Pandemic Priors

Danilo Cascaldi-Garcia
Federal Reserve Board

Real-Time Data Analysis, Methods, and Applications
October 20, 2023

Pandemic Priors

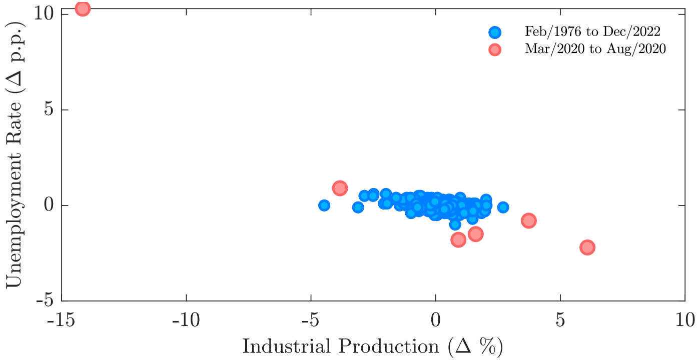
- The pandemic caused macroeconomic variables to display complex patterns that hardly follow any historical behavior

- Bayesian VARs: very low number of extreme pandemic observations bias the estimated persistence of the variables
  - Affect forecasts
  - Myopic view of the economic effects after a structural shock

- **This paper:** Easy and straightforward solution to deal with extreme episodes, that recovers historical relationships and the proper identification and propagation of structural shocks

Data

Figure: Industrial production and unemployment rate variation over time



Implication for Bayesian VARs

- **Challenge:** deal with such unusual behavior and retain historical relationships, produce reliable forecasts, and provide correct interpretations of economic shocks

- Minnesota Prior (Litterman, 1986): computationally feasible estimations of large information sets overcoming the curse of dimensionality

- **Pandemic Priors:** extension to allow for **time dummies**
  - Implementation via dummy observations (Bańbura, Giannone, and Reichlin, 2010)
  - **Optimal** selection of the shrinkage level for extreme observations
  - Nests the boundary cases of Uninformative and Minnesota Priors

Related literature

- Common shift and persistence of the volatility of shocks: Lenza and Primiceri (2021)
- Discarding extreme observations: Schorfheide and Song (2020)
- Extreme observations as random shifts in the stochastic volatility: Carriero, Clark, Marcellino, and Mertens (2022) and Álvarez and Odendahl (2022)
- Non-parametric methods: Huber, Koop, Onorante, Pfarrhofer, and Schreiner (2023)
- VAR with t-distributed errors: Bobeica and Hartwig (2023)
- VAR augmented with the log-differences of the information set during the pandemic: Ng (2021)
- Outliers in the context of dynamic factor models: Antolin-Diaz, Drechsel, and Petrella (2021)

Implementation

- Minnesota Prior (Litterman, 1986) through dummy observations (Bańbura et al., 2010) + **time dummies on extreme observations**

- VAR model with n variables and p lags as in:

$$\mathbf{Y}_t = \mathbf{c} + \mathbb{1}_{t=a}\mathbf{d}_a + ... + \mathbb{1}_{t=a+h}\mathbf{d}_{a+h} + \mathbf{A}_1\mathbf{Y}_{t-1} + ... + \mathbf{A}_p\mathbf{Y}_{t-p} + \mathbf{u}_t,$$

where

- $\mathbb{E}[\mathbf{u}_t\mathbf{u}_t'] = \Psi$
- $\mathbf{d}_a, ..., \mathbf{d}_{a+h}$ are h vectors with n time dummies for periods a through $a+h$ (e.g., the pandemic)
- $\mathbb{1}_{t=i}$ is an indicator function that takes value $\mathbb{1}_{t=i} = 1$ for periods $i = a, ..., a+h$, and 0 otherwise

Implementation

- **Prior:** variables centered around the random walk with a drift + **abnormal period where the relationship between the variables may diverge from history** (e.g., the pandemic)

$$\mathbf{Y}_t = \mathbf{c} + \mathbb{1}_{t=a}\mathbf{d}_a + ... + \mathbb{1}_{t=a+h}\mathbf{d}_{a+h} + \mathbf{Y}_{t-1} + \mathbf{u}_t,$$

- equivalent to shrinking $\mathbf{A}_1$ to the identity and $\mathbf{A}_2, ..., \mathbf{A}_p$ to zero

