The Ronald O. Perelman Center for Political Science and Economics (PCPSE)
133 South 36th Street
Philadelphia, PA 19104-6297

pier@econ.upenn.edu
http://economics.sas.upenn.edu/pier

PIER Working Paper
20-039

Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic

FRANK SCHORFHEIDE
University of Pennsylvania
CEPR, NBER, PIER

DONGHO SONG
Johns Hopkins
Carey Business School

July 8, 2020

https://ssrn.com/abstract=3737641

Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic

Frank Schorfheide*
University of Pennsylvania
CEPR, NBER, PIER

Dongho Song
Johns Hopkins
Carey Business School

This Version: July 8, 2020

Abstract

In this paper we resuscitate the mixed-frequency vector autoregression (MF-VAR) developed in Schorfheide and Song (2015) to generate real-time macroeconomic forecasts for the U.S. during the COVID-19 pandemic. The model combines eleven time series observed at two frequencies: quarterly and monthly. We deliberately do not modify the model specification in view of the recession induced by the COVID-19 outbreak. We find that forecasts based on a pre-crisis estimate of the VAR using data up until the end of 2019 appear to be more stable and reasonable than forecasts based on a sequence of recursive estimates that include the most recent observations. Overall, the MF-VAR outlook is quite pessimistic. The estimated MF-VAR implies that level variables are highly persistent, which means that the COVID-19 shock generates a long-lasting reduction in real activity. Regularly updated forecasts are available at `www.donghosong.com/`.

JEL CLASSIFICATION: C11, C32, C53

KEY WORDS: Bayesian inference; COVID-19; Macroeconomic Forecasting; Minnesota Prior; Real-time data; Vector autoregressions.

*Correspondence: F. Schorfheide: Department of Economics, 133 S. 36th Street, University of Pennsylvania, Philadelphia, PA 19104-6297. Email: schorf@ssc.upenn.edu. D. Song: Johns Hopkins Carey Business School, 100 International Drive, Baltimore, MD 21202. Email: dongho.song@jhu.edu. We thank Frank Diebold, Giorgio Primiceri, and Minchul Shin for helpful comments. Schorfheide gratefully acknowledges financial support from the National Science Foundation under Grant SES 1424843. The views expressed in this paper are those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System.

1 Introduction

Vector autoregressions (VARs) are widely used in empirical macroeconomics. A VAR is a multivariate time series model that can be used, for instance, to forecast individual time series, to predict comovements of macroeconomic or financial variables, to analyze sources of business cycle fluctuations, or to assess the effects of monetary or fiscal policy interventions on the macroeconomy. In this paper, we resuscitate the mixed-frequency VAR, henceforth MF-VAR, developed in Schorfheide and Song (2015). The model combines variables observed at monthly frequency, such as industrial production and unemployment, with variables observed at quarterly frequency, e.g., Gross Domestic Product (GDP).

We consider a MF-VAR for eleven macroeconomic variables, of which three are observed at quarterly frequency and eight are observed at monthly frequency. The quarterly series are GDP, Fixed Investment, and Government Expenditures. The monthly series are the Unemployment Rate, Hours Worked, Consumer Price Index, Industrial Production Index, Personal Consumption Expenditure, Federal Funds Rate, 10-year Treasury Bond Yield, and S&P 500 Index. The MF-VAR can be conveniently represented as a state-space model, in which the state-transition equations are given by a VAR at monthly frequency and the measurement equations relate the observed series to the underlying, potentially unobserved, monthly variables that are stacked in the state vector. To cope with the high dimensionality of the parameter space, the MF-VAR is equipped with a Minnesota prior and estimated using Bayesian methods.

The recent COVID-19 pandemic has triggered long-lasting mobility restrictions in the form of stay-at-home orders across the U.S. In turn, economic activity has collapsed in many sectors and unemployment has soared. This creates a tremendous challenge for macroeconomic forecasting. An important question is how the most recent observations should be treated at the estimation stage. One option is to regard them as outliers generated by large shocks and exclude them from the estimation. Alternatively, one could view them as an indication of structural change and let them influence the parameter estimates in some form. In this paper we examine the accuracy of forecasts from a MF-VAR model that performed well before the pandemic using two sets of estimates: (i) an estimate based on a fixed sample that ends in 2019; (ii) a sequence of estimates based on the most recent data, including those from the pandemic.

We find that forecasts based on a pre-crisis estimate of the VAR using data up until the end of 2019 appear to be more stable and reasonable than forecasts based on a sequence of recursive estimates that include the most recent observations. Overall, the MF-VAR outlook is quite pessimistic. The estimated MF-VAR implies that level variables are highly persistent, which means that the COVID-19 shock generates a long-lasting reduction in real activity.

The pre-COVID forecast performance of the MF-VAR model used in this paper was documented in Schorfheide and Song (2015). We showed that the MF-VAR generates more accurate nowcasts and short-run forecasts than a VAR estimated on time-aggregated quarterly data. The improvement tempers off in the medium and long run. The short-run accuracy gain is largest in the third month of the quarter when a lot of monthly data are available for the current quarter. We also documented that the monthly information helped the MF-VAR track the economic downturn during the 2008-09 (Great) recession period more closely in real time than a VAR estimated on quarterly data only.

There is a rapidly growing literature of applying and adapting existing macroeconomic forecasting techniques to generate forecasts for the current pandemic recession. Most of the techniques involve the combination of time series that are observed at different frequencies, e.g., daily, weekly, monthly, or quarterly, and are released asynchronously. Many of the models, including ours, are cast into state-space form, while others take the form of mixed-data sampling (MIDAS) regressions.

