STAFF
REPORTS

NO. 1027
AUGUST 2022

GDP Solera: The Ideal Vintage Mix

Martín Almuzara | Dante Amengual | Gabriele Fiorentini | Enrique Sentana

FEDERAL RESERVE BANK of NEW YORK

GDP Solera: The Ideal Vintage Mix
Martín Almuzara, Dante Amengual, Gabriele Fiorentini, and Enrique Sentana
Federal Reserve Bank of New York Staff Reports, no. 1027
August 2022
JEL classification: E01, C32

Abstract

We exploit the information in the successive vintages of gross domestic expenditure (GDE) and gross domestic income (GDI) from the current comprehensive revision to obtain an improved, timely measure of U.S. aggregate output by exploiting cointegration between the different measures and taking their monthly release calendar seriously. We also combine all existing overlapping comprehensive revisions to achieve further improvements. We pay particular attention to the Great Recession and the pandemic, which, despite producing dramatic fluctuations, does not generate noticeable revisions in previous growth rates. The estimated parameters of our dynamic state-space model suggest that comprehensive revisions have not changed the long-run growth rate of U.S. GDP.

Key words: cointegration, comprehensive revisions, signal extractions, U.S. aggregate output, vintages

---

Almuzara: Federal Reserve Bank of New York (email: martin.almuzara@ny.frb.org). Amengual: CEMFI (email: amengual@cemfi.es). Fiorentini: Universitá di Firenze and RCEA (email: gabriele.fiorentini@unifi.it). Sentana: CEMFI and CEPR (email: sentana@cemfi.es). The authors thank Boragan Aruoba, António Rua, Dongho Song, and audiences at the RCEA Webinar Series, the 3rd Italian Workshop of Econometrics and Empirical Economics (Rimini), Banco de Portugal, the XII Time Series Econometrics Workshop (Zaragoza), and the North American Summer Meeting of the Econometric Society (Miami) for useful comments and suggestions. They are also thankful to Joël Marbet, Utso Pal Mustafi, and Clara Arroyo for excellent research assistance. In addition, Almuzara, Amengual, and Sentana acknowledge financial support from the Santander Research Chair at CEMFI, while Fiorentini is grateful to MIUR through the PRIN project "High-dimensional time series for structural macroeconomic analysis in times of pandemic."

1 Introduction

Despite the recent interest in alternative measures, such as the Human Development Index or the different Gross National Happiness measures, Gross Domestic Product (GDP) remains the dominant concept to gauge the aggregate performance of an economy over a given period of time. In the United States of America, the estimates of aggregate economic activity that the Bureau of Economic Analysis (BEA) publishes as part of its National Product and Income Accounts (NIPA) are used not only by policy makers and research economists, but also by private sector agents, including households and firms, in making their production and consumption decisions, as well as their financial plans.

The BEA uses a mixture of survey, tax and other business and administrative data, as well as various indicators, which are subject to sampling errors and biases that cannot be directly assessed. As time goes by, though, the BEA acquires more and better information, and for that reason it systematically updates its measures, which results in a sequence of estimates for a given quarter known as revisions. In fact, the whole revision process is rather elaborate, and it is important to distinguish between three types: (i) successive early releases for a given quarter, usually called the “advance”, “second” and “third” estimates; (ii) annual (or “final”) revisions, which simultaneously update all the quarters of several previous calendar years; and (iii) occasional comprehensive revisions, which recompute the entire history of the series after a major methodological change that effectively modifies its definition. The importance of revisions should not be underestimated. For example, Orphanides (2001) convincingly argues that the use of final instead of preliminary GDP measures can lead to different monetary policy recommendations.

While in the last two decades there has been considerable progress in jointly modeling the different vintages of US GDP (see, for example, Aruoba (2008), Jacobs and van Norden (2011) and the references therein), some of these studies have ignored a second important consideration: the BEA produces not just one but two different official measures of real aggregate output and income: Gross Domestic Expenditure (GDE) and Gross Domestic Income (GDI). GDE measures activity as the sum of all final expenditures in the economy, which is reflected in the output side of the NIPAs. In turn, GDI measures activity as the sum of all income generated in production, and is therefore captured on the income side of the NIPAs.[1] In theory, the flows of income and expenditure should be equal, and thus, GDE and GDI should yield the

[1]The value added approach would complete the usual trinity of GDP measurements, but the BEA does not produce quarterly real estimates.

