Working Paper Series

Michele Lenza, Giorgio E. Primiceri How to estimate a VAR
after March 2020

Abstract

This paper illustrates how to handle a sequence of extreme observations—such as those recorded during the COVID-19 pandemic—when estimating a Vector Autoregression, which is the most popular time-series model in macroeconomics. Our results show that the ad-hoc strategy of dropping these observations may be acceptable for the purpose of parameter estimation. However, disregarding these recent data is inappropriate for forecasting the future evolution of the economy, because it vastly underestimates uncertainty.

KEYWORDS: COVID-19, Volatility, Outliers, Density Forecasts

JEL CODES: C32, E32, E37, C11

NON-TECHNICAL SUMMARY

The COVID-19 pandemic is causing unprecedented variation in many key macroeconomic variables. For example, U.S. unemployment increased by approximately 10 percent from March to April 2020, two orders of magnitude more than its typical monthly change. As Frank Diebold put it in his 27 July blog entry, "one wants to fit signal, not noise, and the pandemic recession is in certain respects a massive burst of noise, capable of severely distorting parameter estimates and hence forecasts and nowcasts." How should we deal with these concerns when estimating time-series models? Should we treat the data from the COVID-19 pandemic period as conventional observations or handle them differently? Answering these questions is essential not only now—in the middle of the epidemic—but for many years to come, since these data will "contaminate" any future sample of time-series observations.

In this paper, we address this question by explicitly modeling the change in shock volatility around the time of the pandemic, implicitly weighting observations inversely proportional to their innovation variance. Our approach is simpler than standard models of time-varying volatility because we know the exact timing of the increase in the variance of macroeconomic innovations due to COVID-19. We specialize our analysis to the case of Vector Autoregression models, but our method can be easily adapted to other time-series models.

Our results show that the ad-hoc strategy of dropping the observations corresponding to the COVID-19 pandemic period may be acceptable for the purpose of parameter estimation. However, disregarding these recent data is inappropriate for forecasting the future evolution of the economy, because it vastly underestimates uncertainty.

1 Introduction

The COVID-19 pandemic is devastating the world economy, producing unprecedented variation in many key macroeconomic variables. For example, in March 2020, U.S. unemployment increased by 0.7 percentage points, which is approximately 7 times as much as its typical monthly change. Things got much worse in April, when unemployment reached a record-high level of 14.7 percent, rising by 10 percentage points in a single month. This change was *two orders of magnitude* larger than its typical month-to-month variation. Most other macroeconomic indicators experienced changes of similar proportion, including employment, consumption expenditures, retail sales and industrial production, to name just a few examples. It is too early to tell whether these extremely large shocks will propagate through the economy in a standard way, or whether their transmission mechanism will be altered. Unfortunately, answering this question requires observing a much longer time span of data. But even with an unchanged transmission mechanism, the massive data variation of the last few months constitutes a challenge for the estimation of standard time-series models. When it comes to inference, should we treat the data from the pandemic period as conventional observations? Or will these observations distort the parameter estimates of our models? Should we instead discard these recent data? These questions are not only essential for the estimation of time-series models during the outbreak of COVID-19, but they will remain crucial for many years to come, since data from the pandemic period will "contaminate" any future sample of time-series observations.

This paper provides a simple solution to this inference problem in the context of Vector Autoregressions (VARs), the most popular time-series model in macroeconomics. Our solution consists of explicitly modeling the change in shock volatility, to account for the exceptionally large macroeconomic innovations during the period of the epidemic. What makes our recipe different, and simpler than standard models of time-varying volatility, is the fact that we know the exact timing of the increase in the variance of macroeconomic innovations due to COVID-19. As a result, we can both flexibly model and easily estimate these volatility changes. For example, suppose to work with a monthly VAR based on U.S. macroeconomic data. We know that March 2020 was the first month of abnormal data variation. We can then simply re-scale the standard deviation of the March shocks by an unknown parameter $\bar{s}_0$, and do the same for April and May with two other parameters $\bar{s}_1$ and $\bar{s}_2$. As we show in section 2, the parameters $\bar{s}_0$, $\bar{s}_1$ and $\bar{s}_2$ can be easily estimated using the approach of Giannone et al. (2015), provided that this re-scaling is common to all shocks. This commonality assumption is the same assumption underlying the stochastic volatility model of Carriero et al. (2016), and it should of course only be

interpreted as an approximation. But this approximation seems reasonable in a period in which all variables experienced record variation. This strategy is also surely preferable to assuming $\bar{s}_0 = \bar{s}_1 = \bar{s}_2 = 1$, which would be implicit in a treatment of the data in March, April and May 2020 as conventional observations.

The final step of our simple procedure consists of modeling the likely future evolution of the residual variance, beyond May 2020. This is considerably more challenging, since the data to inform these estimates are not yet available. To tackle this task, we set up a prior centered on the assumption that the residual variance after May will decay at a 20 percent monthly rate. As more data will become available, the researcher will be able to update such prior with the likelihood information. It is important to stress that this step is irrelevant for the estimation of the model on current data, but it plays an important role for using the VAR to derive predictive densities, since the dispersion of these densities depends on the future value of the residual variances.