Carriero, Clark, and Marcellino (2020) compare tail risk nowcasts of economic activity from a variety of macroeconomic forecasting models. Foroni, Marcellino, and Stevanovic (2020) generate GDP growth nowcasts and forecasts for the COVID-19 recession from a variety of mixed-frequency MIDAS models and explore adjustments of these forecasts based on the forecasting experience during the Great Recession. Babii, Ghysels, and Stiaukas (2020) use LASSO techniques to estimate MIDAS regressions for GDP nowcasts that also include text data. There is a related literature on real-activity tracking. Diebold (2020) studies the performance of the Aruoba-Diebold-Scotti (ADS) index over the past decade and documents its behavior since the outbreak of the COVID-19 pandemic. Lewis, Mertens, and Stock (2020) developed a weekly economic index (WEI) to track the rapid economic developments triggered by the coronavirus pandemic.[1]

By historical standards, many of the recent time series observations are (extreme) outliers,

[1]Some authors publish regular updates of their forecasts and activity indices. See, for instance, `www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index` (ADS index), `www.newyorkfed.org/research/policy/weekly-economic-index` (WEI), `www.midasml.com/` (MIDAS LASSO).

meaning they are several standard deviations away from their historical averages or trends. This raises important modeling questions. Primiceri and Tambalotti (2020) forecast the dynamic effect of COVID-19 on the U.S. economy by imposing the assumption that the propagation of the COVID shock is potentially different from other shocks, allowing for a potentially faster recovery than one would predict based on the persistence in the historical series. The presence of outliers in the sample raises the important question of how to weigh them during the estimation. Lenza and Primiceri (2020) introduce breaks in shock variances to essentially down-weigh the recent observations when estimating a VAR.

The remainder of this paper is organized as follows. Section 2 reviews the specification of the MF-VAR. The empirical results are presented in Section 3, and Section 4 concludes. Additional information about the construction of our data set is provided in the Online Appendix. Regularly updated forecasts are available at `www.donghosong.com/`.

2 MF-VAR Specification

We reproduce the model description from Schorfheide and Song (2015) and refer the reader for a detailed discussion of the Bayesian computations to our original paper. We assume that the economy evolves at monthly frequency according to the following VAR(p) dynamics:

$$x_t = \Phi_1 x_{t-1} + \ldots + \Phi_p x_{t-p} + \Phi_c + u_t, \quad u_t \sim iidN\big(0, \Sigma\big). \tag{1}$$

The $n \times 1$ vector of macroeconomic variables x_t can be composed into $x_t = [x'_{m,t}, x'_{q,t}]'$, where the $n_m \times 1$ vector $x_{m,t}$ collects variables that are observed at monthly frequency, e.g., the consumer price index and the unemployment rate, and the $n_q \times 1$ vector $x_{q,t}$ comprises the unobserved monthly variables that are published only at quarterly frequency, e.g., GDP. Define $z_t = [x'_t, \ldots, x'_{t-p+1}]'$ and $\Phi = [\Phi_1, \ldots, \Phi_p, \Phi_c]'$. Write the VAR in (1) in companion form as

$$z_t = F_1(\Phi) z_{t-1} + F_c(\Phi) + v_t, \quad v_t \sim iidN\big(0, \Omega(\Sigma)\big), \tag{2}$$

where the first n rows of $F_1(\Phi)$, $F_c(\Phi)$, and v_t are defined to reproduce (1) and the remaining rows are defined to deliver the identities $x_{q,t-l} = x_{q,t-l}$ for $l = 1, \ldots, p-1$. The $n \times n$ upper-left submatrix of Ω equals Σ and all other elements are zero. Equation (2) is the state-transition equation of the MF-VAR.

We proceed by describing the measurement equation. To handle the unobserved variables we vary the dimension of the vector of observables as a function of time t (e.g., Durbin and Koopman (2001)). Let T denote the forecast origin and let $T_b \leq T$ be the last period that corresponds to the last month of the quarter for which all quarterly observations are available. The subscript b stands for *balanced* sample. Up until period T_b the vector of monthly series $x_{m,t}$ is observed every month. We denote the actual observations by $y_{m,t}$ and write

$$y_{m,t} = x_{m,t}, \quad t = 1, \dots, T_b. \tag{3}$$

Assuming that the underlying monthly VAR has at least three lags, that is, $p \geq 3$, we express the three-month average of $x_{q,t}$ as

$$\tilde{y}_{q,t} = \frac{1}{3}(x_{q,t} + x_{q,t-1} + x_{q,t-2}) = \Lambda_{qz} z_t. \tag{4}$$

For variables measured in logs, e.g., $\ln GDP$, the formula can be interpreted as a log-linear approximation to an arithmetic average of GDP that preserves the linear structure of the state-space model. For flow variables such as GDP, we adopt the NIPA convention and annualize high-frequency flows. As a consequence, quarterly flows are the average and not the sum of monthly flows. This three-month average, however, is only observed for every third month, which is why we use a tilde superscript. Let $M_{q,t}$ be a selection matrix that equals the identity matrix if t corresponds to the last month of a quarter and is empty otherwise. Adopting the convention that the dimension of the vector $y_{q,t}$ is n_q in periods in which quarterly averages are observed and zero otherwise, we write

$$y_{q,t} = M_{q,t}\tilde{y}_{q,t} = M_{q,t}\Lambda_{qz} z_t, \quad t = 1, \dots, T_b. \tag{5}$$

For periods $t = T_b + 1, \dots, T$ no additional observations of the quarterly time series are available. Thus, for these periods the dimension of $y_{q,t}$ is zero and the selection matrix $M_{q,t}$ in (5) is empty. However, the forecaster might observe additional monthly variables. Let $y_{m,t}$ denote the subset of monthly variables for which period t observations are reported by the statistical agency after period T, and let $M_{m,t}$ be a deterministic sequence of selection matrices such that (3) can be extended to

$$y_{m,t} = M_{m,t} x_{m,t}, \quad t = T_b + 1, \dots, T. \tag{6}$$

Notice that the dimension of the vector $y_{m,t}$ is potentially time varying and less than n_m.

The measurement equations (3) to (6) can be written more compactly as

$$y_t = M_t \Lambda_z z_t, \quad t = 1, \ldots, T. \tag{7}$$

Here, M_t is a sequence of selection matrices that selects the time t variables that have been observed by period T and are part of the forecaster's information set. In sum, the state-space representation of the MF-VAR is given by (2) and (7).

