Reconciled Estimates of Monthly GDP in the US*

Gary Koop[1], Stuart McIntyre[2], James Mitchell[3], and Aubrey Poon[4]

[1]University of Strathclyde, UK
[2]University of Strathclyde, UK
[3]Federal Reserve Bank of Cleveland, USA
[4]Örebro University, Sweden

January 13, 2022

Abstract

In the US, income and expenditure-side estimates of GDP ($\mathrm{GDP_I}$ and $\mathrm{GDP_E}$) measure "true" GDP with error and are available at a quarterly frequency. Methods exist for using these proxies to produce reconciled quarterly estimates of true GDP. In this paper, we extend these methods to provide reconciled historical true GDP estimates at a monthly frequency. We do this using a Bayesian mixed frequency vector autoregression (MF-VAR) involving $\mathrm{GDP_E}$, $\mathrm{GDP_I}$, unobserved true GDP, and monthly indicators of short-term economic activity. Our MF-VAR imposes restrictions that reflect a measurement-error perspective (that is, the two GDP proxies are assumed to equal true GDP plus measurement error). Without further restrictions, our model is unidentified. We consider a range of restrictions that allow for point and set identification of true GDP and show that they lead to informative monthly GDP estimates. We illustrate how these new monthly data contribute to our historical understanding of business cycles and we provide a real-time application nowcasting monthly GDP over the pandemic recession.

JEL Codes: C32, E01, E32
Keywords: Mixed frequency; Vector autoregressions; Bayesian methods; Nowcasting; Business cycles; National accounts

---

*We thank an associate editor, three anonymous referees, conference and seminar participants at the 11th ECB forecasting conference, the 2021 ESCoE Economic Measurement conference, the 2021 IAAE annual conference, the Federal Reserve System Committee on Econometrics 2021 Virtual Meetings, Universität Salzburg, as well as Ana Galvao, Florian Huber, Kevin Lee, Simon van Norden, Martin Weale, and Saeed Zaman for helpful comments. The views expressed herein are those of the authors and not necessarily those of the Federal Reserve Bank of Cleveland or the Federal Reserve System.

1 Introduction

Real gross domestic product (GDP) is the most widely used single but comprehensive measure of economic activity. In the US, the Bureau of Economic Analysis (BEA) provides quarterly estimates of real GDP based on expenditure (E) and income (I). This leads to two estimates of GDP: what we call $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$.[1] While theoretically equivalent, these two estimates can in practice differ substantially due to statistical discrepancies. This is because $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are estimated using largely independent and imperfect source data; for example, see Nalewaik (2010). The discrepancy between $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ can have important implications, as examples from each of the last two recessions illustrate. First, while the quarterly annualized growth rate of real $\mathrm{GDP_I}$ turned negative 3 percent in 2007q3, $\mathrm{GDP_E}$ was still growing robustly (at an annualized rate of more than 2 percent).[2] Second, $\mathrm{GDP_I}$ indicated growth of some 15 percent in 2020q4, as opposed to just 4 percent growth in $\mathrm{GDP_E}$. These divergences lead to uncertainty about the timing and nature of these recessions and recoveries.

The desire for a reconciled or blended GDP estimate that combines the information in both estimates and avoids having to choose between $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ inspired Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2016), hereafter ADNSS, to develop an econometric modeling framework for producing historical estimates of "true" GDP. Their measurement-error framework views true GDP as an unobserved variable with $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ being two noisy estimates of it.[3] Estimates of true GDP are then obtained by applying optimal signal-extraction methods. The Federal Reserve Bank of Philadelphia uses the ADNSS model in real time to produce its popular reconciled quarterly measure of true real GDP growth: GDPplus.[4] An attraction of focusing, as we do, on estimation of true GDP is that it avoids having to select either $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$ as the preferred measure of GDP. Previous research has reached mixed conclusions on whether business cycle inference is sensitive to examining $\mathrm{GDP_E}$ or

[1] $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are also often referred to as gross domestic product and gross domestic income (GDI), respectively. We do not use this particular nomenclature to emphasize that both $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are estimates of the same underlying concept (GDP).

[2] This assessment is using end-of-March 2020 vintage data, available at https://apps.bea.gov/histdata/histChildLevels.cfm?HMI=7.

[3] As ADNSS discuss, their model relates to and complements a wider literature on the reconciliation of GDP measures dating back to Stone et al. (1942).

[4] See https://www.philadelphiafed.org/research-and-data/real-time-center/gdpplus.

$\mathrm{GDP_I}$. Chang and Li (2018) find that the common choice to use $\mathrm{GDP_E}$ rather than $\mathrm{GDP_I}$ can have a substantial effect on key empirical conclusions in applied macroeconomic work. Bognanni and Garciga (2016), by contrast, find little systematic difference in terms of how well $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ correlate with macroeconomic indicators. But like Nalewaik (2012), Bognanni and Garciga (2016) do draw attention to the advantages $\mathrm{GDP_I}$ can confer in dating recessions ahead of $\mathrm{GDP_E}$. Our approach, like ADNSS, is to favor reconciled measures of GDP - combination rather than selection.

The present paper builds on ADNSS and the previous literature in several ways. First, ADNSS use quarterly data on $\mathrm{GDP_E}$, $\mathrm{GDP_I}$ and unemployment to produce quarterly estimates of true GDP growth. We develop mixed frequency models that exploit the fact that unemployment data (and many other macroeconomic indicators of short-term economic activity) are available at a monthly frequency. This lets us extend ADNSS to produce monthly estimates of real GDP growth and, we emphasize, measures of uncertainty associated with these estimates. Importantly, these monthly estimates of true GDP are consistent with the published quarterly estimates of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$, but they exploit within-quarter information about economic activity gleaned from monthly indicators. An increasing range of monthly indicators, capturing specific aspects of overall economic activity, are widely consulted by economists interested in timely estimates of the state of the economy. Our methods provide a formal means of aggregating these monthly indicators to produce an estimate of the whole of GDP. While using these methods is less satisfactory than direct measurement of monthly GDP by the BEA, policymakers find monthly estimates of real GDP growth useful.[5] This view is

[5]Indeed, there can be interest in even higher-frequency GDP estimates. In path-breaking work, Evans (2005) developed a methodology to measure GDP, specifically $\mathrm{GDP_E}$, on a daily basis. Our point of departure is to reconcile $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ within a higher-frequency multivariate (VAR) model that like Evans (2005) imposes temporal aggregation constraints but allows for simultaneity between the alternative GDP measures and the higher-frequency indicators. We emphasize in this paper the conceptual importance of reconciling $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ - given that ultimately they measure the same variable - rather than focusing on one measure alone. As discussed further below, seminal work by Aruoba, Diebold, and Scotti (2009) [ADS] has also developed daily index-based measures of economic activity. Our interest, by contrast, is directly with estimation of monthly GDP. Such estimates have, for us, the attraction that when aggregated to a quarterly frequency, they can be compared and evaluated directly against the BEA's own estimates.

supported by the NBER's Business Cycle Dating Committee. On the NBER's website[6] the committee writes: "The committee . . . views real GDP as the single best measure of aggregate economic activity . . . The traditional role of the committee is to maintain a monthly chronology of business cycle turning points. Because the BEA figures for real GDP [GDP_E] and real GDI [GDP_I] are only available quarterly, the committee considers a variety of monthly indicators to determine the months of peaks and troughs." Interest in monthly GDP is also evidenced by the recent Brave, Butters, and Kelley (2019b) index (henceforth BBK) and accompanying monthly GDP_E (MGDP) estimates maintained by the Federal Reserve Bank of Chicago.[7]

Second, mixed frequency vector autoregressions (MF-VARs) involving GDP growth (and many other macroeconomic variables) are enjoying increasing popularity for providing high-frequency nowcasts or forecasts of low-frequency variables (see, among many others, Eraker et al. (2015), Schorfheide and Song (2015), Brave et al. (2019a), and Koop et al. (2020)). Most macroeconomic VARs include a variable reflecting real output growth. But conventionally this variable is based on one of the proxies for GDP, in fact almost always quarterly GDP_E. In this paper, we develop an MF-VAR where the output growth measure is (unobserved) true GDP. In other words, we embed the ADNSS structure within an MF-VAR. Given the growing interest in big data in general, and large VARs in particular, we show how our methods can be used with a large number of variables.

In order to develop our high-dimensional Bayesian MF-VAR, we begin with a low dimensional VAR at a quarterly frequency similar to that used by ADNSS. This allows us to explain the general structure of the ADNSS model and, more importantly, discuss identification and prior elicitation issues. ADNSS consider various models and discuss several different identification schemes. One of these involves an instrumental variable assumption (specifically that the change in the unemployment rate is correlated with true GDP growth but is uncorrelated with the measurement errors in GDP_E and GDP_I). The other involves restricting the variance of true GDP relative to the variance of GDP_E to a specific number. We relax this assumption and, instead, show that bounding this ratio of variances to an interval leads to sensible estimates of true GDP. In other words, we relax the point identification restriction of ADNSS to

[6]See http://www.nber.org/cycles/recessions.html.

[7]See https://www.chicagofed.org/publications/bbki/index.

allow for set identification; empirically this lets us present posterior evidence related to the news/noise restriction. This is a third contribution of this paper. It also sheds light on prior elicitation and allows us to develop a prior for the parameters controlling the relationship between GDP, $\mathrm{GDP_E}$, and $\mathrm{GDP_I}$ that we later use when we move on to the MF-VAR. We emphasize how our set identification approach to measuring true GDP differs from the recent identification strategy of Jacobs et al. (2022) that, generalizing ADNSS, point identifies true quarterly GDP by exploiting multiple data vintages. Our approach also differs in its focus on temporally disaggregating GDP, by embedding the model of ADNSS within an MF-VAR, so as to deliver higher-frequency estimates of true GDP than Jacobs et al. (2022)).

