Cutting-Edge Methods Did Not Improve Inflation Forecasting during the COVID-19 Pandemic

By Amaze Lusompa and Sai A. Sattiraju

Monetary policymakers depend heavily on forecasts about the future state of the economy. Since the beginning of the COVID-19 pandemic, however, the Federal Open Market Committee (FOMC) and economists in general have not been able to accurately forecast inflation. The surge of inflation in 2021–22 caught most experts by surprise, and even economists who predicted a surge in inflation underpredicted the size. Although central bankers' inflation forecasts tend to be fairly accurate during normal times, they do not perform as well during downturns and periods of extreme uncertainty.

To improve this performance gap, researchers over the past 20 years have proposed various innovations to a benchmark class of models known as "time-varying parameter models," which allow the relationships between forecasting variables to change over time. Although these innovations have improved models' forecasting performance during previous recessions, most research on the efficacy of these innovations was conducted prior to the COVID-19 pandemic. A natural question is how these "improved" models have performed during recent extreme events.

In this article, we investigate whether innovations in time-varying parameter models led to improved inflation forecasting during the pandemic. We find that despite their promise prior to the pandemic, forecasting innovations did not improve the accuracy of inflation forecasts relative to a

---

Amaze Lusompa is an economist at the Federal Reserve Bank of Kansas City. Sai A. Sattiraju is a research associate at the bank. This article is on the bank's website at ***www.KansasCityFed.org***

baseline time-varying parameter model during the pandemic. Our results suggest that forecasters may need to develop a new class of forecasting models, introduce new forecasting variables, or rethink how they forecast to yield more effective inflation forecasts during extreme events.

Section I outlines different forecasting innovations of the past 20 years. Section II compares the performance of different forecasting models for U.S. inflation during the pandemic and shows that innovative time-varying parameter models did not outperform a baseline time-varying parameter model.

I. Innovation in Forecasting

Many of the time- and computationally intensive innovations in forecasting over the past several years have resulted from time-varying parameter models. These models are very flexible, as they allow for the relationships between forecasting variables to change over time. However, this flexibility comes with costs. Estimating relationships that are not fixed and can change over time is both time- and computationally intensive. Additionally, researchers may not know exactly which variables to include at a given time. As a result, researchers have combined time-varying parameter models with methods like shrinkage, high dimensionality, and variable selection to maintain the flexibility of time-varying parameter models while minimizing the costs.

Time-varying parameter models

Changes in policy, technology, or economic conditions can all lead the relationship between variables in a regression model to change over time, a quality known as "parameter instability." Both Stock and Watson (1996) and Ang and Bekaert (2002) show that many macroeconomic and financial time series models exhibit parameter instability. Accounting for this quality is important, as models that do not consider parameter instability may yield less accurate forecasts. For example, a researcher may attempt to understand or describe how the FOMC responds to changes in output and inflation by estimating a Taylor rule, which relates the value of the federal funds rate to inflation and economic slack. But the Committee's responses to changes in these variables may depend on the individuals that make up the Committee. Thus, when the Committee changes, the estimated parameters in the Taylor rule may need to change as well to better reflect the Committee. As a result, a researcher seeking to estimate the Taylor rule consistent with the Committee's decisions must account for potential parameter instability to avoid misleading forecasts or analysis.

Time-varying parameter models provide a general way to deal with parameter instability by allowing the parameters of the model to change in each time period in the sample. Intuitively, time-varying parameter models work by discounting information over time, giving more weight to recent information about a particular economic variable than past information for any given time period.

Although time-varying parameter models have existed at least since the 1970s, they were not popular initially due to computational difficulties, and the number of variables included in these models was generally limited to five. In the 2000s, however, Cogley and Sargent (2005) and Primiceri (2005) introduced workhorse time-varying parameter models that could be used for forecasting, and other researchers demonstrated that time-varying parameter models could outperform their constant parameter counterparts.[1] Recently, more efficient estimation methods and approximations have been introduced that can lessen the computational burden of these models.