To illustrate the properties of our procedure, we estimate a monthly VAR including data on employment, unemployment, consumption and prices, and we use the estimated model to perform two exercises. First, we compute the impulse responses to the forecast error in unemployment. To be clear, these responses do not have any structural interpretation, but we use them only as summary statistics of the estimated dynamics. When we attempt to compute these impulse responses using a standard estimation strategy that does not downweight the data from the pandemic period, this experiment produces nonsensical results, as the error bands of the impulse responses explode. This finding suggests that the March, April and May 2020 observations, despite being a tiny fraction of the whole sample, are so wild that they can influence parameter estimates substantially (and not necessarily in a good way). When the VAR is estimated using our proposed procedure, the impulse response are instead very similar to those that we would obtain by estimating the VAR with data until February 2020. This finding suggests that the ad-hoc procedure of dropping the extreme observations from the pandemic is acceptable for the purpose of parameter estimation, at least given the data available at the time of writing of this paper.

This approach based on disregarding the recent data, however, would be inappropriate for generating any type of prediction for the future evolution of the economy, because it vastly underestimates uncertainty. We demonstrate this point in a second application, in which we use the estimated VAR to produce density forecasts of various macroeconomic variables, conditional on the consensus unemployment projection from the latest release of the Blue Chip Forecasts. When we perform this exercise by estimating the model without the data from the pandemic,

the predictions of employment, consumption expenditures and prices appear to be excessively sharp. Instead, our proposed estimation strategy captures the idea that economic fluctuations may be quite volatile for many months to come. As a consequence, our predictions are consistent with a much broader range of possible recovery paths from the COVID-19 crisis.

The literature on time variation in macroeconomic shock volatility is vast, and its comprehensive review is beyond the scope of this paper. However, it is important to contrast the simple volatility modeling strategy we propose here with the more typical time-varying volatility models that have recently been adopted in the context of VARs, unobserved component or DSGE models (e.g. Cogley and Sargent, 2005, Primiceri, 2005, Sims and Zha, 2006, Carriero et al., 2016, Stock and Watson, 2007, Fernandez-Villaverde and Rubio-Ramirez, 2007, Justiniano and Primiceri, 2008, Curdia et al., 2014). A potential issue with these approaches is that the degree of time variation in volatility is always informed by past data. For example, if historically shock volatilities have varied by at most a factor of two or three from month to month, any estimated ARCH, GARCH, Markov-switching or stochastic volatility model would have a hard time capturing the massive increase in volatility associated with the outbreak of the COVID-19 pandemic. The second important difference between our approach and the more typical models in the literature is that the time of the volatility change is known in the case of COVID-19, which simplifies the estimation of our model. This feature distinguishes our strategy from the one adopted by Stock and Watson (2016) to deal with outliers in their univariate model of inflation.

The rest of the paper is organized as follows. Section 2 describes the methodology we propose to handle the extreme observations recorded during the COVID-19 era. Section 3 presents the results of our two empirical applications, and section 4 concludes.

2   The methodology

To account for the large variance of macroeconomic shocks associated with the outbreak of COVID-19, we modify a standard VAR as follows:

$$y_t = C + B_1 y_{t-1} + ... + B_p y_{t-p} + s_t \varepsilon_t \tag{1}$$

$$\varepsilon_t \sim N\left(0, \Sigma\right),$$

where y_t is an $n \times 1$ vector of variables, modeled as a function of a constant term, their own past values, and an $n \times 1$ vector of forecast errors ε_t. In expression (1), the factor s_t is used to scale up

the residual covariance matrix during the period of the pandemic. More precisely, s_t is equal to 1 before the time period in which the epidemic begins, which we denote by t^*. We then assume that $s_{t^*} = \bar{s}_0$, $s_{t^*+1} = \bar{s}_1$, $s_{t^*+2} = \bar{s}_2$, and $s_{t^*+j} = 1 + (\bar{s}_2 - 1)\rho^{j-2}$, where $\theta \equiv [\bar{s}_0, \bar{s}_1, \bar{s}_2, \rho]$ is a vector of unknown coefficients. This flexible parameterization allows for this scaling factor to take three (possibly) different values in the first three periods after the outbreak of the disease, and to decay at a rate $1 - \rho$ after that. This modeling strategy is particularly suitable for monthly and quarterly time series, given that the amount of data variation was substantially different in the months of March, April and May 2020, and will likely be different when comparing 2020:Q1, Q2 and Q3. We note, however, that alternative parameterizations are possible, even though they would not affect the results and their interpretation.[1]

How can we estimate equation (1)? This task is actually relatively easy. To see why, begin by assuming that s_t is known, and rewrite (1) as

$$y_t = X_t \beta + s_t \varepsilon_t,$$

where $X_t \equiv I_n \otimes x_t'$, $x_t \equiv [1, y_{t-1}', ..., y_{t-p}']$ and $\beta \equiv \text{vec}\left([C, B_1, ..., B_p]'\right)$. Dividing both sides of this equation by s_t, we obtain

$$\tilde{y}_t = \tilde{X}_t \beta + \varepsilon_t,$$

in which $\tilde{y}_t \equiv y_t/s_t$ and $\tilde{X}_t \equiv X_t/s_t$. For given s_t, $\tilde{y}_t$ and $\tilde{X}_t$ are simple transformations of our data. Therefore, the parameters β and Σ can be estimated using the transformed data $\tilde{y}_t$ and $\tilde{X}_t$, and the researcher's preferred approach to inference, such as ordinary least squares, maximum likelihood, or Bayesian estimation.