- Standard prior moments:

$$\mathbb{E}\left[(A_k)_{ij}\right] = \begin{cases} \delta_i, & j = i, k = 1 \\ 0, & \text{otherwise} \end{cases}, \mathbb{V}\left[(A_k)_{ij}\right] = \begin{cases} \frac{\lambda^2}{k^2}, & j = i \\ v\frac{\lambda^2}{k^2}\frac{\sigma_i^2}{\sigma_j^2}, & \text{otherwise} \end{cases}.$$

- Prior for the intercept is diffuse

Implementation

- Normal inverse Wishart prior

$$\text{vec}(\mathbf{B})|\Psi \sim \mathcal{N}\left(\text{vec}(\mathbf{B_0}), \Psi \otimes \Omega_0\right) \quad \textit{and} \quad \Psi \sim i\mathcal{W}\left(S_0, \alpha_0\right)$$

  where

- $\mathbf{B}$: reduced-form coefficients from $\mathbf{Y}_t = \mathbf{X}_t\mathbf{B} + \mathbf{U}_t$
- $\mathbf{B}_0$, Ψ_0, S_0, and α_0 are prior expectations
- $\mathbb{E}[\Psi] = \Sigma = diag(\sigma_1^2, ..., \sigma_n^2)$

- In practice: dummy observations

Implementation

$$\mathbf{Y}_d = \begin{pmatrix} \mathrm{diag}(\delta_1\sigma_1, ..., \delta_n\sigma_n)/\lambda \\ \mathbf{0}_{n(p-1)\times n} \\ ... \\ \mathrm{diag}(\sigma_1, ..., \sigma_n) \\ ... \\ \mathbf{0}_{1\times n} \\ \mathbf{0}_{h\times n} \end{pmatrix}$$

$$\mathbf{X}_d = \begin{pmatrix} \mathbf{J}_p \otimes \mathrm{diag}(\sigma_1, ..., \sigma_n)/\lambda & \mathbf{0}_{np\times 1} & \mathbf{0}_{np\times h} \\ ... & ... & ... \\ \mathbf{0}_{n\times np} & \mathbf{0}_{n\times 1} & \mathbf{0}_{n\times h} \\ ... & ... & ... \\ \mathbf{0}_{1\times np} & \epsilon & \mathbf{0}_{1\times h} \\ \mathbf{0}_{h\times np} & \mathbf{0}_{h\times 1} & \mathrm{diag}(\phi_1, ..., \phi_h) \end{pmatrix}$$

- Priors for the h time dummies imposed through $\phi_1, ..., \phi_h = \phi$

Implementation

What is new:

ϕ governs the prior associated with the time dummies

- $\phi \to 0$: prior for the dummies is fairly uninformative, and soak the variance of the pandemic period
- $\phi \to \infty$: full signal is taking from the pandemic period, and that information is treated as any other observation

Pandemic Priors nest the boundary cases of **no-to-full signal** from the pandemic observations through the parameter ϕ

Selection of ϕ: can be arbitrary by the econometrician's taste of how much information to take from those extreme values, or **optimally chosen** (more to follow)

Implementation

- Combine LHS and RHS as $\mathbf{Y}_t^* = [\mathbf{Y}_t', \mathbf{Y}_d']$ and $\mathbf{X}_t^* = [\mathbf{X}_t', \mathbf{X}_d']$

- **Posterior:**

$$
\begin{array}{rcl}
\text{vec}(\mathbf{B})|\Psi, \mathbf{Y}_t & \sim & \mathcal{N}\left(\text{vec}(\tilde{\mathbf{B}}), \Psi \otimes \left(\mathbf{X}_t^{*\prime}\mathbf{X}_t^*\right)^{-1}\right) \\
\Psi|\mathbf{Y}_t & \sim & i\mathcal{W}\left(\tilde{\Sigma}, T_d + 2 + T - m\right),
\end{array}
$$

where

- T is the sample size, T_d is the length of dummy observations, $m = np + 1 + h$
- $\tilde{\mathbf{B}} = \left(\mathbf{X}_t^{*\prime}\mathbf{X}_t^*\right)^{-1}\left(\mathbf{X}_t^{*\prime}\mathbf{Y}_t^*\right)$, and $\tilde{\Sigma} = \left(\mathbf{Y}_t^* - \mathbf{X}_t^*\tilde{\mathbf{B}}\right)'\left(\mathbf{Y}_t^* - \mathbf{X}_t^*\tilde{\mathbf{B}}\right)$
- Possible to also impose a no-cointegration prior by constraining the sum of the coefficients (Bańbura et al., 2010)