same measure of economic activity. In practice, though, they differ not only because of the revisions but also because each is calculated from data from completely different sources (see Landefeld, Seskin, and Fraumeni (2008) for a review). The systematic, and at times noticeable, deviation between them – officially known as *statistical discrepancy*[2]– was traditionally regarded by many academic economists as a curiosity in the NIPAs. However, the Great Recession led to substantially renewed interest in academic and policy circles about the possibility of obtaining more reliable economic activity figures by combining the two measures. As a consequence, various proposals for improved combinations have been discussed (see, e.g. Nalewaik (2010), Nalewaik (2011), Greenaway-McGrevy (2011), Aruoba, Diebold, Nalewaik, Schorfheide, and Song (2016) and Jacobs, Sarferaz, Sturm, and van Norden (2022)). For example, the *GDPplus* measure of Aruoba et al. (2016) is currently released on a monthly schedule by the Federal Reserve Bank of Philadelphia.

The purpose of our paper is to simultaneously tackle all these measurement issues within a single, internally coherent, signal extraction framework.[3] Intuitively, given that GDE and GDI are based on different sources, one would expect to obtain a more accurate estimate of the underlying economic concept by making use of the dynamic and static recurrent patterns in the observed series.

Despite involving a moderately large number of both latent and observed variables, our model is both flexible and parsimonious thanks to the economic and statistical discipline that we impose on the measurement errors. Although the modelling of US GDP as a unit root process rather than as a trend stationary one is now conventional (see Campbell and Mankiw (1987) and the references therein for the earlier debate), our crucial point of departure from the previous literature is that we follow Almuzara, Amengual, and Sentana (2019) and Almuzara, Fiorentini, and Sentana (2022) in imposing that (i) any two aggregate output and income measures (in logs) are cointegrated, with cointegrating vector (1,-1); and (ii) measurement errors are mean-reverting and stationary, although they may be serially correlated. Thus, we are able to focus not only in quarterly growth rates, but also assess the level of US output, which is of considerable interest in itself, particularly in regional or cross-country comparisons.

In addition, the data release calendar is at the core of our model. Specifically, we explicitly take into account that the "advance", "second" and "third" GDE estimates are published one, two and three months after the end of the quarter, respectively. Moreover, we acknowledge the

[2]See Grimm (2007) for a detailed methodological insight.

[3]Stone, Champernowne, and Meade (1942) is the first known reference to the signal extraction framework of our paper. Weale (1992) surveys the early literature; see also Smith, Weale, and Satchell (1998).

fact that the timing of the quarterly releases for GDI is somewhat different, as it incorporates information from the quarterly census of employment and wages. Importantly, we also consider the annual data revisions of both series that are published in the summer of the following and subsequent years, and which typically affect the values for all the quarters of the most recent previous years. For example, the July 2017 annual update revised all quarters for 2014, 2015 and 2016.

The final novel ingredient of our model is the combination of data from different comprehensive revisions, which take place approximately every five years based on an economic census of millions of US businesses. These revisions also incorporate changes in definitions, classifications, and statistical methodology. For example, in 2013 the BEA started counting R&D as an investment rather than as a cost, which "boosted" US GDP by over 2%. The most recent comprehensive revision was published in July 2018, with a detailed analysis in a BEA paper (see Kerry, McCulla, and Wasshausen (2018)). In that report, the U.S. statistical office presented revised annual estimates for 1929-2017 and revised quarterly estimates for 1947-2017.[4] Often, comprehensive revisions reflect either improved or totally new coverage of sectors of the economy that have become increasingly important. In addition, real GDP is usually re-based, with the reference year kept fixed during subsequent annual updates.[5]

Despite these systematic differences, the joint modeling of multiple comprehensive revisions is particularly relevant at the time when a new one is released, which is precisely when there is very little information about the statistical properties of its successive vintages and annual revisions.

The closest paper to ours is Jacobs et al. (2022), which also use the different releases of GDE and GDI to obtain improved real-time estimates of economic activity. Nevertheless, these authors focus on growth rates and abstract from comprehensive redefinitions.[6]

From the point of view of implementation, our model can be cast in linear state-space form and is therefore amenable to the use of Bayesian methods of inference for both parameters and latent variables. In particular, we develop a Gibbs sampling algorithm that tackles estimation and signal-extraction simultaneously, allowing for an efficient and conceptually simple integration of uncertainty coming from different sources. Thus, we obtain a posterior distribution for

[4] The next comprehensive revision is expected in July 2023.

[5] Vintages released in July of both 2011 and 2014 are exceptions because the reference year was also revised. This resulted in a change of the GDP deflator and, in turn, a change of real GDP for the whole series since 1947.

