecasts as black, dotted lines. The fourth set of lines shows the combined SPF-DSGE model forecast when the RMSE of the SPF is scaled by $\kappa = 0.01$, giving a high weight to the SPF. These forecasts are shown as red, dashed lines. These forecasts are point estimates at the posterior mode. The top two panels in Figure 1 are for the SW model, the bottom panels for DNGS.

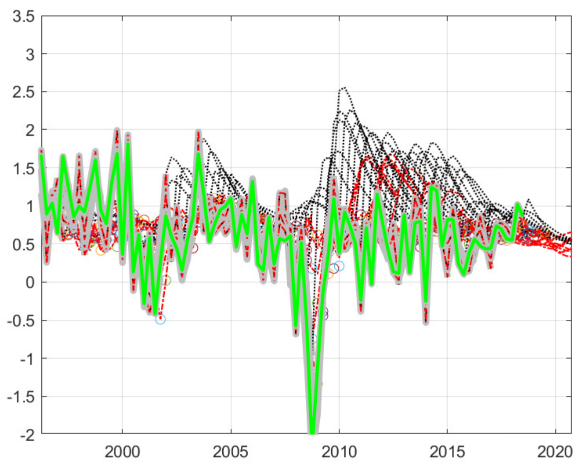
Figure 1(a) clearly shows that there can be systematic departures of the DSGE model forecast from the realized data. For example, the black, dashed lines show that the pure SW model forecast overstated output growth in the aftermath of the Global Financial Crisis (GFC) for several years. In contrast, the red lines show that the combined forecast did so to a lesser degree, improving forecast performance. A similar, but less pronounced, pattern holds for SW inflation forecasts in Figure 1(b).

The picture for the DNGS output forecasts in Figure 1(c) is qualitatively similar. Here, the model forecasts do relatively poorly in the first few years of the sample, understating growth, and modestly overstate growth following the GFC. The combined forecast is closer to the truth. Interestingly, the inflation forecasts are much better even for the pure DSGE model forecast in Figure 1(d).

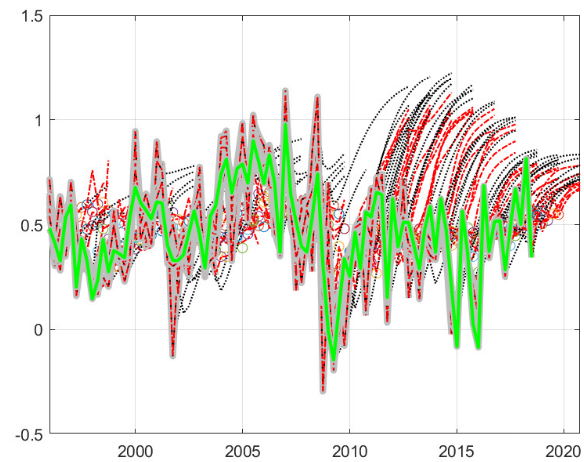
Moving beyond the extreme weights on DSGE model forecast, Figure 2 shows the mean absolute error (MAE) from the nowcast up to 12 quarters out for both models and both variables for a range of scales of the SPF forecast error variances, ranging from $\kappa = 0.01$ to $\kappa = 100$ and the baseline DSGE model forecast (implicitly an infinite scale of the noise). Two different patterns emerge: For the SW inflation forecast and the DNGS output forecast, the combined forecasts are strictly lower at all forecast horizons, and the error increases gradually as κ increases. For example, in Figure 2(c), the nowcast has a mean absolute error just below 0.4pp with error $\kappa \leq 1.0$, that rises to 0.6 with $\kappa = 5.0$ and increases to just above 0.8 for $\kappa = 100$ and the pure DSGE model forecast. Forecast errors stay flat with SW and decline with DNGS in magnitude for longer horizon forecasts, but the same pattern is still visible. A similar pattern also holds for SW inflation forecast errors, which deteriorate at longer horizons. For the DNGS SPF-DSGE forecast, at intermediate frequencies models with a larger κ perform better than those treating the SPF forecast as more precise: At horizons up to two quarters out, giving more weight to the SPF generally lowers the MAE. This flips at horizons of four quarters or longer, and at horizons of eight to 12 quarters, intermediate values of $\kappa = 2$ or $\kappa = 5$ perform best. While less pronounced, the SW SPF-DSGE MAE associated with a κ of 5 also performs better than higher weights on the SPF at horizons of eight to 12 quarters. Figure E.1 in the Appendix shows quarter-by-quarter MAEs and Figure E.2 shows that similar results hold when including the COVID-period.

Table 1 shows that the forecast performance is also statistically better for the forecast combinations, except for inflation forecasts, where the pure DNGS forecast is competitive. Specifically, the table shows the best weight on the external forecast along with a 90% confidence interval of the

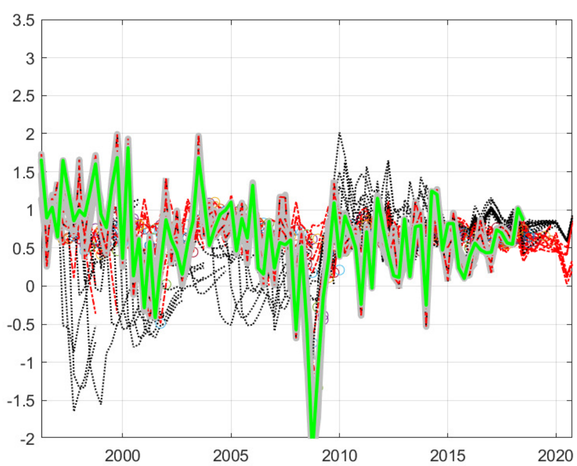
(a) SW Quarterly output growth



(b) SW Quarterly inflation



(c) DNGS Quarterly output growth



(d) DNGS Quarterly inflation

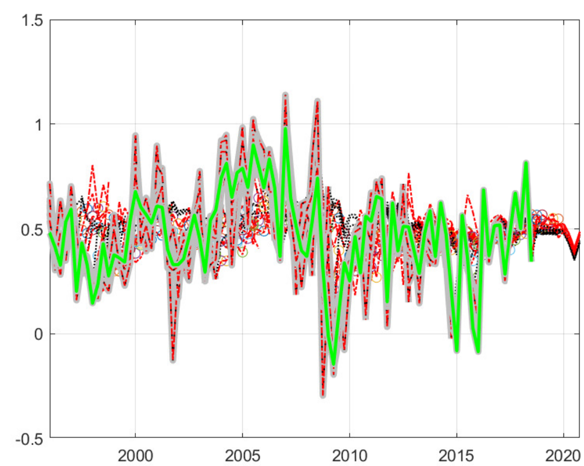


Figure 1: Hair plots: 1997q2 through 2020q4

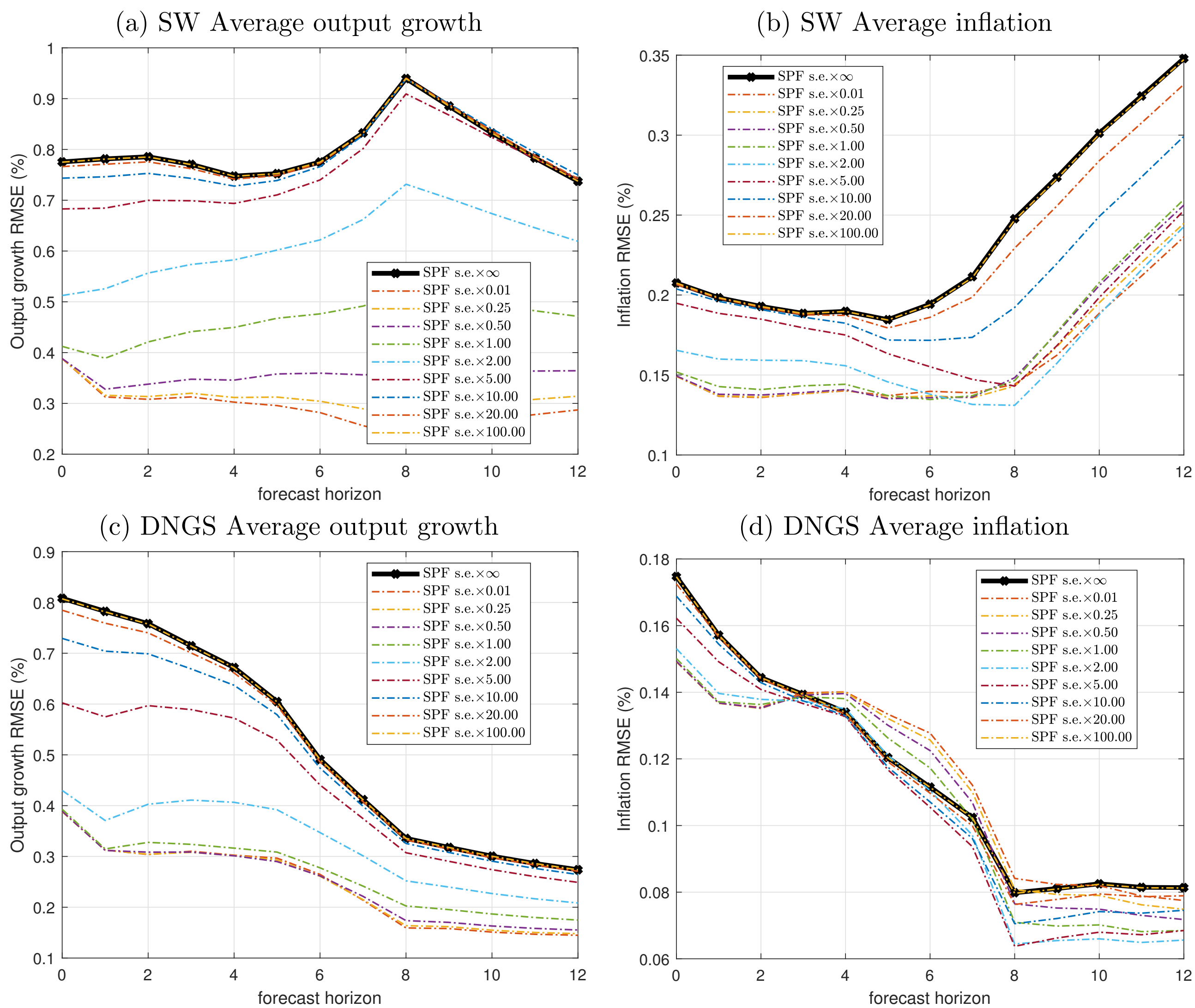


Figure 2: Mean absolute error by variable and horizon: 1997q2 through 2020q4

optimal set of weights based on inverting the Diebold and Mariano (1995) test statistic relative to the best model. The table has four parts, corresponding again to the two models and two variables. Each part shows the optimal weight, the mean absolute error associated with that weight, and the range of weights whose absolute error is not significantly different from the optimum. Each column corresponds to a different horizon, with the average across horizons in the last column.