Optimal selection of ϕ

- Defines how much signal to take from the extreme observations in the system
- Method to select the optimal level of ϕ: maximizing the marginal density of the model
- Adaptation of the optimal overall prior tightness described in Carriero, Kapetanios, and Marcellino (2012) and Carriero, Clark, and Marcellino (2015)

Optimal selection of ϕ

The optimal ϕ^* is defined as

$$\phi^* = \arg\max_{\phi} \ln p_{\phi}(Y),$$

where $p_{\phi}(Y)$ is the marginal density, or marginal likelihood, by integrating the set Θ of coefficients, of

$$p_{\phi}(Y) = p(Y|\phi) = \int p(Y|\Theta, \phi)p(\Theta|\phi)d\Theta$$

Optimal selection of ϕ

Calculated in closed-form (normal inverse Wishart prior):

$$p_{\phi}(Y) = \pi^{-\frac{Tn}{2}} \times \left| \left(\mathbf{I} + \mathbf{X}_t \Omega_0(\phi) \mathbf{X}_t' \right)^{(-1)} \right|^{\frac{n}{2}} \times |S_0|^{\frac{\alpha_0}{2}} \times \left(\frac{\Gamma_n \frac{\alpha_0 + T}{2}}{\Gamma_n \frac{\alpha_0}{2}} \right) \times \ldots$$

$$\ldots \times \left| S_0 + (\mathbf{Y}_t - \mathbf{X}_t \mathbf{B}_0)' \left(\mathbf{I} + \mathbf{X}_t \Omega_0(\phi) \mathbf{X}_t' \right)^{(-1)} (\mathbf{Y}_t - \mathbf{X}_t \mathbf{B}_0) \right|^{-\frac{\alpha_0 + T}{2}},$$

where:

- $\alpha_0 = n + 2$
- Γ_n is the n-variate gamma function
- prior variance expectation $\Omega_0(\phi)$ is now a function of ϕ

In practice: optimal ϕ^* is the one that maximizes the marginal density over a discrete grid of values for ϕ

Pandemic Priors applicability test

Applicability test: evaluate the marginal density of the two boundary cases of a Minnesota Prior model ($\phi^* \to \infty$) and of an uninformative Pandemic Priors model ($\phi^* \to 0$)

Agnostic way of checking if the observations should be treated differently or not

Calculate the ratio $R_{t,w}$ between the marginal density of the boundary cases for every sub-sample with a defined length w:

$$R_{t,w} = \frac{\ln p_{\phi \to \infty}(Y)_{t,w}}{\ln p_{\phi \to 0}(Y)_{t,w}}$$

Pandemic Priors applicability test

The $R_{t,w}$ test:

- If the model favors treating the observations from t to $t+w-1$ as extreme values that should be downplayed by some degree:
  - $\ln p_{\phi \to 0}(Y)_{t,w} > \ln p_{\phi \to \infty}(Y)_{t,w}$
  - $R_{t,w} < 1$
  - Application of the Pandemic Priors is advisable for t to $t+w-1$

- If the model favors a conventional Minnesota Prior:
  - $\ln p_{\phi \to 0}(Y)_{t,w} < \ln p_{\phi \to \infty}(Y)_{t,w}$
  - $R_{t,w} > 1$
  - Time dummies for t to $t+w-1$ will be ineffective

Monte Carlo simulation

- "Abnormal" shocks $e^*_{i,t}$ affecting all variables simultaneously at a pre-defined time, but with different size and persistence

