Abstract
We survey approaches to macroeconomic forecasting during the COVID-19 pandemic. Due to the unprecedented nature of the episode, there was greater dependence on information outside the econometric model, captured through either adjustments to the model or additional data. The transparency and flexibility of assumptions were especially important for interpreting real-time forecasts and updating forecasts as new data were observed. We revisit these themes with a time-varying parameter vector autoregression, which attributes the large jumps primarily to increased volatility rather than changes in the type or propagation of shocks.

Key Words: Macroeconomic Forecasting, COVID-19

1 Introduction
The COVID-19 pandemic and the ensuing recession has posed challenges for forecasters. The size of the fluctuations has been unprecedented, with GDP declining by 9% in 2020Q2, compared to a maximum quarterly decline of 2% in the Great Recession. In addition, the nature of the shock is different than past recessions. Economic activity declined due to social distancing measures rather than, for instance, financial shocks.

In the presence of large and atypical fluctuations in the economy, the challenge for the econometrician is to construct informative forecasts in the absence of data on the new circumstances of the pandemic. There is thus a particularly important role for sources of information outside what one might typically use. The first source is subjective judgement or prior knowledge, typically from economic theory, which is necessary to construct an appropriate model and discipline parameter estimates. For example, one needs to decide whether to model the pandemic as a large realization of a typical shock to the economy or as a new shock that propagates differently from previous shocks.
in the data, and in each case how to model the outlier or unusual propagation. Alternatively, one can consider external sources of data. For instance, epidemiological data was of special interest during the pandemic. However, the value of the additional data depends on knowledge of its joint distribution with the variables of interest, which was not always easy to determine initially.

This essay discusses the literature on forecasting during the COVID-19 pandemic. We divide that literature into the two categories described above: modifications to the model (e.g., adding time-varying volatility or parameters more generally) and the inclusion of additional data (including ways of adjusting forecasts based on past experience). Both approaches find mixed success. Overall, we find ex post that the failures in forecasting reflected differences in the COVID-19 recession that were already discussed at the time the forecasting papers were written, leaving the open question of how one might have formally incorporated these possibilities into the models.

We begin in Section 2 by discussing broad issues and laying out the vector autoregression (VAR) framework that build our discussion around. Next, Section 3 discusses the different forecasting approaches in the literature. We highlight assumptions underlying each approach and discuss how knowledge of the pandemic could have been used to interpret or even improve these forecasts. Finally, in Section 4, we revisit the themes by comparing forecasts in three VARs with varying degrees of time variation in parameters using a single data set and discuss how our forecasts and estimates relate to the various forecasting approaches in the literature.

We limit our scope in several aspects. First, we do not discuss the forecasting of COVID-19 cases in detail. A completely separate article would be required to do justice to the huge literature on the topic (e.g., Bertozzi et al. (2020); Ioannidis et al. (2020)). Next, we focus on the effect of the pandemic on economic variables and do not discuss how the economy may affect the evolution of the pandemic, an issue incorporated into theoretical models (e.g., Eichenbaum et al. (2021); Krueger et al. (2021)). Third, we center our discussion on U.S. data, even though some of the papers we cite do consider data from other countries. Finally, while the economic consequences of the pandemic continue to be felt and there remains work to be done to understand how forecasts could have been improved both in the initial months and through the recovery, we focus on the body of work that arose during the first year of the pandemic as economists developed methods almost in real time to deal with the unprecedented circumstances.

2 The Forecasting Problem

2.1 Adapting Economic Models to Unprecedented Times

In adapting our models to allow for the changes in the economy, one needs to acknowledge the difference in circumstances without going to the extreme of completely ignoring what has been learnt from past experience. In addition, the particular circumstances can make new data either unusually informative or misleading, depending on whether one is able to appropriately discern how the different variables evolve jointly.
Economic Knowledge. The econometrician needs to translate her interpretation of the data’s large fluctuations into model assumptions. This requires subjective judgement informed by economic theory or knowledge outside the model. Is the variation in data explained by large shocks, new shocks, or a different propagation of shocks? These questions have also been asked in attempts to account for previous changes in the behavior of macroeconomic data (e.g., Cogley and Sargent (2005); Sims and Zha (2006); Stock and Watson (2012)).

Assumptions in the model should be easily communicated. Transparent assumptions allow the audience to understand forecasts within the context of the model. For example, if the model assumes an unchanged persistence but it is understood that the current recession is likely to be less persistent than past experience, then one can take the model forecast as providing an overestimate of the depth of the recession in the medium run.

In addition, where possible, assumptions should be imposed probabilistically. For example, even if one believed that past experience was a good benchmark for how the COVID-19 recession may play out, one can impose a prior that the model parameters are “probably close”—rather than identical—to what they were before the pandemic. This allows us to update our parameter estimates as we observe new data. At the time of writing, we are still in the middle of the pandemic, but have over a year of data on the evolution of both economic and epidemiological variables. Even with decades of data, artificially tight prior assumptions can bias model predictions in ways that are often hard to detect (e.g., Canova and Sala (2009); Giannone et al. (2015); Ho (2020)), a problem that is exacerbated when data are relatively scarce. Moreover, these assumptions can give a false sense of precision in forecasts.

The above features are especially important in the context of the COVID-19 pandemic. Given disagreement among economists and policymakers, clear communication is crucial for making quantitative model forecasts useful beyond the narrow audience that fully agrees with the model assumptions. Since individual econometricians were themselves probably less certain about the new structure of the economy, the probabilistic presentation of assumptions results in forecast error bands that more accurately reflect the econometrician’s own degree of confidence. Practically speaking, this may involve averaging across models (Timmermann (2006)) or, from the Bayesian perspective, imposing priors over the parts of the model capturing the changes in the economy.

Additional Data. The COVID-19 pandemic was a striking example of when new data sources became especially useful. The public health dimension of the economic fallout made data on COVID-19 cases, hospitalizations, and deaths important. In addition, the fast-moving nature of the crisis brought a new demand for high frequency data to inform economists and policymakers of the state of the economy before the usual quarterly or even monthly data were published. Beyond raw data, forecasts of the pandemic were and continue to be of interest (e.g., Ho et al. (2020); Atkeson et al. (2021); Li and Linton (2021); Liu et al. (2021)).