The starting point of Bayesian inference for the MF-VAR is a joint distribution of observables $Y_{1:T}$, latent states $Z_{0:T}$, and parameters (Φ, Σ), conditional on a pre-sample $Y_{-p+1:0}$ to initialize lags. The distribution of observables and latent states conditional on the parameters is implied by the above state-space representation of the MF-VAR. For the marginal distribution of the parameters (Φ, Σ) we use a conjugate Minnesota prior. This prior dates back to Litterman (1980) and Doan, Litterman, and Sims (1984). We use the version of the Minnesota prior described in Del Negro and Schorfheide (2011)'s handbook chapter, which in turn is based on Sims and Zha (1998). The main idea of the Minnesota prior is to center the distribution of Φ at a value that implies a random-walk behavior for each of the components of x_t in (1). We implement the Minnesota prior by mixing artificial (or *dummy*) observations into the estimation sample. The artificial observations are computationally convenient and allow us to generate plausible a priori correlations between VAR parameters. The variance of the prior distribution is controlled by a low-dimensional vector of hyperparameters λ.

We generate draws from the posterior distributions of $(\Phi, \Sigma)|Z_{0:T}$ and $Z_{0:T}|(\Phi, \Sigma)$ using a Gibbs sampler. Based on these draws, we are able to simulate future trajectories of y_t to characterize the predictive distribution associated with the MF-VAR and to calculate point and density forecasts.

3 Empirical Results

3.1 Data

We consider a MF-VAR for eleven macroeconomic variables, of which three are observed at quarterly frequency and eight are observed at monthly frequency. The quarterly series are GDP, fixed investment (INVFIX), and government expenditures (GOV). The monthly series are the unemployment rate (UNR), hours worked (HRS), Consumer Price Index (CPI), Industrial Production Index (IP), Personal Consumption Expenditure (PCE), Federal Funds

Table 1: Information at Forecast Origin

January 31

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q4	M12	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q1	M1	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

April 30

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

May 31

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M5	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

June 30

		UNR	HRS	CPI	IP	PCE	FF	TB	SP500	GDP	INVFIX	GOV
Q1	M3	X	X	X	X	X	X	X	X	QAv	QAv	QAv
Q2	M4	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M5	X	X	X	X	X	X	X	X	∅	∅	∅
Q2	M6	∅	∅	∅	∅	∅	X	X	X	∅	∅	∅

Notes: $\emptyset$ indicates that the observation is missing. X denotes monthly observation and QAv denotes quarterly average.

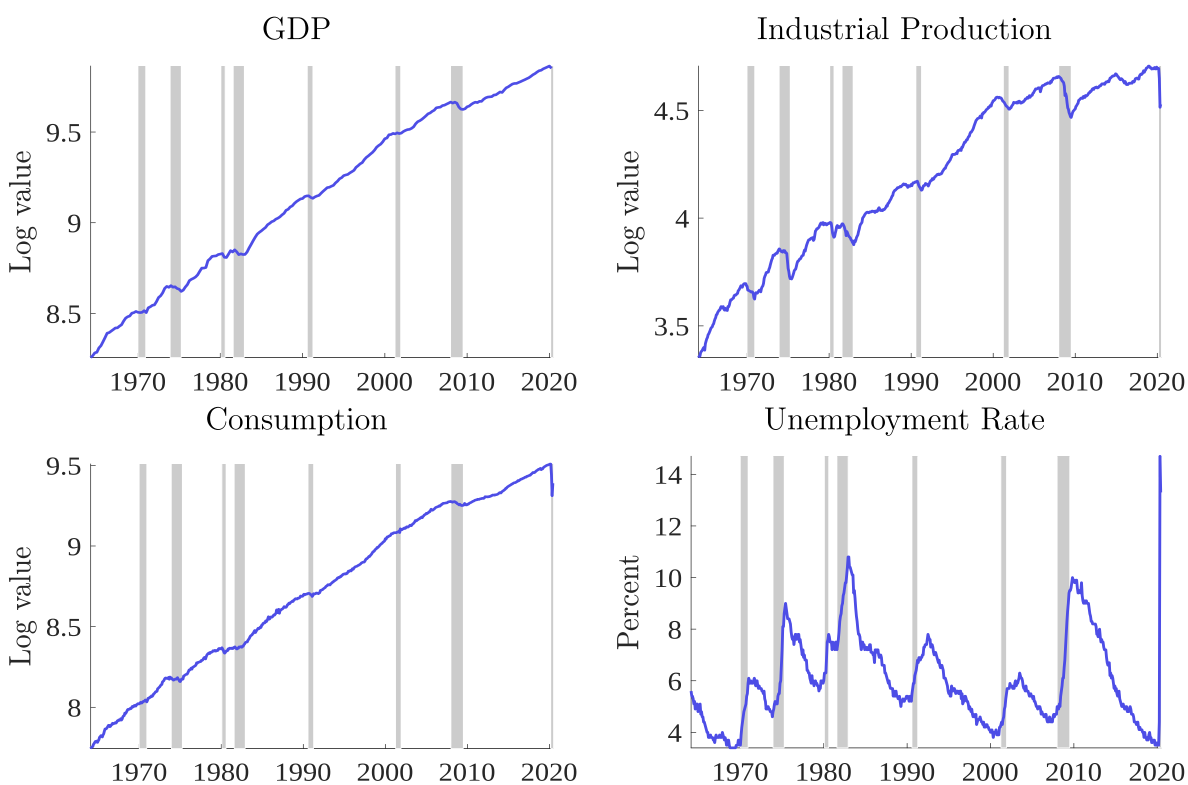
Rate (FF), 10-year Treasury Bond Yield (TB), and S&P 500 Index (SP500). Precise data definitions are provided in the Online Appendix. Series that are observed at a higher than monthly frequency are time-aggregated to monthly frequency. The variables enter the MF-VAR in log levels with the exception of UNR, FF, and TB, which are divided by 100 to make them commensurable in scale to the other log-transformed variables.

Our forecasts are based on real-time data sets, assuming that the econometric analysis is conducted on the last day of each month.[2] The timing convention and the data availability

[2]Due to data revisions by statistical agencies, observations of $Y_{1:T-1}$ published in period T are potentially different from the observations that have been published in period $T-1$. Moreover, some series are published

for each forecast origin are summarized in Table 1. A forecaster on January 31 had access to monthly observations from December, to an initial release of Q4 GDP, investment, and government spending, as well as the January observations for the average federal funds rate, the Treasury bond yield, and the S&P500. Three months later, on April 30, the information set is similar. The only monthly variables available for Q2 are the April financial variables. In May, monthly observations on the April unemployment rate, hours worked, inflation, industrial production, and personal consumption become available. Our last forecast was generated on June 30. At this point, two monthly observations for each non-financial variable are available for the second quarter.