- Stationary system of four variables and two lags

$$\mathbf{D}_0 \begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \\ y_{4,t} \end{bmatrix} = \mathbf{C} + \mathbf{D}_1 \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \\ y_{3,t-1} \\ y_{4,t-1} \end{bmatrix} + \mathbf{D}_2 \begin{bmatrix} y_{1,t-2} \\ y_{2,t-2} \\ y_{3,t-2} \\ y_{4,t-2} \end{bmatrix} + \begin{bmatrix} e_{1,t} \\ e_{2,t} \\ e_{3,t} \\ e_{4,t} \end{bmatrix} + \begin{bmatrix} e^*_{1,t} \\ e^*_{2,t} \\ e^*_{3,t} \\ e^*_{4,t} \end{bmatrix}$$

$$e^*_{i,t} = \begin{cases} 0, & t < t^* \\ e^*_{i,t^*}, & t = t^* \\ \rho_i e^*_{i,t-1}, & t > t^* \end{cases} .$$

- Simulate data for 600 periods, with abnormal shocks at $t^* = 501$

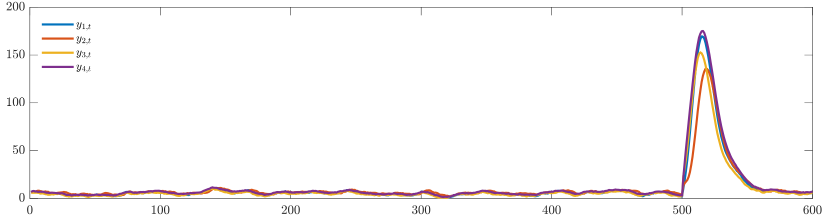
Coefficients

Monte Carlo simulation

Shocks varying from 5 to 20 standard deviations:

$$\begin{bmatrix} \epsilon_{1,t^*} \\ \epsilon_{2,t^*} \\ \epsilon_{3,t^*} \\ \epsilon_{4,t^*} \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 15 \\ 20 \end{bmatrix}, \quad \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} = \begin{bmatrix} 0.6 \\ 0.7 \\ 0.3 \\ 0.9 \end{bmatrix}.$$

Figure: Simulated series

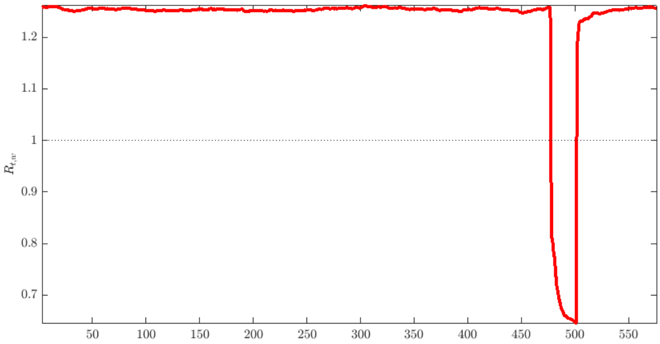


Monte Carlo simulation

First step: $R_{t,w}$ **test**

- Over the 600 observations, with $w = 24$ periods
- $\phi = 5$ for the Minnesota Prior
- $\phi = 0.001$ for the uninformative Pandemic Priors

Figure: Marginal density ratio (Minnesota Prior / uninf. Pandemic Priors)



Monte Carlo simulation

Second step: Comparison of Pandemic Priors with a baseline and the data generating process

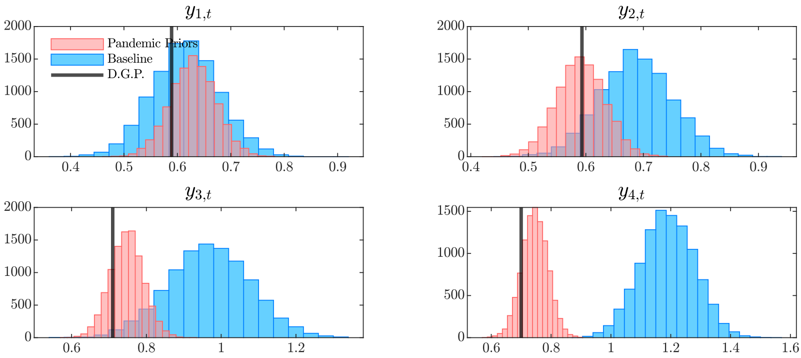
- Bayesian VARs in levels with the **Minnesota Prior** (baseline) and **Pandemic Priors** (time dummies for the first 24 periods from the shock $t = 501, ..., 524$)

- Optimal $\phi^* = 0.075$, distant from the grid boundary cases (0.001 for uninformative Pandemic Priors and 5 for Minnesota Prior)