While the previous insight applies to all estimation procedures (and we will later show how to estimate the model by maximum likelihood), it is now useful to specialize our discussion to the case of Bayesian inference, given the well-known advantages of this approach in the context of heavily parameterized models like VARs. As in Giannone et al. (2015), we focus on prior distributions for VAR coefficients belonging to the conjugate Normal-Inverse Wishart family

$$\Sigma \sim IW\left(\Psi, d\right)$$

$$\beta \sim N\left(b, \Sigma \otimes \Omega\right),$$

where the elements Ψ, d, b and Ω are typically functions of a lower dimensional vector of hyper-

---

[1] In a few months, it may be necessary to extend the model to accommodate a large degree of variation in relation to a second wave of the epidemic.

parameters γ. This class of densities includes the flat prior, the popular Minnesota, Single-Unit-Root and Sum-of-Coefficients priors, as well as the Prior for the Long Run of our earlier work (Litterman, 1980, Doan et al., 1984, Sims and Zha, 1998, Giannone et al., 2019). Giannone et al. (2015) propose a simple method to evaluate the posterior of of β, Σ and γ in a model without s_t. But if we assume that s_t is known and replace y_t and X_t with $\tilde{y}_t$ and $\tilde{X}_t$, we can use the exact same methodology to estimate (1).

Of course, in practice, s_t is unknown and must be estimated as well. Fortunately, the posterior of the parameter vector θ that governs the evolution of s_t can be evaluated like the posterior of γ. More precisely,

$$p\left(\gamma, \theta | y\right) \propto p\left(y | \gamma, \theta\right) \cdot p\left(\gamma, \theta\right), \tag{2}$$

where $y = [y_{p+1}, ..., y_T]'$. In this expression, the first element of the product corresponds to the so-called marginal likelihood, and it can be computed as

$$p\left(y | \gamma, \theta\right) = \int p\left(y | \gamma, \theta\right) p\left(\beta, \Sigma | \gamma\right) d\left(\beta, \Sigma\right),$$

which has an analytical expression. The second density on the right-hand side of (2) is the hyperprior, i.e. the prior on the hyperparameters. Our prior on the elements of γ is the same as in Giannone et al. (2015). As a prior for $\bar{s}_0$, $\bar{s}_1$ and $\bar{s}_2$, we use a Pareto distribution with scale and shape equal to one, which has a very fat right tail, and is thus consistent with possible large increases in the variance of the VAR innovations. For ρ, instead, we impose a Beta prior with mode and standard deviation equal to 0.8 and 0.2, respectively. We stress that the hyperpriors on $\bar{s}_0$, $\bar{s}_1$ and $\bar{s}_2$ are not particularly important for any of the results presented below, because the data are very informative about these parameters. Instead, the hyperprior on ρ determines entirely the shape of its posterior, given that we use data up to May 2020. As more data get released, the researcher will be able to update this prior with the likelihood information.

Appendix A presents some additional technical details of the posterior evaluation procedure. In addition, if a researcher wishes to pursue a frequentist approach, appendix B describes how to easily estimate the full set of unknown parameters in 1, β, Σ and θ, by maximum likelihood. Matlab codes to implement these procedures are also available.

3 Two applications

To illustrate the working and advantages of our modeling approach, we estimate a VAR with some key U.S. macroeconomic indicators. We use the estimated model (i) to track the effects of a surprise change in unemployment; and (ii) to forecast the evolution of the U.S. economy, conditional on the consensus unemployment prediction from the June 2020 release of the Blue Chip Forecasts.

More precisely, the VAR includes six variables available at the monthly frequency: (i) *unemployment*, measured by the civilian unemployment rate; (ii) *employment*, measured by the logarithm of the total number of nonfarm employees; (iii) *PCE*, measured by the logarithm of real personal consumption expenditures; (iv) *PCE: services*, measured by the logarithm of real personal consumption expenditures in services; (v) *PCE (price)*, measured by the logarithm of the price index of personal consumption expenditures; (vi) *PCE: services (price)*, measured by the logarithm of the price index of personal consumption expenditures in services; and (vii) *core PCE*, measured by the logarithm of the price index of personal consumption expenditures excluding food and energy. The VAR has 13 lags and it is estimated on the sample from 1988:12 to 2020:5 using a standard Minnesota prior, whose tightness is chosen as in Giannone et al. (2015). We do not extend the sample before 1988:12 because Del Negro et al. (2020) document a reduced reaction of inflation to fluctuations in real activity since the 1990s, compared to the pre-1990 period.