[6] One additional difference is that Jacobs et al. (2022) propose a framework to separate news from noise in the revision process along the lines of Jacobs and van Norden (2011). In appendix E, we explain how to write their news-noise model as a special case of ours. We could use the expressions we derive there to provide a decomposition of the measurement errors between "news" and "noise", a promising avenue for future research.

the different benchmark definitions of underlying GDP, whence we can obtain not only point estimates but also measures of dispersion that reflect the remaining uncertainty about the true value of aggregate activity. Nevertheless, given that analysts and policy makers typically focus on the evolution of the current GDP definition, we will refer to our point estimate of the most recent benchmark version as *GDPsolera* henceforth. The moniker "solera" arises because the recurrent updating of our signal extraction process is analogous to the criaderas and soleras system of sherry wine aging, whereby the final product is obtained by fractional blending inputs from different vintages over a perennial dynamic procedure that gives sherry its distinctive character.[7]

After estimating our model making the best use of all the available US data, we use it to answer a number of empirically relevant questions. First, do comprehensive revisions modify the empirical characteristics of economic growth, such as its long-term mean or its persistence? Second, what is the contribution of the different estimates (i.e., advance, second, third, etc.) to the precision of signal extraction about economic activity? Our estimates suggest that (i) comprehensive revisions have not led to appreciable changes in the average growth rate, and that (ii) noticeable precision gains in signal extraction occur not only when the advance, second and third estimates of GDE and GDI are released but also when the annual estimates become available in July of the subsequent years.

Finally, we provide several additional empirical exercises, including an assessment of the sensitivity of our improved estimate of economic activity to our identification assumptions, as well as its behavior during the COVID-19 pandemic. In this respect, we find that the real time version of our solera measure provides accurate estimates of the quarters mostly affected by the pandemic, which seem to be in line with the subsequent BEA revised estimates. We also find that despite the dramatic nature of the GDP movements in 2020, our estimates of its growth rate for previous quarters are hardly affected.

The rest of the paper is organized as follows. We begin with a detailed description of the data in section 2. Section 3 introduces the model, while section 4 includes the details of the estimation and filtering algorithms. Section 5 reports the empirical analysis, including the improved *GDPsolera* measure of economic activity produced by our method. Finally, we present our conclusions and directions for further research in section 6, relegating proofs and other technical details to several appendices. Readers mostly interested in the empirical results may safely skip sections 3 and 4 initially.

[7] As explained by agent 007 to M in the 1971 James Bond film Diamonds are forever.

2 Data background

Our empirical analysis uses data on the successive GDE and GDI vintages from the BEA. To get a better sense of the data, it is instructive to review the timing of the release process as it happens regularly over a typical year. Table 1 exemplifies the process in a recent period. Estimates for quarterly GDP are released in the following order:[8]

(A) Advance estimate, based on source data incomplete or subject to further revision by the source agency, and released near the end of the first month after the end of the quarter.

(B) Second/third estimates, which use broader and more detailed data, and are released near the end of the second and third months, respectively.

(C) Latest estimates, which reflect the results of both annual and comprehensive updates.

For GDI only second, third and latest estimates are prepared because of data availability, except for the fourth quarter of each year, for which only third and latest estimates are released.

Normally, a single estimate for the latest quarter is added to the GDE/GDI series at a time, but there are two kinds of updates where multiple quarters are simultaneously updated:

(a) Annual updates, usually done in July, which cover at least the three most recent calendar years (e.g. the July 2017 annual update revised all quarters for 2014, 2015 and 2016). They incorporate newly available annual source data, and minor methodological changes.

(b) Comprehensive (or benchmark) updates, which are done approximately every 5 years (the actual updates took place in December 2003, July 2009, July 2013 and July 2018). They incorporate periodic data released at frequencies lower than 1 year, such as the quinquennial US Economic Census, and some major methodological changes.

In our main empirical analysis, we use the available seasonally adjusted[9] GDE and GDI vintages over the period 1984Q1-2021Q4, including the five benchmark versions of US economic activity resulting from the comprehensive revisions in 2003, 2009, 2013 and 2018.

We depict the series (in levels) of different comprehensive revision releases in Figure 1, where we also plot data produced by early and annual revisions for the periods between two consecutive benchmark revisions. As we explained in the introduction, the vertical differences

[8] Before 2009Q2, the BEA used the terminology "advance", "preliminary" and "final" for what it now calls "advance", "second" and "third", respectively.

[9] Since July 2018, BEA also publishes non-seasonally adjusted data.

TABLE 1. GDE and GDI release schedule for the period 2016Q1-2018Q2.