Grid: $[0.001, 0.01, 0.025, 0.050, 0.075, 0.10, 0.15, 0.20, 0.25, 0.30, 0.35, 0.40, 0.45, 0.50, 0.75, 1, 2, 5]$

Monte Carlo simulation

Figure: Posterior draws for the autoregressive coefficients



1. Large and persistent shocks → d.g.p. lies outside of the baseline
2. Considerably more uncertainty with the baseline
3. Optimal ϕ^* matters for coverage

Uninformative prior

Empirical example

- Monthly Bayesian VAR in levels, from January 1975 through December 2022, 12 lags

Table: Information set

	Name	Description
1	EBP	Excess bond premium as computed by Gilchrist and Zakrajšek (2012).
2	S&P 500	S&P 500 stock index in log levels.
3	Shadow Rate	Fed funds rate shadow rate as computed by Wu and Xia (2016).
4	Consumption (PCE)	Real consumption in log levels.
5	Price index	PCE Price Index in log levels.
6	Employment	PCE Total nonfarm payroll in log levels
7	Ind. production	Real industrial output in log levels.
8	Unemployment rate	Number of unemployed as a percentage of the labor force.

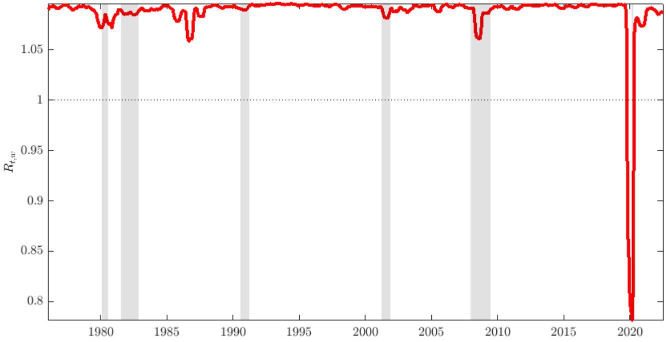
Note: All for the January 1975 to March 2022 period, retrieved on February 2023.

Empirical example

First step: $R_{t,w}$ **test**

- January 1975 to December 2022, with $w = 6$ months
- Pandemic should be downplayed, but no other period
- $R_{t,w}$ seems to drop near recessions

Figure: Marginal density ratio (Minnesota Prior / uninf. Pandemic Priors)



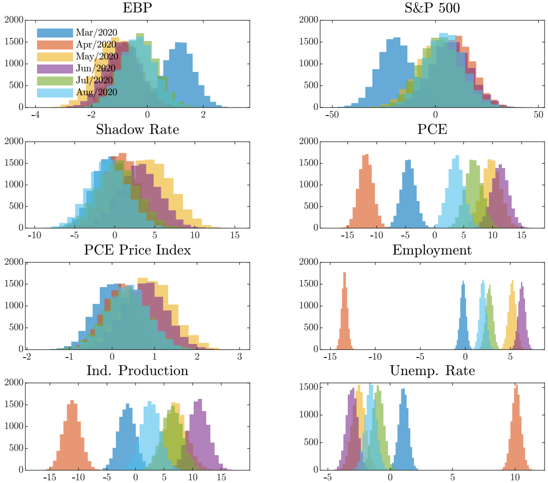
Empirical example

- Fixed overall prior tightness $\lambda = 0.2$
- Six individual dummies for March 2020 through August 2020 $\rightarrow$ onset of the pandemic, extreme observations in unemployment rate and industrial production
- Optimal $\phi^* = 0.05$, distant from the grid boundary cases (0.001 for uninformative Pandemic Priors and 5 for Minnesota Prior)

Pandemic Priors matter for estimation, ...