The question of how to integrate these data into model forecasts of the economy returns us to the question of model specification. Assumptions need to be made about the relationship between
the additional data series and the variables one was looking at previously. If these assumptions are misguided, incorporating the additional data can lead one to misleading conclusions about the current and future state of the economy.

2.2 Formal Framework

Model. To formally organize our discussion, we consider the VAR:

\[ X_t = BX_{t-1} + C \varepsilon_t, \]

where \( X_t \) is a vector of current and lagged macroeconomic variables, and \( \varepsilon_t \sim \mathcal{N}(0, I) \) is a vector of orthogonal shocks normalized to each have variance one. The parameter \( C \) determines how each shock affects the economy on impact, and the parameter \( B \) captures how these shocks propagate through the economy. We can include more lags into (1) by writing the system in companion form.

Many forecasting models can be cast as special cases or extensions of this. For example, dynamic stochastic general equilibrium (DSGE) models and factor models take the form (1) with restrictions on \( B \) and \( C \) that arise from either economic or statistical restrictions derived from the respective models. The time-varying parameter (TVP) VAR (e.g., Primiceri (2005); Lubik and Matthes (2015)) takes the form (1) but allows \( B \) and \( C \) to vary over time. More involved methods, especially from the field of machine learning, have been used for macroeconomic forecasting before and during the COVID-19 recession (e.g., Coulombe et al. (2021); Huber et al. (2020)). While we do not provide a detailed description of these methods here, we note that these methods introduce a high degree of flexibility for the variables \( X_t \) to evolve differently during contrasting episodes, which can be thought of as less parametric approaches to the ideas in Section 3.1.\(^1\)

Forecasting. Suppose all the parameters are known. Then the optimal forecast \( h \) periods ahead is \( B^h X_t \). One can also condition on a future path for a subset of variables in \( X_t \) to derive forecasts for the rest of \( X_t \).

Uncertainty in the forecast arises from two sources. First, there is uncertainty about future \( \varepsilon_t \). The coefficient \( C \) directly scales the level of this uncertainty by changing the variance of the errors \( C \varepsilon_t \). The coefficient \( B \) also impacts uncertainty by influencing the propagation of the shocks \( \varepsilon_t \). For example, if \( B \) implies a very persistent system, then shocks today can have a lasting impact on macroeconomic outcomes far in the future. As a result, the value of \( X_t \) not only depends on the contemporaneous shock, but is also heavily driven by past shocks, thus increasing the variance. Second, there is uncertainty about the parameter values for \( B \) and \( C \). Even if one knew all future realizations of \( \varepsilon_t \), the way these impact \( X_t \) depends on the parameter values.

\(^1\)While Coulombe et al. (2021) find that machine learning methods improved forecasting performance over certain benchmark linear models similar to (1) during the pandemic, it is not clear how these methods compare to parametric extensions to (1), such as those discussed in Section 3.1. In addition, machine learning methods and nonlinear models more generally have not been found to have systematic benefits over linear models before the pandemic (see Ferrara et al. (2015) and Coulombe et al. (2021) for details).
Allowing $B$ and $C$ to vary over time, as the literature has done for both pre- and post-COVID periods, adds an additional layer to the estimation and forecasting process. For estimation, less weight will be placed on periods with large shocks (captured by $C$). We will see that this can influence both the point estimate and the variance of $B$, which in turn affects the point forecasts and the uncertainty around them. Time variation in $B$ gives the model greater flexibility to match the data. This flexibility can reduce shock uncertainty by attributing variation in the data to changes in $B$ rather than shocks $C \varepsilon_t$. However, it can simultaneously introduce greater uncertainty about the value of $B$ both in the current period and in the future.

**Unusual Times.** The challenge in uncertain times such as the COVID-19 pandemic, is that the recent parameters are probably not the ones governing the evolution of the data $X_t$ in the near future. The large share of temporary layoffs made the initial shock to unemployment less persistent than in the past (i.e., $B$ was closer to 0). The speed with which social distancing measures were adopted (both by policy and individual decisions) resulted in larger and more sudden changes in the economy (i.e., a larger scale of $C$). The fact that the public health outcomes started playing a role in business cycle fluctuations that they never had in recent history suggests that there may be new data series or shocks to consider (i.e., additional relevant series in $X_t$ or new shocks in $\varepsilon_t$).

The challenge of adapting forecasting models to an evolving economic environment is not new (see, e.g., Clements and Hendry (2008)), but the speed and size of the pandemic’s impact dwarfed similar recent episodes. As a result, macroeconomic forecasters had to be judicious in the amount of structure they imposed in their models. As discussed by Galvao (2021), the adaptability was key for models to fit the evolving data, but excessive adaptability led to overfitting and inaccurate forecasts. From a Bayesian perspective, one could view the looseness of the prior as capturing the degree of adaptability. Whereas one might typically rely on the data to update our estimates, the fast-moving nature of the pandemic required forecasts before sufficient data had been observed to reliably update the parameters, i.e. the data had not yet swamped the prior. On the other hand, the uniqueness of the crisis increases uncertainty about what is known before seeing data, i.e. an excessively tight prior might understate uncertainty. For instance, while there is a rich literature of DSGE models studying typical recessions, quantitative structural models of the public health dimension of the COVID-19 crisis had not been widely studied until the onset of the pandemic.\(^2\) The literature has grappled real-time with this—not only is there qualitative disagreement about economic events, but the high dimensionality of many forecasting models presents challenges to expressing qualitative views and knowledge quantitatively.\(^3\)

---

\(^2\)The pandemic triggered an explosion of papers connecting macroeconomic and epidemiological outcomes (e.g., Acemoglu et al. (2020); Alvarez et al. (2021); Eichenbaum et al. (2020)). However, there remains work to be done to explicitly connect these theories with statistical forecasting models.

\(^3\)For instance, in the model (1), if $X_t$ contained 12 lags of 6 variables, which is typical with monthly data, $B$ would have 432 elements to be estimated.
3 Forecasting Approaches During the COVID-19 Pandemic

3.1 Models

We first focus on extensions to (1). We discuss how the results may have been influenced by particular assumptions that each paper makes about how the COVID-19 economy differs from the pre-pandemic economy.