Release Month	**Estimate**	**GDE**		**GDI**	
		New	**Updated**	**New**	**Updated**
Jan 2017	Advance	2016Q4			
Feb 2017	Second		2016Q4		
Mar 2017	Third		2016Q4	2016Q4	
Apr 2017	Advance	2017Q1			
May 2017	Second		2017Q1	2017Q1	
Jun 2017	Third		2017Q1		2017Q1
Jul 2017	Advance	2017Q2	2014Q1-2016Q4		2014Q1-2016Q4
Aug 2017	Second		2017Q2	2017Q2	
Sep 2017	Third		2017Q2		2017Q2
Oct 2017	Advance	2017Q3			
Nov 2017	Second		2017Q3	2017Q3	
Dec 2017	Third		2017Q3		2017Q3
Jan 2018	Advance	2017Q4			
Feb 2018	Second		2017Q4		
Mar 2018	Third		2017Q4	2017Q4	
Apr 2018	Advance	2018Q1			
May 2018	Second		2018Q1	2018Q1	
Jun 2018	Third		2018Q1		2018Q1
Jul 2018	Advance	2018Q2	1947Q1-2017Q4	2018Q2	1947Q1-2017Q4

NOTES. [*] Annual update, [**] Comprehensive update, [†] 13 quarters, i.e. last 3 years

partly reflect different base years for the deflators. In turn, Figure 2 zooms in on two three-year subperiods to illustrate in closer detail the different measures of economic activity. The release of the July 2018 comprehensive revision led to a thorough revision of the GDE and GDI figures for the first subperiod (2015Q1-2017Q4), which explains the marked differences in levels between the advance, second and third releases, and the annual ones. In contrast, no such differences appear in the second subperiod (2019Q1-2021Q4), which nevertheless shows the dramatic effects of the COVID-19 pandemic. We will return to the analysis of this second period in subsection 5.4.

3 Model

Let x_t be an aggregate quantity of interest — in our empirical analysis, US aggregate economic output (in logs) during quarter t. As most of the literature that followed Stone et al. (1942), we treat x_t as a latent variable of which only noisy measurements y_t are available. The task is to construct rules mapping measurements into inferences about the latent x_t.[10]

Next, we develop the framework that will allow us to combine multiple y_t's for the purposes

[10]For background on output measurements, see Landefeld et al. (2008), Nalewaik (2010), and Nalewaik (2011).

of obtaining an improved estimate of economic activity. For the sake of clarity, we begin in subsection 3.1 with a version of our model that has no comprehensive revisions, adding them in subsection 3.2.

3.1 Modeling early and annual estimates

Let y_{it}^m be a noisy measurement of x_t, where the index i denotes type (e.g., GDE and GDI estimates) while the index m denotes release (e.g., early and annual estimates). This distinction is important because we will assume orthogonality of measurement errors along i but we will permit correlation over m for measurements with the same i. Orthogonality between the measurement errors of the expenditure and income estimates is not only plausible because they are based on completely different data sources, but also useful to achieve identification of the serial dependence in x_t. In contrast, correlation between the measurement errors of different releases of the same measure is to be expected, as they share revised versions of the same data sources.

The model is given by the set of measurement equations

$$y_{it}^m = x_t + v_{it}^m, \;\; m = 1, \dots, M_i, \quad i = 1, \dots, N,$$

where v_{it}^m is the measurement error in y_{it}^m. For each i, we collect $y_{it}^1, \dots, y_{it}^{M_i}$ into the vector y_{it} and stack $y_{1t}, \dots, y_{Nt}$ into y_t. Defining v_{it}, for each i, and v_t likewise, we obtain,

$$y_t = 1_{M\times 1} x_t + v_t, \tag{1}$$

where $M = \sum_{i=1}^N M_i$ and $1_{M\times 1}$ is an M-dimensional vector of ones.

In this context, we assume that the following conditions hold:

Assumption 1.

(a) Δx_t *is I(0);*

(b) $v_{1t}, \dots, v_{Nt}$ *are I(0);*

(c) $\Delta x_t, v_{1t}, \dots, v_{Nt}$ *are orthogonal across blocks at all leads and lags.*

We make assumption 1(a) because y_t measures economic activity in (log) levels.[11] Together

[11] We take the definition of I(0) process from the multivariate generalization of the one in Stock (1994): Consider a time series $\omega_t = \sum_{\ell=0}^{\infty} \Theta_\ell \varepsilon_{t-\ell}$, with Θ_ℓ an $n \times n$ matrix and ε_t and n-dimensional vector. Then, ω_t is I(0) if (i) ε_t is a weakly stationary vector martingale difference sequence, (ii) $\sum_{\ell=0}^{\infty} \Theta_\ell$ is nonsingular, and (iii) $\sum_{\ell=0}^{\infty} \ell \|\Theta_\ell\| < \infty$.