- Substantial heterogeneity: size, timing, and persistence

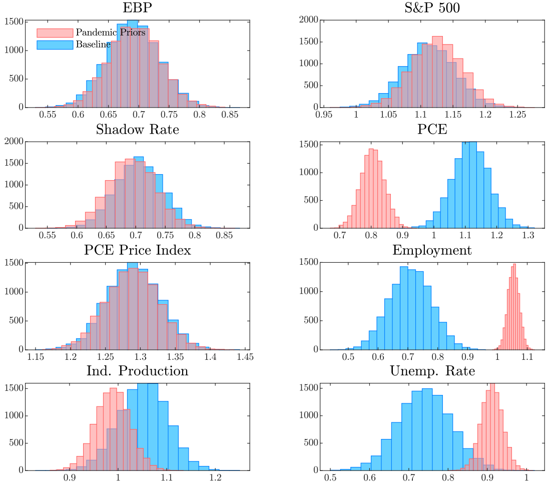
Figure: Posterior draws for the intercept and pandemic dummies



Pandemic Priors matter for estimation, ...

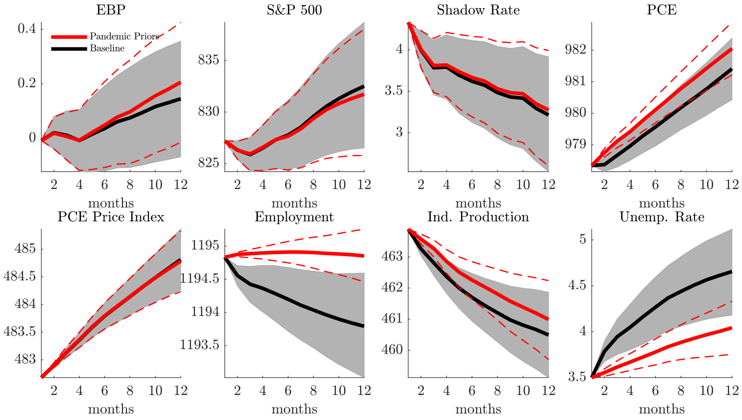
- **Minnesota Prior** (baseline) and **Pandemic Priors**

Figure: Posterior draws for the autoregressive coefficients



..., for forecasts, ...

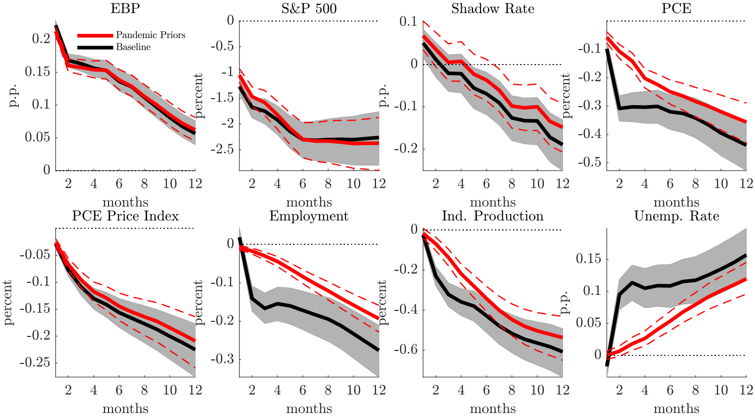
Figure: Unconditional forecasts as of December 2022



..., and for the identification of structural shocks

- Excess Bond Premium shock (Gilchrist and Zakrajšek, 2012 and Caldara, Fuentes-Albero, Gilchrist, and Zakrajšek, 2016)

Figure: Impulse responses to a 1 s.d. EBP shock



Comparison to alternative methods

Schorfheide and Song (2020):

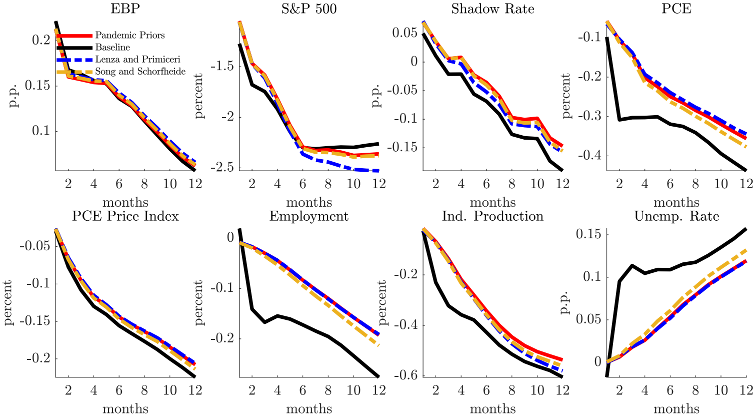
- Advocate for excluding the extreme observations from March to June 2020
- Pandemic Priors nest the procedure with uninformative priors ($\phi \to 0$) as a boundary case