(a) Output growth: Smets and Wouters (2007)

Horizon	Current	1 qtr	2 qtr	3 qtr	4 qtr	8 qtr	12 qtr	Avg
Best weight	0.25	0.01	0.01	0.01	0.01	0.01	0.01	0.01
Value	0.39	0.31	0.31	0.31	0.30	0.24	0.29	0.29
CI	[0.01, 0.5]	[0.01, 0.25]	[0.01, 0.25]	[0.01, 0.25]	[0.01, 0.01]	[0.01, 0.01]	[0.01, 0.01]	[0.01, 0.01]

(b) Output growth: Del Negro et al. (2015)

Horizon	Current	1 qtr	2 qtr	3 qtr	4 qtr	8 qtr	12 qtr	Avg
Best weight	0.01	0.50	0.01	0.25	0.50	0.01	0.01	0.01
Value	0.39	0.31	0.30	0.31	0.30	0.16	0.14	0.24
CI	[0.01, 1]	[0.01, 1]	[0.01, 0.5]	[0.01, 1]	[0.01, 1]	[0.01, 0.01]	[0.01, 0.01]	[0.01, 1]

(c) Inflation: Smets and Wouters (2007)

Horizon	Current	1 qtr	2 qtr	3 qtr	4 qtr	8 qtr	12 qtr	Avg
Best weight	0.25	0.01	0.01	0.25	0.01	2.00	0.01	0.01
Value	0.15	0.14	0.14	0.14	0.14	0.13	0.24	0.16
CI	[0.01, 1]	[0.01, 0.5]	[0.01, 1]	[0.01, 1]	[0.01, 1]	[0.01, 2]	[0.01, 5]	[0.01, 0.25]

(d) Inflation: Del Negro et al. (2015)

Horizon	Current	1 qtr	2 qtr	3 qtr	4 qtr	8 qtr	12 qtr	Avg
Best weight	0.01	0.01	0.25	5.00	5.00	5.00	2.00	2.00
Value	0.15	0.14	0.14	0.14	0.13	0.06	0.07	0.10
CI	[0.01, 2]	[0.01, 2]	[0.01, Inf]	[0.01, Inf]	[0.01, Inf]	[1, 5]	[1, 5]	[0.01, Inf]

Table 1: Optimal forecast weight

For output growth in Table 1(a) and (b), a very low degree κ of measurement error always yields the best forecast performance, and the highest κ is not significantly worse than the best scale of $\kappa = 1.0$ for DNGS and of $\kappa = 0.5$ for SW. The same holds for SW inflation forecasts, where the highest κ that is not statistically worse is again 1.0. However, for inflation forecasts, the DNGS model does well. At horizons beginning at three quarters out, κ of 2.0 or 5.0 is the best, and the infinite scale often is no worse in statistical terms. This corroborates the finding in DNGS that financial frictions improve the forecast performance.

To interpret this finding, note that $\kappa \geq 2$ implies a low weight on the SPF forecast. For example, Figure E.3 in the Appendix shows the forecast revisions in SW when only an output growth nowcast and one-quarter-ahead forecast is added to the model. For $\kappa = 2$, the forecast is revised by less than half of the external forecast revision; for $\kappa = 5$, the revision is just one tenth.[3]

4.2 External forecasts and shock decompositions

While many methods exist for forecast averaging, the advantage of incorporating SPF forecasts in structural models as proposed here is that researchers can still structurally decompose forecasts or

[3]Figure E.4 shows the analogous experiment for the inflation forecast, and the numbers are very similar.

histories. Here, I show how incorporating external forecasts changes the inferred role of shocks in explaining historical data and shaping forecasts.

Figure 3 shows the historical and forecast decomposition of output growth from 2007q1 through 2009q4 from the SW perspective using 2008q1 information. The top panel shows the pure DSGE results, the bottom panel incorporates SPF forecasts using the measurement error scale of $\kappa = 0.01$. The history and forecast are shown as the black line, which is decomposed in colored bars corresponding to the seven structural shocks and the exogenous population growth and initial conditions.

The shock decomposition shows: (1) Since the ABCD condition holds, the past history of shocks is left unrevised by the new data. Only the (unobserved) current quarter and future quarters are affected – which is why the news approach of Monti (2010) can also work via the channel of revealing the current state. (2) Current and future shocks can change a lot with the forecasts that I condition on. The SPF-information leads to a more persistent drag from IST shocks, a large role for risk premium shocks, and overall a more pessimistic forecast.

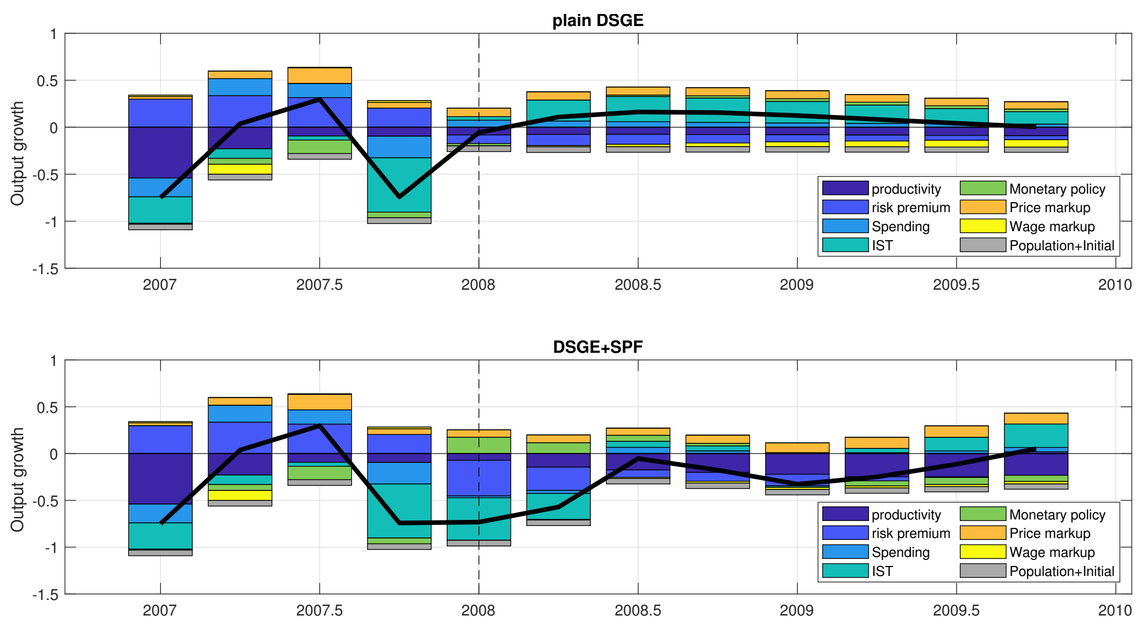


Figure 3: Historical and Forecast Shock Decomposition of Output Growth from 2007q1 through 2009q4 from the perspective of 2008q1 in the SW model.

5 Conclusion

Treating external forecasts as noisy realizations of the truth is a computationally simple and flexible approach to improve the accuracy of forecast without giving up on their structural interpretability. Model-based forecasts can receive a meaningful weight with quantitatively successful models. Given that the results here are based on any model with a linear state-space representation, the results could also be applied to structural VARs, including those that are factor-augmented or estimated on mixed-frequency data.

References

Stéphane Adjemian, Houtan Bastani, Michel Juillard, Fréderic Karamé, Junior Maih, Ferhat Mihoubi, George Perendia, Johannes Pfeifer, Marco Ratto, and Sébastien Villemot. 2011. *Dynare: Reference Manual Version 4*. Dynare Working Papers 1. CEPREMAP. https://ideas.repec.org/p/cpm/dynare/001.html

Board of Governors of the Federal Reserve System. n.d.. Industrial Production and Financial Market Data. FRED, Federal Reserve Bank of St. Louis [distributor]. URLs listed in the Data Appendix. (Accessed: 8/25/2022).

Todd E. Clark and Michael W. McCracken. 2010. Averaging Forecasts from VARs with Uncertain Instabilities. *Journal of Applied Econometrics* 25, 1 (2010), 5–29. https://doi.org/10.1002/jae.1127 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.1127

Marco Del Negro, Marc P. Giannoni, and Frank Schorfheide. 2015. Inflation in the Great Recession and New Keynesian Models. *American Economic Journal: Macroeconomics* 7, 1 (January 2015), 168–96. https://doi.org/10.1257/mac.20140097

Marco Del Negro and Frank Schorfheide. 2013. Chapter 2 - DSGE Model-Based Forecasting. In *Handbook of Economic Forecasting*, Graham Elliott and Allan Timmermann (Eds.). Handbook of Economic Forecasting, Vol. 2. Elsevier, 57–140. https://doi.org/10.1016/B978-0-444-53683-9.00002-5

F.X. Diebold and R. Mariano. 1995. Comparing Predictive Accuracy. *Journal of Business and Economic Statistics* 13 (1995), 253–265. https://doi.org/10.1080/07350015.1995.10524599

Thorsten Drautzburg and Harald Uhlig. 2015. Fiscal Stimulus and Distortionary Taxation. *Review of Economic Dynamics* 18, 4 (2015), 894–920. https://doi.org/10.1016/j.red.2015.09.003

Federal Reserve Bank of Atlanta. n.d.. Wu-Xia Shadow Federal Funds Rate. https://www.atlantafed.org/-/media/documents/datafiles/cqer/research/wu-xia-shadow-federal-funds-rate/WuXiaShadowRate.xlsx. (Accessed 8/18/2022).