The remainder of our paper is structrured as follows. Section 2 discusses the quarterly and monthly data. Section 3 introduces the structural VAR modeling framework used throughout. Section 4 then sets out and estimates various quarterly data reconciliation models. Having discussed identification and prior elicitation in these quarterly VARs, we move on to the MF-VAR in Section 5. We explore various versions of this model, comparing their historical estimates of true monthly GDP growth and examining their time-series properties. We illustrate the utility of our new estimates of reconciled monthly GDP by analyzing their historical properties and providing a real-time application over the pandemic recession. We show how our model can be adapted to accommodate revisions-driven uncertainty about the most recent data when interest lies with nowcasting monthly GDP. Supplementary results summarized in the main paper but available in full in the online appendix evaluate the ability of our models, in-sample and out-of-sample, to capture historical US business cycles as identified by the NBER. We find that our reconciled estimates of GDP better date recessions than the use of either $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$ data alone. Section 6 concludes. Online appendices include a full description of the data and our econometric methods as well as tables of additional empirical results.

2 Quarterly and Monthly Data

Our models all make use of quarterly real $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ data from the BEA. We supplement these data, in some of our models, with monthly data on unemployment, hours worked, the consumer price index, the industrial production index, personal consumption expenditure

(PCE), the federal funds rate, the Treasury bond yield, and the S&P 500 index. These 8 monthly variables are those considered by Schorfheide and Song (2015), although they add quarterly $\mathrm{GDP_E}$, but not $\mathrm{GDP_I}$, into their MF-VAR model. All of these variables provide monthly information on underlying economic activity. Indeed, some constitute the monthly source data used by the BEA to estimate quarterly $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$; for example, emphasizing its utility in measurement specifically of underlying monthly GDP, monthly PCE includes roughly 70 percent of real $\mathrm{GDP_E}$. We also experiment, to demonstrate the utility of our methods with Big Data, with an even larger set of 48 monthly indicators (as summarized in the online Data Appendix) also believed to be helpful when tracking the evolution of the economy. This includes variables such as monthly real personal income (which typically amounts to more than 80 percent of $\mathrm{GDP_I}$) that we should expect to track GDP closely.[8] Later, to help establish the properties of our monthly GDP estimates, we compare them to a range of monthly business cycle indicators and alternative estimates of monthly GDP.

Following the argument in ADNSS that measurement errors are best modeled as *iid* in growth rates rather than in levels, we work in a stationary model with the $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ data, and the other non–stationary macroeconomic indicators, modeled in growth rates. Appendix B details data sources and the specific data transformations taken. Specifically, we use the log difference growth rate transformation.[9] We emphasize that, following the practice at the BEA and at the Federal Reserve Banks of Chicago and Philadelphia when publishing MGDP and GDPplus, respectively, we present monthly (and quarterly) GDP estimates as quarterly (quarter-over-quarter) annualized percentage changes.

Following ADNSS, when presenting historical estimates of reconciled GDP, we focus on the consideration of latest vintage $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ data. At the time of writing, these were (near the end of) June 2021 vintage data; matching vintage data are used for the 8 and 48 indicator variables and for GDPplus. To allow for (some) revisions even to recent data, the

[8]Personal income equals national income minus corporate profits with inventory valuation and capital consumption adjustments, taxes on production and imports less subsidies, contributions for government social insurance, net interest and miscellaneous payments on assets, business current transfer payments (net), current surplus of government enterprises, and wage accruals less disbursements, plus personal income receipts on assets and personal current transfer receipts.

[9]We note that our model would work equally well using exact growth rates. But the temporal aggregation constraint introduced below would require modification as discussed, for example, in Koop et al. (2020).

historical sample period runs from 1960q1/1960m1 through 2019q4/2019m12 (rather than 2021q1/2021m5, as available from the June 2021 vintage data). But we do consider real-time data vintages and accommodate revisions-driven data uncertainty when nowcasting GDP over the pandemic recession: to mimic real-time use of our models, we use the data available at the time and consider models estimated in the first and second (rather than latest vintage) releases of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$.[10] These $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ vintages are combined with monthly vintages of our monthly indicators from McCracken and Ng's (2016) FRED-MD database.

3 Overview of the Econometrics

All of the models used in this paper are either VARs or have a VAR as one of their main components. Accordingly, we establish some general notation that we use repeatedly in the remainder of the paper. We always work with VARs in structural form:

$$Ay_t = By_{t-1} + \epsilon_t, \epsilon_t \sim N(0, \Sigma), \tag{1}$$

for $t = 1, \ldots, T$ where y_t is a vector of N dependent variables, A is a lower triangular matrix with ones on the diagonal and Σ is a diagonal matrix.[11] For future reference, we denote the individual coefficients in A and B by a_{ij} and b_{ij}. This form for the VAR is of particular use for computational reasons, since the diagonality of Σ allows for equation-by-equation estimation of the model. As stressed, for example, in Carriero et al. (2019), this leads to large reductions in the computational burden, which can be particularly useful in high-dimensional models. But the structural VAR form is also useful, since some of the key data reconciliation relationships we use relate to the contemporaneous relationships between GDP, $\mathrm{GDP_E}$, $\mathrm{GDP_I}$ and unemployment and these all appear in A.

[10]Recent real-time data vintages for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are extracted from https://apps.bea.gov/histdata/histChildLevels.cfm?HMI=7. We extend these data vintages back to 1991 by making use of the real-time data vintages for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ available from Garciga and Knotek II (2019).

[11]For simplicity, but also to nest ADNSS, we write the VAR with one lag (a value we also use in our empirical work), and no intercepts or exogenous variables. Allowing any or all of these or more lags is straightforward. We also stress that the assumption that A is lower triangular is used only as an estimation device, not as a way of identifying structural shocks. The results are invariant to re-ordering of the variables in the sense described in Sub-section 3.1 of Carriero et al. (2019).

Bayesian estimation and forecasting for VARs involve choosing priors for A, B, and Σ and then developing a Markov chain Monte Carlo (MCMC) method for posterior and predictive simulation. We will discuss prior elicitation below in the context of the individual models. We provide only a brief description of our MCMC methods here, since these are standard. Additional details are given in online Appendix A. In our models, some of the elements of y_t are unobserved latent states (that is, true GDP is such a state and in the MF-VAR the unobserved monthly values of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are states). In the context of Gaussian linear state space models such as we use in this paper, standard Bayesian MCMC methods exist for drawing the states. Accordingly, we do not describe these in any detail either. In sum, we use MCMC algorithms that provide draws of the VAR coefficients (conditional on the states) using standard methods and draws of the states (conditional on the parameters) using standard methods.

4 Econometric Methods at a Quarterly Frequency

We start at a quarterly frequency and, thus, in this section $t = 1, \ldots, T$ in (1) denotes quarters.

4.1 Models Involving Only GDP

4.1.1 Theory

Many of the ADNSS results are obtained using the following model involving only the three GDP measures: expenditure-side, GDP_{Et}; income-side, GDP_{It}; and true latent GDP, GDP_t. It is worth stressing again that all of these GDP measures enter in growth rates (for example, GDP_{Et} is the growth rate of $\mathrm{GDP_E}$ constructed using the log-difference). ADNSS write their model in dynamic factor form as:[12]

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{2}$$

$$GDP_t = \rho \; GDP_{t-1} + \epsilon_{Gt}, \tag{3}$$

[12]For expositional simplicity we omit intercepts, although ADNSS include one in the GDP equation, but not in the other equations.

where:

$$\begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \sim iidN \left[\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_{GG} & \sigma^2_{GE} & \sigma^2_{GI} \\ \sigma^2_{GE} & \sigma^2_{EE} & \sigma^2_{EI} \\ \sigma^2_{GI} & \sigma^2_{EI} & \sigma^2_{II} \end{pmatrix} \right]. \tag{4}$$

Note that ADNSS adopt a measurement-error perspective and (2) specifies a model for the measurement errors in expenditure- and income-side GDP with true GDP itself following an AR(1) process. We emphasize that in this paper, like ADNSS, when interested in producing historical estimates of GDP, we work with a given (the latest) data vintage. Since this means that data near the end of the sample have undergone fewer revisions than older data, as discussed above we caution against extending our historical estimates to the present day. But, in the nowcasting application below, when interest resides with the latest GDP estimate (nowcast), we recursively update the data vintage used to mimic real-time application. We also allow for data uncertainty by modeling the first and second releases of GDP rather than the latest estimates as in ADNSS.[13]

It is straightforward to show that this model can be written as the VAR defined in (1) with $y_t = (GDP_t, GDP_{Et}, GDP_{It})'$ and all the elements of B zero except for b_{11}.[14] The fact that the error covariance matrix in (4) is unrestricted implies that A is unrestricted (other than being restricted to be lower triangular). This model is not identified.

ADNSS consider various ways of ensuring identification. First, they show that identification is achieved if $\sigma^2_{GI} = \sigma^2_{GE} = 0$. In words, the measurement errors in GDP_{Et} and GDP_{It} are uncorrelated with ϵ_{Gt}. This restriction can be shown to imply a VAR as in (1) with $a_{21} = -1$, $a_{31} = -1 + \sigma_{EI}$ and $a_{32} = -\sigma_{EI}$. Adopting the terminology of Mankiw and Shapiro (1986), we will refer to this restriction as the "noise" restriction - since this specification ensures that the volatility of true GDP is less than the volatility of $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$. Thus, measurement error is purely noise, as opposed to the idiosyncratic variation in $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ containing "news" or information about the true state of the economy. If the measurement error is pure news, true GDP is more volatile than either $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$. As emphasized by Fixler and Nalewaik (2010), noise implies that more volatile GDP measures should be weighted less

[13]The ADNSS model regards the data as given and does not allow for the reality that, especially recent, GDP_{Et} and GDP_{It} values are likely to be revised.

[14]See online Appendix A for further details.

when reconciling alternative measures of true GDP; in contrast, news implies they should be weighted more heavily. Although there is some empirical evidence against the noise restriction (for example, see Fixler and Nalewaik (2010) and Nalewaik (2010)), some variants of our models include this restriction; in others, we use it to center the prior (that is, the prior mean satisfies the noise restriction).