3.1.1 Increasing Volatility

Lenza and Primiceri (2021) interpret the volatility in the economy during the pandemic as arising from a sequence of large shocks. In particular, they replace (1) with

$$X_t = BX_{t-1} + s_t C \varepsilon_t,$$

where the additional variable $s_t$ determines how the variance of the shock varies over time. They assume $s_t = 1$ before the pandemic, estimate $s_t$ for March, April, and May 2020, and assume a gradual (autoregressive) decay back to one subsequently. The idea of incorporating time-varying volatility in models is not new (e.g., Engle (1982); Jacquier et al. (2004); Stock and Watson (2016)). In the forecasting context, Clark (2011) shows that including time-varying volatility improves macroeconomic forecasting performance. There are benefits of incorporating the change in variance captured in (2) for both the estimation and forecasting steps.

For estimation, acknowledging that recent data is subject to a large amount of noise ensures that the estimates for $B$ and $C$ are not overly impacted by the small number of outliers. Indeed, Lenza and Primiceri (2021) find that when they estimate (1) instead of (2) using data that includes the COVID-19 periods, the model produces unreasonable forecasts with explosive responses to shocks that can predict that the recession continually deepen without any end, emphasizing that the modeling of variances affects not just forecast uncertainty but also point forecasts.

Using a similar VAR model, Bobeica and Hartwig (2021) propose $t$-distributed instead of Gaussian errors as an alternative to the use of $s_t$ by Lenza and Primiceri (2021) to account for the early pandemic outliers. Their model imposes less structure on the evolution of volatility than Lenza and Primiceri (2021) and postulates that large shocks are more likely to happen not only during the pandemic but in any period. In addition, they show that a sufficiently informative prior can help produce more stable estimates. On the other hand, Hartwig (2021) argues that if one uses a prior that is scaled using the data, as is common practice for the frequently used Minnesota prior (Litterman (1986)), the COVID-19 observations can lead to a tight prior on real variables but loose prior on price variables, reducing the in-sample fit. Beyond the VAR framework, Huber et al. (2020) provide an alternative approach to dealing with outliers using a nonparametric model. Their approach recognizes the increased uncertainty when the data look different than in the past, as was the case in 2020Q2 and 2020Q3. While Lenza and Primiceri (2021) attribute the difference to a large shocks, Huber et al. (2020) allow for a change in parameters.
For forecasting, since a large part of forecast uncertainty arises from the possibility of future disturbances, acknowledging the increased variance of shocks captures the increased uncertainty in forecasts. Carriero et al. (2021a) show that assumptions about the time-varying volatility $s_t$ matter for forecasts. They distinguish between allowing $s_t$ to fluctuate gradually (as a random walk both before and after the COVID-19 pandemic) and allowing for occasional one-off jumps in $s_t$. Because the March and April 2020 data were such huge outliers, allowing for occasional one-off jumps in the shock variances tightens forecast error bands, arguably making them more reasonable by allowing volatility to quickly return to closer to historical levels after the initial sharp adjustments to the pandemic. In related work, Carriero et al. (2021b) estimate a time-varying volatility VAR using data through June 2020 and find a substantial role for such one-off increases in volatility in a number of variables, including employment and industrial production (IP).

There are two attractive features of the Lenza and Primiceri (2021) approach. First, the assumptions are transparent to communicate. Even if one does not believe the exact model, the forecasts can be informative if one has a clear idea of the direction in which they are biased. Second, estimation is a straightforward application of generalized least squares.

There are two implicit assumptions that are potentially problematic:

1. Since $B$ remains unchanged, the effect of a one time shock $\varepsilon_t$ propagates through the economy the same way before and during the pandemic.

2. The variance of all shocks increases proportionally.

In other words, (2) treats the pandemic as merely a period with unusually high volatility even though there was an understanding early on in the pandemic that the COVID-19 recession would differ from previous recessions in terms of job losses, revenue losses, and consumer behavior. As a result, both the propagation of and the relative variance of shocks likely differed from the pre-pandemic data. These assumptions are dropped, for instance, in the TVP-VAR of Primiceri (2005) or Lubik and Matthes (2015).

Lenza and Primiceri (2021) present forecasts conditional on actual unemployment data through May 2021 and Blue Chip consensus forecasts for unemployment in subsequent months. They compare the forecasts from (2) to an estimation without accounting for the change in variance (using only pre-pandemic data for the estimation of the parameters). The conditional forecasts from June 2020, shown in Figure 1, show that the inclusion of the time-varying variance substantially widens the forecast error bands, as the model predicts a future sequence of large shocks (i.e., large future $s_t$). In contrast, when a fixed variance is assumed, the forecast error bands are much narrower, with realized data falling outside or on the edge of the 95% credible region. Figure 2 shows that the time-varying variance no longer has a significant effect on the size of the error bands when the forecast is made using data through May 2021. By this time, the estimated variance is closer to normal and parameter uncertainty is the main source of forecast uncertainty.

While the approach of Lenza and Primiceri (2021) helps account for uncertainty, it leaves the point forecasts relatively unchanged, reflecting a relatively unchanged coefficient $B$. Even though
Figure 1: Figure 3 from Lenza and Primiceri (2021). Forecasts as of June 2020 conditional on path for unemployment (actual realizations until May 2021, Blue Chip June 2021 release consensus forecasts subsequently), 68% and 95% bands. Crosses are data realizations. **Left panel:** Estimation with change in volatility using data through June 2020; **Right panel:** Estimation with constant volatility using data through February 2020.
Figure 2: Figure 4 from Lenza and Primiceri (2021). Forecasts as of May 2021 conditional on path for unemployment (actual realizations until May 2021, Blue Chip June 2021 release consensus forecasts subsequently), 68% and 95% bands.
data from the COVID-19 period can potentially alter the estimate of $B$, the estimation places very low weight on these data due to their large variance, captured by $s_t$. For the conditional forecasts in Figures 1 and 2, this turns out to be a reasonable assumption, as the distribution of employment, consumption, and prices conditional on unemployment did not change too much. We will show later that this was not true for other variables. In addition, the assumption of a constant $B$ would have likely resulted in misleading unconditional forecasts. The highly persistent response of the economy to an unemployment shock, shown in Figure 3, suggests that the model would not have predicted the relatively rapid decline in unemployment shortly after the start of the pandemic. The response is consistent with the historical behavior of employment but not necessarily the specific circumstances of the COVID-19 recession.
3.1.2 Introducing a New Shock