Figure 1 plots the posterior distribution of four (hyper)parameters of the model. The left panel presents the posterior of the overall standard deviation of the Minnesota prior (denoted by λ), which provides information on the appropriate degree of shrinkage on the β coefficients. The other panels of the figure, present instead the posterior distribution of the volatility scaling factors in March, April and May 2020, i.e. $\bar{s}_0$, $\bar{s}_1$ and $\bar{s}_2$. These posteriors peak around 17, 70 and 20, suggesting that the innovation standard deviation in these three months were one and two orders of magnitude larger than in the pre-COVID-19 period. For comparison, figure 1 also reports the posterior of λ when the VAR is estimated in a standard way, without s_t. When the latest observations are excluded from the estimation sample (essentially assuming $\bar{s}_0 = \bar{s}_1 = \bar{s}_2 = \infty$), the posterior of λ is similar to the one implied by our baseline model, consistent with the fact that our inferential procedure assigns comparatively less weight to these most recent observations. When instead these observations are included in the estimation sample and treated as conventional data (essentially assuming $\bar{s}_0 = \bar{s}_1 = \bar{s}_2 = 1$), the posterior of λ exhibits a large
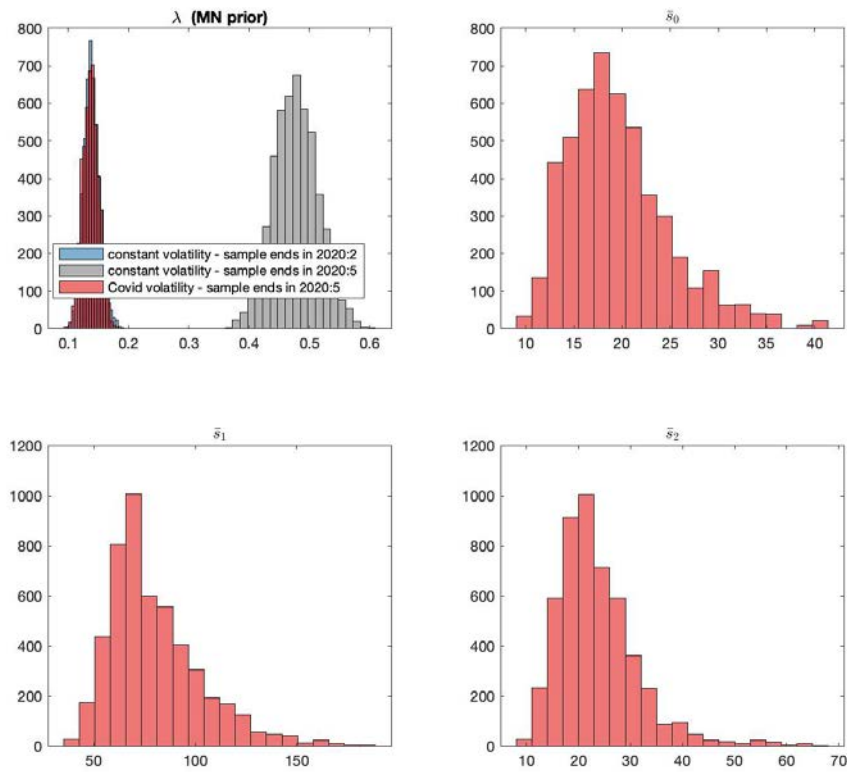
Figure 1: Posterior distribution of the overall standard deviation of the Minnesota prior, and the March, April and May 2020 volatility scaling factors.

shift to the right, implying much less shrinkage for the β coefficients. This reduction in shrinkage is the necessary cost to pay to fit the large variability of the latest data with a change in the estimated β.

We now illustrate the implications of these estimation results in two empirical applications. In our first application, we study the dynamic response of the variables in the VAR to a positive shock to unemployment, when it is ordered first in a Cholesky identification scheme. Therefore, this shock corresponds to the typical linear combination of structural disturbances that drives the one-step-ahead forecast error of unemployment. We do not assign a structural interpretation to these impulse responses, but just use them as summary statistics of the estimated dynamics. Figure 2 presents their posterior when the VAR is estimated using the procedure outlined in section 2. The real economy (employment, unemployment and consumption) initially slows down and then partially recovers. Prices also experience some downward pressure, consistent with a "demand" interpretation of this shock.

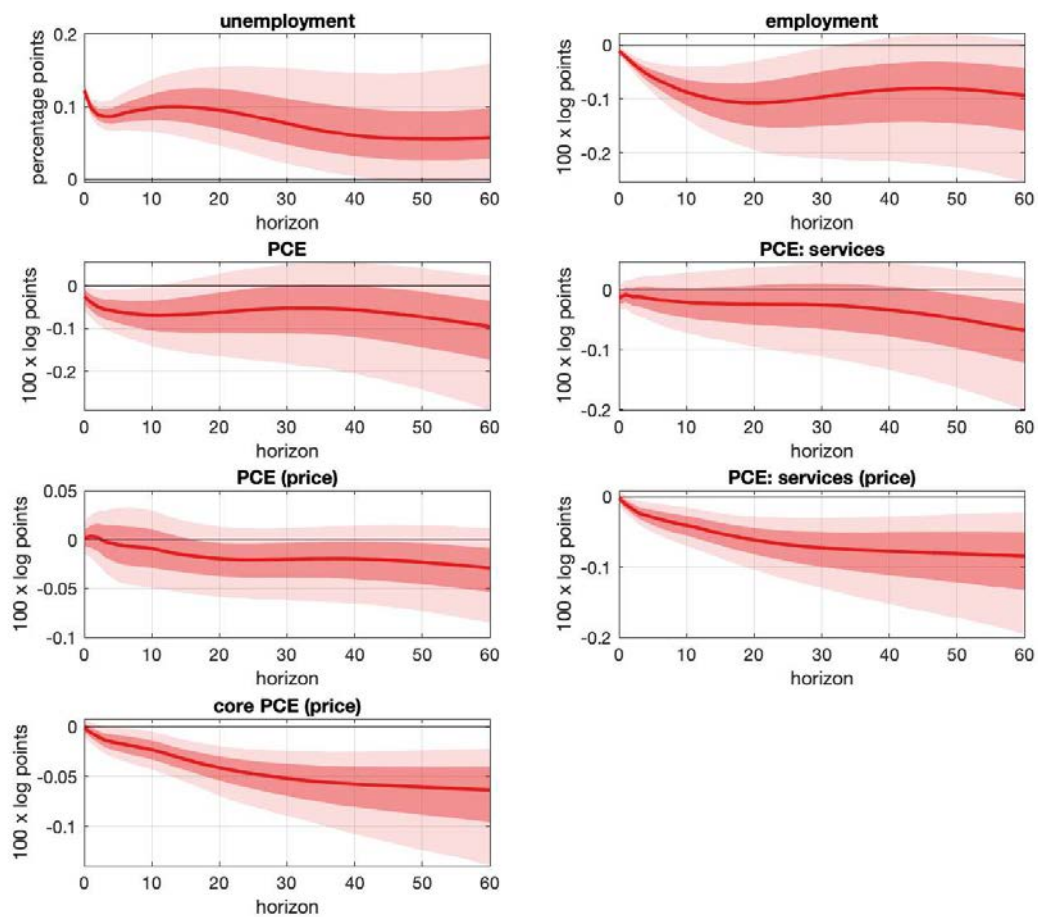As we did earlier, it is useful to compare these impulse responses to those obtained when