Second, ADNSS introduce what they call a "useful re-parameterization" and introduce a new parameter they call ξ, which is the ratio of the variance of GDP to the variance of $\mathrm{GDP_E}$. They show that restricting ξ to a specific value identifies the model. They present empirical results for a range of values of ξ. In a similar spirit, we introduce the parameters ξ_E and ξ_I, which are the ratios of the variances of GDP to $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$, respectively. Posterior inference about these parameters can also be used to shed light on whether the measurement errors are purely noise or whether they contain news as well. That is, the noise restriction implies that ξ_E and ξ_I are both less than one. When working with a model that does not impose the noise restriction, we can calculate the posterior probability that either or both are greater than one.

We first emphasize that, although fixing ξ_E or ξ_I to a specific value suffices to identify the model, identification may not be necessary to ensure sensible inference about GDP. That is, identification is not necessarily required for the Bayesian econometrician. Combining an unidentified likelihood with a proper prior will yield a proper posterior. If a parameter is completely unidentified (that is, does not appear in the likelihood function) and prior independence is assumed, then the posterior for the unidentified parameter equals its prior. However, in cases where the parameters are not completely unidentified and prior independence is not assumed, then posterior learning can occur even in unidentified models. Intuitively, posterior updating of the identified parameters can spill over into unidentified parameters via the assumed prior links between them. See Poirier (1998) for a theoretical discussion of these points.

In our case, even if prior independence is assumed about the parameters in (2), (3) and (4), ξ_E, and ξ_I are nonlinear functions of parameters and it is possible that learning about them can occur even in this unidentified model. Furthermore, a prior that bounds ξ_E and ξ_I can be used to set-identify the model. In the following sub-section we demonstrate that set identification can be used to estimate true GDP and that there is no need to fix ξ_E and/or ξ_I to specific values.

4.1.2 Empirics

We estimate the unrestricted quarterly VAR with latent GDP in (1), with $y_t = (GDP_t, GDP_{Et}, GDP_{It})'$, using a prior that is similar in spirit to ADNSS's. That is, we begin with priors for error variances that are relatively non–informative inverse-Gamma distributions (see the online appendix for complete details of the priors for all of the parameters in the model). The priors for the error variances are assumed to be independent of one another. To such a prior, ADNSS add a restriction that ξ_E is a specific value (for example, $\xi_E = 0.8$). They show that this identifies the model. It also means the priors for the error variances are no longer independent. This makes the actual prior used by ADNSS quite different from the apparently independent relatively non–informative prior they begin with. Instead of doing this, we achieve set identification by restricting ξ_E and ξ_I to lie within the interval $[0.55, 1.15]$.[15] This interval is fairly wide, expressing a range of different views about likely values for these two parameters accommodating both news and noise. ADNSS choose 0.8 as their benchmark and argue that ξ_E is likely less than one (implying noise). Our choice of bounds reflects such beliefs. Posterior computation proceeds by using MCMC methods to draw from the unrestricted posterior (that is, the posterior based on the VAR in (1) and the prior specified earlier in this paragraph) and discarding all draws that imply values of ξ_E or ξ_I outside the interval $[0.55, 1.15]$. Following ADNSS, we include an intercept in the GDP equation, but not in the equations for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$.

Figures 1a and 1b plot the priors and posteriors, respectively, for ξ_E and ξ_I. It can be seen that the priors are sensible, allocating weight across the interval $[0.55, 1.15]$, but with more weight allocated to values less than one. This is because, following ADNSS, we view it as more likely, but not certain, that the measurement errors in $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are noise. If we compare priors to posteriors, the key point to note is that they are different. Despite the fact that this model is not identified, data-based learning about ξ_E and ξ_I occurs. It is also worth noting that the probability that $\xi_E < 1$ is very close to one, suggesting that the measurement

[15] Results are robust to extending this interval to [0.5,1.5].

error in $\mathrm{GDP_E}$ is mainly noise. For $\mathrm{GDP_I}$, most of the posterior evidence also supports the noisy-measurement-errors conclusion, but it is not as strong in indicating a news component to $\mathrm{GDP_I}$. Fixler and Nalewaik (2010) found similar evidence, but based on modeling revisions to $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$.[16] Almost 5 percent of the posterior probability for ξ_I lies in the region above one, indicating some probability that measurement errors are news.

Inspection of the posterior parameter estimates in this model reveals all of the posterior means to be reasonable (in the sense they are similar to those given in ADNSS) and the credible intervals to be fairly narrow, indicating relatively precise inference despite the lack of identification.[17] Finally, Figure 2 plots our quarterly estimates of true GDP (posterior medians) along with a 68 percent credible interval. The relatively narrow credible interval shows that true GDP is precisely estimated. Figure 2 compares these estimates of true GDP against the BEA's quarterly estimates of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$. It shows that our quarterly estimates of true GDP do balance those of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ and that they are smoother than both proxies, although it should be emphasized that the posterior median estimates of true GDP do not always lie between the BEA's estimates of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$. They can be higher or lower than both. Over the sample period 1960q1-2019q4, the posterior median estimate of true GDP is more highly correlated with $\mathrm{GDP_I}$ (correlation coefficient of 0.97) than $\mathrm{GDP_E}$ (correlation coefficient of 0.91). This is consistent with the evidence in ADNSS that $\mathrm{GDP_I}$ contributes more to true GDP than $\mathrm{GDP_E}$. It also fits with the fact that the posterior median estimates of true GDP, plotted in Figure 2, are very highly correlated (at 0.97) with ADNSS's estimates as measured by the published quarterly GDPplus series.[18] This is as we should hope, given that the one aim of this paper was to embed ADNSS's quarterly measurement-error model within a Bayesian VAR model with set identification. In short, set

[16]As Fixler and Nalewaik (2010) show, exploiting data revisions (for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$) provides an alternative means of identification (to ADNSS) in models of data reconciliation that allow measurement errors to contain both news and noise components. Jacobs et al. (2022) develop this idea and propose a model to reconcile $\mathrm{GDP_E}$ and $\mathrm{GDP_E}$ data that exploits multiple data vintages. In Section 5.2.3 below, we extend our monthly GDP model to model data revisions; but given our use of set identification, we do not need to impose additional restrictions.

[17]See Table **??** in the online appendix.

[18]Table **??** in online Appendix D provides these and other supplementary details on the time-series properties of our quarterly estimates of true GDP.

identification suffices to produce reasonable estimates of true GDP at a quarterly frequency, even in a model involving only the two proxies for GDP.

We have also produced results for this model with the noise restriction imposed (that is, imposing $a_{21} = -1$ and $a_{31} + a_{32} = -1$). This restriction identifies the model and, thus, our prior is simply a prior rather than a means of imposing set identification. For the sake of brevity, we will not present empirical results for this case here. They are very similar to the set-identified results. This is not surprising, since the point estimates of a_{21}, a_{31}, and a_{32} (in Table **??** in the online appendix) come close to satisfying the noise restriction.

4.2 An Identified Model Involving GDP and Unemployment

4.2.1 Theory

ADNSS also work with a model that is identified by adding the change in the unemployment rate, U_t, to the model. They provide a convincing argument that unemployment can be treated as an instrument for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$. Their argument is based on the fact that unemployment is constructed using household surveys (by the Bureau of Labor Statistics), whereas GDP measures are independently constructed (by the BEA) using business surveys and, thus, the measurement errors in the two estimates should be uncorrelated with one another.

Their model comprises (2), (3), and (4) with an additional equation for U_t that says U_t depends on GDP_t, but not on $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$. It can be shown that this leads to a VAR representation based on (1) with the variables ordered as $y_t = (U_t, GDP_t, GDP_{Et}, GDP_{It})'$ where:

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ a_{21} & 1 & 0 & 0 \\ 0 & a_{32} & 1 & 0 \\ 0 & a_{42} & a_{43} & 1 \end{bmatrix} B = \begin{bmatrix} b_{11} & b_{12} & 0 & 0 \\ b_{21} & b_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{5}$$

Note that this specification, in essence, breaks the model into two parts. One part is a bivariate VAR for unemployment and true GDP that nests the AR(1) structure for GDP seen in ADNSS and their assumption that unemployment depends only on the contemporary true value of GDP. The other part is a structure inspired by ADNSS linking GDP to its two proxies. It captures the idea that GDP belongs in the macroeconomic VAR and, once GDP is included, $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ provide no additional explanatory power for any variable in the VAR other than GDP. The noise restriction now becomes $a_{32} = -1$, $a_{42} = -1 + \sigma_{EI}$ and $a_{43} = -\sigma_{EI}$.

4.2.2 Empirics

We have estimated three versions of the model with A and B restricted as in (5). The three versions impose the noise restriction, use a prior that is centered over this restriction, and use a prior that is centered over zero, respectively. They give very similar results.[19] We present results here using the version of the model with a prior centered over the restriction.

Again we find that the point estimates indicate that the noise restriction nearly holds.[20] The posteriors of ξ_E and ξ_I allocate slightly more weight to larger values than in the model without unemployment,[21] but the point estimates are nearer the benchmark choice of ADNSS. As before, we find almost no evidence that $\xi_E > 1$. However, for ξ_I, more than 10 percent of the probability is above one. Thus the evidence for measurement errors being noise is strong, but not overwhelmingly so for $\mathrm{GDP_I}$.

Another important comparison is between our quarterly estimates of true GDP and the GDPplus estimates produced by the Philadelphia Fed using this model. We plot both of these estimates in Figure 3. It can be seen that they match each other closely, with a correlation coefficient of 0.94 and with the GDPplus series falling within the 68 percent credible interval 89 percent of the time.[22] Given it is the variant of the ADNSS model actually used for production of GDPplus, the model presented in Section 4.1, whose GDP estimates are plotted

[19]We have also estimated an unrestricted version of the model that does not assume unemployment is an instrument and, thus, the model is only set-identified using the bounded prior on ξ_E and ξ_I. Results for this case were reasonable (that is, as defined before, the estimates of GDP were broadly consistent with those in ADNSS), but credible intervals were wider. Accordingly, we use both the prior bounds on ξ_E and ξ_I and assume unemployment is an instrument in the remainder of this paper.