Figure 1: Data



Notes: The data are obtained from the June 2020 vintage, staring in 1964. The shaded bars indicate the NBER recession dates.

We plot four of the eleven time series in Figure 1. GDP is available only at quarterly frequency. In June 2020, we have the Q1 observation, which does not show any substantial decline. Industrial Production, Personal Consumption Expenditures, and the Unemployment Rate are available at monthly frequency and exhibit large end-of-sample movements. Both

with a delay of several periods.

IP and PCE drop considerably, almost to Great Recession levels, whereas the unemployment rate soars to 15%, a level that is unprecedented in the past six decades. The graph suggests that the monthly variables are likely to be very influential for the forecasts of the quarterly variables.

3.2 Forecasts

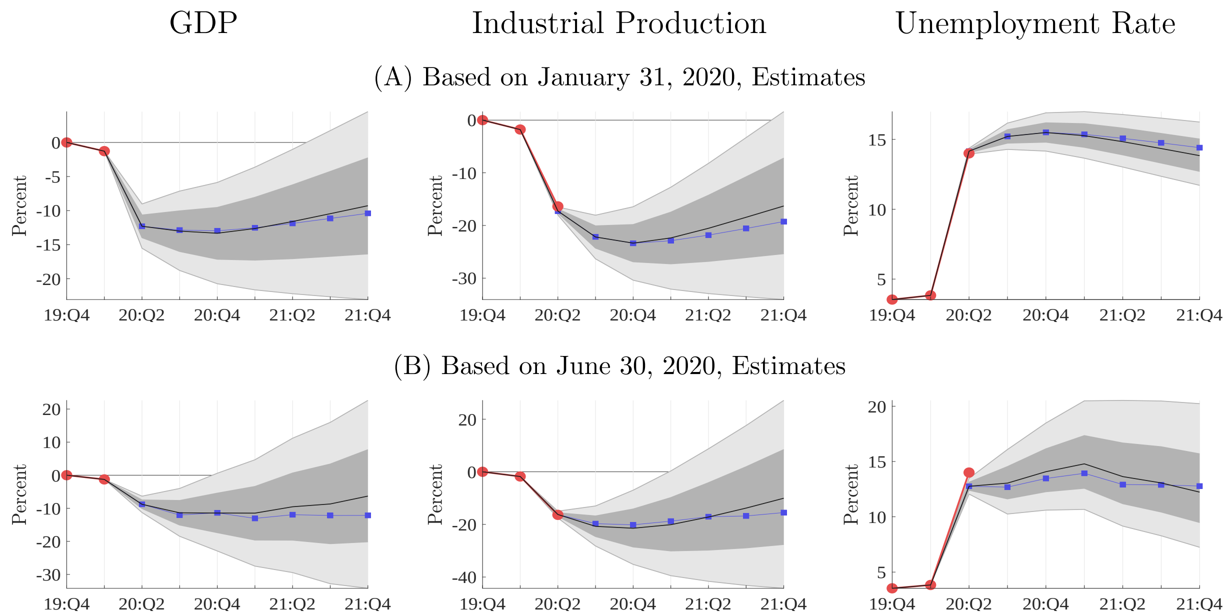
We estimate the MF-VAR using $p = 6$ lags based on the following 2020 data vintages and forecast origins: January 31, April 30, May 31, and June 30. Our estimation samples start in 1964. We optimize the hyperparameters based on the January vintage and keep them fixed for subsequent estimation samples. For each of the eleven variables we aggregate the monthly forecasts from the MF-VAR to quarterly averages. Results are summarized in Figures 2, 3, 4, and 5.

We begin by examining the effect of choosing the endpoint of the estimation sample on the forecasts. The COVID-19 pandemic generated unprecedented job losses and unemployment benefit claims. Moreover, industrial production dropped by more than 15% in April. This raises the question of whether to include observations past January 2020 in the estimation sample. If the pandemic was a shock to the economy that was unusually large, indeed several standard deviations in magnitude, but did not change the fundamental workings of the aggregate economy, then it is best to exclude the most recent observations from the estimation. Unless explicitly modeled, the outliers will simply distort the parameter estimates.

If, on the other hand, the pandemic created a fundamental change in how macroeconomic variables interact with each other, then the most recent observations should be included in the estimation, and earlier observations should be discounted. Unfortunately, the discounting is currently infeasible, because the MF-VAR has a large number of parameters and only observations starting in April 2020 reflect the impact of the pandemic. We begin the empirical analysis by comparing June 30, 2020, forecasts in Figure 2 that are based on two different parameter estimates: (i) the monthly-updated posterior distribution of (Φ, Σ) based on the most recent data vintage, and (ii) the posterior distribution of (Φ, Σ) conditional on the January 2020 vintage.

The panels of Figure 2 (and the subsequent figures) show actual values from the most

Figure 2: Effect of Estimation Sample on Forecasts, Origin is June 30, 2020



Notes: We forecast quarterly averages. Actual values (solid red) and forecasts: median (solid black), 60% bands (dark grey), and 90% bands constructed from the posterior predictive distribution. Solid blue line represents point forecasts obtained by fixing the federal funds rate at 5 basis points. For GDP and Industrial Production we depict percentage change relative to December 31, 2019.

recent vintage (solid red) and forecasts generated from the posterior predictive distribution.[3] The solid black line corresponds to the median and the bands are constructed to have posterior coverage probabilities of 60% and 90%, respectively. We also show a second posterior median forecast that is obtained by setting the federal funds rate to 5 basis points (bp) over the forecast horizon. This is a crude way of imposing the effective lower bound (ELB) on the federal funds rate. Because we are forecasting quarterly averages, the tick marks on the x-axis of the panels correspond to quarters.

In view of the large drop in Industrial Production (by about 18% from the beginning of the year) and the large spike in the unemployment rate (from 3% to almost 15%), the MF-VAR forecasts an equally large drop in GDP. According to the posterior median forecast, GDP will be 15% lower by the end of the second quarter than it was at the beginning of 2020. It is well-known that macroeconomic aggregates are very persistent and exhibit (near) unit-root dynamics. This persistence is reflected in the forecasts. For none of the three

[3] Discrepancies between the red solid lines (actuals) and the black lines prior to the forecast origin may arise due to data revisions. Discrepancies at the nowcast horizon are due to the fact that the red line is constructed by averaging the available observations, e.g., from the first and second month of the quarter, to obtain a quarterly value. The black line represents a nowcast from the MF-VAR.

variables do the posterior median forecasts imply a recovery over the next six months.