Figure 2: Impulse responses to a one standard deviation shock to the unemployment equation. The shock is identified using a Cholesky strategy, with unemployment ordered first. The solid lines are posterior medians, while the shaded areas correspond to 68- and 95-percent posterior credible regions.

estimating the model in a more conventional way. Unfortunately, if we include the latest observations in the estimation sample without any special treatment, the implied impulse responses become explosive and unreasonable (for this reason, we do not report them). This finding illustrates the importance of explicitly modeling the change in shock volatility during the COVID-19 era. Instead, the responses in figure 2 are very similar to those implied by a standard VAR estimation that excludes the March, April and May 2020 data (given this similarity, we do not report them either). We summarize the lesson from this first application as follows: For the purpose of estimating the parameters β and Σ, we recommend to adopt the procedure we described in section 2, which involves a minimal deviation from the conventional VAR estimation. However, if a researcher still wishes to estimate a VAR "as usual," it is much better to exclude the data from the pandemic rather than including them and treating them as any other observation in the sample. It is likely that the latter approach will produce meaningless results, as it did in our application.

We now illustrate a second empirical application in which the gains of adopting our approach to inference are even more evident. In this application, we conduct a scenario analysis to highlight the impact of the current change in shock volatility and its expected future evolution on the U.S. economic outlook and the uncertainty surrounding it. More precisely, we compute the most likely evolution of the macroeconomic indicators included in our VAR under the consensus unemployment projection from the latest release of the Blue Chip Forecasts. This consensus unemployment projection is plotted in the first panel of figure 3.[2] According to the Blue Chip projections, unemployment will likely decline steadily over the next few years, reaching 9 percent by the end of 2020, and 7 percent by the end of 2021. The other panels of the figure, show a slow and gradual recovery for employment and consumption of services, and a more sudden rebound of total consumption The three measures of prices seem to be affected much less from the deep recession and, after a short-lived downturn and a rebound, they return to their historical trend.

How would these conditional forecasts look like if we estimated the model in a more standard way? Once again, the answer to this question depends on whether the last few observations from the COVID-19 period are included or excluded from the estimation sample. If they are included, the estimates of β and Σ are unduly affected, and these conditional forecasts become unreasonable, as in the case of the impulse responses above. If instead the estimation sample ends in February 2020, the implied conditional forecasts are more in line with those produced by our model with time-varying volatility. But despite a broad similarity in the qualitative features, the

---

[2]The Blue Chip Forecasts report quarterly projections. To translate them into monthly projections, we simply interpolated them using a smooth exponential curve.
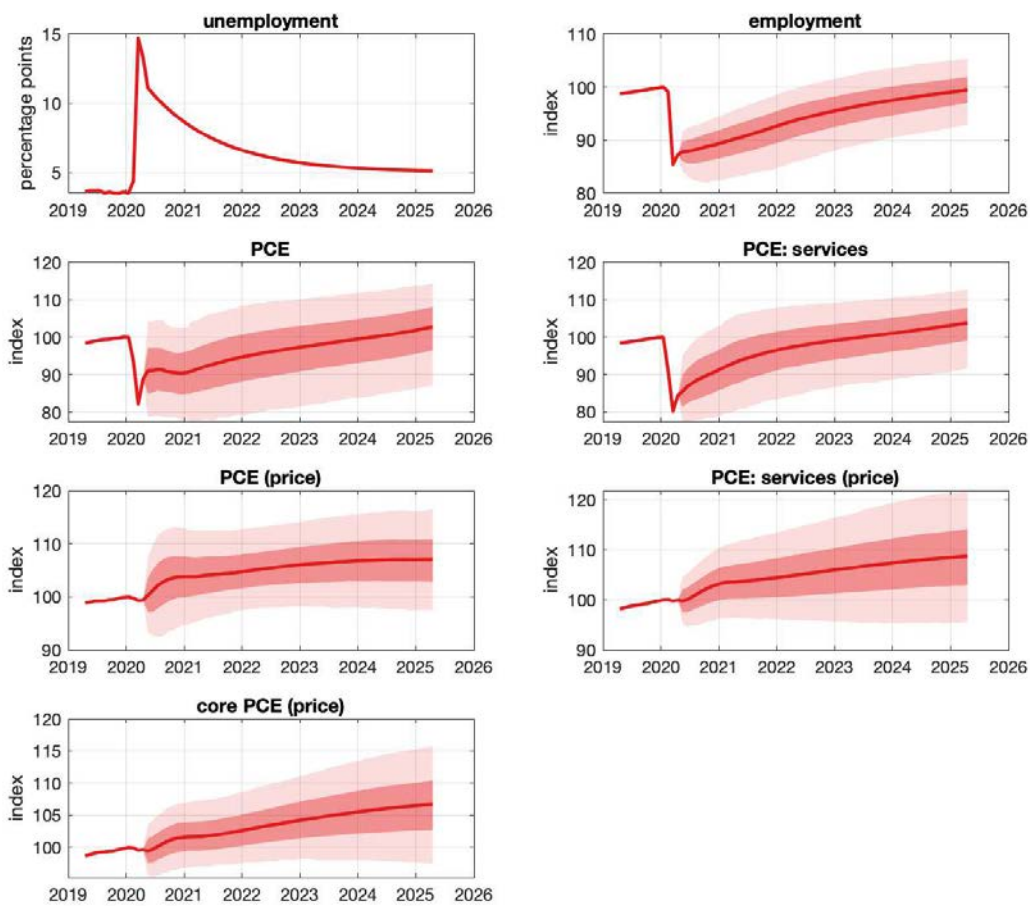
Figure 3: Forecasts of the variables in the VAR conditional on unemployment following the path in the first subplot. The solid lines are posterior medians, while the shaded areas correspond to 68- and 95-percent posterior credible regions.