[20]See Table **??** in the online appendix.

[21]Table **??** in the online appendix.

[22]Table **??** in online Appendix D provides these and other supplementary details on the time-series properties of our quarterly estimates of true GDP.

in Figure 2, is, as discussed above, even more closely correlated (correlation coefficient of 0.97) with GDPplus; GDPplus falls within its 68 percent credible intervals on 91 percent of occasions between 1960q1-2019q4. This all serves to reassure us that our Bayesian approach to estimation of the ADNSS model and our identification and prior elicitation strategy are mimicking ADNSS's estimation approach under exact identification. It also indicates that the restrictions imposed (but not tested) by ADNSS are supported empirically.

In summary, we have shown how to embed the data reconciliation models of ADNSS within a structural VAR framework where one of the variables (true GDP) is an unobserved latent variable. We have used this framework to show how true GDP can be identified using either an instrumental variables approach or set identification, with little consequence for the time-series properties of true GDP. Finally, we have used insights from this exercise to discuss prior elicitation. In particular, we have demonstrated that it is useful either to impose the noise restriction or to use a prior that is centered over this restriction. With this framework established, we now turn to the main goal of the paper: estimating monthly true GDP using quarterly GDP_{E} and GDP_{I} and various monthly indicator variables.

5 The MF-VAR with a Quarterly/Monthly Mixed Frequency

In this section, $t = 1, \ldots, T$ in (1) denotes time at a monthly frequency.

5.1 Theory

We return to the VAR model of Section 4.2, except that the model is now specified at a monthly frequency and we include additional monthly indicator variables in the VAR. Hence, $y_t = (X_t', U_t, GDP_t, GDP_{Et}, GDP_{It})'$ where X_t is a vector containing other monthly indicator variables. X_t and U_t are observed, but the other elements of y_t are not. True monthly GDP is never observed. For GDP_{E} and GDP_{I} we observe quarterly values, but not monthly values. Thus, we have an MF-VAR. If it were not for the inclusion of true GDP, this would be a conventional MF-VAR as in, for example, Schorfheide and Song (2015). The model we develop in this section combines the MF-VAR of Schorfheide and Song (2015) with the model of ADNSS to produce monthly estimates of true GDP. A side benefit is that we can also produce monthly estimates of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ that are temporally consistent with the quarterly estimates published by the BEA.

The MF-VAR treats the VAR in (1) as state equations in a state space model. The measurement equations link what we observe (for example, quarterly observations of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$) to the unobserved states (for example, monthly values of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$) via an inter-temporal restriction. For the case of log-differenced data, for a generic quarterly variable, $y_t^Q = \Delta_3 \ln Y_t^Q$ where Y_t^Q is the quarterly variable in levels (which we observe every third month), the link with its underlying monthly observations, $y_t^M = \Delta \ln Y_t$ where Y_t is the monthly variable in levels, is approximately:[23]

$$y_t^Q = \frac{1}{3}y_t^M + \frac{2}{3}y_{t-1}^M + y_{t-2}^M + \frac{2}{3}y_{t-3}^M + \frac{1}{3}y_{t-4}^M. \tag{6}$$

Another ingredient in the measurement equations for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ describes when they are observed. That is, quarterly variables are not observed in the first two months of the quarter, only for the third month (for example, statistical agencies produce these data for the calendar quarter January, February, March, but not for February, March, April). Thus, the measurement equations for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are given by (6) in the third month of each quarter and do not exist in the first and second months. The equations are formally set out in online Appendix A.

For true GDP there is no measurement equation, since it is never observed either at monthly or quarterly frequencies. For the monthly variables, the measurement equation simply reiterates that they are observed every month. These measurements, along with the monthly VAR of (1), define the likelihood function. It is a Gaussian linear state space model and, when combined with the priors used in this paper, standard Bayesian MCMC methods can be used for posterior and predictive simulation.

The MF-VAR just described is completely unrestricted (that is, A and B have no restrictions placed on them) and is not identified. In practice, we impose the (zero) restrictions in (5), which involve the assumption that U_t is an instrument. These are characterized by the features discussed at the end of Section 4.2.1 and the noise restriction remains the same as

[23]See Mariano and Murasawa (2003, 2010) and Mitchell et al. (2005).

described there. We also face the issue of whether we want to place any restrictions on how the other monthly indicator variables enter the model. We consider two treatments of this issue. The first of these follows the common practice of treating GDP, unemployment, and other monthly variables as defining a VAR independent of other sources of information. In other words, after controlling for GDP, the measurement errors in $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ do not have explanatory power for the other variables and do not belong in the VAR. This means all of the monthly indicator variables are instruments in the same way as U_t, and the coefficients in the A matrix corresponding to X_t in the equations for $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are set to zero. The second of these simply works with an unrestricted A matrix, except for the restriction that implies U_t is an instrument. The precise forms for the A matrices that result are given in Appendix A.

Finally, we consider versions of our models that do not include X_t, to investigate whether including additional monthly indicators affects monthly estimation of true GDP. As discussed in Section 2, we consider both the 8 monthly variables considered in Schorfheide and Song (2015) and a larger set of 48 monthly indicators; these are denoted X^8 and X^{48}, respectively. In turn, let ADNSS+SS denote the ADNSS structure embedded within the MF-VAR of Schorfheide and Song (2015) with X^8, and let ADNSS+SS^+ denote the SS model augmented with the larger set of 48 predictors.

Summarizing, we entertain models that involve four restrictions (the noise restriction, the restriction that unemployment alone is an instrument, the restriction that all of the monthly variables are instruments, and the restriction that additional monthly predictors are excluded from the MF-VAR). We always impose the restriction that unemployment is an instrument, even though we could relax this and rely on set identification instead. We do so given the aforementioned evidence that imposing unemployment as an instrument sharpens our estimates of GDP. To assess the empirical relevance of the remaining restrictions, we produce empirical results from models that consider various combinations of them.

As for the prior, we break the coefficients into two groups. The first of these are the parameters of the small quarterly VAR of Sub-section 4.2. For these, we use the prior developed previously, which involves centering the prior over the noise restriction and bounding ξ_E and ξ_I to the interval $[0.55, 1.15]$. The second group is all of the remaining parameters associated with the role of the potentially high-dimensional vector X_t in the VAR. For these we use a Dirichlet-Laplace prior. This is a popular global-local shrinkage prior that requires minimal prior hyperparameter choice and can automatically sort through the large number of VAR coefficients and decide which to shrink to zero. It has been used successfully with large VARs (see Kastner and Huber (2020)) and MF-VARs (see Koop et al. (2020)). Full details are given in Appendix A. Bayesian inference and prediction can be carried out in the MF-VAR with the Dirichlet-Laplace prior using MCMC methods as described in Koop et al. (2020).

5.2 Empirics

The main goal of this paper is to produce and analyze historical monthly estimates of true GDP growth. Given that the BEA produces neither monthly estimates of GDP, whether via the income or expenditure approach, nor quarterly estimates of true GDP against which we can evaluate our estimates, we analyze the estimates produced by our models in alternative ways. These are detailed in the following sub-sections.

5.2.1 Model Comparison

To assess the empirical relevance of the different restrictions, we produce empirical results from seven models that consider various combinations of them. Table 1 summarizes the features of these seven models. Full results for each model are in online Appendix C. Here we focus on our preferred model but summarize relevant cross-model differences.

Our preferred model, as selected by the deviance information criterion (DIC) but maintaining a preference for a parsimonious model, is the ADNSS+SS model; see Table 1. Only the ADNSS+SS^+ model delivers a lower DIC, but the properties of monthly GDP are similar to those from the smaller ADNSS+SS model, hence our focus on it here. In fact, the dynamics of monthly GDP are similar across all seven models. But the importance of news versus noise components can vary, as shown by the posterior estimates that $p(\xi_E > 1)$ and $p(\xi_I > 1)$ reported in Table 1. The models preferred by the DIC favor both news and noise. While the noise restriction tends to hold for $\mathrm{GDP_E}$, there is stronger probabilistic evidence that the measurement error in $\mathrm{GDP_I}$ is at least in part news. This result is consistent with the quarterly analysis in Fixler and Nalewaik (2010), which uses evidence from data revisions to identify the news and noise components. Like Fixler and Nalewaik (2010), but using our MF-VAR and set identification, our preference is for a model that allows for both news and noise.

Aware of the macroeconomic evidence that the real-time forecasting accuracy of BVAR models is improved when temporal changes in macroeconomic volatility are accommodated (see Clark (2011)), we also considered a variant of the ADNSS+SS model that allows for stochastic volatility (SV). As shown in online appendix C (Section **??**), the historical properties of the monthly GDP estimates from the ADNSS+SS model are little affected by the inclusion of SV (cf. Figure C3).

However, accommodating SV does introduce time variation in the posterior estimates for $p(\xi_E > 1)$ and $p(\xi_I > 1)$. This evidence of temporal instabilities in the interpretation given to news versus noise is also confirmed when our preferred ADNSS+SS model is estimated on sub-samples of our data. Summarizing the results tabulated in the online appendix (see Table **??** in Section **??**), we note that estimation over more recent samples of data tends to increase the news component to $\mathrm{GDP_E}$, even though the properties of true monthly GDP (our focus) are indistinguishable. Nevertheless, this sensitivity in interpretation does help in understanding the mixed evidence found in previous research. Using a similar sample period (post-2000), Jacobs et al. (2022) also find a larger news share in $\mathrm{GDP_E}$ than in $\mathrm{GDP_I}$.

5.2.2 Historical Properties of True Monthly GDP

We start by summarizing the statistical properties of the historical monthly estimates of true GDP produced by the ADNSS+SS model. For a fuller discussion and cross-model comparison, see online Appendix **??**.