Federal Reserve Bank of Philadelphia. 2022. Survey of Professional Forecasters. URLs provided in the Appendix. (Accessed 8/25/2022).

Jesús Fernández-Villaverde, Juan F. Rubio-Ramírez, Thomas J. Sargent, and Mark W. Watson. 2007. ABCs (and Ds) of Understanding VARs. *American Economic Review* 97, 3 (June 2007), 1021–1026. https://doi.org/10.1257/aer.97.3.1021

Jordi Galí. 2008. *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework*. Princeton University Press, Chapter 3.

John Geweke and Gianni Amisano. 2012. Prediction with Misspecified Models. *American Economic Review* 102, 3 (May 2012), 482–86. https://doi.org/10.1257/aer.102.3.482

Elmar Mertens, Todd E. Clark, and Gergely Ganics. 2022. Constructing the Term Structure of Uncertainty from the Ragged Edge of SPF Forecasts. (2022). unpublished, Bundesbank.

Francesca Monti. 2010. Combining Judgment and Models. *Journal of Money, Credit and Banking* 42, 8 (December 2010), 1641–1662. https://doi.org/10.1111/j.1538-4616.2010.00357.x

Moody's. n.d.. Moody's Seasoned Baa Corporate Bond Yield [BAA]. https://fred.stlouisfed.org/series/BAA. (Accessed 8/25/2022).

Frank Smets, Anders Warne, and Rafael Wouters. 2014. Professional Forecasters and Real-Time Forecasting with a DSGE Model. *International Journal of Forecasting* 30, 4 (2014), 981–995. https://doi.org/10.1016/j.ijforecast.2014.03.018

Frank Smets and Rafael Wouters. 2007. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review* 97, 3 (June 2007), 586–606. https://doi.org/10.1257/aer.97.3.586

James H. Stock and Mark W. Watson. 2004. Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting* 23, 6 (2004), 405–430. https://doi.org/10.1002/for.928 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.928

U.S. Bureau of Economic Analysis. n.d.. National Income and Product Account data. FRED, Federal Reserve Bank of St. Louis [distributor]. URLs listed in the Data Appendix. (Accessed: 8/25/2022).

U.S. Bureau of Labor Statistics. n.d.. Labor Market Data. FRED, Federal Reserve Bank of St. Louis [distributor]. URLs listed in the Data Appendix. (Accessed: 8/25/2022).

Jonathan H. Wright. 2009. Forecasting US Inflation by Bayesian Model Averaging. *Journal of Forecasting* 28, 2 (2009), 131–144. https://doi.org/10.1002/for.1088 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/for.1088

Jing Cynthia Wu and Fan Dora Xia. 2016. Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking* 48, 2-3 (March 2016), 253–291. https://doi.org/10.1111/jmcb.12300

A Properties of external forecast errors

A.1 VAR(1)-structure of forecast errors across horizons

Lemma 1. *Let the data-generating process be given by the canonical state-space model in* (2.1)*. Let the information available to forecasters at time* t *be denoted* $\mathcal{F}_t^x$*. If* $\mathbf{D}^{-1}$ *exists and* $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=\mathbf{0}$*, then the forecast errors* $\mathbf{u}_t^h \equiv \mathbf{y}_{t+h} - \mathbb{E}[\mathbf{y}_{t+h}|\mathcal{F}]$ *follow a VAR(1) structure across horizons* h*. A sufficient condition for* $\mathbf{A}-\mathbf{B}\mathbf{D}^{-1}\mathbf{C}=\mathbf{0}$ *is that* $\mathbf{y}_t \subset \mathbf{x}_t$ *and that* $\mathbf{D}^{-1}$ *exists.*

Proof. As a preliminary step, substitute recursively the VAR(1) law of motion for the state $\mathbf{x}_t$ in (2.1):

$$\mathbf{x}_{t+h} = \mathbf{A}^{h+1}\mathbf{x}_{t-1} + \sum_{s=0}^{h} \mathbf{A}^s \mathbf{B} \boldsymbol{\varepsilon}_{t+h-s}$$

$$\mathbf{y}_{t+h} = \mathbf{D}\boldsymbol{\varepsilon}_{t+h} + \mathbf{C}\mathbf{A}^h \mathbf{x}_{t-1} + \sum_{s=0}^{h-1} \mathbf{C}\mathbf{A}^s \mathbf{B} \boldsymbol{\varepsilon}_{t+h-1-s}$$

Apply the expectation operator and then use the VAR(1) law of motion for $\mathbf{x}_t$ again to write:

$$\begin{aligned}
\mathbb{E}[\mathbf{y}_{t+h}|\mathcal{F}_t^x] &= \mathbf{D}\mathbb{E}[\epsilon_{t+h}|\mathcal{F}_t^x] + \mathbf{C}\mathbb{E}[\mathbf{x}_{t+h-1}|\mathcal{F}_t^x] \\
&= \mathbf{D}\mathbb{E}[\boldsymbol{\varepsilon}_{t+h} + \mathbf{C}\mathbf{B}\mathbb{E}[\boldsymbol{\varepsilon}_{t+h-1}|\mathcal{F}_t^x] + \mathbf{C}\mathbf{A}\mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x]
\end{aligned}$$

Using that $\mathbb{E}[\mathbf{y}_{t+h-1}|\mathcal{F}_t^x] = \mathbf{D}\mathbb{E}[\epsilon_{t+h-1}|\mathcal{F}_t^x] + \mathbf{C}\mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x]$ from the first equality above for $h-1$ and using the second equality for h , I can compute the forecast errors:

$$\begin{aligned}
\mathbf{u}_t^{h-1} &\equiv \mathbf{y}_{t+h-1} - \mathbb{E}[\mathbf{y}_{t+h-1}|\mathcal{F}_t^x] = \mathbf{D}(\boldsymbol{\varepsilon}_{t+h-1} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h-1}|\mathcal{F}_t^x]) + \mathbf{C}(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x]) \\
\mathbf{u}_t^{h} &\equiv \mathbf{y}_{t+h} - \mathbb{E}[\mathbf{y}_{t+h}|\mathcal{F}_t^x] = \mathbf{D}(\boldsymbol{\varepsilon}_{t+h} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h}|\mathcal{F}_t^x]) + \mathbf{C}\mathbf{B}(\boldsymbol{\varepsilon}_{t+h-1} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h-1}|\mathcal{F}_t^x]) + \mathbf{C}\mathbf{A}(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x])
\end{aligned}$$

Since $\mathbf{D}$ is invertible:

$$\boldsymbol{\varepsilon}_{t+h-1} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h-1}|\mathcal{F}_t^x] = \mathbf{D}^{-1}(\mathbf{u}_t^{h-1} - \mathbf{C}(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x]))$$

Plugging in the equation for $\mathbf{u}_t^h$ yields:

$$\begin{aligned}
\mathbf{u}_t^h &= \mathbf{D}(\boldsymbol{\varepsilon}_{t+h} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h}|\mathcal{F}_t^x]) + \mathbf{C}\mathbf{B}\mathbf{D}^{-1}(\mathbf{u}_t^{h-1} - \mathbf{C}(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x])) + \mathbf{C}\mathbf{A}(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x]) \\
&= \mathbf{D}(\boldsymbol{\varepsilon}_{t+h} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h}|\mathcal{F}_t^x]) + \mathbf{C}\mathbf{B}\mathbf{D}^{-1}\mathbf{u}_t^{h-1} + \mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})(\mathbf{x}_{t+h-2} - \mathbb{E}[\mathbf{x}_{t+h-2}|\mathcal{F}_t^x])
\end{aligned}$$

If $\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C} = \mathbf{0}$, then the result holds immediately. A sufficient condition guaranteeing this equality is if $\mathbf{C} = \mathbf{S}\mathbf{A}, \mathbf{D} = \mathbf{S}\mathbf{B}$ and $\mathbf{D}$ is invertible. We then have that $\mathbf{B}\mathbf{D}^{-1}\mathbf{C} = \mathbf{B}\mathbf{B}^{-1}\mathbf{S}^{-1}\mathbf{S}\mathbf{A} = \mathbf{A}$ so that $(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C}) = \mathbf{0}$. Consequently, the forecast errors have a VAR(1) structure in the forecast horizon h.

$$\mathbf{u}_t^h = \mathbf{D}(\boldsymbol{\varepsilon}_{t+h} - \mathbb{E}[\boldsymbol{\varepsilon}_{t+h}|\mathcal{F}_t^x]) + \mathbf{CBD}^{-1}\mathbf{u}_t^{h-1} \qquad h \geq 1,$$
$$\mathbf{u}_t^0 = \mathbf{D}(\boldsymbol{\varepsilon}_t - \mathbb{E}[\boldsymbol{\varepsilon}_t|\mathcal{F}_t^x]).$$

□

Assume that the external forecasters have access at time t to signals about shocks to the system (2.1) at horizon $t+h$ given by the vector $\boldsymbol{\nu}_t^h$. Let the signal structure underlying $\mathcal{F}_t^x$ be given by:

$$\boldsymbol{\nu}_t^h = \mathbf{E}\boldsymbol{\varepsilon}_{t+h} + \mathbf{f}_t^h, \qquad \mathbb{E}_{t-1}[\mathbf{f}_t^h] = 0,$$

where $\mathbf{E}$ is a loading matrix (not to be confused with the expectation operator $\mathbb{E}$). Since $\boldsymbol{\varepsilon}_t$ is *iid*, external signals are uncorrelated ($\text{Cov}[\boldsymbol{\nu}_t^0, \boldsymbol{\nu}_t^1] = \mathbf{0}$), if $\mathbf{f}_t^0, \mathbf{f}_t^1$ are uncorrelated.