conditional forecasts show now important quantitative differences, as we can see by comparing figures 3 and 4. In particular, the VAR estimated with our proposed procedure incorporates a higher innovation variance not only in March, April and May 2020, but also in the subsequent few months.[3] As a consequence, the estimated conditional forecasts in figure 3 exhibit a higher degree of uncertainty about the future prospects of the U.S. economy, relative to the forecasts in figure 4 obtained by excluding the latest data. For example, the predictions of consumption produced by the standard VAR at short horizons are quite sharp, presumably too much. The density forecast in figure 3, instead, not only entails more uncertainty, but it also implies a faster rebound, since it is consistent with the idea that the size of fluctuations may be larger for some time. This implication is reasonable, given that consumption has declined considerably more relative to other variables at the onset of the COVID-19 recession.

---

[3]As mentioned earlier, the expected future evolution of the shock volatility depends only on the prior at the moment, but the availability of a longer sample will soon allow to update this prior with more likelihood information.
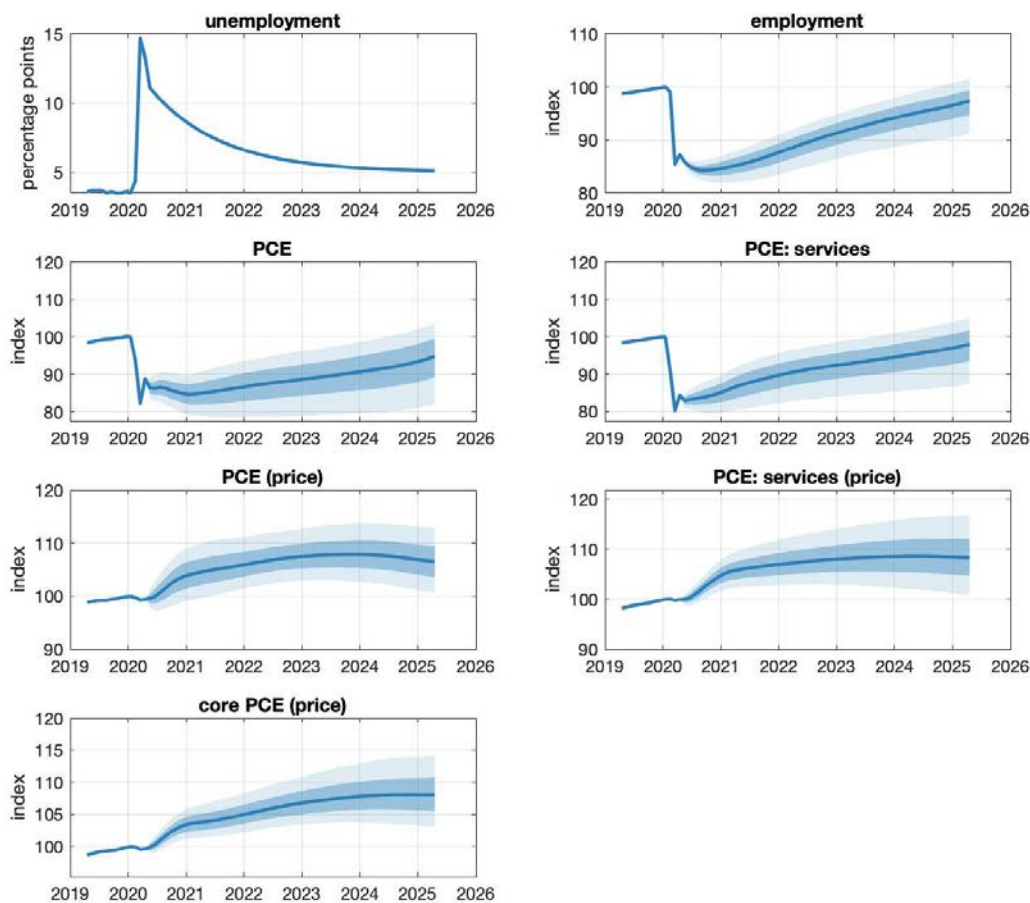
Figure 4: Forecasts of the variables in the VAR conditional on unemployment following the path in the first subplot. The solid lines are posterior medians, while the shaded areas correspond to 68- and 95-percent posterior credible regions.