Primiceri and Tambalotti (2020) provide a framework that acknowledges that the COVID-19 shock is different than previous shocks, weakening the assumptions in Lenza and Primiceri (2021). We present a simplified version of their approach using the VAR model (1). Noting that (1) can be rewritten as

\[ X_t = \sum_{s=0}^{\infty} B^s C \xi_{t-s} + \sum_{s=0}^{\infty} G^s H \circ v_{t-s}, \]

where \( v_t \) is the COVID-19 shock that takes the value zero before March 2020 and \( \circ \) denotes element-wise multiplication.

At its core, the above procedure is forecasting conditional on an unobserved variable \( v_t \). The path of shocks is hypothetical but can in principle be derived from external data such as mobility measures or case numbers, as is the case in Ng (2021). How successful and reasonable the forecasts are depend on the assumptions made about the path of the COVID-19 shock and its propagation through the economy. Primiceri and Tambalotti (2020) consider a version of the model:

\[ X_t = \sum_{s=0}^{\infty} B^s C \xi_{t-s} + \sum_{s=0}^{\infty} G^s H \circ v_{t-s}, \]

where \( v_t \) is the COVID-19 shock that takes the value zero before March 2020 and \( \circ \) denotes element-wise multiplication.\(^4\)

Assumption 1 is a reasonable approximation ex ante because COVID-19 has overshadowed all other typical variation in the economy. It is the typical assumption made in event studies.

Assumptions 2 and 3 are more controversial. Both assumptions treat the COVID-19 shock as typical even though the nature of the COVID-19 shock to the economy is vastly different from any recent experience. Moreover, the rigidity of the assumptions imply that incoming data have limited capacity to revise model estimates. Nevertheless, completely discarding past experience would result in overly uninformative forecasts. Even though the shock is unusual, the structure of the economy does not completely change overnight.

Forecasters should acknowledge the unprecedented situation while allowing the model to be informed by data. For the propagation of shocks, we consider the pre-pandemic value of \( B \) is a good benchmark but expect the economy to recover more rapidly than in previous recessions. One alternative would be to impose this probabilistically, replacing Assumption 2, so that we do not force our beliefs on the data but instead allow the forecasts to combine our beliefs with the available data. For example, one could allow \( G = B + \eta \), where \( \eta \) is a parameter to be estimated. The prior

\(^4\)More precisely, Primiceri and Tambalotti (2020) take \( v_t = \sum_{s=0}^{\infty} r_s u_{t-s} \), where \( u_t \) is a scalar that takes value one in March 2020 and zero for all other months.
on $\eta$ then captures the degree of confidence that $G$ is close to $B$. For the evolution of the shock, epidemiological model forecasts can replace Assumption 3 by providing informed predictions about the future path of the pandemic and uncertainty about that path.\footnote{For example, Meza (2020) implements the Primiceri and Tambalotti (2020) methodology for Mexican data, conditioning forecasts on the path of the pandemic predicted by a simple epidemiological model.}

Through the lens of a TVP-VAR, the new shock in Primiceri and Tambalotti (2020) can be captured through a change in parameters. Starting with an initial belief about the parameter values that is centered at the pre-pandemic estimates, the TVP-VAR allows the data to inform the estimates of a potentially new set of parameters. The model has the flexibility to allow for changes in comovements, persistence, and variances.

Figure 4 shows that Primiceri and Tambalotti (2020) forecast a more persistent decline in economic activity than occurred. Their forecast conditions on $v_t$ being close to zero from around

Figure 4: Figure 2 from Primiceri and Tambalotti (2020). Forecasts from April 2020 under baseline scenario for $v_t$, 68% and 95% bands.
January 2021, which has turned out to be an overly optimistic prediction. Nevertheless, even the 95% forecast error bands do not include the rapid rebound in unemployment and consumption that we have seen since the early weeks of the pandemic. Allowing for the widely discussed possibility of the COVID-19 shock having a more transitory effect than previous recessions (i.e., $G$ closer to 0 than $B$) would have widened the error bands, providing a quantitative forecast that more accurately reflected the debate at the time of the forecast and one that is ex post more accurate.

An alternative approach to identify the COVID-19 shock is proposed by Chudik et al. (2021), who attribute revisions in the International Monetary Fund (IMF) 2020Q1:2020Q4 GDP projections after the start of the pandemic solely to the shock, comparing projections in April 2020 to those made at end of 2019. Using a version of (1) that allows for a change in intercept during periods of high volatility, they estimate the impact of the COVID-19 shock on real GDP in a panel of countries. Their forecasts for the U.S. perform well, with the data realizations falling within the 90% error bands. It is hard to interpret the forecasts or precisely attribute the sources of success because the median forecasts are forced to match the IMF projections that are formed outside the model. Nevertheless, the results do suggest that by April 2020, economic forecasters had the data needed to forecast GDP relatively accurately, a theme that we elaborate on below.

3.2 Data

Whereas the approaches above keep the same variables but modify the model, we now discuss approaches to utilize recent and past data to improve forecasts. We emphasize that the ability to leverage information in the data depends on valid assumptions about the comovement of variables. In addition, we argue that from the perspective of the currently available data vintages, it is unlikely that data revisions had a significant role in the overall performance of forecasts.

3.2.1 High Frequency Data

Given how rapidly the effects of the pandemic were felt and how swiftly policymakers have had to respond, it has been crucial to incorporate high frequency data to ensure that we are using all available information for our forecasts. To that end, several papers have used mixed-frequency (MF) models for forecasting. In the context of (1), these papers allow for the econometrician to only observe $X_t$ or averages of $X_t$ in particular periods, which provides the possibility of incorporating data of different frequencies in $X_t$ (e.g., monthly IP and quarterly GDP). Intuitively, the joint behavior of $X_t$ allows us to infer the values of unobserved elements of $X_t$. During the pandemic, it was reasonable to expect that these joint distributions might have evolved (i.e., $B$ and $C$ in equation (1) changed). Without accounting for these changes, one may draw misleading conclusions from the high frequency data.