with assumption 1(b), it implies that y_t is cointegrated with cointegration rank $M-1$.[12] Cointegration is a very plausible assumption for aggregate measurement problems. In fact, assuming that the growth rates in y_t follow a strictly invertible covariance stationary process necessarily implies that the different measures of x_t would diverge in the long run, which is implausible (see Almuzara et al. (2022) for additional discussion). On the other hand, Assumption 1(c), which allows for dynamic dependence within blocks but rules out dependence between shocks to the signal and the different measurement errors, is key for identification, as asserted in the following proposition, whose proof can be found in appendix A:

Proposition 1. *Under assumption 1, if $N > 1$, the autocovariance matrices of $\Delta x_t, v_{1t}, \ldots, v_{Nt}$ are nonparametrically identified from the autocovariance matrices of Δy_t.*

Our empirical analysis features $N = 2$, as we use GDE and GDI measurements of output.[13]

3.2 Modeling comprehensive revisions

Our approach to modeling comprehensive revisions is to treat each version of the variable of interest introduced by the revision process as a different latent variable, while at the same time allowing for strong dependence among them.

Let C be the number of benchmark versions. Rather than a single variable, our extended model makes x_t a vector, namely $x_t = (x_{1t}. \ldots, x_{Ct})'$. Here x_{ct} represents the hypothetical value of economic output that could be measured with the definitions and methods adopted for the comprehensive revision c if the data sources and measuring tools were perfect. For example, the first three elements of x_t would treat R&D as a cost while the last two as an investment, as we explained in the introduction.

While analysts and policy makers typically focus on the latest version x_{Ct}, there are important reasons for jointly modeling $x_{1t}, \ldots, x_{Ct}$: first, older definitions of economic activity are important from a historical perspective because, after all, those were the only ones available at the time; second, understanding the impact of comprehensive revisions on the static and dynamic characteristics of the growth rates in aggregate economic activity is particularly relevant too; finally, there is also substantial interest in quickly learning about the dynamics of the measurement errors in the most recent version, which might lead to improved inferences about x_{Ct} itself.

[12] Any set consisting of $M-1$ pairwise differences among the y_{it}^m is a potential basis for the cointegration space.

[13] $N = 1$ may be relevant for other applications. In those cases, identification can be achieved by imposing restrictions on the cross-dependence among $v_{1t}^1, \ldots, v_{1t}^{M_1}$ (e.g., assuming $v_{1t}^{m_1}$ and $v_{1t}^{m_2}$ orthogonal at all leads and lags), or by a sufficiently tight parametric structure.

Measurement equation. Let δ_{it}^m be a $1 \times C$ array that has 1 in entry c if y_{it}^m measures x_{ct} and 0 otherwise. The array δ_{it}^m is deterministically time-varying but known, and can be easily computed by comparing the year of the comprehensive revisions and the exact release date of y_{it}^m. Our model postulates that

$$y_{it}^m = \delta_{it}^m x_t + v_{it}^m, \quad i = 1, \ldots, N, \; m = 1, \ldots, M_i.$$

Concatenating δ_{it}^m vertically to conform with y_{it} and y_t, we obtain the $M_i \times C$ array δ_{it} and the $M \times C$ array δ_t, which lead to the measurement equation

$$y_t = \delta_t x_t + v_t. \tag{2}$$

Equation (2) generalizes (1) into a deterministically time-varying measurement equation. We also note that some of the entries of y_t may be missing because, for example, the release protocol stipulates so or old methods are not applied to the computation of new estimates.

Thus, our framework generalizes naturally the multiple measurements - single latent variable models in the extant literature (e.g., Weale (1992), Smith et al. (1998), Aruoba et al. (2016), Almuzara et al. (2019), and Almuzara et al. (2022)) to a situation in which there are multiple latent variables of interest.

Identification revisited. We adopt assumption 1 without much change, except that Δx_t is a vector process now. Because the measurement equation is time-varying, the spectrum of y_t depends on t. However, given that the time-variation is deterministic, this entails a trivial form of non-stationarity from the point of view of identification. In our empirical analysis, moreover, there is a subvector of y_t that is stationary since there is a time-invariant block in δ_t. This allows us to establish identification through a generalization of proposition 1 applied to the time-invariant block. We state sufficient conditions for non-parametric identification in proposition 2, whose proof is also in appendix A.