4 Concluding Remarks

The sequence of wild macroeconomic variation experienced during the COVID-19 pandemic constitutes a challenge for the estimation of macro-econometric models in general, and VARs in particular. In this paper, we propose a simple solution to this problem, which consists of explicitly modeling the large change in shock volatility during the outbreak of the disease. We also show that estimating such a model is quite straightforward, because the time of the volatility change is known. Our empirical results show that the ad-hoc procedure of dropping the extreme observations from the pandemic era is acceptable for the purpose of parameter estimation—at least given the data available at the time of writing of this paper—but it is inappropriate for

forecasting the future evolution of the economy, because it vastly underestimates uncertainty.

A Posterior Evaluation

This appendix describes the technical details of the MCMC algorithm that we use to evaluate the posterior of the model parameters. This algorithm is a standard Metropolis algorithm, almost identical to that in Giannone et al. (2015), consisting of the following steps:

1. Initialize the hyperparameters γ and θ at their posterior mode, which requires a preliminary numerical maximization of their marginal posterior (whose analytical expression is derived below).

2. Draw a candidate value $[\gamma^*, \theta^*]$ of the hyperparameters from a Gaussian proposal distribution, with mean equal to $\left[\gamma^{(j-1)}, \theta^{(j-1)}\right]$ and variance equal to $c \cdot W$, where $\left[\gamma^{(j-1)}, \theta^{(j-1)}\right]$ is the previous draw of $[\gamma, \theta]$, W is the inverse Hessian of the negative of the log-posterior of the hyperparameters at the peak, and c is a scaling constant chosen to obtain an acceptance rate of approximately 25 percent.

3. Set
$$
\left[\gamma^{(j)}, \theta^{(j)}\right] = \left\{
\begin{array}{ll}
[\gamma^*, \theta^*] & \text{with pr. } \alpha^{(j)} \\
\left[\gamma^{(j-1)}, \theta^{(j-1)}\right] & \text{with pr. } 1 - \alpha^{(j)},
\end{array}
\right.
$$
where
$$
\alpha^{(j)} = \min\left\{1, \frac{p\left(\gamma^*, \theta^*|y\right)}{p\left(\gamma^{(j-1)}, \theta^{(j-1)}|y\right)}\right\}
$$

4. Draw $\left[\beta^{(j)}, \Sigma^{(j)}\right]$ from $p\left(\beta, \Sigma|y, \gamma^{(j)}, \theta^{(j)}\right)$, which is a Normal-Inverse-Wishart density (see details below).

5. Increment j to $j + 1$ and go to 2.

In step 3, the density $p\left(\gamma, \theta|y\right)$ is given by

$$
p\left(\gamma, \theta|y\right) \propto p\left(y|\gamma, \theta\right) \cdot p\left(\gamma, \theta\right),
$$

where the second term of the product corresponds to the hyperprior. The first term is instead the marginal likelihood, and it can be computed analytically as in Giannone et al. (2015). If we

condition on the initial p observations of the sample, which is a standard assumption, we obtain

$$p\left(y|\gamma,\theta\right) = \prod_{t=p+1}^{T} p\left(y_t|X_t,\gamma,\theta\right) = \prod_{t=p+1}^{T} \frac{p\left(\tilde{y}_t|\tilde{X}_t,\gamma,\theta\right)}{s_t^n},$$

where the denominator on the right-hand side captures the Jacobian of the transformation $\tilde{y}_t = y_t/s_t$. From the results in Giannone et al. (2015), it follows immediately that

$$p\left(y|\gamma,\theta\right) = \left(\prod_{t=p+1}^{T} s_t^{-n}\right) \left(\frac{1}{\pi}\right)^{\frac{n(T-p)}{2}} \frac{\Gamma_n\left(\frac{T-p+d}{2}\right)}{\Gamma_n\left(\frac{d}{2}\right)} \cdot$$

$$|\Omega|^{-\frac{n}{2}} \cdot |\Psi|^{\frac{d}{2}} \cdot \left|\tilde{x}'\tilde{x} + \Omega^{-1}\right|^{-\frac{n}{2}} \cdot$$

$$\left|\Psi + \hat{\tilde{\varepsilon}}'\hat{\tilde{\varepsilon}} + \left(\hat{\tilde{B}} - \hat{b}\right)' \Omega^{-1} \left(\hat{\tilde{B}} - \hat{b}\right)\right|^{-\frac{T-p+d}{2}},$$

where $\tilde{x}_t \equiv \left[1, y'_{t-1}, ..., y'_{t-p}\right]/s_t$, $\tilde{x} \equiv [\tilde{x}_{p+1}, ..., \tilde{x}_T]'$, $\tilde{y} \equiv [\tilde{y}_{p+1}, ..., \tilde{y}_T]'$, $\hat{\tilde{B}} \equiv \left(\tilde{x}'\tilde{x} + \Omega^{-1}\right)^{-1}\left(\tilde{x}'\tilde{y} + \Omega^{-1}\hat{b}\right)$, $\hat{\tilde{\varepsilon}} \equiv \tilde{y} - \tilde{x}\hat{\tilde{B}}$, and $\hat{b}$ is a matrix obtained by reshaping the vector b in such a way that each column corresponds to the prior mean of the coefficients of each equation (i.e. $b \equiv \text{vec}\left(\hat{b}\right)$).