An illustration of the role of high frequency data in informing forecasts is provided by Schorfheide and Song (2020), who present real-time macroeconomic forecasts during the pandemic using a relatively standard model from Schorfheide and Song (2015). They include three quarterly series (GDP, investment, and government expenditures) and eight monthly series (unemployment rate,
hours worked, Consumer Price Index, IP, Personal Consumption Expenditure, Federal Funds Rate, 10-year Treasury Bond Yield, and S&P 500 Index). For the results we will discuss, the underlying parameters $B$ and $C$ are estimated using data available on January 31, 2020, and forecasts are updated purely from new information about the realized path of the data rather than changes in parameter values.\footnote{When the data after January 2020 were used for estimation, Schorfheide and Song (2020) run into the challenge highlighted by Lenza and Primiceri (2021), as the pandemic era outliers result in unreasonable parameter estimates, leading to poor forecasts.}

Figure 5 shows that we find mixed success comparing the realizations of the 2020Q2 data to what is forecast by the model using data available on June 30, 2020. By this time, the quarter had come to an end, but the GDP numbers had not yet been announced. GDP fell to 10% below the 2019Q4 level, close to the Schorfheide and Song (2020) median forecast of 12% and within the 60% predictive bands. On the other hand, 2020Q2 investment fell 9% relative to 2019Q4, substantially less than the June 30 median forecast of 22% and even the 95% quantile of 16%. Note that the quantiles already capture the inherent relative volatility of the respective series, with the error bands for investment double the width of those for GDP.

The fact that the GDP nowcast (i.e., the forecast of current quarter GDP before it has been reported) performs so well relative to the investment nowcast suggests a change in the relationship between the monthly and quarterly variables. This is a similar issue to that raised in our discussion of Lenza and Primiceri (2021) and Primiceri and Tambalotti (2020)—there were features of the data during the pandemic that we might have expected ex ante to have changed, but the model forecasts did not take these into account. Economic theory suggests that investment should fall less for more transitory shocks. By June 30, the stock market also suggested that the recession might be short-lived, as the S&P500 index had already recovered a large fraction of its losses after reaching its trough by mid-March, just one month after the initial decline in February 2020. In contrast, during the Great Recession the S&P 500 only reached its lowest point in February 2009, more than a year after the start of the recession in December 2007. Given the likelihood that the COVID-19 recession would be more short-lived than past recessions, one might have expected that the response of investment relative to the movement in the monthly series could be muted.

The forecasts updated dramatically as monthly data was observed during the initial decline of economic activity between April to June. On April 30, May 31, and June 30, Schorfheide and Song (2020) forecast 2020Q2 GDP of 4%, 15%, and 12% below the 2019Q4 level, respectively. These updates are large not only relative to past data, but also relative to the error bands of the forecasts. Moreover, none of these updates are arising from changes in the underlying parameters. To justify such large updates, one needs to interpret the shocks in May and June as being extremely low probability events given the April 30 data. Given the uniqueness of the period, the plausibility of such an interpretation depends on subjective judgement.

A natural extension is to include time variation in the parameters of the Schorfheide and Song (2020) MF-VAR model. Koop et al. (2021) study a similar model to Schorfheide and Song (2020) and ask how allowing for different forms of time-varying volatility, in the spirit of Lenza and Primiceri (2021).
Figure 5: Figure 3 from Schorfheide and Song (2020). Forecasts from January 31, April 30, May 31, and June 30, 2020, realized data (solid red) and median forecast (solid black) with 60% and 90% bands. Solid blue line represents point forecasts obtained by fixing the federal funds rate at 5 basis points. All variables are plotted relative to their 2019Q4 levels.
Primiceri (2021), affects nowcasts for GDP. While the point nowcasts do not differ substantially, time-varying volatility leads to larger nowcast error bands. In addition, imposing time-invariant variances leads to a more rapid update of the nowcasts as high frequency data are observed. Given the short sample, evaluating which of these nowcasts are preferred requires subjective judgement. Götz and Hauzenberger (2021) finds that including time-varying intercepts has a relatively small effect on forecasts once time-varying volatility is included. The information value of high frequency data is further emphasized by Berger et al. (2020), who find that a constant coefficient MF-VAR similar to Schorfheide and Song (2020) produces nowcasts of the output gap given April data that are consistent with ex-post estimates from a VAR, the Congressional Budget Office, and the Hodrick-Prescott filter.

Antolin-Diaz et al. (2021) incorporate high frequency data into a version of the factor model:

\[ X_t = c + H f_t + u_t \]  

where \( X_t \) is a large panel of monthly data (potentially with missing observations differing frequencies or sample sizes) and \( f_t \) is an unobserved factor that captures their comovement, representing an index of economic activity. They allow for time-varying parameters and outliers in (4) and produce GDP nowcasts. Including high frequency data of up to daily frequency results in more accurate nowcasts in the initial weeks of the pandemic.\(^7\) In 2020Q1, the GDP nowcast dips below zero by the middle of March 2020 (shortly after the declaration of a National Emergency) to match the actual decline of 5%. In contrast, Carriero et al. (2020) use weekly data to produce GDP nowcasts that are declining but above zero at the end of the quarter. These results emphasize how rapidly the economy evolved in the initial weeks of the pandemic, with substantial information flowing in at the daily frequency. In 2020Q2, Antolin-Diaz et al. (2021) find that the gap between the nowcasts with and without the high frequency data shrinks substantially following the release of the monthly IP and retail sales data on April 15, 2020, consistent with the accurate GDP nowcasts of Schorfheide and Song (2020).

3.2.2 Additional Sources of Data

An alternative to increasing the time series dimension through high frequency data is to focus on the cross section studying panel data. For example, Aaronson et al. (2020) and Larson and Sinclair (2021) find some success using the panel of U.S. states (together with data on Google Trends and state of emergency declarations) to improve nowcasts of unemployment insurance (UI) claims.

Raw data may not contain information that is known to forecasters or modelers, including predictions from theoretical models that may be difficult to formally incorporate into our model.