To see the value added of these signals, compare the forecasts knowing the initial state $\mathbf{x}_{t-1}$ only and a signal about the current shocks $\boldsymbol{\nu}_t^0$:

$$\mathbb{E}[\mathbf{y}_t|\mathbf{x}_{t-1}] = \mathbf{C}\mathbf{x}_{t-1} \tag{A.1}$$
$$\mathbb{E}[\mathbf{y}_t|\mathbf{x}_{t-1}, \boldsymbol{\nu}_t^0] = \mathbf{C}\mathbf{x}_{t-1} + \mathbf{D}\boldsymbol{\omega}_0\boldsymbol{\nu}_t^0 \tag{A.2}$$
$$\text{Cov}[\boldsymbol{\varepsilon}_{t+h}, \boldsymbol{\nu}_t^h] = \text{Var}[\boldsymbol{\varepsilon}_{t+h}]\mathbf{E}' = \boldsymbol{\omega}_h(\mathbf{E}\,\text{Var}[\boldsymbol{\varepsilon}_{t+h}]\mathbf{E}' + \text{Var}[\mathbf{f}_t^h]) \tag{A.3}$$
$$\Rightarrow \boldsymbol{\omega}_h = \text{Var}[\boldsymbol{\varepsilon}_{t+h}]\mathbf{E}'(\mathbf{E}\,\text{Var}[\boldsymbol{\varepsilon}_{t+h}]\mathbf{E}' + \text{Var}[\mathbf{f}_t^h])^{-1} \tag{A.4}$$

As the measurement error variance $\text{Var}[\mathbf{f}_t^h]$ disappears and if $\mathbf{E}$ is invertible, $\boldsymbol{\omega}_0 \to \mathbf{I}$: The shocks are exactly revealed by the signals.[4]

Under this information structure, the forecast errors satisfy:

$$\mathbf{u}_t^0 = \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_0)\boldsymbol{\varepsilon}_t - \mathbf{D}\omega_0\mathbf{f}_t^0 \tag{A.5}$$
$$\begin{aligned}\mathbf{u}_t^1 &= \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_1)\boldsymbol{\varepsilon}_{t+1} - \mathbf{D}\omega_1\mathbf{f}_t^1 + \mathbf{CB}(\mathbf{I} - \boldsymbol{\omega}_0)\boldsymbol{\varepsilon}_t - \mathbf{CB}\omega_0\mathbf{f}_t^0 \\ &= \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_1)\boldsymbol{\varepsilon}_{t+1} - \mathbf{D}\omega_1\mathbf{f}_t^1 + \mathbf{CBD}^{-1}\mathbf{u}_t^0 \qquad \text{if } \mathbf{D}^{-1} \text{ exists}\end{aligned} \tag{A.6}$$
$$\begin{aligned}\mathbf{u}_t^2 &= \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_2)\boldsymbol{\varepsilon}_{t+2} - \mathbf{D}\omega_2\mathbf{f}_t^2 + \mathbf{CB}(\mathbf{I} - \boldsymbol{\omega}_1)\boldsymbol{\varepsilon}_{t+1} - \mathbf{CB}\omega_1\mathbf{f}_t^1 + \mathbf{CAB}(\mathbf{I} - \boldsymbol{\omega}_0)\boldsymbol{\varepsilon}_t - \mathbf{CAB}\omega_0\mathbf{f}_t^0 \\ &= \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_2)\boldsymbol{\varepsilon}_{t+2} - \mathbf{D}\omega_2\mathbf{f}_t^2 + \mathbf{CBD}^{-1}(\mathbf{u}_t^1 - \mathbf{CBD}^{-1}\mathbf{u}_t^0) + \mathbf{CABD}^{-1}\mathbf{u}_t^0 \\ &= \mathbf{D}(\mathbf{I} - \boldsymbol{\omega}_2)\boldsymbol{\varepsilon}_{t+2} - \mathbf{D}\omega_2\mathbf{f}_t^2 + \mathbf{CBD}^{-1}\mathbf{u}_t^1 + \mathbf{C}(\mathbf{A} - \mathbf{BD}^{-1}\mathbf{C})\mathbf{BD}^{-1}\mathbf{u}_t^0\end{aligned} \tag{A.7}$$

A.2 Kalman gain for noisy forecasts

Let $y_{i,t} = \mathbf{c}_i'\mathbf{x}_t + \mathbf{d}_i'\boldsymbol{\varepsilon}_t$ be a model observable. Assume $\text{Var}_t[\mathbf{x}_t] = \mathbf{0}$ – which is implied by the ABCD condition.

[4] In that case, the forecast error below converges to $\mathbf{f}_t^h$, which, in turn, collapses to $\mathbf{0}$.

What do we learn from $y_{i,t}$?

$$\mathbb{E}_t[\mathbf{x}_t|\mathbf{y}^{t-1}, y_{i,t}] = \boldsymbol{\beta}(y_{i,t} - \mathbf{c}_i\mathbf{x}_{t-1}), \qquad \boldsymbol{\beta} = \mathrm{Cov}_{t-1}[\mathbf{x}_t, y_{i,t}]\,\mathrm{Var}_{t-1}[y_{i,t}]^{-1} = \mathbf{D}\mathbf{c}_i(\mathbf{c}_i\mathbf{c}_i')^{-1}$$

Updates to future $\mathbf{x}_{t+h}, h \geq 1$ then follow by tracing out the dynamics given by (2.1).

Now, the SPF forecast of variable $y_{i,t}$ is the truth plus white noise $u_{i,t}$. Multiplying and dividing by $\mathrm{Var}_{t-1}[y_{i,t}]$ allows us to write the Kalman gain as a product of the Kalman gain without measurement error and a scaled-down variable:

$$\mathbb{E}_t[\boldsymbol{\varepsilon}_t|\mathbf{y}^{t-1}, y_{i,t} + u_{i,t}] = \boldsymbol{\beta}\gamma(y_{i,t} - \mathbf{c}_i\mathbf{x}_{t-1} + u_{i,t}), \qquad \boldsymbol{\beta} = \mathbf{D}\mathbf{c}_i(\mathbf{c}_i\mathbf{c}_i')^{-1}, \quad \gamma = \frac{\mathrm{Var}_{t-1}[y_{i,t}]}{\mathrm{Var}_{t-1}[y_{i,t}] + \mathrm{Var}_{t-1}[u_{i,t}]}.$$

Thus, in this simple case, the Kalman gain of the SPF forecast is a scaled-down version of the Kalman gain of observing the true model.

Note that in practice, when multiple forecasts for different horizons are provided, the gain is more complicated, because the forecast at longer horizons also leads to updates at shorter horizons as $\mathbf{x}_t$ is not perfectly revealed by noisy observations.

B Extra observation equations

To allow for noisy external forecasts across multiple horizons with correlated forecast errors requires two sets of equations: (1) the definition of the observed forecasts as noisy measures of the model observables, and (2) the definition of the noise as the sum of a correlated component and *iid* noise. A third set of two equations allows moving from real, per-capita forecasts to simply real forecasts.

(1) Definition of observed forecasts

$$\text{Nowcast}$$

$$FC_t(\mathbb{E}^x_{t-}[v_t]) = v^{obs}_t + u^{v,0}_t \tag{B.1a}$$

$$\text{1-qtr ahead}$$

$$FC_t(\mathbb{E}^x_{t-1-}[v_t]) = v^{obs}_t + u^{v,1}_t \tag{B.1b}$$

$$\text{4-qtr ahead (Q4/Q4)}$$

$$FC_t(\mathbb{E}^x_{t-4-}[v_t]) = \sum_{s=0}^{3} v^{obs}_{t-s} + u^{v,4}_t \tag{B.1c}$$

$$\text{Xor} \tag{B.1d}$$

$$\text{4-qtr ahead (Q4 lvl)}$$

$$FC_t(\mathbb{E}^x_{t-4-}[v_t]) = v^{obs}_t + u^{v,4}_t \tag{B.1e}$$

$$\text{Xor} \tag{B.1f}$$

$$\text{4-qtr ahead (annual avg)}$$

$$FC_t(\mathbb{E}^x_{t-4-}[v_t]) = \frac{1}{4}\sum_{s=0}^{3} v^{obs}_{t-s} + u^{v,4}_t \tag{B.1g}$$

1-y ahead (gr of annual avg)

$$FC_t(\mathbb{E}^x_{t-h-}[v_t]) = \frac{1}{16}\left(v_{t-6} + 2v_{t-5} + 3v_{t-4} + 4v_{t-3} + 3v_{t-2} + 2v_{t-1} + v_t\right) + u^{v,1y}_t \qquad h \geq 4 \tag{B.1h}$$

2-y ahead (gr of annual avg)

$$FC_t(\mathbb{E}^x_{t-h-}[v_t]) = \frac{1}{16}\left(v_{t-10} + 2v_{t-9} + 3v_{t-8} + 4v_{t-7} + 3v_{t-6} + 2v_{t-5} + v_{t-4}\right) + u^{v,1y}_t \qquad h \geq 8 \tag{B.1i}$$

3-y ahead (gr of annual avg)

$$FC_t(\mathbb{E}^x_{t-h-}[v_t]) = \frac{1}{16}\left(v_{t-14} + 2v_{t-12} + 3v_{t-12} + 4v_{t-11} + 3v_{t-10} + 2v_{t-9} + v_{t-8}\right) + u^{v,1y}_t \qquad h \geq 12 \tag{B.1j}$$

Here, v corresponds to the change in log real GDP levels times 100, the change in the log GDP deflator times 100, or the Federal Funds Rate (at a quarterly rate).

For the calendar-based forecasts, the code generating the observations takes care of the variation of h with the calendar year – so that the variable is always populated such that the non-missing value corresponds to a fourth quarter.

(2) Definition of forecast errors Transcribing time t forecasts into noisy realizations means translating correlation across forecast horizons to correlations over time. Below, $q(t)$ denotes the calendar quarter associated with date t. Thus, $t + q(t) - 4$ accounts for the fact that in the first quarter of a given year, the 1-year ahead forecast and the 4-quarter ahead forecast have a maximum difference in forecast horizons of three quarters, whereas the maximum forecast horizons coincide in the fourth quarter.