Relation with $\mathbf{GDP_I}$ and $\mathbf{GDP_E}$

While our modeling approach produces historical estimates of monthly GDP that, when aggregated to a quarterly frequency, closely track the quarterly $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ data published by the BEA, there are important statistical differences even when only looking at the true GDP estimates after aggregation to a quarterly frequency. As shown by the probability integral transform plots of Figure **??**, discussed in Appendix **??**, the densities of true GDP differ from those of $\mathrm{GDP_I}$ and GDP_E. These differences are especially marked for $\mathrm{GDP_E}$, indicating that true GDP has a closer relationship with $\mathrm{GDP_I}$ than with $\mathrm{GDP_E}$, an issue we explore further below. Our historical estimates also continue to correlate highly with the quarterly GDPplus estimates from the Federal Reserve Bank of Philadelphia. But, as we shall see in the pandemic update below, this does not continue to hold in a real-time application over the pandemic.

Moving on to the focus of this paper, namely, monthly GDP, Figure 4 plots the monthly estimates of true GDP from the ADNSS+SS model against its implied monthly estimates of $\mathrm{GDP_I}$ and GDP_E, which of course aggregate to the BEA's quarterly estimates. Over the sample, 1960m1-2019m12, the posterior median estimate of true monthly GDP falls between $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$ 85 percent of the time. Of the 15 percent of "misses," 71 percent occur during NBER recessionary periods. This reminds us that true GDP is not always simply an average of $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$. True GDP can paint a different picture of either BEA estimate, and these differences tend to happen during recessions, presumably when policymakers especially crave accurate economic measurement. For example, looking at Figure 4, we see that our true GDP point estimates are lower than both $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$ during the global financial crisis. As Figure 5 shows, a side benefit of our model is the production of monthly estimates of $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$, again particularly helpful when tracking economic turning points. Figure 5 plots 68 percent interval estimates for monthly $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$ against the quarterly estimates of the BEA. Note that the credible intervals are quite narrow, indicating precise estimation.

True GDP is more negatively skewed than either $\mathrm{GDP_E}$ or $\mathrm{GDP_I}$. True GDP and $\mathrm{GDP_I}$ exhibit slightly more persistence (as measured by sample autocorrelations) than $\mathrm{GDP_E}$ (see Table **??**). True GDP and $\mathrm{GDP_I}$ have smaller AR(1) innovation variances and greater predictability as measured by the R^2 than $\mathrm{GDP_E}$.

Relative Contributions of $\mathbf{GDP_E}$ and $\mathbf{GDP_I}$ to True GDP

As at a quarterly frequency, our monthly estimates of true GDP are more highly correlated with our estimates of monthly $\mathrm{GDP_I}$ than monthly $\mathrm{GDP_E}$ (see the final column of Table **??**).

This is also understood by inferring the relative contributions of $\mathrm{GDP_I}$ and $\mathrm{GDP_E}$ to true GDP. Following ADNSS, and in the spirit of the least squares minimizations used in the data reconciliation literature (for example, see Weale (1985)), we can estimate the weight, λ, of $\mathrm{GDP_I}$ in our monthly estimates of true GDP:

$$\lambda^* = \underset{\lambda}{\operatorname{argmin}} \sum_{t=1}^{T} [(\lambda GDP_{E,t} + (1-\lambda) GDP_{I,t}) - GDP_t]^2. \tag{7}$$

Table **??** reports these weights and confirms that $\mathrm{GDP_I}$ is more important than $\mathrm{GDP_E}$ in explaining true GDP, explaining up to two-thirds of its variation. This new monthly result is consistent with the quarterly evidence in Fixler and Nalewaik (2010). Table **??** does indicate some modest differences across models in the combination weight. The weight on $\mathrm{GDP_E}$ rises when the noise restriction is imposed.

To shed further light on the relative contributions of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ to true GDP, we look at the Kalman filter gains. We do so not just using the latest vintage ("full-sample" or "balanced-sample") data considered above, but for a "ragged edge sample" that reflects the staggered release of data in real time. An attraction of our MF-VAR approach to measurement of monthly GDP is that current-month true GDP can be estimated even when not all observations exist up to the end of the sample, due to delays in data release. Here we contrast the Kalman filter gain estimates when we continue to assume that the quarterly data for both $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are known, with the estimates obtained when we condition only on the latest $\mathrm{GDP_E}$ data, given the reality that the Q1-Q3 $\mathrm{GDP_I}$ data are published by the BEA at an additional one-month lag (there is a two-month lag for Q4 $\mathrm{GDP_I}$). The posterior median gain estimates of 0.07 and 0.13 on $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ for the "full-sample" are consistent with the λ estimates we presented above. But the gain on $\mathrm{GDP_E}$ rises to 0.12 when the "ragged edge sample" is used.[24] This indicates that, in the absence of $\mathrm{GDP_I}$ data, the $\mathrm{GDP_E}$ data are informative. We comment further on the evolving relative contributions of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ in the pandemic update below, when we both consider real-time estimation of the ADNSS+SS model and allow for data revisions.

[24]The credible intervals around these posterior median gain estimates indicate precise estimation.

Informational Content

To further evidence the utility of our estimates at a monthly frequency, we find our GDP estimates to be highly correlated with a commonly used set of monthly business cycle indicators (see Table **??** in the appendix).[25] This result is robust across the seven model specifications of Table 1.

In online Appendix **??** we assess the historical ability of our new monthly GDP estimates to date business cycle turning points as identified (*ex post*) by the NBER. We find that our true monthly GDP estimates provide superior classification performance to GDP_I and GDP_E. Out-of-sample, in a real-time case study revisiting the 2007-9 recession, we further illustrate the utility of our monthly GDP density estimates at tracking the US business cycle in the face of staggered data releases, acknowledging that quarterly GDP_I data are published by the BEA at a lag to their estimates of GDP_E. But as our focus in this paper is on the measurement of (monthly) true GDP, rather than the dating of business cycle turning points (a task for which there already exist specialized models, for example, Chauvet and Piger (2008)), we return to measuring true GDP but now in real time.

5.2.3 Nowcasting True Monthly GDP during the Coronavirus Pandemic

Both to showcase the use of our models in practice and to turn attention to estimation of current rather than historical GDP, we estimate our models monthly through 2020 and 2021.

To mimic use in real time, we now make use of the real-time monthly data vintages. We acknowledge the staggered release of data in real time (the ragged edge) due to differing publication lags. These monthly variables are aligned with real-time monthly data vintages of quarterly GDP_I and GDP_E. Data vintages are organized so that our y_t^Q estimates of GDP for month t are produced near the end of month $t+1$, using monthly and quarterly indicator

[25] The indicators considered are: the industrial production index, the change in the unemployment rate, the Institute for Supply Management's Purchasing Managers Index for manufacturing, employment growth, the S&P500 index, and the Aruoba, Diebold, and Scotti (ADS) business conditions index (aggregated to a monthly frequency from the underlying daily index data). In addition, we consider the correlations against four alternative direct estimates of monthly GDP computed by Stock and Watson (2014), IHS Markit, the OECD, and BBK's estimates published at the Federal Reserve Bank of Chicago.

data available at this point in time. Given GDP_I data are published more slowly than GDP_E data, this means that while at the end of the first month of each calendar quarter the previous quarter's GDP_E estimate is known, the BEA has yet to publish GDP_I. Then, at the end of the second and third months of each calendar quarter, we use our MF-VAR to produce monthly GDP estimates in the absence of quarterly GDP data relating to the previous month. But we do condition on the latest monthly indicators for month t. Hence, our model fills in the intra-quarter data gaps.

Aware of the particular importance of GDP data revisions, we introduce a variant of the ADNSS+SS model that models not the latest-available data vintage as above, but the time series of first and second data releases. Clements and Galvao (2020) emphasize the utility of these "real-time vintages" when forecasting with BVAR models. Jacobs et al. (2022) show how incorporating information from multiple releases like this can deliver more precise quarterly estimates of true GDP. In our MF-VAR context, extending the ADNSS structure seen in (2), we assume that both the first and second releases of GDP_{It} and GDP_{Et} relate to true GDP_t, but we make no further assumptions. See online Appendix A9 for the precise model specification. This ADNSS+SS model (with revisions) thereby accommodates data uncertainty about the most recent data and delivers monthly estimates of true GDP that reconcile early (rather than later, revised) releases of GDP_E and GDP_I.[26] Unlike Jacobs et al.'s (2022) model, our model does not allow for separate identification of news and noise shocks but our use of set identification obviates the need for this.

Figures 6 and 7 plot the recursively computed real-time estimates of monthly GDP from the ADNSS+SS model and the ADNSS+SS (revisions) model. They plot the latest (current) posterior median estimates of monthly GDP with 68 percent credible intervals. These are denoted as first estimates, given that they are computed at the end of the month indicated. We also plot the posterior median of the second estimates of monthly GDP, computed at the end of the following month (when, notably, the latest GDP_I data become available for Q1-Q3), and the final or latest estimates computed using end-of-sample (2021m6) information.

[26] We re-emphasize that data availability (see Section 2) means that this revisions model is estimated on a sample beginning in 1991 rather than 1960 as with the benchmark ADNSS+SS model. We note that the time-series properties of the monthly GDP estimates from the ADNSS+SS model are similar if the ADNSS+SS model is estimated on the shorter sample starting in 1991.

Alongside, we plot the BEA's (the Federal Reserve Bank of Philadephia's) first, second, and final estimates of quarterly GDP_E and GDP_I (GDPplus).

Figures 6 and 7 show that both ADNSS+SS models rapidly detected the collapse in economic activity caused by the lockdowns designed to contain the spread of the coronavirus. Filling in the data gaps after publication of the BEA's (first) estimate that GDP_E fell by nearly 5 percentage points in Q1, the ADNSS+SS model assesses true GDP to have declined by 10 percentage points in the three months ending in April and 19 percentage points in the three months ending in May; see Figure 6. True monthly GDP reaches its trough in the three months ending in June. Notably this trough in true GDP is less severe than the trough in either GDP_E or GDP_I, reminding us once again that true GDP need not lie between GDP_E and GDP_I. Looking at the underlying month-on-month estimates of true GDP, y_t^M, we observe the biggest falls in May 2020. This contrasts slightly with the NBER's assessment that the trough (of the business cycle) was April 2020, but less so with the weekly economic indicator of Lewis et al. (2021), which is lowest in the last week of April. In any case, we are explicitly measuring true GDP rather than the "business cycle" or "real activity."