\(^7\)The high frequency data are aggregated to a monthly frequency. Many of these series start only in 2020, leading to a relatively short sample for estimation. To overcome the lack of data, the authors estimate linear relationship between each high frequency series with a traditional series that is available for a longer sample, with a tight prior on an intercept of zero and slope of one. The high frequency data are used to predict the values of the traditional series before their values are announced. The information is then incorporated into the GDP nowcasts through inclusion of the traditional series in the factor model (4).
specifications, estimates and forecasts. Bobeica and Hartwig (2021) incorporate this information by conditioning forecasts on projections from economic institutions and professional forecasters, taken from the European (System of) Central Banks Macroeconomic Projection Exercise and the Survey of Professional Forecasters, respectively. They find that information from these external projections help to make forecasts more well-behaved. The finding in Bobeica and Hartwig (2021) that incorporating external projections can improve forecasts provides support for the use of IMF projections by Chudik et al. (2021).

3.2.3 Learning from Past Crises

One response to the shortcomings in Schorfheide and Song (2020) is to ask if there are historical episodes that can suggest where our current estimated model might make errors. To that end, Foroni et al. (2020) adjust forecasts from various MF models using information from the Great Recession. The class of models used can be viewed as minor modifications of (1). In what follows, we reference our benchmark VAR model (1) instead of the actual models in Foroni et al. (2020) in order to economize on notation and simplify the exposition.

First, they compute similarity adjusted forecasts in which they increase the weight on observations from the Great Recession. The reweighting captures the idea that the COVID-19 economy is more similar to the Great Recession than the average time period. In the notation of (1), this changes the estimates of $B$ and $C$ to be closer to the values that best fit the Great Recession period.

Second, they produce intercept adjusted forecasts by correcting forecasts based on forecast errors made during the Great Recession. The correction accounts for systematic bias that may occur during recessions. Formally, rather than take $B^hX_t$ as the $h$-period-ahead forecast, they use $B^hX_t + u^*_t$, where $u^*_t$ is the forecast error from the corresponding period in the Great Recession.

Foroni et al. (2020) generate forecasts using data available in April 2020, which includes monthly observations for the beginning of the pandemic but only includes quarterly data through 2019Q4. They focus on quarterly forecasts of GDP and investment but include monthly data on IP, employment, or the VIX as additional predictors. They consider a range of models that differ based on the number of lags included and assumptions about the parameters.

The results for GDP and investment are shown in Figures 6 and 7, respectively. The forecast error correction substantially changes the forecasts but the reweighting of observations does not. However, with the benefit of hindsight, we see that the model substantially underpredicts the initial decline in GDP but overpredicts the persistence of the decline in GDP growth, with forecast annualized declines of approximately 10% and 5% in 2020Q2 and 2020Q3, respectively, in contrast to the 33% decline and 38% increase in the data. Foroni et al. (2020) forecast a decline in investment

---

8 More precisely, Foroni et al. (2020) consider $y_{t+h} = \sum_{k=0}^{K} \beta_k y_{t-k} + \sum_{l=0}^{L} \delta_l x_{t-l} + \sum_{m=0}^{M} \gamma_m u_{t-m}$, where $y_t$ is the variable of interest, $x_t$ contains additional predictors, and $u_t$ captures the forecast error. The variety of models involve different assumptions on $\beta_k$ and $\delta_l$ as well as different methods of estimation.

9 These results are consistent with Stock and Watson (2012) who, using a different model, find that the evolution of the macroeconomic variables during the Great Recession was due to large shocks rather than structural changes in the economy.
Figure 6: Figure 6 from Foroni et al. (2020). Forecasts for GDP, unadjusted (magenta), intercept adjusted (red), and similarity adjusted (blue). Solid lines average across models, dotted lines indicate individual models.
that had a trough level close to the data but overpredict the persistence of this decline. Their forecasts are roughly in line with the Schorfheide and Song (2020) April 2020 forecasts. Like Schorfheide and Song (2020), Foroni et al. (2020) overpredict the decline of investment relative to GDP, reflecting once again the differences in the COVID-19 recession compared with previous recessions.

Instead of using the Great Recession, Ludvigson et al. (2021) use data on past natural disasters to predict the economic effects of the COVID-19 pandemic. They also find mixed success depending on the variable, performing relatively well for IP but poorly for initial claims and service sector employment. Similarly, Aaronson et al. (2020) use the comovement between UI claims and Google Trends during recent hurricanes in the U.S. to forecast UI claims during the pandemic.

Even if one believed that the COVID-19 recession would evolve differently than the Great Recession, the the results from Foroni et al. (2020) could still be useful as bounds for forecasts if one had a prior about what these differences might be. For conditional forecasting, Foroni et al. (2020) also provide a way to generate plausible bounds on the paths of variables that we choose to condition our forecasts on. More generally, even though we may perceive clear differences between the current situation and previous recessions, knowing the direction of these differences can allow us to productively incorporate past experience to refine our forecasts.

### 3.2.4 Data Revisions

An additional aspect of the data is measurement error and data revisions. Mismeasured data can contaminate estimates and reduce the accuracy of forecasts (see, e.g., Croushore (2006, 2011)). Data measurement was a particular challenge early in the pandemic not only because of the size of the shocks and variation, but also because of the physical challenge of collecting some of the data.
Figure 8: IP and PCE price index: April 2020, June 2020, August 2020, and September 2021 vintages. **Left:** Indices in levels; **Right:** Annualized month-over-month growth rates.


Since our forecast evaluation has focused on published statistics, we evaluate the severity of these data disruptions by considering the size of data revisions over the course of the pandemic. Figure 8 plots April 2020, June 2020, August 2020, and September 2021 vintages of IP and the PCE price index for the U.S. in levels and annualized month-over-month growth rates. The revisions are non-trivial but remain an order of magnitude smaller than the fluctuations of the data and the forecast errors we have discussed. For example, IP growth and PCE inflation for March 2020 have been revised from \(-19.9\%\) and \(-3.2\%\) to \(-14.5\%\) and \(-2.9\%\), respectively, since the initial data releases in April 2020.