Nowcast error

$$u_t^{v,0} = \sigma_{v0}\epsilon_t^{v,0} \qquad v \in \{dy, \pi\} \tag{B.2a}$$

$$u_t^{r,0} = \sum_{v \in \{dy,\pi\}} cov_{r,v,0} u_t^{v,1} + \sigma_{r0}\epsilon_t^{r,0} \tag{B.2b}$$

1-qtr ahead

$$u_t^{v,1} = ar_{v,01} u_{t-1}^{v,0} + \sigma_{v1}\epsilon_t^{v,1} \qquad v \in \{dy, \pi\} \tag{B.2c}$$

$$u_t^{r,1} = ar_{r,01} u_{t-1}^{r,0} + \sum_{v \in \{dy,\pi\}} cov_{r,v,1} u_t^{v,1} + \sigma_{r1}\epsilon_t^{r,1} \tag{B.2d}$$

4-qtr ahead

$$u_t^{v,4} = ar_{v,14} u_{t-3}^{v,1} + \sigma_{v4}\epsilon_t^{v,4} \qquad v \in \{dy, \pi\} \tag{B.2e}$$

$$u_t^{r,4} = ar_{r,14} u_{t-3}^{r,1} + \sum_{v \in \{dy,\pi\}} cov_{r,v,4} u_t^{v,4} + \sigma_{r4}\epsilon_t^{r,4}$$ 1-cy ahead

$$u_t^{v,1y} = ar_{v,41y} u_{t+q(t)-4}^{v,4} + \sigma_{v1y}\epsilon_t^{v,1y} \qquad v \in \{dy, \pi\} \tag{B.2f}$$

$$u_t^{r,1y} = ar_{r,41y} u_{t+q(t)-4}^{r,4} + \sum_{v \in \{dy,\pi\}} cov_{r,v,1y} u_t^{v,1y} + \sigma_{r1y}\epsilon_t^{r,1y} \tag{B.2g}$$

2-cy ahead

$$u_t^{dy,2y} = ar_{dy,1y2y} u_{t-4}^{dy,1cy} + \sigma_{dy2y}\epsilon_t^{dy,2y} \tag{B.2h}$$

3-cy ahead

$$u_t^{dy,3y} = ar_{dy,2y3y} u_{t-4}^{dy,2cy} + \sigma_{dy3y}\epsilon_t^{dy,3y} \tag{B.2i}$$

(3) Population growth

$$\Delta \ln pop_t = \mu_{pop} + \rho_{pop}\Delta \ln pop_{t-1} + \sigma_{pop}\epsilon_t^{pop}. \tag{B.3a}$$

$$\Delta \ln y_t = \Delta \ln y_t^{pc} + \Delta \ln Pop_t \tag{B.3b}$$

In practice, this is implemented by inserting code modules in each block of original Dynare code for (1) endogenous variables, (2) endogenous variables, (3) the calibration, (4) the model equations, (5) the steady state, and (6) the shock processes. These take the following form:

```
@#include "SPF_shocks.mod"
```

C Data appendix

C.1 Download links for composite citations

1. Table C.1 provides the download links for Board of Governors of the Federal Reserve System (nd).

2. Table C.2 provides the download links for U.S. Bureau of Labor Statistics (nd).

3. Table C.3 provides the download links for U.S. Bureau of Economic Analysis (nd).

4. Table C.4 provides the download links for Federal Reserve Bank of Philadelphia (2022).

Source	Series name	Link
Board of Governors of the Federal Reserve System (US) (n.d.)	Market Yield on U.S. Treasury Securities at 10-Year Constant Maturity, Quoted on an Investment Basis [GS10]	https://fred.stlouisfed.org/series/GS10
Board of Governors of the Federal Reserve System (US) (n.d.)	Industrial Production: Total Index [INDPRO]	https://fred.stlouisfed.org/series/INDPRO
Board of Governors of the Federal Reserve System (US) (n.d.)	Federal Funds Effective Rate [FEDFUNDS]	https://fred.stlouisfed.org/series/FEDFUNDS

Table C.1: Board of Governors Industrial and Financial Market data (Board of Governors of the Federal Reserve System, nd) retrieved from FRED, Federal Reserve Bank of St. Louis: data links

Source	Series name	Link
U.S. Bureau of Labor Statistics (n.d.)	Nonfarm Business Sector: Average Weekly Hours Worked for All Employed Persons [PRS85006023]	https://fred.stlouisfed.org/series/PRS85006023
U.S. Bureau of Labor Statistics (n.d.)	Nonfarm Business Sector: Hourly Compensation for All Employed Persons [COMPNFB]	https://fred.stlouisfed.org/series/COMPNFB
U.S. Bureau of Labor Statistics (n.d.)	Average Weekly Hours of Production and Nonsupervisory Employees, Total Private [AWHNONAG]	https://fred.stlouisfed.org/series/AWHNONAG
U.S. Bureau of Labor Statistics (n.d.)	Unemployment Rate [UNRATE]	https://fred.stlouisfed.org/series/UNRATE
U.S. Bureau of Labor Statistics (n.d.)	Employment Level [CE16OV]	https://fred.stlouisfed.org/series/CE16OV
U.S. Bureau of Labor Statistics (n.d.)	Population Level [CNP16OV]	https://fred.stlouisfed.org/series/CNP16OV

Table C.2: BLS Labor Market data (U.S. Bureau of Labor Statistics, nd) retrieved from FRED, Federal Reserve Bank of St. Louis: data links

Source	Series name	Link
U.S. Bureau of Economic Analysis (n.d.)	Personal Consumption Expenditures: Services [PCESV]	https://fred.stlouisfed.org/series/PCESV
U.S. Bureau of Economic Analysis (n.d.)	Personal Consumption Expenditures: Nondurable Goods [PCND]	https://fred.stlouisfed.org/series/PCND
U.S. Bureau of Economic Analysis (n.d.)	Personal Consumption Expenditures: Durable Goods [PCDG]	https://fred.stlouisfed.org/series/PCDG
U.S. Bureau of Economic Analysis (n.d.)	Real Gross Domestic Product [GDPC1]	https://fred.stlouisfed.org/series/GDPC1
U.S. Bureau of Economic Analysis (n.d.)	Gross Domestic Product: Implicit Price Deflator [GDPDEF]	https://fred.stlouisfed.org/series/GDPDEF
U.S. Bureau of Economic Analysis (n.d.)	Gross Domestic Investment [GPDI]	https://fred.stlouisfed.org/series/GPDI

Table C.3: BEA NIPA data (U.S. Bureau of Economic Analysis, nd) retrieved from FRED, Federal Reserve Bank of St. Louis: data links

Source	Series name	Link
Federal Reserve Bank of Philadelphia (2022)	Media Forecast Data for Levels	https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/survey-of-professional-forecasters/historical-data/medianlevel.xlsx
Federal Reserve Bank of Philadelphia (2022)	Media Forecast Data for Growth	https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/survey-of-professional-forecasters/historical-data/mediangrowth.xlsx
Federal Reserve Bank of Philadelphia (2022)	Additional 10-Year-Ahead Inflation Forecasts from Other Sources	https://www.philadelphiafed.org/-/media/frbp/assets/surveys-and-data/survey-of-professional-forecasters/historical-data/additional-cpie10.xlsx

Table C.4: *Survey of Professional Forecasters* data (Federal Reserve Bank of Philadelphia, 2022): data links.

C.2 Variable definitions and transformations

Using the series mnemonics from FRED from Tables C.1 through C.3 as well as the mnemonics for Moody's (nd) and defining the shadow rate Federal Reserve Bank of Atlanta (nd) as `WuXia rate`, the data series used in the estimation are defined as follows:

- Population pop_t is exponentially smoothed prior to computations.

$$\ln pop_t = (1 - 1/50) \ln pop_{t-1} + \frac{1}{50} \ln \texttt{CNP16OV}_t, \qquad \ln pop_0 = \ln \texttt{CNP16OV}_0.$$

- Consumption level:

$$\ln \frac{\texttt{PCESV}_t + \texttt{PCND}_t}{\texttt{GDPDEF}_t}$$

- Investment level:

$$\ln \frac{\texttt{PCDG}_t + \texttt{GPDI}_t}{\texttt{GDPDEF}_t}$$

- Hours per capita:

$$\ln(\texttt{CE16OV}_t \times \texttt{AWHNONAG}_t) - \ln \texttt{CNP16OV}_t$$

- Real wage:

$$\ln(\texttt{COMPNFB}_t) - \ln(\texttt{GDPDEF}_t)$$

- Effective short rate:

$$r_t = \begin{cases} \texttt{WuXia rate}/4 & \texttt{FEDFUNDS} < 0.4/4 \\ \texttt{FEDFUNDS}/4. & \texttt{FEDFUNDS} \geq 0.4/4 \end{cases}$$

- Spread:

$$\texttt{BAA}_t - \texttt{GS10}_t.$$

- Inflation:

$$100 \times \ln(\texttt{GDPDEF}_t/\texttt{GDPDEF}_{t-1})$$

D DSGE model estimation

Parameter estimation is based on the pure DSGE model independent of whether the forecasts include SPF forecasts as added information. This is because, as Del Negro and Schorfheide (2013) observe, the few extra observations at the end of the estimation sample leave the likelihood function approximately unaffected.

The estimation of SW closely follows the original paper, except that the parameters of the population growth process (an AR(1) process with intercept) are also estimated. Table D.5 shows the distribution of the posterior mode estimates from 1997q2 through 2018q4.

In the estimation of DNGS, intercepts of observation equations as well as the persistence and volatility of the population growth process are set to the sample moments prior to estimation. The intertemporal elasticity of substitution is fixed at unity prior to estimation. Table D.6 shows the mean and standard deviation of posterior mode estimates from 1997q2 through 2018q4 as well as prior means and standard deviations.