While the posterior mean forecasts for the two sets of estimates in rows (A) and (B) of the figure are quite similar, the predictive bands are markedly different. The bands obtained from the June 30 estimates are considerably wider. The increase in width is mainly driven by the estimates of Σ which increased due to the extreme observations in the second quarter of 2020. Going forward, we will focus on forecasts generated based on the January 31 parameter estimates because after the initial adjustment of the economy to the COVID-19 pandemic we expect the magnitude of subsequent shocks to be more similar to the pre-2020 experience. Nonetheless, even the predictive intervals for GDP and IP reported in row (A) of Figure 2 are wide. Based on the January 31, 2020, estimates, the 90% credible interval for 2021:Q4 ranges from approximately 1% to 30% below the 2019:Q4 level.

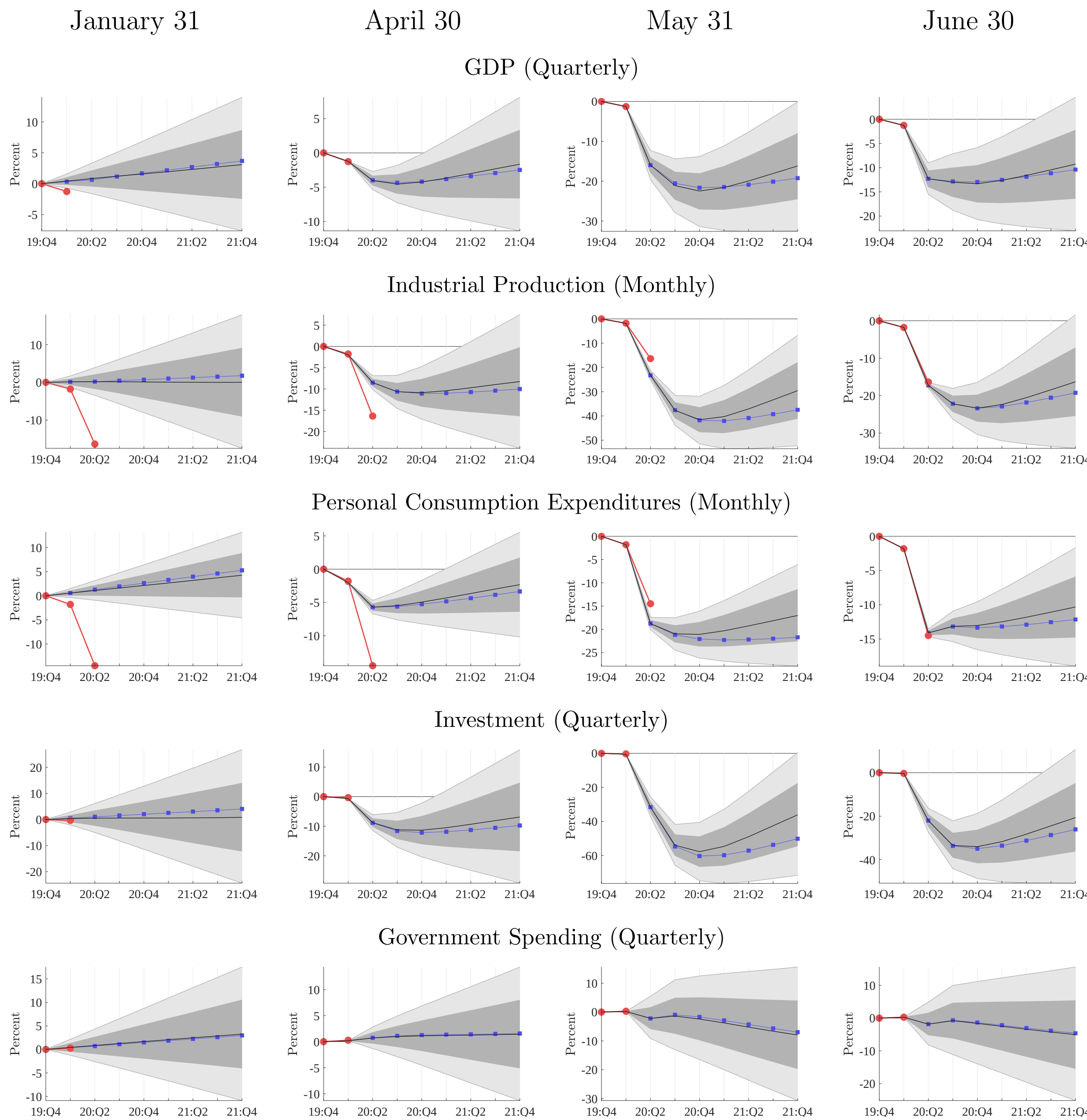
In Figure 3 we display forecasts for real activity measures that are generated based on four different information sets: January 31, April 30, May 31, and June 30. To the extent that actuals are available, they are represented by the solid red lines. The forecasts are conditional on the January posterior distribution for the parameters (Φ, Σ) and represent % changes relative to the December 31, 2019 level of the variables. Each row in the table corresponds to a different series and we report whether the series is observed at monthly or quarterly frequency. For the quarterly series a Q2 actual is not yet available. For the monthly series we generate a quarterly value by averaging the available monthly observations; see Table 1.

While the January posterior median forecast for GDP was essentially flat, bad news about industrial production and unemployment in April led to a downward revision. According to the point forecast, GDP is expected to be about 5% lower by 2021:Q4 relative to 2019:Q4. The 90% band ranges from -15% to +3%. The monthly observations released in May generate an even bleaker outlook. The 18-month-ahead posterior median forecast falls below -20%. The June data contain a bit of "good" news in that the median forecast rises to -15% with a 90% interval from -30% to 0% change relative to 2019:Q4.

The pessimistic outlook for GDP is driven by the large drop in Industrial Production and Personal Consumption Expenditures in April and May. Because the system has very little mean reversion of the levels, the model translates the fall in the production and consumption measures into a long-lasting depression of real activity.