Proposition 2. *Suppose there are indices i_1, i_2 ($i_1 \neq i_2$) and matrices E_{i_1}, E_{i_2} such that (a) $E_{i_1} y_t$ and $E_{i_2} y_t$ are nonempty subvectors of $y_{i_1,t}$ and $y_{i_2,t}$, respectively, (b) $E_{i_1}\delta_t$ and $E_{i_2}\delta_t$ are time-invariant, and (c) $\operatorname{rank}(E_{i_1}\delta_t) = \operatorname{rank}(E_{i_2}\delta_t) = C$. Then, under assumption 1, the autocovariances of $\Delta x_t, v_{1t}, \ldots, v_{Nt}$ are nonparametrically identified from those of Δy_t.*

As an example, consider a model with $C = 2$ versions of economic activity. Suppose $N = 2$

with $M_1 = M_2 = 2$ and $\delta_t = (I_2 \;\; I_2)'$ for all t. The measurement equation is

$$\begin{pmatrix} y_{1t}^1 \\ y_{1t}^2 \\ y_{2t}^1 \\ y_{2t}^2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x_{1t} \\ x_{2t} \end{pmatrix} + \begin{pmatrix} v_{1t}^1 \\ v_{1t}^2 \\ v_{2t}^1 \\ v_{2t}^2 \end{pmatrix}.$$

This setup clearly satisfies the conditions of proposition 2 with $i_1 = 1, i_2 = 2$, $E_{i_1} = (I_2 \;\; 0_{2\times 2})$, and $E_{i_2} = (0_{2\times 2} \;\; I_2)$. Consequently, the autocovariance matrices of $\Delta x_t, v_{1t}, v_{2t}$ are identified from those of Δy_t. Some intuition can be gained by first considering the measurement sub-systems

$$\begin{pmatrix} y_{1t}^c \\ y_{2t}^c \end{pmatrix} = 1_{2\times 1} x_{ct} + \begin{pmatrix} v_{1t}^c \\ v_{2t}^c \end{pmatrix}, \quad c = 1, 2.$$

Proposition 1 can be applied and immediately delivers the marginal serial dependence structure of the processes $\Delta x_{1t}, \Delta x_{2t}, v_{1t}^1, v_{1t}^2, v_{2t}^1, v_{2t}^2$.

Next, it is possible to recover the cross-autocovariances of the two signals by observing that

$$\mathrm{Cov}\Big(\Delta x_{1t}, \Delta x_{2,t-\ell}\Big) = \mathrm{Cov}\Big(\Delta y_{1t}^c, \Delta y_{2,t-\ell}^c\Big)$$

holds for $c = 1, 2$ and all ℓ. Finally, for $i = 1, 2$ and all ℓ, we have

$$\mathrm{Cov}\Big(\Delta v_{it}^1, \Delta v_{i,t-\ell}^2\Big) = \mathrm{Cov}\Big(\Delta y_{it}^1, \Delta y_{i,t-\ell}^2\Big) - \mathrm{Cov}\Big(\Delta x_{1t}, \Delta x_{2,t-\ell}\Big).$$

In our empirical analysis we rely on $C = 5$ versions of both GDE and GDI, in addition to their early and latest estimates. This implies that, for all t, δ_t contains two distinct blocks which are equal to I_C corresponding to GDE and GDI measurements, respectively, so the conditions in proposition 2 are automatically satisfied. Consequently, the joint dynamics of Δx_t are non-parametrically identified.[14]

Transition equation. Although the spectrum of x_t is non-parametrically identified, to implement our empirical analysis we specify a parametric model for the spectra of $\Delta x, v_1, \ldots, v_N$ that satisfies assumption 1 and, at the same time, is amenable to estimation by Bayesian methods. We

[14]One qualification worth making is that because past benchmark versions are discontinued, we are learning about the joint autocorrelation structure of x_t within the period in which they overlap. This amounts to a long period in our sample, spanning 1947Q1 to 2003Q2 (the time of the first comprehensive revision), yet a period that excludes the instabilities originating with the Great Recession or the COVID-19 pandemic.

adopt a Bayesian approach because it allows us to easily integrate both estimation uncertainty and filtering uncertainty in performing signal extraction, our main objective.