Estimated Parameter	Prior Mean (s.d.)	Posterior modes Mean over time (s.d.)
ea	0.100 (2.000)	0.707 (0.013)
eb	0.100 (2.000)	0.182 (0.045)
eg	0.100 (2.000)	0.402 (0.018)
eqs	0.100 (2.000)	1.291 (0.227)
em	0.100 (2.000)	0.296 (0.015)
epinf	0.100 (2.000)	0.236 (0.022)
ew	0.100 (2.000)	0.256 (0.057)
epop	0.005 (0.200)	0.003 (0.000)
crhoa	0.500 (0.200)	0.988 (0.002)
crhob	0.500 (0.200)	0.302 (0.220)
crhog	0.500 (0.200)	0.982 (0.006)
crhoqs	0.500 (0.200)	0.503 (0.135)
crhoms	0.500 (0.200)	0.042 (0.005)
crhopinf	0.500 (0.200)	0.992 (0.002)
crhow	0.500 (0.200)	0.978 (0.003)
cmap	0.500 (0.200)	0.921 (0.008)
cmaw	0.500 (0.200)	0.917 (0.031)
csadjcost	4.000 (1.500)	5.099 (0.548)
csigma	1.500 (0.375)	1.369 (0.097)
chabb	0.700 (0.100)	0.797 (0.067)
cprobw	0.500 (0.100)	0.758 (0.018)
csigl	2.000 (0.750)	2.238 (0.185)
cprobp	0.500 (0.100)	0.654 (0.026)
cindw	0.500 (0.150)	0.512 (0.056)
cindp	0.500 (0.150)	0.285 (0.038)
czcap	0.500 (0.150)	0.369 (0.083)
cfc	1.250 (0.125)	1.513 (0.047)
crpi	1.500 (0.250)	1.914 (0.062)
crr	0.750 (0.100)	0.815 (0.007)
cry	0.125 (0.050)	0.126 (0.007)
crdy	0.125 (0.050)	0.201 (0.008)
constepinf	0.625 (0.100)	0.675 (0.027)
constebeta	0.250 (0.100)	0.134 (0.017)
constelab	0.000 (2.000)	2.511 (0.402)
ctrend	0.400 (0.100)	0.496 (0.046)
cgy	0.500 (0.250)	0.215 (0.009)
calfa	0.300 (0.050)	0.192 (0.003)
constepop	0.000 (0.200)	0.005 (0.001)
crhopop	0.500 (0.200)	0.987 (0.002)

SPF f'cast parameter	Mean over time (s.d.)
se dy now	0.005 (0.000)
se dy plus1q	0.005 (0.000)
se dy plus4q	0.003 (0.000)
se dy plus1cy	0.001 (0.000)
se r now	0.009 (0.001)
se r plus1q	0.001 (0.000)
se r plus4q	0.004 (0.001)
se r plus1cy	0.001 (0.000)
se pinf now	0.002 (0.000)
se pinf plus1q	0.002 (0.000)
se pinf plus4q	0.001 (0.000)
se pinf plus1cy	0.001 (0.000)
ar dy now1q	0.299 (0.158)
ar dy 1q4q	0.481 (0.068)
ar dy 4q1cy	0.725 (0.122)
ar pinf now1q	0.410 (0.098)
ar pinf 1q4q	0.568 (0.107)
ar pinf 4q1cy	0.725 (0.122)
ar dy 1cy2cy	0.000 (0.000)
ar dy 2cy3cy	0.000 (0.000)
se dy plus2cy	6.431 (4.811)
se dy plus3cy	6.891 (4.647)
cov r pinf now	0.000 (0.000)
cov r dy now	0.000 (0.000)
ar r now1q	0.965 (0.015)
cov r pinf plus1q	0.000 (0.000)
cov r dy plus1q	0.000 (0.000)
ar r 1q4q	0.787 (0.135)
cov r pinf plus4q	0.000 (0.000)
cov r dy plus4q	0.000 (0.000)
ar r 4q1cy	0.732 (0.130)
cov r pinf plus1cy	0.000 (0.000)
cov r dy plus1cy	0.000 (0.000)

Calibrated parameter	Mean over time (s.d.)
cg	0.180 (0.000)
clandaw	1.500 (0.000)
crhoas	1.000 (0.000)
crhols	0.993 (0.000)
ctou	0.025 (0.000)
curvp	10.000 (0.000)
curvw	10.000 (0.000)

Table D.5: SW parameter estimates: Distribution of posterior modes from 1997q2 through 2018q4

Estimated Parameter	Prior Mean (s.d.)	Posterior modes Mean over time (s.d.)
psi z	0.100 (2.000)	0.009 (0.000)
psi b	0.100 (2.000)	0.006 (0.000)
psi g	0.100 (2.000)	0.022 (0.001)
psi mu	0.100 (2.000)	0.032 (0.008)
psi rm	0.100 (2.000)	0.009 (0.000)
psi laf	0.100 (2.000)	0.007 (0.000)
psi law	0.100 (2.000)	0.011 (0.001)
psi pist	0.100 (2.000)	0.008 (0.000)
psi sigw	0.050 (4.000)	0.005 (0.000)
epop	0.005 (0.200)	0.003 (0.000)
rho sigw	0.750 (0.150)	0.691 (0.080)
rho z	0.500 (0.200)	0.956 (0.017)
rho b	0.500 (0.200)	0.693 (0.076)
rho g	0.500 (0.200)	0.958 (0.011)
rho mu	0.500 (0.200)	0.940 (0.062)
rho rm	0.500 (0.200)	0.403 (0.054)
rho laf	0.500 (0.200)	0.761 (0.020)
rho law	0.500 (0.200)	0.086 (0.248)
s2	4.000 (1.500)	1.100 (0.331)
h	0.700 (0.100)	0.151 (0.012)
zeta w	0.500 (0.100)	0.855 (0.167)
iota w	0.500 (0.150)	0.406 (0.135)
zeta p	0.500 (0.100)	0.954 (0.004)
iota p	0.500 (0.150)	0.050 (0.006)
alp	0.300 (0.050)	0.216 (0.007)
ppsi	0.500 (0.150)	0.703 (0.084)
psi1	1.500 (0.061)	1.217 (0.055)
psi2	0.125 (0.050)	0.547 (0.057)
psi3	0.125 (0.050)	0.379 (0.042)
zeta spb	0.050 (0.005)	0.028 (0.005)

Calibrated parameter	Mean over time (s.d.)
Bigphi	1.250 (0.000)
Fom	0.030 (0.000)
clandaw	1.500 (0.000)
constelab	-1.800 (0.000)
constepinf	0.965 (0.080)
constepop	0.005 (0.001)
conster	1.559 (0.168)
constespread	2.000 (0.000)
crhopop	0.985 (0.003)
ctrend	0.438 (0.039)
del	0.025 (0.000)
epsp	10.000 (0.000)
epsw	10.000 (0.000)
eta gz	0.874 (0.000)
eta laf	0.714 (0.000)
eta law	0.572 (0.000)
gammstar	0.990 (0.000)
gstar	0.180 (0.000)
nu l	2.673 (0.000)
rho	0.675 (0.000)
rho pist	0.990 (0.000)
sigmac	1.010 (0.000)

SPF f'cast parameter	Mean over time (s.d.)
se dy now	0.005 (0.000)
se dy plus1q	0.005 (0.000)
se dy plus4q	0.003 (0.000)
se dy plus1cy	0.001 (0.000)
se r now	0.009 (0.001)
se r plus1q	0.001 (0.000)
se r plus4q	0.004 (0.001)
se r plus1cy	0.001 (0.000)
se pinf now	0.002 (0.000)
se pinf plus1q	0.002 (0.000)
se pinf plus4q	0.001 (0.000)
se pinf plus1cy	0.001 (0.000)
ar dy now1q	0.299 (0.158)
ar dy 1q4q	0.481 (0.068)
ar dy 4q1cy	0.725 (0.122)
ar pinf now1q	0.410 (0.098)
ar pinf 1q4q	0.568 (0.107)
ar pinf 4q1cy	0.725 (0.122)
ar dy 1cy2cy	0.000 (0.000)
ar dy 2cy3cy	0.000 (0.000)
se dy plus2cy	6.431 (4.811)
se dy plus3cy	6.891 (4.647)
cov r pinf now	0.000 (0.000)
cov r dy now	0.000 (0.000)
ar r now1q	0.965 (0.015)
cov r pinf plus1q	0.000 (0.000)
cov r dy plus1q	0.000 (0.000)
ar r 1q4q	0.787 (0.135)
cov r pinf plus4q	0.000 (0.000)
cov r dy plus4q	0.000 (0.000)
ar r 4q1cy	0.732 (0.130)
cov r pinf plus1cy	0.000 (0.000)
cov r dy plus1cy	0.000 (0.000)

Table D.6: DNGS parameter estimates: Distribution of posterior modes from 1997q2 through 2018q4