Figure 4 presents forecasts for two labor market measures: the unemployment rate and hours worked. The evolution of the forecasts and the actuals is similar to that of the real

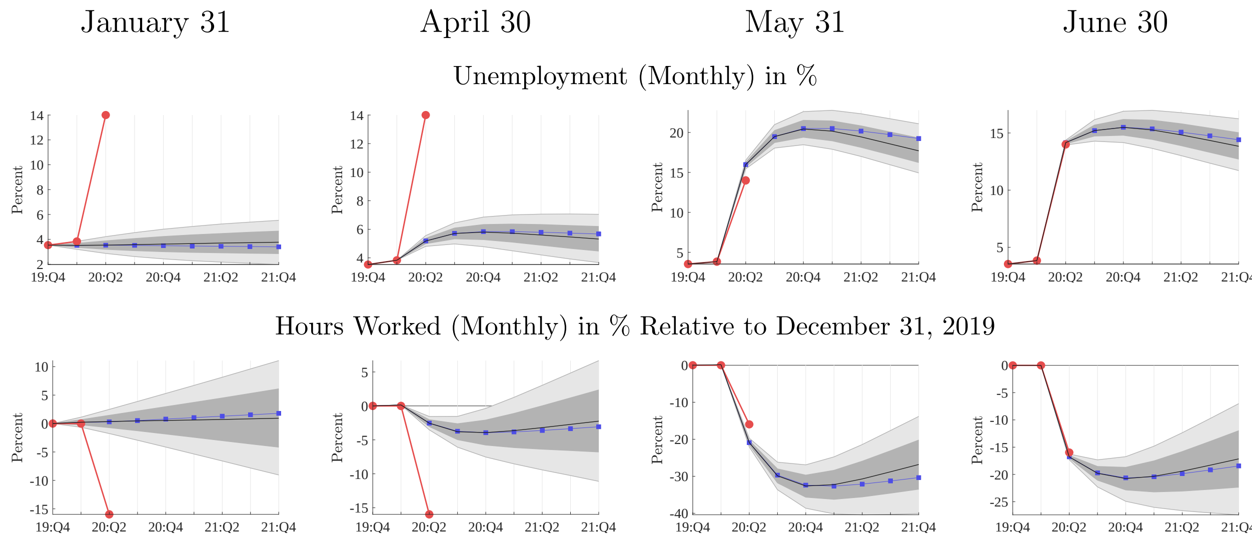
Figure 3: Evolution of Forecasts: Real Activity in % Relative to December 31, 2019



Notes: We forecast quarterly averages. Actual values (solid red) and forecasts: median (solid black), 60% bands (dark grey), and 90% bands constructed from the posterior predictive distribution. Solid blue line represents point forecasts obtained by fixing the federal funds rate at 5 basis points. Forecasts are based on the posterior of (Φ, Σ) from the January 31, 2020, vintage.

activity measures, except that the predictive intervals are generally tighter. Based on the March data that are available by the end of April, there is no indication yet of the upcoming

Figure 4: Evolution of Forecasts: Labor Market



Notes: We forecast quarterly averages. Actual values (solid red) and forecasts: median (solid black), 60% bands (dark grey), and 90% bands constructed from the posterior predictive distribution. Solid blue line represents point forecasts obtained by fixing the federal funds rate at 5 basis points. Forecasts are based on the posterior of (Φ, Σ) from the January 31, 2020, vintage.

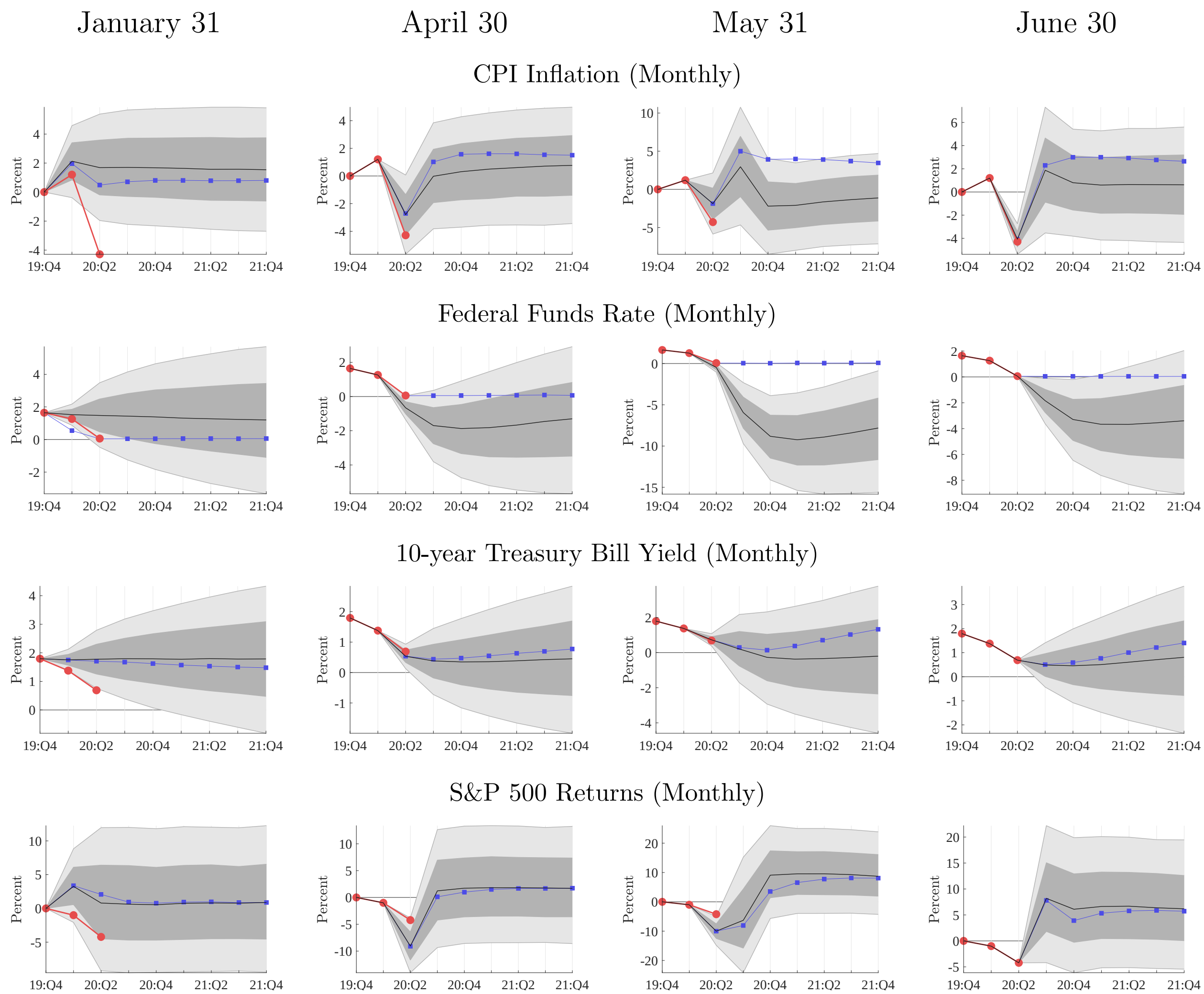
depression. The long-run forecast of unemployment is between 0% and 7% and hours worked are expected to change by between -12% and +3%. The outlook becomes significantly worse in May and slightly improves in June.