The posterior of (β, Σ) in step 4 is given by

$$\Sigma|Y \sim IW\left(\Psi + \hat{\tilde{\varepsilon}}'\hat{\tilde{\varepsilon}} + \left(\hat{\tilde{B}} - \hat{b}\right)' \Omega^{-1} \left(\hat{\tilde{B}} - \hat{b}\right), T - p + d\right)$$

$$\beta|\Sigma, Y \sim N\left(\text{vec}\left(\hat{\tilde{B}}\right), \Sigma \otimes \left(\tilde{x}'\tilde{x} + \Omega^{-1}\right)^{-1}\right).$$

B   Maximum Likelihood Estimation

This appendix describes how to estimate our model using a frequentist, maximum likelihood method. The likelihood function of model (1) is given by

$$p\left(y|\beta,\Sigma,\theta\right) \propto \prod_{t=p+1}^{T} \left|s_t^2\Sigma\right|^{-\frac{1}{2}} \cdot \exp\left\{-\frac{1}{2}\sum_{t=p+1}^{T}\left(y_t - X_t'\beta\right)'\left(s_t^2\Sigma\right)^{-1}\left(y_t - X_t'\beta\right)\right\}$$

$$\propto \left(\prod_{t=p+1}^{T} s_t^{-n}\right) \cdot |\Sigma|^{-\frac{T-p}{2}} \cdot \exp\left\{-\frac{1}{2}\left(\tilde{Y} - \tilde{X}\beta\right)'\left(\Sigma \otimes I_{T-p}\right)^{-1}\left(\tilde{Y} - \tilde{X}\beta\right)\right\}, \quad (3)$$

where $\tilde{Y} \equiv \text{vec}(\tilde{y})$ and $\tilde{X} \equiv I_n \otimes \tilde{x}$, and we are conditioning on the initial p observations, as usual.

The maximum likelihood estimators of β and Σ, as functions of θ, are given by

$$\hat{B}_{mle}(\theta) = (\tilde{x}'\tilde{x})^{-1}\tilde{x}'\tilde{y} \tag{4}$$

$$\hat{\beta}_{mle}(\theta) = \text{vec}\left[\hat{B}_{mle}(\theta)\right] \tag{5}$$

$$\hat{\Sigma}_{mle}(\theta) = \frac{\hat{\tilde{\varepsilon}}'_{mle}\hat{\tilde{\varepsilon}}_{mle}}{T-p}, \tag{6}$$

where $\hat{\tilde{\varepsilon}}_{mle} \equiv \tilde{y} - \tilde{x}\hat{B}_{mle}(\theta)$. Recal that all variables with a "" depend on θ, although we do not make this dependence explicit to streamline the notation. Substituting (4)-(6) into (3) yields the concentrated likelihood, which, after some algebraic manipulations, can be written as

$$p\left(y|\hat{\beta}_{mle}(\theta), \hat{\Sigma}_{mle}(\theta), \theta\right) \propto \prod_{t=p+1}^{T} s_t^{-n} \cdot \left|\frac{\hat{\tilde{\varepsilon}}'_{mle}\hat{\tilde{\varepsilon}}_{mle}}{T-p}\right|^{-\frac{T-p}{2}}. \tag{7}$$

The parameters θ can then be estimated by numerically maximizing (7), and the estimates of β and Σ can be obtained using (4)-(6).

References

CARRIERO, A., T. E. CLARK, AND M. MARCELLINO (2016): "Common Drifting Volatility in Large Bayesian VARs," *Journal of Business & Economic Statistics*, 34, 375–390.

COGLEY, T. AND T. J. SARGENT (2005): "Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S," *Review of Economic Dynamics*, 8, 262–302.

CURDIA, V., M. D. NEGRO, AND D. L. GREENWALD (2014): "Rare Shocks, Great Recessions," *Journal of Applied Econometrics*, 29, 1031–1052.

DEL NEGRO, M., M. LENZA, G. E. PRIMICERI, AND A. TAMBALOTTI (2020): "What's up with the Phillips Curve?" *Brookings Papers on Economic Activity*, 51, forthcoming.

DOAN, T., R. LITTERMAN, AND C. A. SIMS (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

FERNANDEZ-VILLAVERDE, J. AND J. F. RUBIO-RAMIREZ (2007): "Estimating Macroeconomic Models: A Likelihood Approach," *Review of Economic Studies*, 74, 1059–1087.

GIANNONE, D., M. LENZA, AND G. E. PRIMICERI (2015): "Prior Selection for Vector Autoregressions," *The Review of Economics and Statistics*, 97, 436–451.

——— (2019): "Priors for the Long Run," *Journal of the American Statistical Association*, 114, 565–580.

JUSTINIANO, A. AND G. E. PRIMICERI (2008): "The Time-Varying Volatility of Macroeconomic Fluctuations," *American Economic Review*, 98, 604–641.

LITTERMAN, R. B. (1980): "A Bayesian Procedure for Forecasting with Vector Autoregression," Working paper, Massachusetts Institute of Technology, Department of Economics.

PRIMICERI, G. E. (2005): "Time Varying Structural Vector Autoregressions and Monetary Policy," *Review of Economic Studies*, 72, 821–852.

SIMS, C. A. AND T. ZHA (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39, 949–68.

——— (2006): "Were There Regime Switches in U.S. Monetary Policy?" *American Economic Review*, 96, 54–81.

STOCK, J. H. AND M. W. WATSON (2007): "Why Has U.S. Inflation Become Harder to Forecast?" *Journal of Money, Credit and Banking*, 39, 3–33.

——— (2016): "Core Inflation and Trend Inflation," *The Review of Economics and Statistics*, 98, 770–784.

Michele Lenza
European Central Bank, Frankfurt am Main, Germany; ECARES; email: michele.lenza@ecb.europa.eu

Giorgio E. Primiceri
Northwestern University, Evanston, United States; CEPR; NBER; email: g-primiceri@northwestern.edu