E Additional results

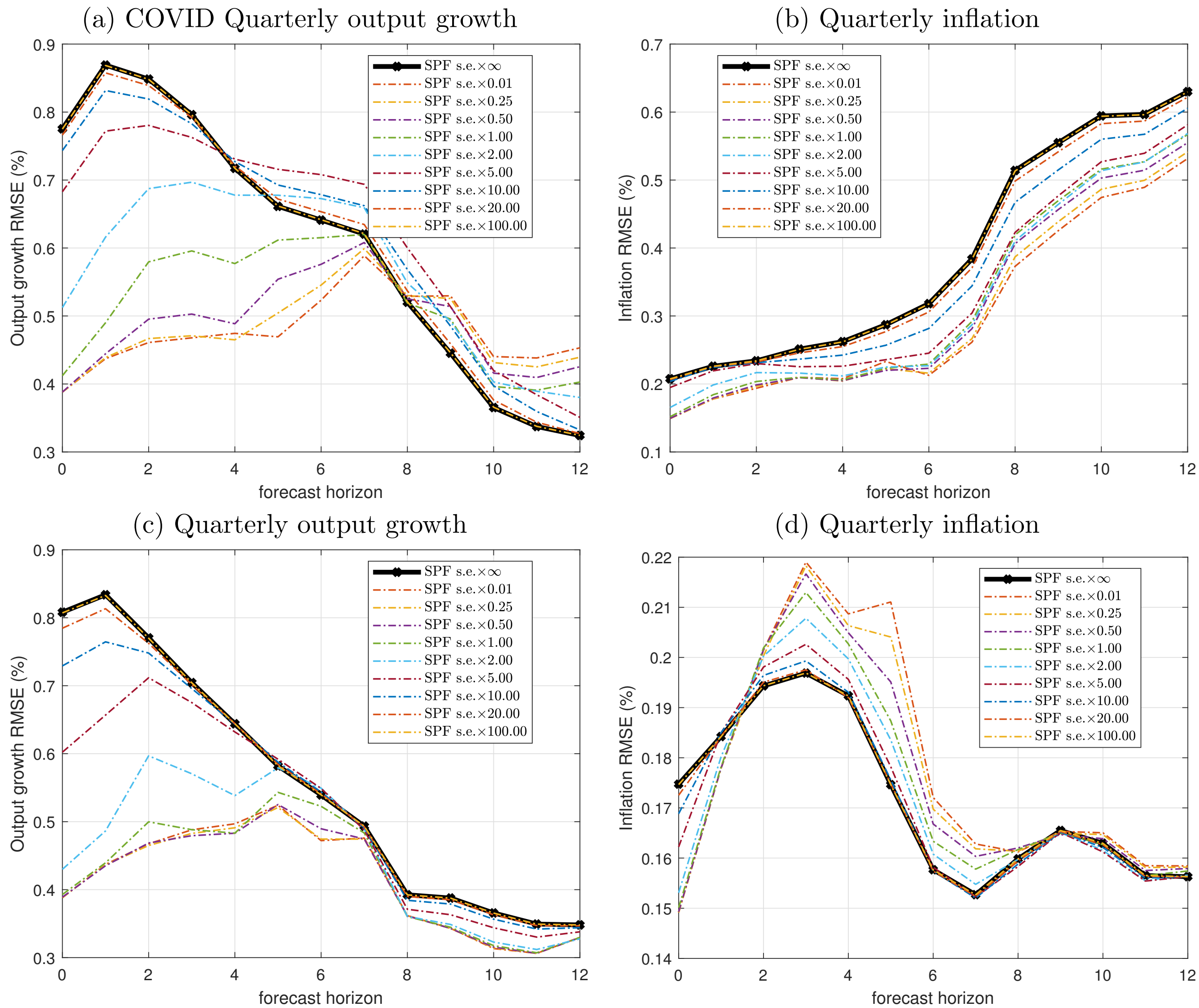


Figure E.1: Non-cumulative mean absolute forecast errors by variable and horizon. Forecasts from 1997q2 through 2018q4.

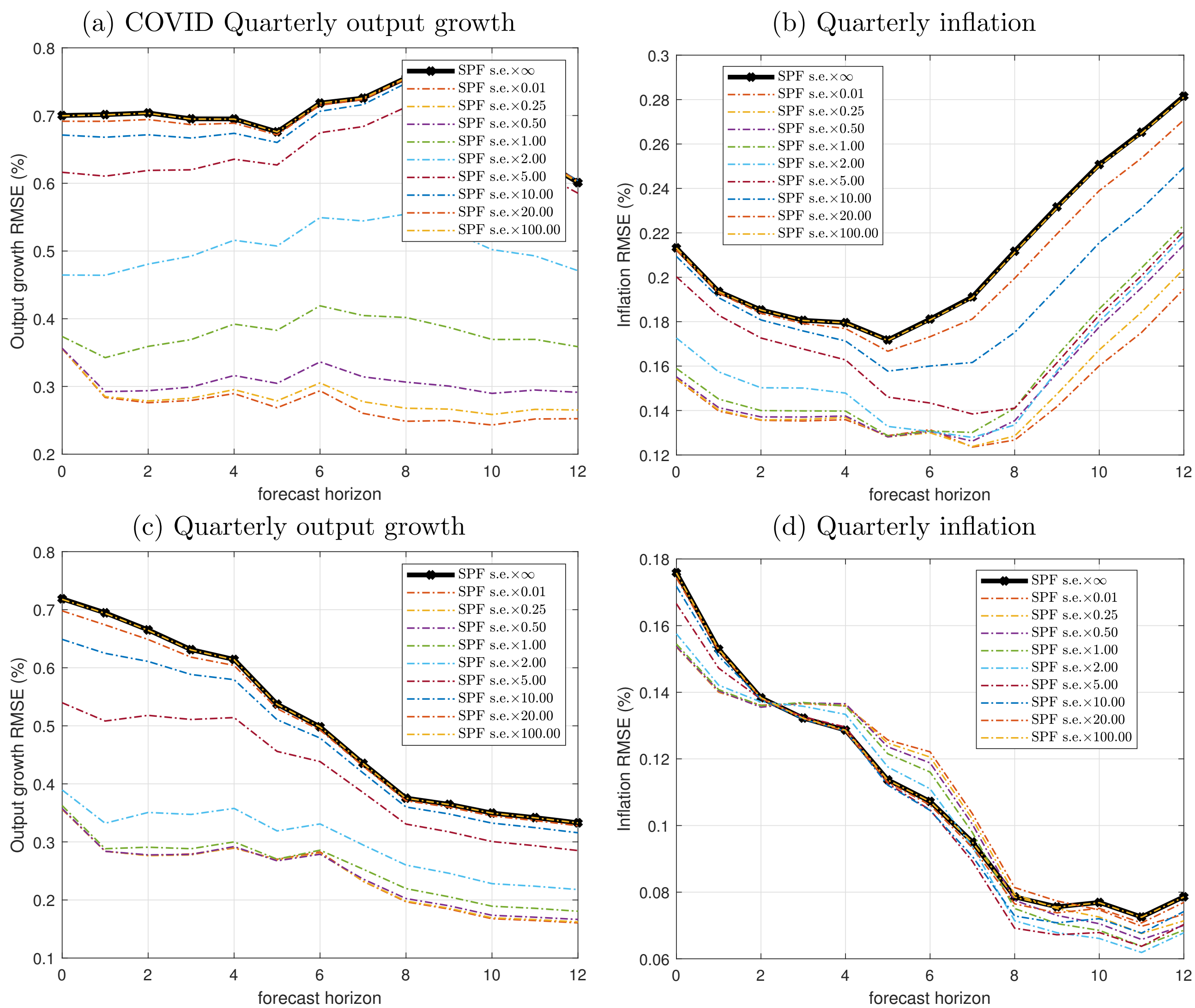
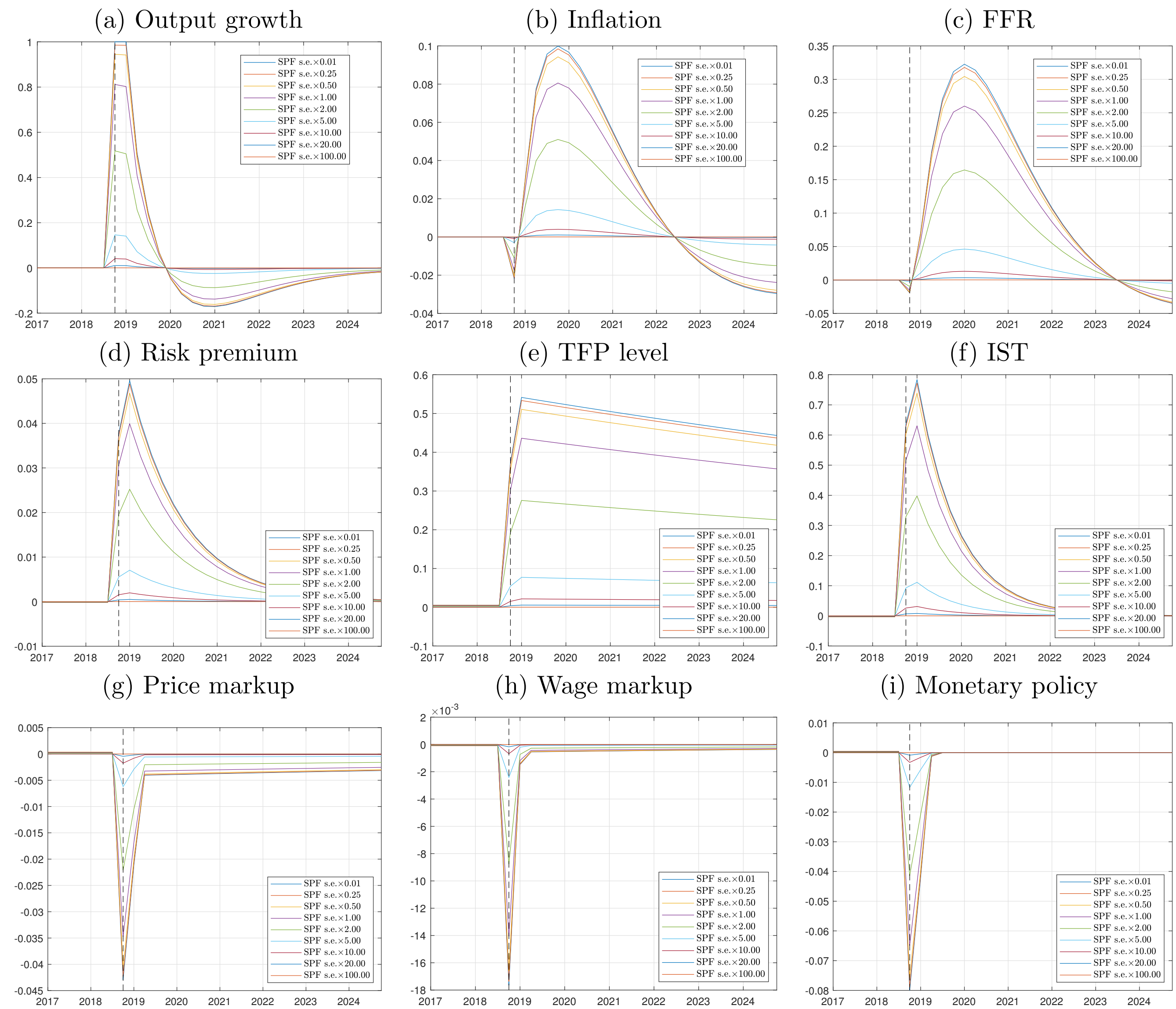
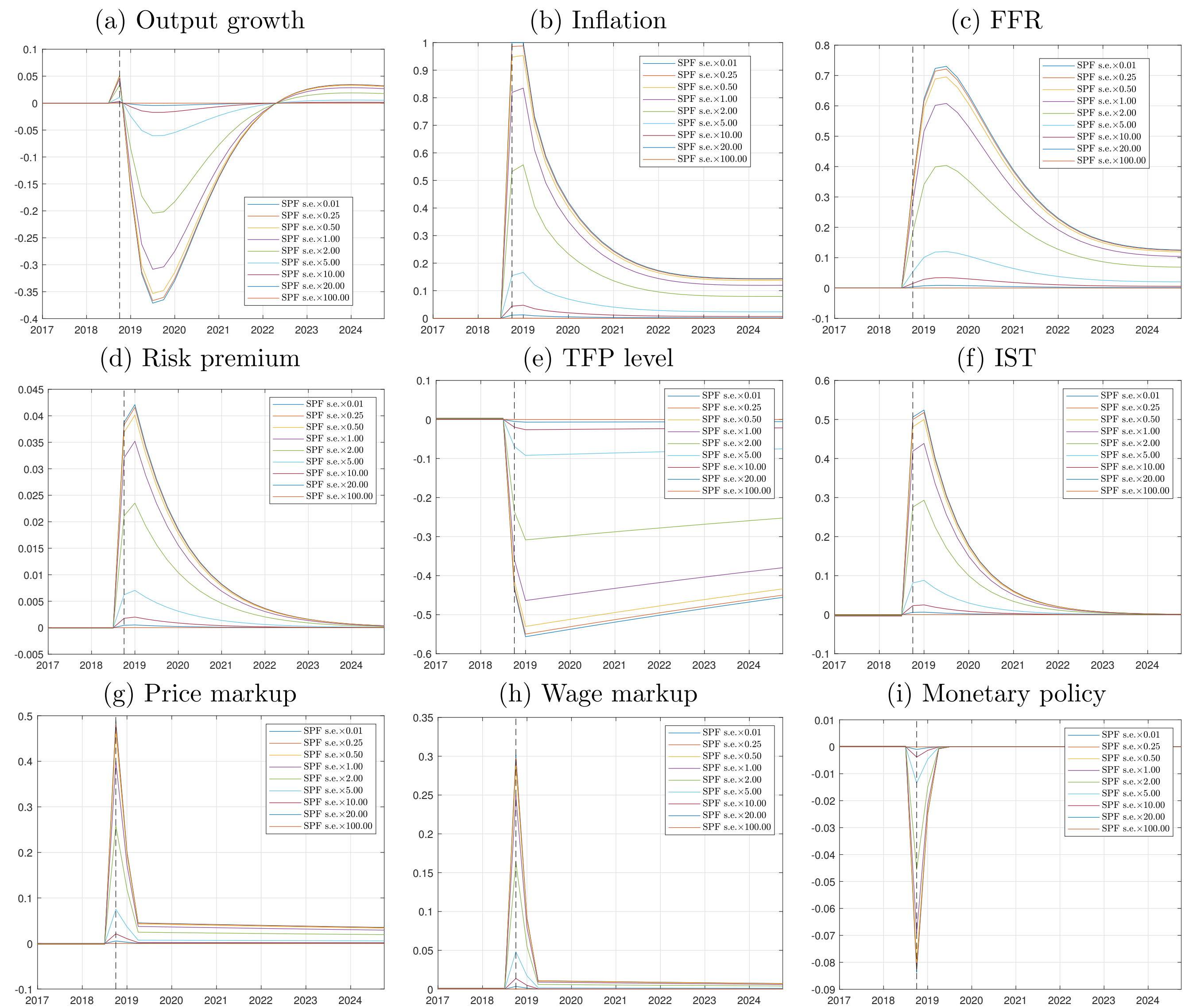