Turning to the ADNSS+SS (revisions) model plotted in Figure 7, we see greater uncertainty about the first estimates of true GDP, as evidenced by wider credible intervals especially around turning points, than when the latest-vintage data are modeled. Comparing Figures 6 and 7, we also see when modeling the first and second estimates that the fall in GDP in June and July 2020 is initially - as judged by the first estimate of true GDP - perceived to be greater than in Figure 6 when the latest-vintage data are modeled. But this difference is revised away, as the final estimates of true GDP indicate a less severe fall and subsequent rebound in GDP than indicated by the first estimates.

Compared with GDPplus, the revisions to estimates from both of our ADNSS+SS models are mild. The revisions to true GDP as measured by GDPplus are particularly pronounced for 2020q2, as the final estimate suggests a fall in true GDP of only 14 percent, compared to a first estimate of -26 percent. In contrast, the estimates from the ADNSS+SS models exhibit fewer revisions and are far more in-line with the BEA's own estimates of -38 percent and -40 percent. We emphasize again that the estimates of true GDP from the ADNSS+SS model do not always lie in between the estimates of GDP_E and GDP_I. We also see once more the argument for reconciliation: in 2020q4 the divergence between the $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ estimates is very large, at close to 9 percentage points. While the first estimate of true GDP from the ADNSS+SS model is close to the BEA's first estimate of $\mathrm{GDP_E}$, understandable in the absence at this point in time of any data on $\mathrm{GDP_I}$ for 2020q4, subsquent estimates of true GDP are revised upward strongly toward $\mathrm{GDP_I}$ once this estimate is published. This is explained by the Kalman gain estimates placing as much weight on $\mathrm{GDP_I}$ as on $\mathrm{GDP_E}$, once $\mathrm{GDP_I}$ is published by the BEA. Interestingly, once both first and second estimates of $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ are available, the Kalman gain estimates are highest on the first estimate of $\mathrm{GDP_E}$ but the second estimate of $\mathrm{GDP_I}$. The first estimate of $\mathrm{GDP_I}$ is zero-weighted once the second estimate is available.

Despite these extreme GDP observations seen in 2020, we find the historical estimates of monthly reconciled GDP from 1960m1 through 2019m12 from this ADNSS+SS model estimated on the pandemic samples to be virtually identical to those seen in Figure 4. That is, re-estimating the ADNSS+SS model on augmented data through 2020 and 2021 does not change the historical path of true GDP (the posterior median estimates are correlated 0.98 over the common sample up to 2019m12). This stability, in comparison to recent evidence showing that parameter estimates from MF-VAR models can change abruptly in the face of extreme observations and that nowcasts and forecasts can be affected (for example, see Schorfheide and Song (2015) and Lenza and Primiceri (2020)), will be due to the structure imposed via both the measurement-error model of ADNSS and the temporal aggregation constraint, (6). Diagnostic tests for the Gaussian assumption for the disturbances ϵ_{Gt}, ϵ_{Et}, and ϵ_{It} only slightly deteriorate when we include the pandemic observations. These have p-values for the Kolomogorov-Smirnov Gaussian tests of 0.13, 0.17, and 0.02, respectively, using data through 2019. Adding the pandemic observations led to these values changing to 0.02, 0.11, and 0.02, respectively.

6 Conclusions

GDP remains the most informative and readily interpretable single measure of economic activity. But arguably its measurement, in the US at least, is confused by separate and disparate point estimates from the BEA on the expenditure and income sides. $\mathrm{GDP_E}$ and

GDP_I estimates can and often do differ and in economically important ways. Moreover, the quarterly frequency of the BEA's estimates impedes both historical economic analysis, such as the within-quarter impact of historical events, and timely tracking of the evolution of economic activity. Accordingly, this paper embeds the quarterly GDP_E and GDP_I data reconciliation model of ADNSS within a Bayesian MF-VAR model with temporal aggregation constraints. The argument for reconciliation, not just of quarterly GDP_E and GDP_I data but also for exploiting the wealth of monthly indicators that take the pulse of the economy, is that the reconciled monthly GDP estimates incorporate more information. Unlike index-based measures of economic activity, such as those developed by ADS or Lewis et al. (2021), estimates of higher-frequency (here monthly) GDP have a natural interpretation: when aggregated to a quarterly frequency, they can be compared (and evaluated) directly against the BEA's own estimates.

Having explained identification and prior elicitation issues and established the validity of the proposed Bayesian approach, we estimate different variants of the model to produce reconciled historical estimates of monthly GDP, and its uncertainty, from 1960 to the present day. These new reconciled estimates of monthly GDP are consistent with the BEA's published quarterly estimates of GDP_E and GDP_I, but they exploit within-quarter information about economic activity gleaned from many monthly indicator variables.

Our Bayesian modeling approach, which relies on set rather than point identification, allows us to present new posterior evidence on the relative importance of news and noise components to the GDP measurement error and on the relative importance of GDP_E and GDP_I to measurement of true GDP. Our results favor models that allow for both news and noise components, and we find that interpretation of the relative importance of GDP_E and GDP_I to true GDP measurement is sensitive to modeling assumptions, such as over the set of monthly variables to include in the model, how identification is achieved, the sample period used for estimation, and whether the ragged edge is accommodated. Reassuringly, however, we find that historical estimates of reconciled monthly GDP are robust to these modeling choices.

Our new monthly reconciled density estimates of true GDP are found to better align with historical US business cycles than separate estimates of GDP_E and GDP_I. Our historical estimates of monthly GDP are largely unaffected when we update our sample to include the 2020 pandemic period and its extreme data realizations. Interesting future applications of our model will involve using it to forecast, as well as to measure historical and current, true GDP.

References

Aruoba, S. Borağan, Francis X. Diebold, Jeremy Nalewaik, Frank Schorfheide, and Dongho Song (2016). "Improving GDP Measurement: A Measurement-Error Perspective." *Journal of Econometrics*, 191(2), pp. 384–397. doi:10.1016/j.jeconom.2015.12.009.

Aruoba, S. Borağan, Francis X. Diebold, and Chiara Scotti (2009). "Real-Time Measurement of Business Conditions." *Journal of Business & Economic Statistics*, 27(4), pp. 417–427. doi:10.1198/jbes.2009.07205.

Bognanni, Mark and Christian Garciga (2016). "Does GDI Data Change Our Understanding of the Business Cycle?" *Economic Trends (Federal Reserve Bank of Cleveland)*. URL https://fraser.stlouisfed.org/title/3952/item/529755.

Brave, Scott A., R. Andrew Butters, and Alejandro Justiniano (2019a). "Forecasting Economic Activity with Mixed Frequency BVARs." *International Journal of Forecasting*, 35(4), pp. 1692–1707. doi:10.1016/j.ijforecast.2019.02.010.

Brave, Scott A., R. Andrew Butters, and David Kelley (2019b). "A New "Big Data" Index of U.S. Economic Activity." *Economic Perspectives*, (1), pp. 1–30. doi:10.21033/ep-2019-1.

Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino (2019). "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors." *Journal of Econometrics*, 212(1), pp. 137–154. doi:10.1016/j.jeconom.2019.04.024.

Chang, Andrew C. and Phillip Li (2018). "Measurement Error in Macroeconomic Data and Economics Research: Data Revisions, Gross Domestic Product, and Gross Domestic Income." *Economic Inquiry*, 56(3), pp. 1846–1869. doi:10.1111/ecin.12567.

Chauvet, Marcelle and Jeremy Piger (2008). "A Comparison of the Real-Time Performance of Business Cycle Dating Methods." *Journal of Business & Economic Statistics*, 26(1), pp. 42–49. doi:10.1198/073500107000000296.

Clark, Todd E. (2011). "Real-Time Density Forecasts From Bayesian Vector Autoregressions With Stochastic Volatility." *Journal of Business & Economic Statistics*, 29(3), pp. 327–341. doi:10.1198/jbes.2010.09248.

Clements, Michael P. and Ana Beatriz Galvao (2020). "Density Forecasting with BVAR Models under Macroeconomic Data Uncertainty." EMF Research Papers 36, Economic Modelling and Forecasting Group. URL https://ideas.repec.org/p/wrk/wrkemf/36.html.

Eraker, Bjørn, Ching Wai (Jeremy) Chiu, Andrew T. Foerster, Tae Bong Kim, and Hernán D. Seoane (2015). "Bayesian Mixed Frequency VARs." *Journal of Financial Econometrics*, 13(3), pp. 698–721. doi:10.1093/jjfinec/nbu027.

Evans, Martin D. D. (2005). "Where Are We Now? Real-Time Estimates of the Macroeconomy." *International Journal of Central Banking*, (2). URL https://www.ijcb.org/journal/ijcb05q3a4.htm.

Fixler, Dennis J. and Jeremy Nalewaik (2010). "News, Noise, and Estimates of the "True" Unobserved State of the Economy." BEA Working Papers 0068, Bureau of Economic Analysis. URL https://ideas.repec.org/p/bea/wpaper/0068.html.

Garciga, Christian and Edward S. Knotek II (2019). "Forecasting GDP Growth with NIPA Aggregates: In Search of Core GDP." *International Journal of Forecasting*, 35(4), pp. 1814–1828. doi:10.1016/j.ijforecast.2019.03.024.

Jacobs, Jan P. A. M., Samad Sarferaz, Jan-Egbert Sturm, and Simon van Norden (2022). "Can GDP Measurement Be Further Improved? Data Revision and Reconciliation." *Journal of Business & Economic Statistics*, 40(1), pp. 423–431. doi:10.1080/07350015.2020.1831928.

Kastner, Gregor and Florian Huber (2020). "Sparse Bayesian Vector Autoregressions in Huge Dimensions." *Journal of Forecasting*, 39(7), pp. 1142–1165. doi:10.1002/for.2680.