High dimensionality and shrinkage

Forecasters often have to make difficult choices about how much information to include in their models. Generally, forecasters want to include as much relevant information as possible to maximize the accuracy of their forecasts. To do so, they can estimate a high-dimensional model—that is, a model with many independent variables. However, including too much information in a model can lead to imprecise parameter estimates and therefore imprecise forecasts. For example, if a model includes too many variables relative to the sample size, the parameters may not be estimated accurately; this could lead variables that are relatively less important to have disproportionate influence on the forecast, thereby distorting the forecast. A model with too many independent variables is often described as being "overparameterized."

Overparameterization can be dealt with in several ways, including using "shrinkage." Shrinkage is simply a method that "shrinks" an estimate of a parameter toward a pre-specified value. More precisely, shrinkage can be used so that parameters that are less relevant shrink toward zero, while parameters that are more important are left alone (or have minimal shrinkage). Note that shrinkage is not only used in models with many variables. Researchers often use shrinkage to restrict certain parameter values in models with only a few variables. For example, forecast models, including time-varying parameter models, often include past values of a variable, such as inflation, to predict what future values of that variable will be. Because more recent values of inflation are assumed to be more important for predicting future inflation than older values, researchers may elect to shrink the parameters for older values of inflation toward zero so that they are weighted less heavily in the forecast.

Sparsity or variable selection

Shrinkage is often combined with sparsity or variable selection methods to prevent issues such as overparameterization. Despite advances in high dimensional models, computational constraints or practical considerations may still limit the number of variables researchers can include in a model. For example, a time-varying parameter model estimates a different parameter for each period, so for a sample size of T, the model would have T times the number of parameters of a model where the parameters do not change. If a constant parameter model covering 200 periods (for example, 50 years of quarterly data) has 12 parameters, a time-varying parameter model covering the same period would have 2,400.

One way to limit the number of variables in a model is to use sparsity or variable selection methods. As their name suggests, these methods can reduce the number of potential variables in a model to a smaller set that ideally includes enough information to generate accurate forecasts. Although variable selection has been used since at least the early 1990s, algorithms and computing power have only recently evolved to the point where researchers can perform variable selection without a supercomputer.[2]

Dynamic variable selection is a particular form of variable selection that can be especially useful when combined with time-varying parameter models. Dynamic variable selection accounts for the fact that some variables may be helpful in forecasting during certain time periods but not others—as is often the case for macroeconomic and financial variables (see, for example, Korobilis and Koop 2020). For example, expected shipping times are not generally used when predicting inflation, as shipping logjams have historically not been large enough to measurably affect inflation. During the pandemic, however, the increase in expected shipping times is thought to have led to higher shipping costs and hence higher prices, so incorporating an expected shipping time variable could improve inflation forecasts. Dynamic variable selection allows forecasters to incorporate variables in their models only when they are likely to be relevant, thus providing an alternative to estimating high dimensional models.

In general, sparsity can also be used to prevent overfitting of a model. An overfitted model is one that does a good job explaining random variation in one dataset but that performs relatively poorly when used with other datasets. As an analogy, consider a student preparing for a test not by studying the material holistically but by spending too much time on one-off questions used on previous versions of the test. In this case, the student will be prepared only for the one-off questions rather than more general material likely to appear on the test. Similarly, an overfitted model is one that is adapted too closely to "one-off" data (for example, an outlier), which may worsen its ability to forecast. To prevent overfitting, forecasters often use a mechanism such as sparsity or shrinkage that prevents the model from adapting too well to the initial sample.