Lenza and Primiceri (2021):

- Shock volatilities scaled up by the same constant, and same persistence (commonality assumption)
- Pandemic Priors allow for heterogeneous shifts (timing and size) and rate of decay $\to$ similar results, but simpler and more flexible

Comparison to alternative methods

Figure: Comparison of impulse responses to a 1 s.d. EBP shock

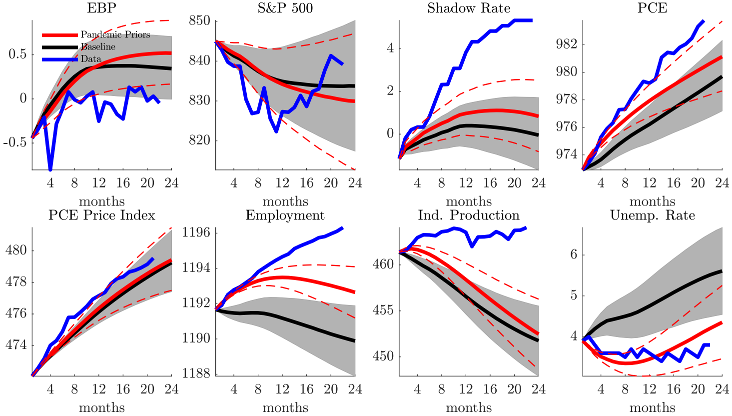


- Distortion comes from the early months of the pandemic
- IRFs are similar, but not the same

A reality check

- What happened since December 2021?

Figure: Unconditional forecast as of December 2021 *vs* Data



Conclusion

- Extreme observations blur our interpretation of macroeconomic historical relationships and economic effects of shocks
- Easy and straightforward way of dealing with such episodes, accepting that there is uncertainty about their potential outcome
- Pandemic Priors...
  - ...recover historical relationships and the proper identification of structural shocks
  - ...accommodates any state-of-the-art structural identification
  - allow policymakers to make well-informed decisions about responses to economic shocks going forward

MATLAB, Julia, and Python codes at www.danilocascaldigarcia.com

Pandemic Priors

Danilo Cascaldi-Garcia
Federal Reserve Board

Real-Time Data Analysis, Methods, and Applications
October 20, 2023

Monte Carlo simulation

Monte Carlo simulation - Coefficients

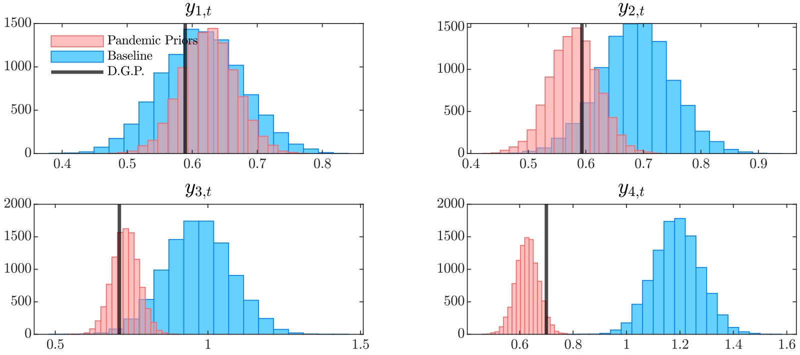
$$\mathbf{C} = \begin{bmatrix} 0.10 \\ 0.15 \\ 0.05 \\ 0.20 \end{bmatrix}, \quad \mathbf{D}_0 = \begin{bmatrix} 1 & 0.20 & -0.15 & -0.1 \\ 0 & 1 & -0.15 & 0.20 \\ 0 & 0 & 1 & -0.30 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{D}_1 = \begin{bmatrix} 0.65 & -0.10 & 0.10 & 0.05 \\ 0.20 & 0.60 & 0.10 & -0.10 \\ -0.10 & -0.20 & 0.65 & 0.15 \\ -0.05 & -0.15 & 0.20 & 0.80 \end{bmatrix}, \quad \mathbf{D}_2 = \begin{bmatrix} 0.15 & 0 & 0.05 & 0 \\ 0.10 & 0.10 & 0.05 & 0 \\ 0 & -0.01 & 0.10 & 0.05 \\ 0 & -0.05 & 0.10 & 0.10 \end{bmatrix},$$