Specifically, we model Δx_t as a diagonal VAR with a single factor structure in the error term,

$$\begin{aligned}\Delta x_t &= \mu_x + \text{diag}(\rho_x)\,(\Delta x_{t-1} - \mu_x) + (\lambda_x \eta_{xt} + \text{diag}(\sigma_x)\varepsilon_{xt})\,, \\ \eta_{xt} &\overset{iid}{\sim} N(0,1) \text{ independent of } \varepsilon_{xt} \overset{iid}{\sim} N(0_{C\times 1}, I_C).\end{aligned} \tag{3}$$

The single factor structure parsimoniously captures the strong cross-sectional dependence in the innovations of the signals of the different comprehensive revisions. We then collect the unknown parameters of the Δx_t process in $\theta_x = (\mu_x, \rho_x, \lambda_x, \sigma_x)$. In principle, there could be differences in the mean, persistence and variance of economic growth across versions, which will allow us to empirically test whether comprehensive revisions had any impact on the static or dynamic properties of US output.

The initial condition for the level is modelled as independent of $\eta_{xt}, \varepsilon_{xt}$ for all t as:

$$x_1 \sim N(\mu_1, \Sigma_1).$$

This accommodates potential differences in levels between versions x_t, which adequately captures the use of deflators with a different base year, among other things.[15]

For the measurement errors of type i we postulate the following parsimonious diagonal VAR(1) model with a single factor structure in the error too:

$$\begin{aligned}v_{it} &= \text{diag}(\rho_i) v_{i,t-1} + (\lambda_i \eta_{it} + \text{diag}(\sigma_i)\varepsilon_{it})\,, \\ \eta_{it} &\overset{iid}{\sim} N(0,1) \text{ independent of } \varepsilon_{it} \overset{iid}{\sim} N(0_{M_i\times 1}, I_{M_i}),\end{aligned} \tag{4}$$

and place the unknown parameters of this process into $\theta_i = (\rho_i, \lambda_i, \sigma_i)$. Autocorrelated measurement errors in levels capture the persistent but stationary serial dependence observed in the statistical discrepancies. We also allow for variation in the autocorrelations and volatilities across different releases.

State-space representation. The parameter vector of the model is $\theta = (\theta_x, \theta_1, \ldots, \theta_N)$. Given θ, we can cast equations (2), (3) and (4) ($i = 1, \ldots, N$) in state-space form as

$$y_t = H_t X_t,$$

[15] We will treat μ_1 and Σ_1 as known and take Σ_1 to reflect a diffuse prior over x_1. A relatively easy-to-implement alternative would be to estimate μ_1.

$$
\begin{aligned}
X_t &= C(\theta) + F(\theta)X_{t-1} + G(\theta)U_t, \\
U_t &\overset{iid}{\sim} N(0_{(C+M+N+1)\times 1}, I_{C+M+N+1}), \\
\text{where} \qquad X_t &= (x_t, x_{t-1}, v_{1t}, \ldots, v_{Nt})', \\
U_t &= (\eta_{xt}, \varepsilon_{xt}, \eta_{1t}, \varepsilon_{1t}, \ldots, \eta_{Nt}, \varepsilon_{Nt})', \\
H_t &= \left(\delta_t \;\; 0_{M\times C} \;\; I_M\right), \\
C(\theta) &= \begin{pmatrix} [I_C - \mathrm{diag}(\rho_x)]\mu_x \\ 0_{C\times 1} \\ 0_{M\times 1} \end{pmatrix}, \\
F(\theta) &= \mathrm{diag}\left(\begin{bmatrix} I_C + \mathrm{diag}(\rho_x) & -\mathrm{diag}(\rho_x) \\ I_C & 0_{C\times C} \end{bmatrix}, \mathrm{diag}(\rho_1), \ldots, \mathrm{diag}(\rho_N) \right), \\
G(\theta) &= \mathrm{diag}\left(\begin{bmatrix} \lambda_x & \mathrm{diag}(\rho_x) \end{bmatrix}, \begin{bmatrix} \lambda_1 & \mathrm{diag}(\rho_1) \end{bmatrix}, \ldots, \begin{bmatrix} \lambda_N & \mathrm{diag}(\rho_N) \end{bmatrix} \right).
\end{aligned}
$$

For the initial condition we have $X_1 \sim N(\tilde{\mu}_1, \tilde{\Sigma}_1)$ where $\tilde{\mu}_1, \tilde{\Sigma}_1$ are compatible with μ_1, Σ_1 and the covariance-stationarity of $v_{1t}, \ldots, v_{Nt}$.

This linear state-space representation with Gaussian errors is important because it implies that $X_{1:T}, U_{1:T}$ will be jointly normally distributed conditional on $y_{1:T}, \theta$, so that we can rely on the algorithm of Durbin and Koopman (2002) to efficiently simulate from the conditional distribution of the latent variables given the observed ones (see subsection 4.2).