II. Estimating the Performance of Innovative Time-Varying Parameter Models during the Pandemic

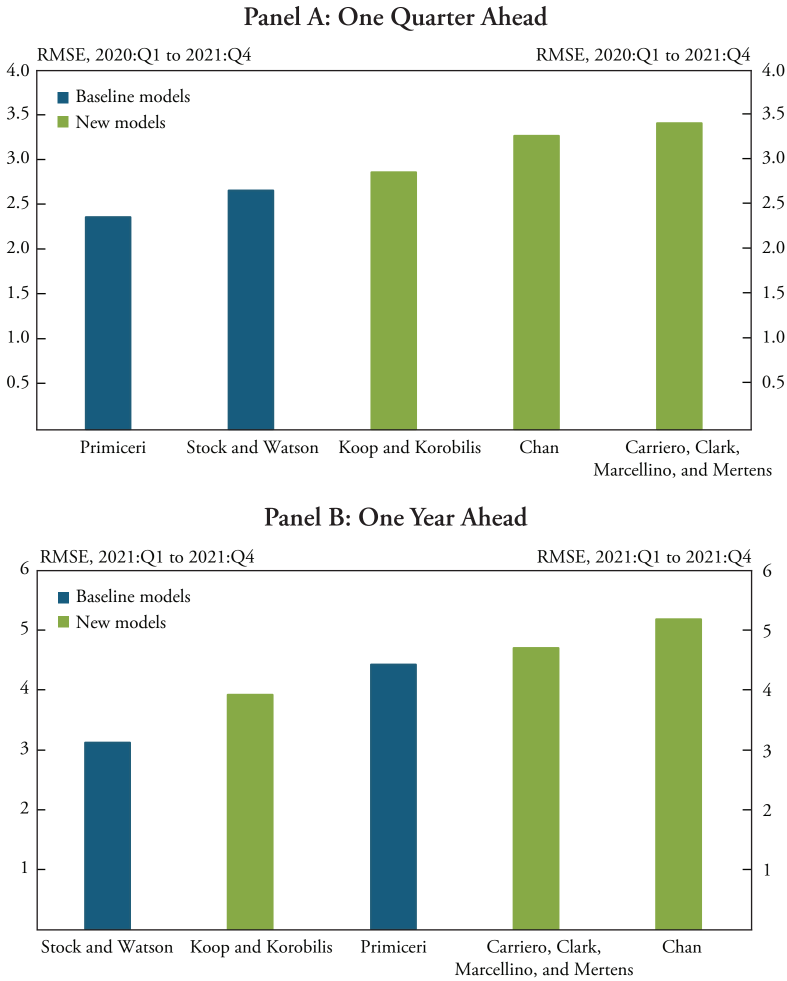
To determine whether the forecasting innovations of the past two decades improved inflation forecasting during the COVID-19 pandemic, we conduct a forecasting exercise that compares the performance of two simple time-varying parameter models generally used in inflation forecasting as well as three newer models that incorporate some of the innovations from the previous section. In particular, we forecast inflation as measured by the price index for personal consumption expenditures (PCE), as PCE inflation is the Federal Reserve's preferred measure (Bernanke 2015). Our two baseline models are the unobserved components model from Stock and Watson (2007), which estimates a time-varying mean of inflation and includes no other predictors, and the time-varying parameter model from Primiceri (2005), which has a small number of variables. Our three newer time-varying parameter models are the model from Carriero and others (2021), which is a moderate-sized dimensional model that incorporates shrinkage and is designed to handle outliers; the model from Chan (2021), which includes many predictors and uses shrinkage; and the model from Korobilis and Koop (2020), which incorporates many predictors, dynamic variable selection, and shrinkage.[3] In summary, the two baseline models have only time-varying parameters, while the three newer models combine time-varying parameters with high dimensionality, shrinkage, or variable selection. Additional details on each of these models as well as their implementation are available in the appendix.