Koop, Gary, Stuart McIntyre, James Mitchell, and Aubrey Poon (2020). "Regional Output Growth in the United Kingdom: More Timely and Higher Frequency Estimates from 1970." *Journal of Applied Econometrics*, 35(2), pp. 176–197. doi:10.1002/jae.2748.

Lenza, Michele and Giorgio Primiceri (2020). "How to Estimate a VAR after March 2020." Working Paper 27771, National Bureau of Economic Research. doi:10.3386/w27771.

Lewis, Daniel J., Karel Mertens, James H. Stock, and Mihir Trivedi (2021). "Measuring Real Activity Using a Weekly Economic Index." *Journal of Applied Econometrics*, p. jae.2873. doi:10.1002/jae.2873.

Mankiw, N. Gregory and Matthew D. Shapiro (1986). "News or Noise? An Analysis of GNP Revisions." *Survey of Current Busines*, 66(5), pp. 20–25. URL https://apps.bea.gov/scb/pdf/1986/0586cont.pdf.

Mariano, Roberto S. and Yasutomo Murasawa (2003). "A New Coincident Index of Business Cycles Based on Monthly and Quarterly Series." *Journal of Applied Econometrics*, 18(4), pp. 427–443. doi:10.1002/jae.695.

Mariano, Roberto S. and Yasutomo Murasawa (2010). "A Coincident Index, Common Factors, and Monthly Real GDP." *Oxford Bulletin of Economics and Statistics*, 72(1), pp. 27–46. doi:10.1111/j.1468-0084.2009.00567.x.

McCracken, Michael W. and Serena Ng (2016). "FRED-MD: A Monthly Database for Macroeconomic Research." *Journal of Business & Economic Statistics*, 34(4), pp. 574–589. doi:10.1080/07350015.2015.1086655.

Mitchell, James, Richard J. Smith, Martin R. Weale, Stephen Wright, and Eduardo L. Salazar (2005). "An Indicator of Monthly GDP and an Early Estimate of Quarterly GDP Growth." *The Economic Journal*, 115(501), pp. F108–F129. doi:10.1111/j.0013-0133.2005.00974.x.

Nalewaik, Jeremy J. (2010). "The Income- and Expenditure-Side Estimates of U.S. Output Growth [with Comments and Discussion]." *Brookings Papers on Economic Activity*, 2010(Spring), pp. 71–127. doi:10.1353/eca.2010.0002.

Nalewaik, Jeremy J. (2012). "Estimating Probabilities of Recession in Real Time Using GDP and GDI." *Journal of Money, Credit and Banking*, 44(1), pp. 235–253. doi:10.1111/j.1538-4616.2011.00475.x.

Poirier, Dale J. (1998). "Revising Beliefs in Non-Identified Models." *Econometric Theory*, 14(4), pp. 483–509. doi:10.1017/S0266466698144043.

Schorfheide, Frank and Dongho Song (2015). "Real-Time Forecasting With a Mixed-Frequency VAR." *Journal of Business & Economic Statistics*, 33(3), pp. 366–380. doi:10.1080/07350015.2014.954707.

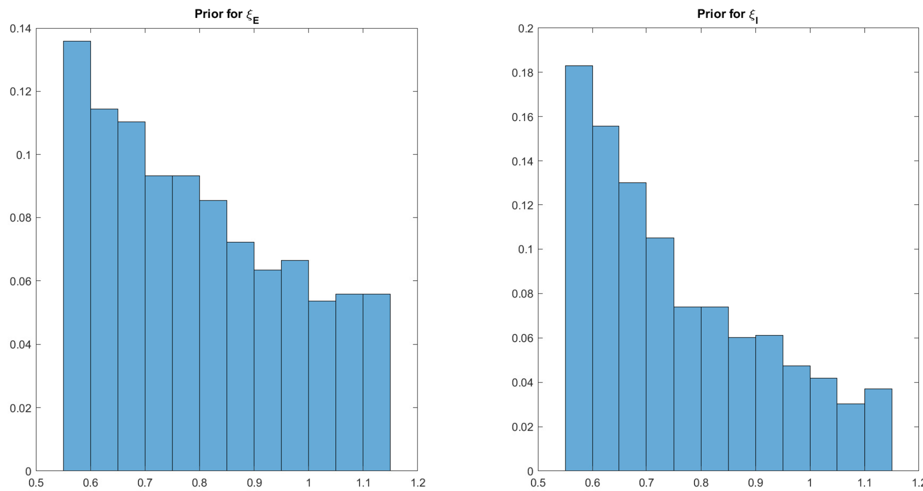
Stock, James H. and Mark W. Watson (2014). "Estimating Turning Points Using Large Data Sets." *Journal of Econometrics*, 178, pp. 368–381. doi:10.1016/j.jeconom.2013.08.034.

Stone, Richard, D. G. Champernowne, and J. E. Meade (1942). "The Precision of National Income Estimates." *The Review of Economic Studies*, 9(2), pp. 111–125. doi:10.2307/2967664.

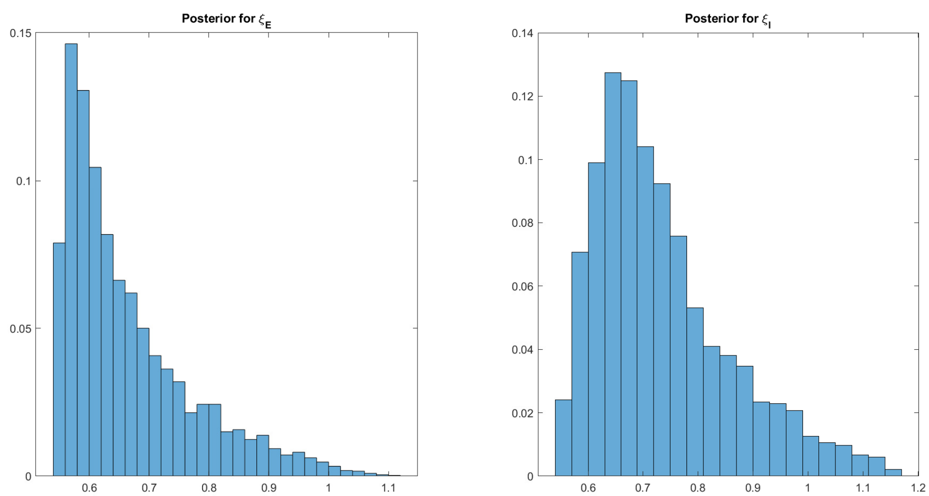
Weale, Martin (1985). "Testing Linear Hypothesis on National Account Data." *The Review of Economics and Statistics*, 67(4), pp. 685–689. doi:10.2307/1924815.

Figures and tables

Figure 1: Probabilities of ξ_E and ξ_I (the ratios of the variances of GDP to $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$, respectively)

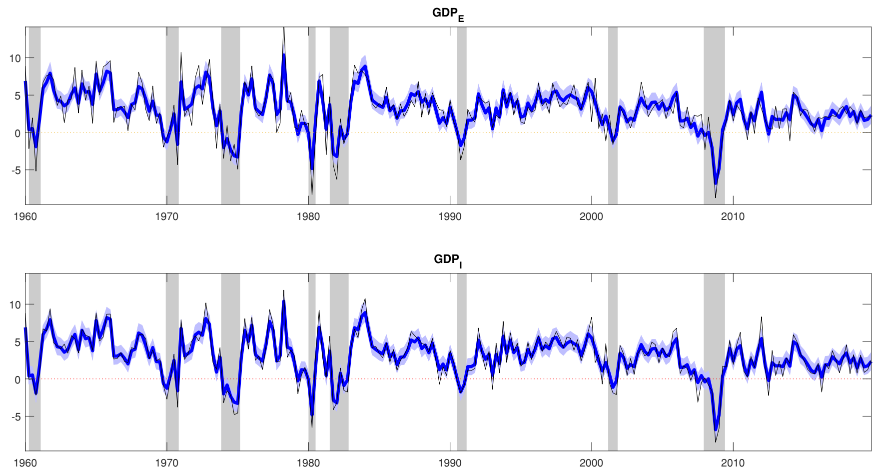


(a) Prior distribution



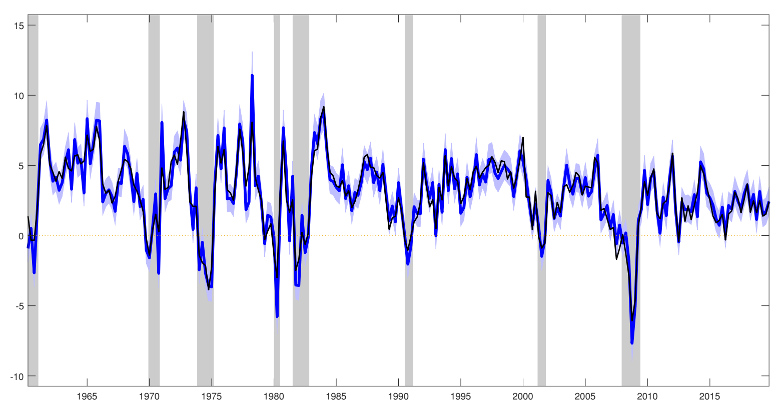
(b) Posterior distribution

Figure 2: Quarterly posterior median estimates of US real GDP growth



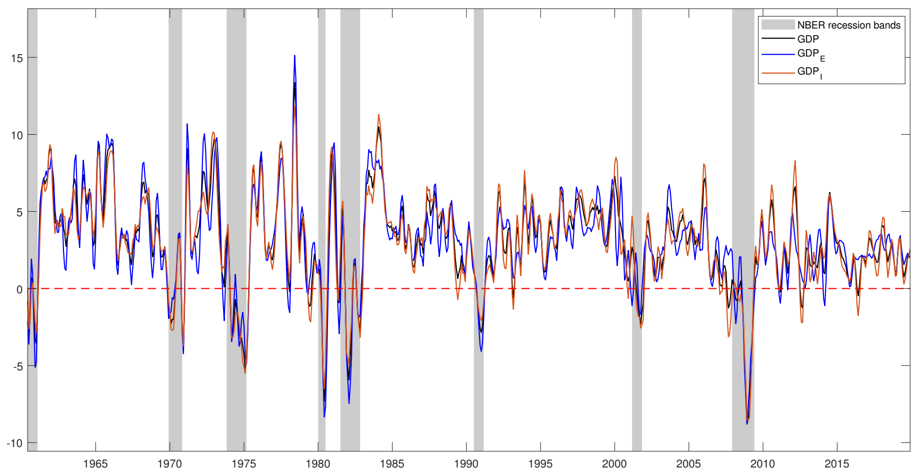
Notes: GDP growth in quarterly annualized percent changes from 1960q1-2019q4 (blue line) from the VAR model in only GDP_E and GDP_I, as seen in Section 4.1.2. Shaded blue region is the interval between the 16th and 84th percentiles of the posterior density of true GDP. Black line shows the BEA's quarterly estimates of GDP_E (top panel) and GDP_I (bottom panel) growth. Vertical shaded areas represent NBER-defined recessions.