The size of the fluctuations relative to the revisions suggest that forecasting models need to be relatively sensitive to small differences in tail realizations for the data revisions to matter substantially for forecasts or the forecast errors. Nevertheless, as pointed out in Croushore (2006), the data may continue to be revised further in coming years. We have no guarantee that our conclusions

---

10 Each vintage contains data through the month before the vintage release date.
about the dependence of the forecasts on the data revisions or the size of the forecast errors will remain unchanged. In addition, we have assumed that the September 2021 vintage incorporates any necessary adjustments to account for pandemic-related data collection disruptions.

4 Perspective From a Time-Varying Parameter VAR

In Section 3, we discussed various approaches of forecasting in the literature and challenges faced by each of them. A recurring question is which parameters need to be allowed to vary over time in order to accommodate the large COVID-19 shock and help forecasts to adapt to the unusual circumstances of the pandemic.

4.1 Model

To dig deeper, we use Bayesian methods to estimate three versions of (1) using a common data set to make the results comparable across models. Following Primiceri (2005) and Lubik and Matthes (2015), we decompose $C$ as $C = \Lambda \Sigma$, where $\Lambda$ is a lower triangular matrix with ones on the diagonal and $\Sigma$ is diagonal, and consider models nested by:

$$X_t = B_t + \Lambda_t \Sigma_t \varepsilon_t$$

where $\tilde{B}_t, \tilde{\Lambda}_t$, and $\tilde{\Sigma}_t$ are vectors containing all variable elements of $B_t, \Sigma_t$, and $\Lambda_t$ respectively. The variances $\Omega^B, \Omega^\Lambda$, and $\Omega^\Sigma$ are estimated. The coefficients $B_t$ captures changes to the propagation of shocks, $\Lambda_t$ captures changes in the correlation between shocks, and $\Sigma_t$ captures the variances of these shocks. See Lubik and Matthes (2015) for further details.

The three models we consider are:

1. **Time-varying parameter VAR (TVP-VAR):** $B, \Lambda$, and $\Sigma$ vary over time.

2. **Stochastic volatility VAR (SV-VAR):** Only $\Sigma$ varies over time ($\Omega^B = 0$ and $\Omega^\Lambda = 0$).

3. **Constant coefficient VAR (VAR):** $B, \Lambda$, and $\Sigma$ are constant over time ($\Omega^B = 0, \Omega^\Lambda = 0$, and $\Omega^\Sigma = 0$).

We use quarterly data from 1975Q1 through 2021Q2 for year-over-year GDP growth, unemployment, year-over-year PCE inflation, year-over-year core PCE inflation, and the federal funds rate. Given the finding in Figure 8 that the fluctuations in the data swamp the data revisions, we focus on the most recently available vintage of data.
4.2 Results

Figure 9 plots the three models’ out-of-sample forecasts for 2020Q1, 2020Q2, and 2020Q3, with 95% error bands. Despite the flexibility of the TVP-VAR, we find mixed success as its forecasts suffer some of the same issues the approaches in the literature face. Nevertheless, the TVP-VAR produces more accurate forecasts than both the SV-VAR and constant coefficient VAR over the first year of the pandemic.

As we found in both Schorfheide and Song (2020) and Foroni et al. (2020), the forecasting accuracy varies substantially across variables. The realized data for GDP growth lies only just outside the 95% error bands and realized inflation lies within the 95% error bands of the TVP-VAR 2020Q2 forecasts. However, the model projects that unemployment will increase or remain elevated with high probability, in line with historical experience, in stark contrast to the realized 4.3% decline in 2020Q3. This decline was seven times the size of the next largest quarterly decline in unemployment in the sample. We can now attribute the rapid rise and fall in unemployment to factors such as the large number of short-term layoffs (Cajner et al. (2020); Gallant et al. (2020); Hall and Kudlyak (2020)) and decreased labor force participation (Coibion et al. (2020)).
Figure 10: Time variation in coefficients, correlations, and variances. **Top:** Change in coefficients ($\| \hat{B}_t - \hat{B}_{t-1} \|_2$); **Middle:** Change in correlations ($\| \hat{\Lambda}_t - \hat{\Lambda}_{t-1} \|_2$); **Bottom:** Level of volatility ($\frac{1}{N} \sum_{i=1}^{N} \log \hat{\sigma}_{i,t}^2$).

It is important to note, however, that these possibilities were considered as early as May 2020 by Petrosky-Nadeau and Valletta (2020), who produce a range of projections based on a theoretical model of employment flows, some of which include steep declines in 2020Q3. This is an example in which the theoretical literature provided forecasters information that could have been incorporated real time either in the forecasting process or in the interpretation of the model forecasts. Despite the challenges faced by the model in 2020Q2, the forecasts quickly adjust in 2020Q3 to a path that is much closer to the realized data.

The forecasts for the SV-VAR are similar to the TVP-VAR with slightly smaller error bands, but the constant coefficient VAR produces forecasts that are significantly further from the realized data. The poor performance of the constant coefficient VAR emphasizes the previous discussion
about how it is important to incorporate the possibility of large shocks in order for the model to deal with the large fluctuations at the start of the pandemic. However, the flexibility of the coefficients to vary over time does not have a big impact on the forecasts, echoing the findings in Götz and Hauzenberger (2021). These results suggest that to improve the forecasts, one would have to model the variation in the parameters more explicitly or impose further information through a prior on the coefficients around 2020Q2, although admittedly neither of these are straightforward.

Estimating the model using data through 2021Q2, we once again see the importance of stochastic volatility relative to time-varying coefficients \( B \). Specifically, Figure 10 gives a sense of how much of the model attributes to time-variation in \( B \) and \( \Lambda \) as opposed to increases in volatility \( \Sigma_t \). Time variation in \( B \) and \( \Lambda \) are measured by \( \| \hat{B}_t - \hat{B}_{t-1} \|_2 \) and \( \| \hat{\Lambda}_t - \hat{\Lambda}_{t-1} \|_2 \), respectively, where \( \hat{B}_t \) and \( \hat{\Lambda}_t \) are the posterior means for \( \vec{B}_t \) and \( \vec{\Lambda}_t \), respectively.\(^{11}\) The volatility is summarized by the \( \frac{1}{N} \sum_{i=1}^{N} \log \hat{\sigma}_{i,t}^2 \), where \( \hat{\sigma}_{i,t}^2 \) is posterior mean of the \( i \)th element on the diagonal of \( \Sigma_t \) and \( N \) is the number of series.