To judge the forecasting performance of the models, we compare their root mean square errors (RMSE). The RMSE quantifies how much a model prediction deviates from the actual data, with smaller RMSEs indicating better forecast performance. In addition, a model's RMSE can help reflect the influence of outliers, in that an inaccurate prediction in one period will have a greater effect on the model's RMSE than an accurate prediction. This quality makes RMSEs especially useful for policymakers. Because one inaccurate inflation forecast can lead to the wrong policy prescription, policymakers may care more about avoiding especially "bad" predictions than about achieving "good" predictions most of the time.

To capture our models' accuracy in forecasting both short-term and longer-term inflation, we examine both one-quarter-ahead and one-year-ahead forecasts. We begin our one-quarter-ahead forecasts in 2020:Q1, at the start of the pandemic, and forecast inflation for the next quarter based on information known up until the previous quarter. For example, our one-quarter-ahead forecast for 2020:Q2 inflation is based on data from up until 2020:Q1. Similarly, our one-year-ahead forecast is based on information known up until the previous year. For this reason, we begin our one-year-ahead forecast in 2021:Q1, as earlier years would not reflect any information from the pandemic.

Panels A and B of Chart 1 show that the baseline models forecast at least as well as the newer models. Panel A shows that the two baseline models (blue bars) have lower RMSEs than the newer models (green bars) for one-quarter-ahead inflation forecasting. Similarly, Panel B shows that

Chart 1
RMSE for Inflation Forecasts during the Pandemic



Sources: Stock and Watson (2007), Primiceri (2005), Korobilis and Koop (2020), Chan (2021), Carriero and others (2021), and FRED (Federal Reserve Bank of St. Louis).

the unobserved components model from Stock and Watson (2007) has the lowest RMSE for one-year-ahead forecasting. Together, the panels suggest that the baseline models yield inflation forecasts at least as accurate—if not more so—than newer models for both time horizons.

The superior performance of the baseline models is somewhat disconcerting. In general, the newer models are more flexible versions of the baseline models; given their increased flexibility, newer models should be able to perform at least as well as the baseline models. Particularly concerning is that the unobserved components model from Stock and Watson (2007), which simply estimates a time-varying mean of inflation and includes no other predictors, outperforms models with a larger number of predictors. The Stock and Watson model has the lowest RMSE for the one-year-ahead forecasts, and the second-lowest RMSE for the one-quarter-ahead forecasts, eclipsed only by the baseline Primiceri (2005) time-varying parameter model. Thus, including additional information does not appear to improve inflation forecasts during the pandemic for the models and data sets we consider.