Figure 5 shows forecasts of inflation and the financial variables. The model predicts a short-lived deflation and a counterfactual negative federal funds rate. To impose the ELB we generate an alternative set of forecasts in which we simply set the short-term nominal interest rate to zero. This can be viewed as a poor-man's version of a VAR with censoring; see Aruoba, Schorfheide, and Villalvazo (2020). Notice that imposing the ELB onto the forecasts by fixing the nominal interest rate at zero leads to an upward revision of the inflation forecast by roughly 100 bp and a decrease of the production and consumption forecasts. The unemployment and hours worked forecasts remain essentially unaffected.

4 Conclusion

The goal of this paper was to resuscitate the mixed-frequency vector autoregression (MF-VAR) developed in Schorfheide and Song (2015) to generate real-time macroeconomic fore-

Figure 5: Evolution of Forecasts: Inflation and Financial Variables in Annualized %



Notes: We forecast quarterly averages. Actual values (solid red) and forecasts: median (solid black), 60% bands (dark grey), and 90% bands constructed from the posterior predictive distribution. Solid blue line represents point forecasts obtained by fixing the federal funds rate at 5 basis points. Forecasts are based on the posterior of (Φ, Σ) from the January 31, 2020, vintage.

casts for the U.S. during the COVID-19 pandemic. The model combines eleven time series observed at two frequencies: quarterly and monthly. We find that it is preferable to exclude the most recent observations from the estimation sample. The estimated MF-VAR implies that level variables are highly persistent. This means that the shock triggered by the COVID-19 outbreak will lead to a long-lasting depression of economic activity. Time will tell whether this prediction is accurate, or whether it is possible to re-start the economy quickly, shortening the duration of the recessionary effect that the shock has on the economy, and to recover by the end of 2021.

References

Aruoba, B., F. Schorfheide, and S. Villalvazo (2020): "SVARs with Occasionally-Binding Constraints," *Manuscript, University of Pennsylvania.*

Babii, A., E. Ghysels, and J. Stiaukas (2020): "Machine Learning Time Series Regressions with an Application to Nowcasting," *arXiv Working Paper*, 2005.14057v2.

Carriero, A., T. E. Clark, and M. Marcellino (2020): "Nowcasting Tail Risks to Economic Activity with Many Indicators," *FRB Cleveland Working Paper*, 20-13.

Del Negro, M., and F. Schorfheide (2011): "Bayesian Macroeconometrics," in *The Oxford Handbook of Bayesian Econometrics*, ed. by J. Geweke, G. Koop, and H. van Dijk, pp. 293–389. Oxford University Press.

Diebold, F. X. (2020): "Real-Time Real Economic Activity: Exiting the Great Recession and Entering the Pandemic Recession," *Manuscript, University of Pennsylvania.*

Doan, T., R. Litterman, and C. A. Sims (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3(1), 1–100.

Durbin, J., and S. J. Koopman (2001): *Time Series Analysis by State Space Methods.* Oxford University Press.

Foroni, C., M. Marcellino, and D. Stevanovic (2020): "Forecasting the Covid-19 Recession and Recovery: Lessons from the Financial Crisis," *Cirano Working Paper*, 2020S-32.

Lenza, M., and G. E. Primiceri (2020): "How to Estimate a VAR after March 2020," *Manuscript, Northwestern University.*

Lewis, D., K. Mertens, and J. Stock (2020): "U.S. Economic Activity During the Early Weeks of the SARS-Cov-2 Outbreak," *FRB Dallas Working Paper*, 20-11.

Litterman, R. B. (1980): "Techniques for Forecasting with Vector Autoregressions," Ph.D. thesis, University of Minnesota.

Primiceri, G. E., and A. Tambalotti (2020): "Macroeconomic Forecasting in the Time of COVID-19," *Manuscript, Northwestern University.*

Schorfheide, F., and D. Song (2015): "Real-Time Forecasting with a Mixed-Frequency VAR," *Journal of Business & Economic Statistics*, 33(3), 366–380.

Sims, C. A., and T. Zha (1998): "Bayesian Methods for Dynamic Multivariate Models," *International Economic Review*, 39(4), 949–968.