Figure E.2: MAE by variable and horizon: COVID



Note: If the forecasts were mapped into the DSGE model as the sum of the truth and pure white noise and the ABCD condition holds, then the external forecasts would result in updated paths that are just scaled down versions of the updates following the same measurement-error free observations. See Appendix A.2. Because forecast errors here exhibit cross-correlations, this does not hold exactly.

Figure E.3: Effects of a unit forecast revision in the 2018q4 output growth nowcast and 1-quarter-ahead forecast on quarterly output, inflation, FFR forecasts as well as shocks in SW



Note: If the forecasts were mapped into the DSGE model as the sum of the truth and pure white noise and the ABCD condition holds, then the external forecasts would result in updated paths that are just scaled down versions of the updates following the same measurement-error free observations. See Appendix A.2. Because forecast errors here exhibit cross-correlations, this does not hold exactly.

Figure E.4: Effects of a unit forecast revision in the 2018q4 inflation nowcast and 1-quarter-ahead forecast on quarterly output, inflation, FFR forecasts as well as shocks in SW

F NK3 Monte Carlo study

The environment builds on the standard, three equation New Keynesian model as in Galí (2008): A New Keynesian Phillips Curve (NKPC) links inflation today to the expected future inflation and the output gap. A New Keynesian Intertemporal Substitution (NKIS) equation links the output gap today to the expected future output gap and the real interest rate gap, that is, the nominal short rate minus expected inflation and minus the natural rate of interest. The nominal short rate follows a Taylor rule, which, unlike in Galí (2008), exhibits persistence.

The solution to the linearized version of the model has the canonical state space representation (2.1). If shocks are persistent, the external states are part of $\mathbf{x}$. Otherwise, only the lagged interest rate is part of $\mathbf{x}$. $\mathbf{y}_t$ contains the observables of the model, as well as internal, static or purely forward-looking variables.

The model has three structural shocks:

1. A shock to the natural rate of interest.
2. A monetary policy shock.
3. A cost push shock, ie, an ad hoc shock to the NKPC.

In addition, I will be adding measurement error to observables, or a large shock that is a stylized representation of the shocks buffeting the economy during the COVID pandemic.

Below, I simulate data for $T = 1,108$ periods from the NK3 model. I then create 1,000 data sets of size $t_n \in \{100, 101, \ldots, 1000\}$ with an extra 8 observations of "external forecasts". I do so by adding an AR(1) process of noise to the actual future realizations, with quarterly persistence of 0.5 and $\mathcal{N}(0, \kappa^2 \operatorname{Var}_{t-1}[y_t])$ innovations – measurement error is proportional to the one-step-ahead forecast error variance. I consider a precise version with $\kappa^2 = 4^{-2}$ and a noisy version with $\kappa^2 = 1$. The external expectations are observed at horizons $h \in \{1, 4, 8\}$. I also consider a version where Dynare assumes that the measurement error is small ($\kappa^2 = 40^{-2}$).

Table F.7 shows the forecast performance of the plain DSGE model and the model augmented with external "forecasts" for varying scaled κ of measurement error. The simulation also speaks to the case of misspecification – when the underlying external forecasts are created to have a higher measurement error variance than the calibration allows for. This misspecification reduces the gains from incorporating the external forecasts, but only slightly.

(a) Plain NK3 forecast RMSE

Variable	Horizon 1	2	3	4	5	6	7	8
y	239.6	275.7	292.4	300.6	304.1	306.3	307.5	308.1
infl	73.5	87.5	95.2	100.2	103.1	105.7	107.9	109.5
i	88.7	103.4	112.6	117.6	121.8	125.2	128.2	131.4

(b) NK3 forecast with external forecasts treated as noisy data RMSE − $\kappa = \frac{1}{4}$ calibration

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	120.3	241.9	276.4	118.9	241.8	277.1	293.6	120.2
infl	35.0	74.9	88.3	34.5	74.5	88.4	96.0	34.4
i	47.6	91.2	105.8	47.5	91.7	105.6	114.2	47.8

(c) Improvement of (b) over (a) [in %] − $\kappa = \frac{1}{4}$ calibration

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	-49.8	-12.3	-5.5	-60.4	-20.5	-9.5	-4.5	-61.0
infl	-52.4	-14.4	-7.3	-65.5	-27.8	-16.4	-11.0	-68.5
i	-46.4	-11.8	-6.1	-59.6	-24.7	-15.7	-10.9	-63.6

(d) Improvement over (a) [in %] − $\kappa = \frac{1}{4}$ calibration, treated as more precise ($\kappa = 1/40$)

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	-47.2	-11.7	-5.5	-60.5	-20.6	-9.6	-4.4	-60.8
infl	-48.8	-12.8	-6.5	-64.8	-27.2	-15.8	-10.3	-68.0
i	-46.3	-10.9	-4.7	-59.6	-23.7	-14.5	-10.3	-63.8

(e) NK3 forecast with external forecasts treated as noisy data RMSE − $\kappa = 1$ calibration

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	173.6	255.1	282.1	177.1	256.6	283.2	296.9	180.6
infl	54.5	81.0	91.5	59.7	83.3	93.7	100.4	60.9
i	69.1	98.1	110.8	80.8	106.6	116.0	121.4	83.6

(f) Improvement of (d) over (a) − $\kappa = 1$ calibration [in %]

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	-27.6	-7.5	-3.5	-41.1	-15.6	-7.5	-3.4	-41.4
infl	-25.9	-7.5	-3.8	-40.4	-19.2	-11.3	-7.0	-44.4
i	-22.1	-5.1	-1.7	-31.3	-12.5	-7.4	-5.3	-36.4

(g) Improvement over (a) − $\kappa = 1$ calibration [in %], treated as more precise ($\kappa = 1/40$)

Variable	Horizon (* indicates added observation) 1*	2	3	4*	5	6	7	8*
y	3.9	3.5	0.1	-31.2	-12.3	-6.2	-2.1	-31.2
infl	15.5	24.7	18.3	-27.8	-4.4	1.1	3.7	-34.6
i	19.4	28.3	30.1	-16.2	5.7	10.7	10.7	-26.8

Table F.7: RMSE and RMSE reduction for plain NK3 and NK3 with noisy versions of the true data added as forecasts.