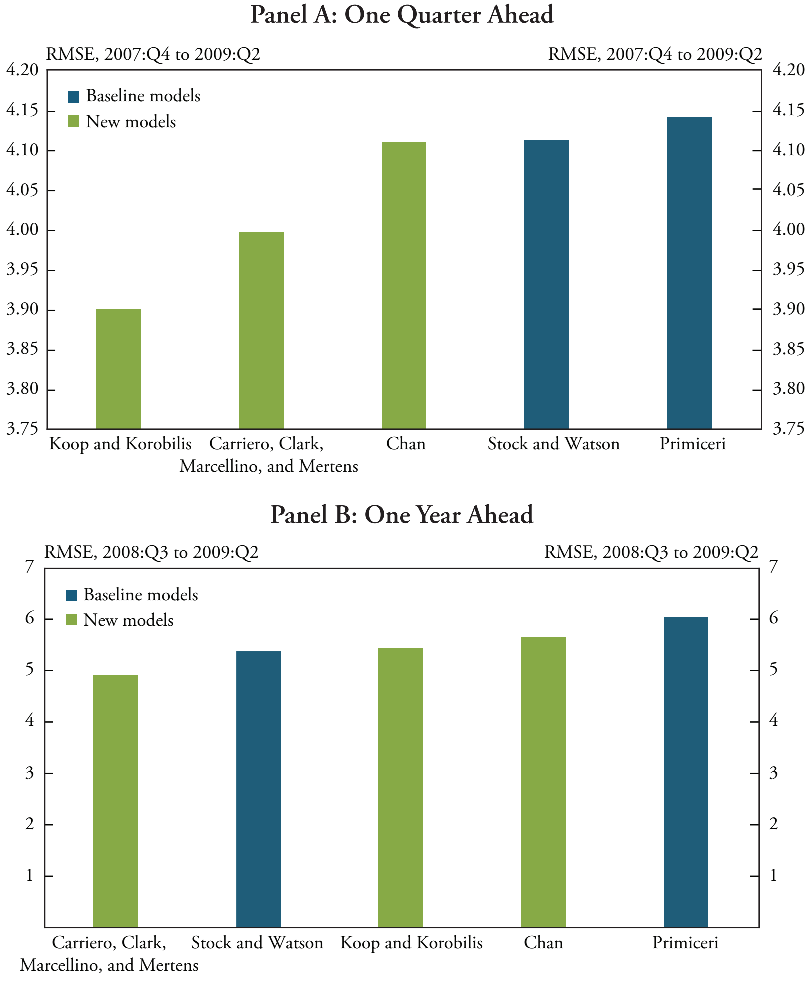
However, these results do not suggest that newer, more sophisticated models should be abandoned entirely. During the Great Recession, for example, these models showed improved forecasting performance against the baseline models. Panel A of Chart 2 shows that for one-quarter-ahead forecasts during the Great Recession, the newer models have a lower RMSE than the baseline models. Although the results are more mixed for the one-year-ahead forecasts during the Great Recession, Panel B of Chart 2 shows that one of the newer models has the lowest RMSE.

To show how the models' forecasting performance evolved over the full Great Recession period, Chart 3 compares the models' forecast errors—the difference between the actual and predicted values of inflation—from 2007:Q4 to 2009:Q2. Values closer to zero indicate a smaller forecast error and therefore better performance. Panel A of Chart 3 shows that no one model dominates for the one-quarter-ahead forecasts. Similarly, Panel B of Chart 3 shows that no one model dominates for the one-year-ahead forecasts, though the relative performance of each model tends to stay consistent across the sample with the exception of the unobserved component model of Stock and Watson (2007). Together, the results from Charts 2 and 3 show that newer models outperformed baseline models during the Great Recession, suggesting they may yet have some benefits in times of distress outside of the pandemic.

Moreover, it may be the case that additional information would have improved inflation forecasts during the pandemic, but that our newer models included the wrong information. Macroeconomic forecasting models in general use macroeconomic and financial variables to forecast. During the COVID-19 pandemic, however, the standard

Chart 2

RMSE for Forecasts during the Great Recession

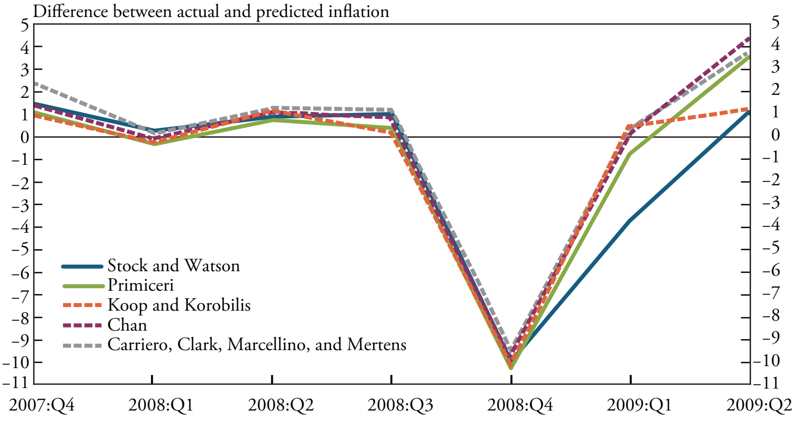


Sources: Stock and Watson (2007), Primiceri (2005), Korobilis and Koop (2020), Chan (2021), Carriero and others (2021), and FRED (Federal Reserve Bank of St. Louis).