Supplemental Online Appendix to "Real-Time Forecasting with a (Standard) Mixed-Frequency VAR During a Pandemic"

Frank Schorfheide and Dongho Song

A Data Set

The eleven real-time macroeconomic data series are obtained from the ALFRED database maintained by the Federal Reserve Bank of St. Louis. Table A-1 summarizes how the series used in this paper are linked to the series provided by ALFRED. The recent vintages of

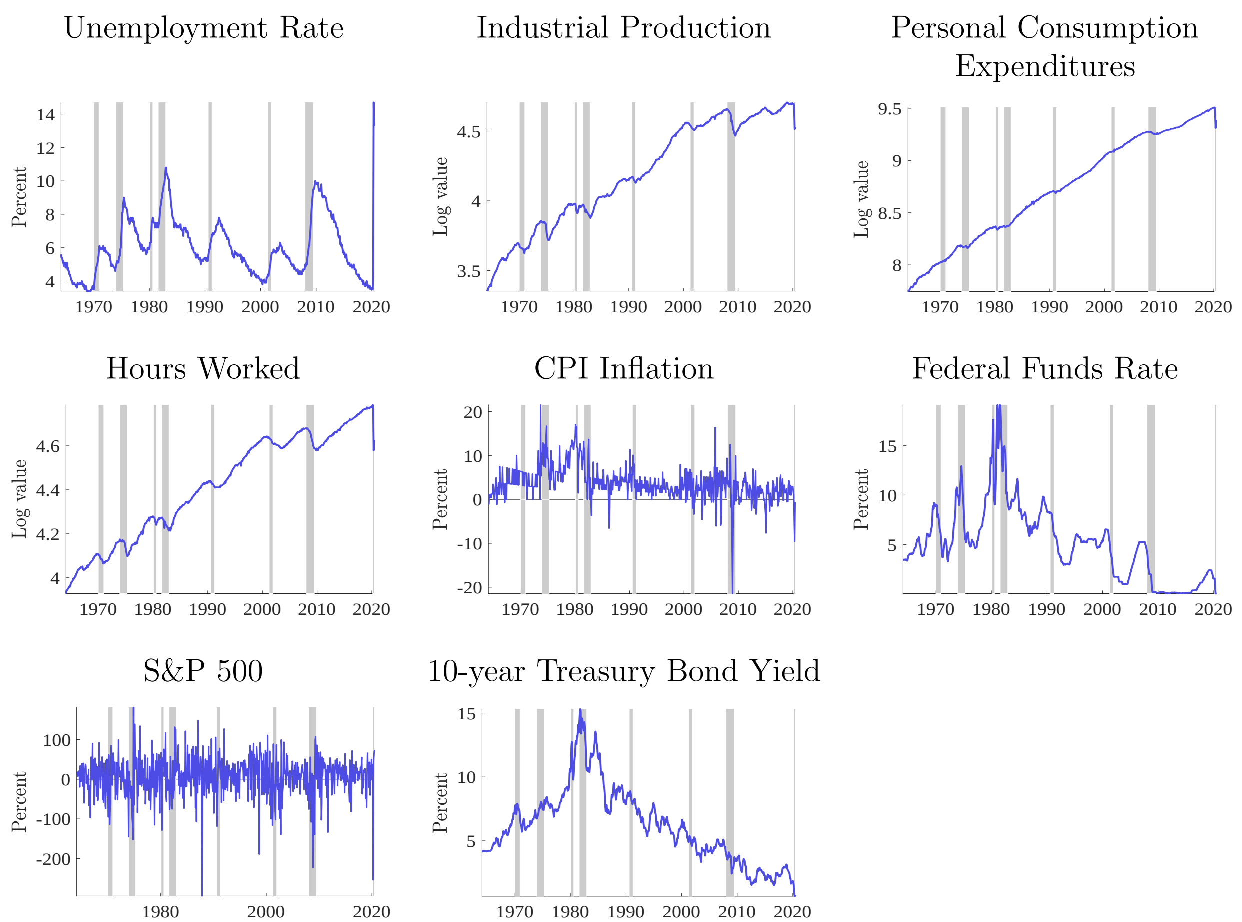
Table A-1: ALFRED Series Used in Analysis

Time Series	ALFRED Name
Gross Domestic Product (GDP)	GDPC1
Fixed Investment (INVFIX)	FPIC1
Government Expenditures (GOV)	GCEC1
Unemployment Rate (UNR)	UNRATE
Hours Worked (HRS)	AWHI
Consumer Price Index (CPI)	CPIAUCSL
Industrial Production Index (IP)	INDPRO
Personal Consumption Expenditure (PCE)	PCEC96
Federal Fund Rate (FF)	FEDFUNDS
10-year Treasury Bond Yield (TB)	GS10
S&P 500 (SP500)	SP500

PCE and INVFIX from ALFRED do not include data prior to 2002. However, the most recent data for PCE and INVFIX can be obtained from BEA or NIPA Tables. Specifically, we download "Table 2.8.3. Real Personal Consumption Expenditures by Major Type of Product, Monthly, Quantity Indexes" for PCE and "Table 5.3.3. Real Private Fixed Investment by Type, Quantity Indexes" for INVFIX, which are available from 1/1/1959 and 1/1/1948 to current periods, respectively. First, we compute the growth rates from the quantity indexes. Based on the computed growth rates, we can backcast historical series up to 1/1/1964 using the 1/1/2002 data points as initializations. We think this is a reasonable way to construct the missing points for PCE and INVFIX.

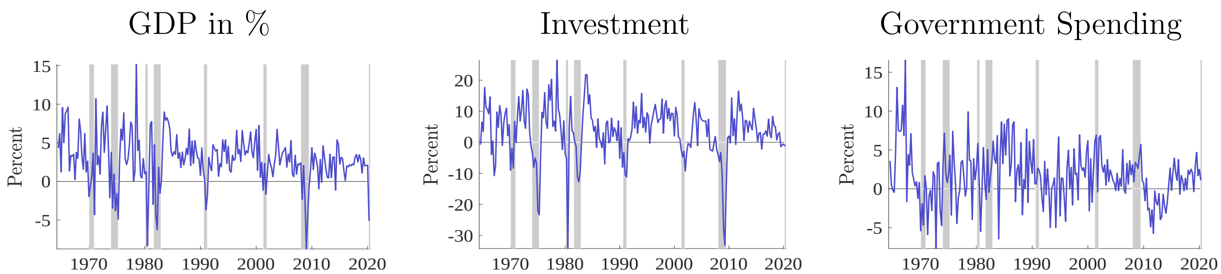
Figures A-1 and A-2 provide the time series plot of our eleven macroeconomic variables obtained from the June 2020 vintage.

Figure A-1: Monthly Data



Notes: M-o-M percentage changes are annualized. The data are obtained from the June 2020 vintage, starting from 1964. The shaded bars indicate the NBER recession dates.

Figure A-2: Quarterly Data, Q-o-Q Growth Rates in Annualized %



Notes: The data are obtained from the June 2020 vintage, starting in 1964. The shaded bars indicate the NBER recession dates.