Figure 3: Quarterly posterior median estimates of true US real GDP growth (blue line) versus the Philadelphia Fed's GDPplus (black line)



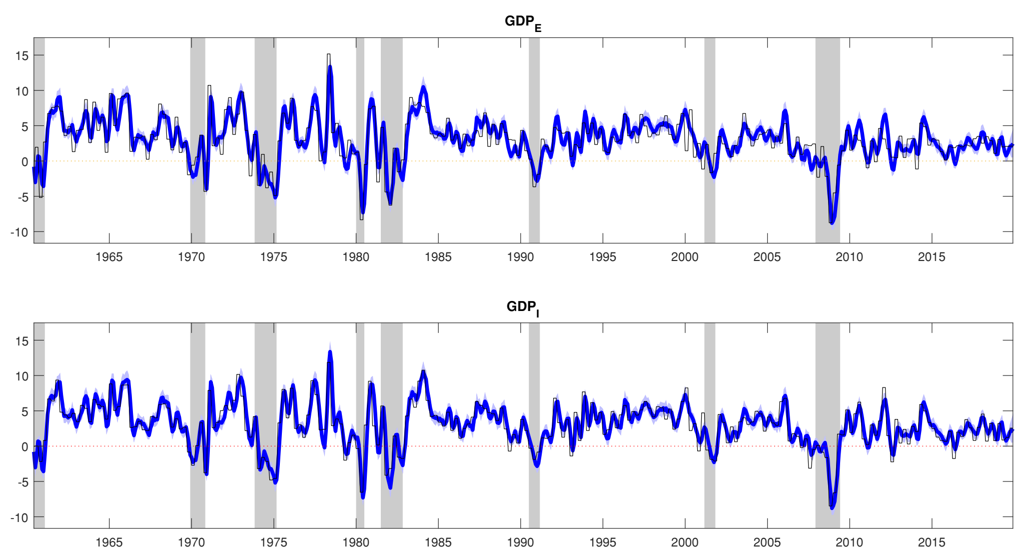
Notes: GDP growth in quarterly annualized percent changes from 1960q1-2019q4 (blue line) from the VAR model in GDP_E, GDP_I and unemployment, as seen in Section 4.2.2. Blue shaded region is the 16th and 84th percentile interval of the posterior density of true GDP. Vertical shaded areas represent NBER-defined recessions.

Figure 4: Monthly estimates (posterior medians) of GDP (black line), GDP_E (blue line) and GDP_I (red line) from the ADNSS+SS MF-VAR model



Notes: GDP growth in quarterly annualized percent changes from 1960m1-2019m12 (blue line) from the ADNSS+SS model. Vertical shaded areas represent NBER-defined recessions

Figure 5: Monthly estimates (posterior medians) of GDP_E and GDP_I growth (with 68 percent credible intervals) from the ADNSS+SS MF-VAR model versus BEA's quarterly estimates



Notes: GDP growth in quarterly annualized percent changes from 1960m1-2019m12 (blue line) from the ADNSS+SS model. Black line shows the BEA's quarterly estimates of GDP growth (constant over a calendar quarter). Shaded areas represent NBER-defined recessions

Figure 6: Real-time monthly GDP estimates from the ADNSS+SS model alongside quarterly GDPplus and quarterly GDP outturns from the BEA

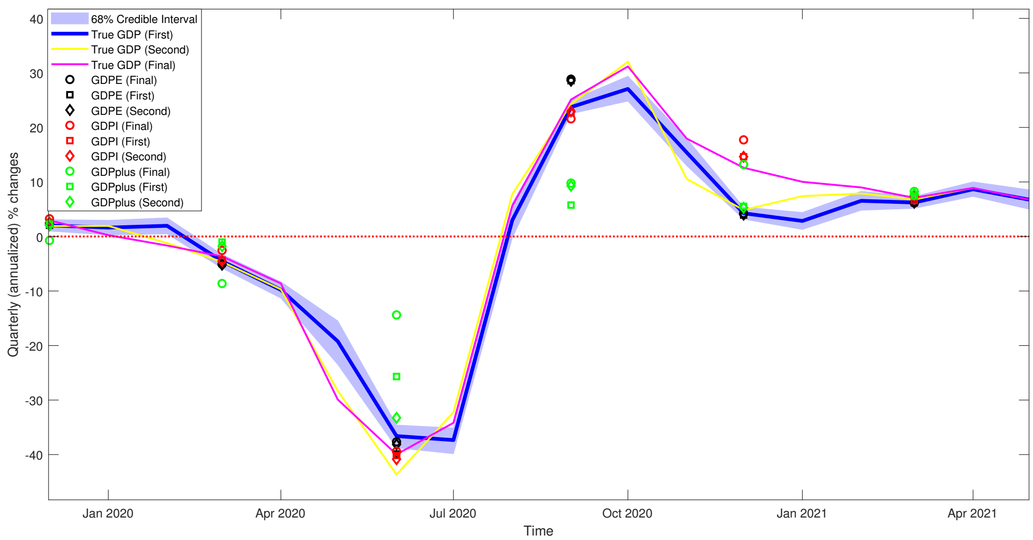


Figure 7: Real-time monthly GDP estimates from the ADNSS+SS (revisions) model alongside quarterly GDPplus and quarterly GDP outturns from the BEA

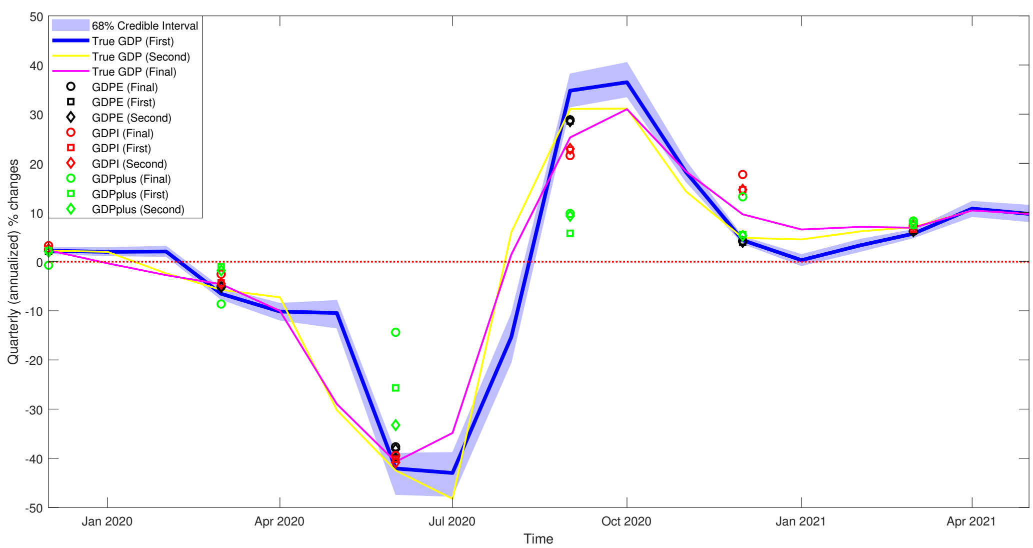


Table 1: Summary of MF-VAR models used to estimate monthly GDP

Model	Monthly Variables	Noise	IV	DIC1	DIC2	$p(\xi_E > 1)$	$p(\xi_I > 1)$	$\hat{\lambda}^*$
ADNSS+SS(IV)	$X^8, U, GDP, GDP_E, \mathrm{GDP_I}$	No	X^8, U	13500	6.7	0.00	0.01	0.37
ADNSS+SS(IV+N)	$X^8, U, GDP, GDP_E, \mathrm{GDP_I}$	Yes	X^8, U	13853	9.2	0.00	0.00	0.44
ADNSS+SS	$X^8, U, GDP, GDP_E, \mathrm{GDP_I}$	No	U	13270	3.6	0.01	0.37	0.35
ADNSS+SS(N)	$X^8, U, GDP, GDP_E, \mathrm{GDP_I}$	Yes	U	13872	4.3	0.00	0.00	0.40
ADNS S+SS^+	$X^{48}, U, GDP, GDP_E, \mathrm{GDP_I}$	No	U	12364	2.3	0.00	0.51	0.27
ADNSS	$U, GDP, GDP_E, \mathrm{GDP_I}$	No	U	14290	6.9	0.00	0.01	0.34
ADNSS(N)	$U, GDP, GDP_E, \mathrm{GDP_I}$	Yes	U	14337	11.0	0.00	0.00	0.44

Notes: Noise indicates whether the noise restriction is imposed or not imposed. IV denotes the instruments. DIC1 is the conditional deviance information criterion calculated based on the unemployment (U), GDP, $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ equations. DIC2 is the conditional DIC ($\times 10^6$) based on only the $\mathrm{GDP_E}$ and $\mathrm{GDP_I}$ equations. $p(\xi_E > 1)$ and $p(\xi_I > 1)$ are the posterior probabilities that ξ_E and ξ_I are greater than one, implying news. $\hat{\lambda}^*$ is the proportionate contribution of $\mathrm{GDP_E}$ in explaining GDP.