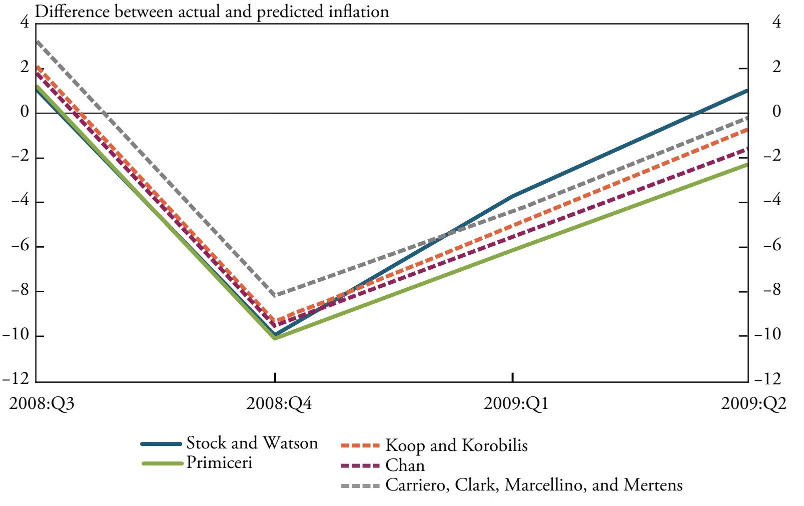
Chart 3

Forecast Errors during the Great Recession

Panel A: One Quarter Ahead



Panel B: One Year Ahead



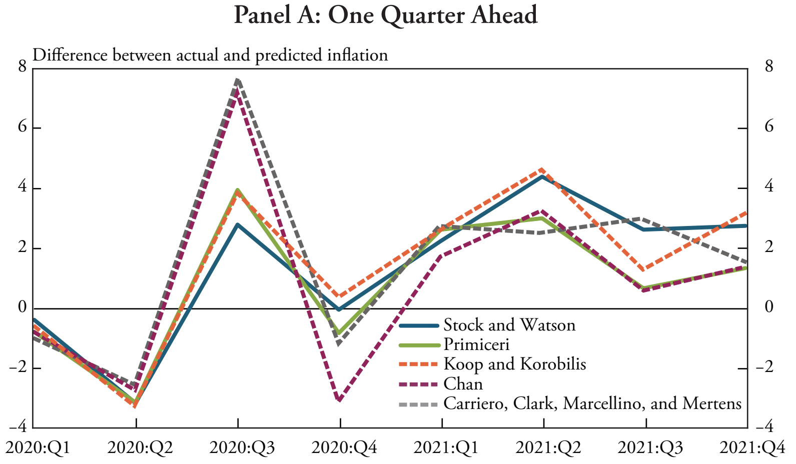
Note: Solid lines represent baseline models and dashed lines represent new models.
Sources: Stock and Watson (2007), Primiceri (2005), Korobilis and Koop (2020), Chan (2021), Carriero and others (2021), and FRED (Federal Reserve Bank of St. Louis).