Back

Monte Carlo simulation

Figure: Posterior draws for the autoregressive coefficients with uninformative ϕ prior

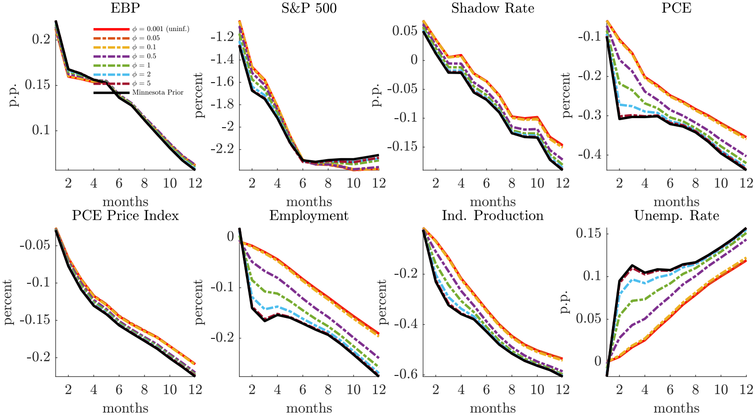


Back

Sensitivity to different levels of ϕ

Sensitivity to different levels of ϕ

Figure: Impulse responses to a 1 s.d. EBP shock under different ϕ levels



Pandemic Priors nest any setup for the prior belief about how much information one wants to be stemmed from the pandemic period

Álvarez, L. J., and Odendahl, F. (2022). Data outliers and Bayesian VARs in the Euro Area. Tech. rep.

Antolin-Diaz, J., Drechsel, T., and Petrella, I. (2021). Advances in nowcasting economic activity: The role of heterogeneous dynamics and fat tails. Tech. rep.

Bańbura, M., Giannone, D., and Reichlin, L. (2010). Large bayesian vector auto regressions. *Journal of Applied Econometrics*, *25(1)*, 71–92.

Bobeica, E., and Hartwig, B. (2023). The COVID-19 shock and challenges for inflation modelling. *International Journal of Forecasting*, *39*(1), 519–539.

Caldara, D., Fuentes-Albero, C., Gilchrist, S., and Zakrajšek, E. (2016). The macroeconomic impact of financial and uncertainty shocks. *European Economic Review*, *88*, 185–207.

Carriero, A., Clark, T. E., and Marcellino, M. (2015). Bayesian VARs: specification choices and forecast accuracy. *Journal of Applied Econometrics*, *30(1)*, 46–73.

Carriero, A., Clark, T. E., Marcellino, M., and Mertens, E. (2022). Addressing COVID-19 Outliers in BVARs with Stochastic Volatility. *The Review of Economics and Statistics*, (pp. 1–38).

Carriero, A., Kapetanios, G., and Marcellino, M. (2012). Forecasting government bond yields with large bayesian vector autoregressions. *Journal of Banking & Finance*, *36*(7), 2026–2047.

Gilchrist, S., and Zakrajšek, E. (2012). Credit spreads and business cycle fluctuations. *American Economic Review*, *102(4)*, 1692–1720.

Huber, F., Koop, G., Onorante, L., Pfarrhofer, M., and Schreiner, J. (2023). Nowcasting in a pandemic using non-parametric mixed frequency VARs. *Journal of Econometrics*, *232*(1), 52–69.

Lenza, M., and Primiceri, G. E. (2021). How to Estimate a VAR after March 2020. *Journal of Applied Econometrics (forthcoming)*.

Litterman, R. B. (1986). Forecasting with bayesian vector autoregressions: five years of experience. *Journal of Business & Economic Statistics*, *4(1)*, 25–38.

Ng, S. (2021). Modeling macroeconomic variations after COVID-19. Tech. rep., National Bureau of Economic Research.

Schorfheide, F., and Song, D. (2020). Real-time forecasting with a (standard) mixed-frequency VAR during a pandemic. *International Journal of Central Banking (forthcoming)*.

Wu, J. C., and Xia, F. D. (2016). Measuring the Macroeconomic Impact of Monetary Policy at the Zero Lower Bound. *Journal of Money, Credit and Banking*, *48*(2-3), 253–291.