We see a large spike in volatility around the start of 2020 but do not estimate greater fluctuations in the coefficients or correlations.\(^{12}\) The spike in volatility is larger even than the spikes in 1980 and 2008 and persists through the end of the sample. Carriero et al. (2021b) similarly report a peak macroeconomic uncertainty level during the COVID-19 pandemic that is higher than the Great Recession peak. The lack of variation in the correlations goes against the model of Primiceri and Tambalotti (2020), which attributes the data fluctuations to a new COVID-19 shock.

### 4.3 Comparison to a Factor Model with Epidemiological Data

Our findings are broadly consistent with Ng (2021), who studies related questions using a factor model similar to (4), but taking \( f_t \) to be a vector and accounting for the pandemic using moving averages of the growth in COVID-19 cases, hospitalizations, or deaths. Ng (2021) finds a large spike in volatility measured using the methodology of Jurado et al. (2015). However, there is a high correlation of over 0.95 between most of the factors estimated with pre-COVID data and those estimated with data through December 2020, suggesting that the correlation across variables did not change substantially once one controls for epidemiological data.\(^{13}\) These results corroborate the finding that the unusual variation in the pandemic era data can be attributed to large shocks rather than large changes in the TVP-VAR coefficients.

That the factors remained relatively unchanged suggests that simply incorporating COVID-19 data can reverse the result in Lenza and Primiceri (2021) that the pandemic era macroeconomic data can severely impact parameter estimates. It further suggests, for instance, that the COVID-19 data can capture much of the change in the fluctuations of GDP and investment relative to other

---

\(^{11}\)We use the Euclidean norm \( \| \cdot \|_2 \), defined by \( \| (z_1, \ldots, z_n)' \| = \sqrt{\sum_{i=1}^{n} z_i^2} \) for an \( n \)-dimensional vector \( (z_1, \ldots, z_n)' \).

\(^{12}\)The estimated volatility starts increasing before 2020, an issue discussed by Prüser (2021).

\(^{13}\)An important caveat in the results is that the data used by Ng (2021) are cleaned to fit the factor structure. First, the data are demeaned, with different means for data before and after March 2020. Second, a dummy is included for March 2020 that absorbs the initial large variation. The results should be interpreted as capturing the variation around these two different means after ignoring the initial large impact rather than the full extent of fluctuations.
variables in the economy, which we highlighted in our discussion of Schorfheide and Song (2020) and Foroni et al. (2020). A possible explanation is that the moving averages have large spikes in March 2020, but quickly decay and remain at low levels from May 2020, thus providing a way to account for the atypical variation in the initial months of the pandemic. In the context of the TVP-VAR, these fluctuations are just accounted for by a sequence of large shocks.

Ng (2021) also argues that incorporating COVID-19 data is necessary to correctly identify economic shocks when the data sample includes the pandemic period. First, she compares the response of unemployment and IP to an unemployment shock for a sample ending before the start of the pandemic and another ending in December 2020. When COVID-19 data are included in the model, the responses are similar. Without the COVID-19 data, the responses differ substantially. Next, she finds a response to the COVID-19 shock (identified recursively) that differs substantially from the unemployment shock. These findings support the inclusion of a new shock in Primiceri and Tambalotti (2020), contradicting the TVP-VAR and Miescu and Rossi (2021), who interpret their COVID-19 shock (identified recursively or by sign restrictions) as a structural uncertainty shock. These findings should be interpreted with caution given the short sample, with further analysis required to distinguish the use of the COVID-19 data from large shocks such as those in Lenza and Primiceri (2021), Carriero et al. (2021a), or Bobeica and Hartwig (2021) to account for the fluctuations early in the pandemic.

5 Conclusion

We have discussed the various ways the literature has sought to use external knowledge or new data sources to construct informative forecasts in the absence of precedent, with several key takeaways:

1. The first approach was to modify the model by introducing stochastic volatility, a new shock, or time-varying parameters more generally. While the additional flexibility improved forecasts, the forecasting performance of these models remained mixed. For example, the models tended to predict much greater persistence in unemployment than was realized in the data.

2. The second approach was to include additional data. Again, despite the information conveyed by the supplementary data, the forecasting performance was mixed. Adjusting forecasts using information from the Great Recession yielded similar issues.

3. Incorporating knowledge about the particular circumstances could have led to more accurate forecasts. For example, employment recovered more quickly than in past recessions and investment fell relatively little given the large decline in GDP. These could arguably have been considered possibilities at the time of the forecasts.

4. Data revisions were small relative to the overall fluctuations and are unlikely to have played a large role in driving forecast errors.
These conclusions are made with the benefit of hindsight and reflect the struggle to formally model possible changes in the economy in real time. As an econometrician, one has to balance feasibility, transparency, and sophistication in one’s models. As a consumer of these forecasts, one has to then interpret each forecast in the context of the assumptions. However, until the assumptions have been modeled or incorporated probabilistically through a prior, it is difficult to know what quantitative adjustment one should make to one’s forecasts to account for changes in the economy.

This is not straightforward even after the event. For instance, while Sims and Zha (2006) argue that the Great Moderation is best explained by a reduction in volatility, Primiceri (2006) and Sargent et al. (2006) argue that the decline in inflation should be explained by policymakers learning and adapting policy. A similar debate is found in Stock and Watson (2012) and the accompanying discussion. In our setting, we have shown mixed evidence about whether the COVID-19 shock was a combination of existing shocks or a completely new shock. The difficulty in establishing what changed in the economy ex post emphasizes the substantial role for parameter uncertainty in the middle of an unprecedented episode.

While this essay has used the COVID-19 pandemic as a case study, the themes apply to other rare events and even normal times. Assumptions often mask uncertainty and bias forecasts. Additional data is informative insofar as it is appropriately modeled. Models should have the flexibility for parameter estimates to adjust probabilistically to incoming data, and forecasters should transparently communicate limitations in this learning process.
References