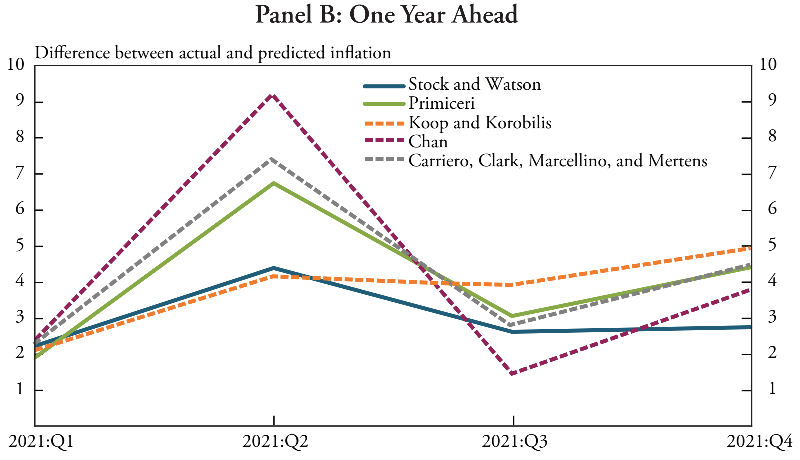
macro and financial variables may have been less useful in forecasting inflation due to the unique combination of strong demand and persistent supply shocks; instead, variables such as U.S. hospitalization rates for COVID-19, expected shipping logjam times at U.S. ports, and some type of production indicator for the countries exporting to the United States may have been more relevant to inflation and thus may have improved inflation forecasts.

Finally, our comparison only accounts for the performance of these models in forecasting inflation—newer models may offer improvements over the baseline models during the pandemic for other variables of interest. Even though the baseline models perform slightly to somewhat better overall, no one model dominates every period. Indeed, Panels A and B of Chart 4, which plot the difference between actual inflation and predicted inflation for the different forecasting models, show a wide variation in the performance of these models across different periods.

Overall, our results suggest that forecasters should not focus on only one model but rather continuously monitor multiple models. One way to do this systematically is by using model averaging, or averaging the predictions of a set of models. Importantly, this method can be combined with time-varying parameter models: dynamic model averaging allows the "importance" or influence of each model on the average prediction to change over time. Some studies have shown that model averaging or combining forecasts can outperform any one model by safeguarding against a bad forecast from a single model (Hoeting and others 1999; Faust and Wright 2013). As the results from this article intimate, model averaging might be a useful tool for forecasting during future extreme events.

Chart 4
Forecast Errors during the Pandemic





Note: Solid lines represent baseline models and dashed lines represent new models.
Sources: Stock and Watson (2007), Primiceri (2005), Korobilis and Koop (2020), Chan (2021), Carriero and others (2021), and FRED (Federal Reserve Bank of St. Louis).

Conclusion

In this article, we investigate whether forecasting innovations in time-varying parameter models led to improved inflation forecasting during the pandemic. We find that despite their promise prior to the pandemic (including during the Great Recession), these innovations did not improve the accuracy of inflation forecasts relative to a baseline model during the pandemic. Considering that forecasting inflation is more important during times of duress than normal times, researchers may need to continue developing models that can perform well during all periods or develop a different set of models specifically for times of duress.

Appendix

Model Specifications

All models use four lags, and all the samples start around 1960, with some slight variation due to data availability. For the Primiceri (2005) we use PCE inflation, the three-year Treasury yield constant maturity, and the unemployment rate as variables and obtain data from the FRED series PCEPI, DGS3, and UNRATE, respectively. We use the three-year Treasury yield to avoid issues with the zero lower bound (see Swanson and Williams 2014). For the Primiceri (2005) model, we use a Minnesota prior with code derived from Chan (2021). For the Stock and Watson (2007) model, we use the non-centered parameterization and priors from Chan (2018) and use PCE inflation data from FRED (PCEPI). For the Carriero and others (2021) model, we obtain the input data from FRED-MD, a monthly macroeconomic database. The code from Carriero and others constructs quarterly averages based on this monthly data series, and we use the "SVOt" specification to run the model. For the Chan (2021) and Korobilis and Koop (2020) models, we use the same priors and variables as in the papers and obtain the input data from FRED-QD, a quarterly macroeconomics database.

Endnotes

[1]For example, Granger (2008) shows that time-varying parameter models can even approximate nonlinearities in general (in the conditional mean).

[2]Advancements in algorithms and computing power were necessary for variable selection due to the sheer number of variables considered in this method. For example, if a researcher wanted to consider p different variables in their inflation forecasting model, then they would need to consider 2^p different model combinations with those predictions. Thus, if $p = 20$, the researcher would need to estimate and compare the performance of 1,048,576 different models. Estimating and comparing all those models would be impractical, so methods were developed to allow researchers to estimate a small number of models and decide which model to estimate next based on the forecast performance of the previously estimated models.

[3]In the Carriero and others (2021) model, only the covariance matrix is time-varying.

References

Ang, Andrew, and Geert Bekaert. 2002. "Regime Switches in Interest Rates." *Journal of Business and Economic Statistics*, vol. 20, no. 2, pp. 163–182. Available at https://doi.org/10.1198/073500102317351930

Bernanke, Ben S. 2015. "The Taylor Tule: A Benchmark for Monetary Policy?" Brookings, April 28.

Carriero, Andrea, Todd E. Clark, Massimiliano Marcellino, and Elmar Mertens. 2021. "Addressing COVID-19 Outliers in BVARs with Stochastic Volatility." Federal Reserve Bank of Cleveland, working paper no. 21-02R, August. Available at https://doi.org/10.26509/frbc-wp-202102r

Chan, Joshua C. C. 2021. "Minnesota-Type Adaptive Hierarchical Priors for Large Bayesian VARs." *International Journal of Central Banking*, vol. 37, no. 3, pp. 1212–1226. Available at https://doi.org/10.1016/j.ijforecast.2021.01.002

———. 2018. "Specification Tests for Time-Varying Parameter Models with Stochastic Volatility." *Econometric Reviews*, vol. 37, no. 8, pp. 807–823. Available at https://doi.org/10.1080/07474938.2016.1167948

Cogley, Timothy, and Thomas J. Sargent. 2005. "Drifts and Volatilities: Monetary Policies and Outcomes in the Post-WWII U.S." *Review of Economic Dynamics*, vol. 8, no. 2, pp. 262–302. Available at https://doi.org/10.1016/j.red.2004.10.009

Faust, Jon, and Jonathan H. Wright. 2013. "Forecasting Inflation," in Graham Elliott and Allan Timmermann, eds., *Handbook of Forecasting: Volume 2*, pp. 2–56. Amsterdam: Elsevier.

Granger, Clive W. J. 2008. "Non-Linear Models: Where Do We Go Next—Time Varying Parameter Models?" *Studies in Nonlinear Dynamics Econometrics*, vol. 12, no. 3. Available at https://doi.org/10.2202/1558-3708.1639

Hoeting, Jennifer A., David Madigan, Adrian E. Raftery, and Chris T. Volinsky. 1999. "Bayesian Model Averaging: A Tutorial." *Statistical Science*, vol. 14, no. 4, pp. 382–417. Available at https://doi.org/10.1214/ss/1009212519

Korobilis, Dimitris, and Gary Koop. 2020. "Bayesian Dynamic Variable Selection in High Dimensions." Munich Personal RePEc Archive, paper no. 100164, May 5.

Primiceri, Giorgio E. 2005. "Time Varying Structural Vector Autoregressions and Monetary Policy." *Review of Economics Studies*, vol. 72, no. 3, pp. 821–852. Available at https://doi.org/10.1111/j.1467-937X.2005.00353.x

Stock, James H., and Mark W. Watson. 2007. "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, vol. 39, no. s1, pp. 3–33. Available at https://doi.org/10.1111/j.1538-4616.2007.00014.x

———. 1996. "Evidence on Structural Instability in Macroeconomic Time Series Relations." *Journal of Business and Economic Statistics*, vol. 14, no. 1, pp. 11–30. Available at https://doi.org/10.1080/07350015.1996.10524626

Swanson, Eric T., and John C. Williams. 2014. "Measuring the Effect of the Zero Lower Bound on Medium- and Longer-Term Interest Rates." *American Economic Review*, vol. 104, no. 10, pp. 3154–3185. Available at https://doi.org/10.1257/aer.104.10.3154