4 Inference for parameters and latent variables

Our objective is to conduct inference on parameters θ and latent variables $x_{1:T}$. As we have already mentioned, a Bayesian approach offers a convenient option to perform both tasks, integrating estimation and signal-extraction uncertainties in a unified, conceptually natural framework.

4.1 Estimation

Prior. We start by specifying $N+1$ independent priors for $\theta_x, \theta_1, \ldots, \theta_N$. The family of priors we describe is fairly standard and permits a simple implementation of the Gibbs sampler when the priors are conjugate conditional on the latent variables. It can also accommodate a flat prior for certain values of the hyperparameters.

Specifically, for the parameters of the signals process we use

- $\pi_x = 1/\sigma_x^2 \sim \Gamma_C(d_x/2, p_x/d_x)$, with divisions understood elementwise and Γ_C representing a vector of independent gamma-distributed random variables, and

- $\beta_x = ((I_C - \text{diag}(\rho_x)\mu_x, \rho_x, \lambda_x)|\sigma_x \sim N(b_x, R_x \otimes \text{diag}(\sigma_x^2))$.

The hyperparameters p_x and b_x control the prior mean of π_x and β_x, while d_x and R_x govern the informativeness of the prior distributions. In particular, higher d_x and R_x produce tighter priors while $d_x = 0_{C\times 1}$ and $R_x = 0_{3\times 3}$ yield a flat prior over π_x and β_x, which is not necessarily flat for θ_x.

In turn, for the parameters of the measurement errors process we use for each $i = 1, \ldots, N$

- $\pi_i = 1/\sigma_i^2 \sim \Gamma_{M_i}(\nu_i/2, p_i/\nu_i)$, and

- $\beta_i = (\rho_i, \lambda_i)|\sigma_i \sim N(b_i, R_i \otimes \text{diag}(\sigma_i^2))$.

The same considerations we made for p_x, b_x, d_x, R_x above apply to p_i, b_i, d_i, R_i too.

Gibbs sampler. Let $p(\cdot)$ denote a generic density (with respect to an appropriate dominating measure). Although the prior $p(\theta)$ and the likelihood $p(y_{1:T}|\theta)$ are readily available because the latter is an output of the Kalman filter applied to the state-space representation of the model, the posterior $p(\theta|y_{1:T})$ is not. Bayesian estimation can instead be performed via Markov Chain Monte Carlo (MCMC), which effectively draws a Markov chain $\{\theta^s\}_{s\geq 1}$ whose unconditional distribution coincides with the desired posterior.

As we mentioned before, a convenient approach to MCMC in our model is Gibbs sampling, which draws from the posterior of a block of variables or parameters conditional on previous draws from the other blocks in a sequential manner. We describe the algorithm in detail in appendix B.

4.2 Filtering

Signal extraction of x_t is a natural by-product of our estimation procedure. The latent variable draws we obtain in step (1) from iteration over the Gibbs sampler algorithm $(x_{0:T}^s)_{s\geq 1}$ have the desired distribution $p(x_{0:T}|y_{1:T})$. Moreover, the Gibbs sampler already integrates estimation uncertainty because

$$p(x_{0:T}\big|y_{1:T}) = \int_\Theta p(x_{0:T}\big|\theta, y_{1:T})\ p(\theta\big|y_{1:T})\ d\theta,$$

where Θ denotes the parameter space.

It is worth noting that while $p(x_{0:T}|\theta, y_{1:T})$ is a normal density, $x_{1:T}$ need not be normal given $y_{1:T}$ once θ is integrated out. In particular, Var $(x_t|y_{1:T})$ may depend on the data through the posterior density of θ, in contrast to Var $(x_t|\theta, y_{1:T})$, which is constant in $y_{1:T}$.

Finally, the Markov chain $(x^s_{0:T}, \theta^s)_{s\geq 1}$ is all that we need to approximate by simulation the posterior distribution of the different objects of interest that we will study in the next section.

In practice, we estimate our model using the prior described above running the Gibbs sampler for 105,000 iterations with a burn-in of 5,000 and a thinning of 1 every 5 iterations. The result is a Markov chain $(X^s_{1:T}, \theta^s)^S_{s=1}$, with $S = 20,000$ and low autocorrelation across draws that by all accounts appears to have converged. Our empirical analysis in the next section is based on it.

5 GDP solera: empirical analysis

5.1 Parameter estimates and their stability across comprehensive revisions

Table C.1 in Appendix C summarizes the posterior distributions of the model parameters, while Figures C.1, C.2 and C.3 in the same appendix compare those distributions to their priors. Although the annual revision process was extended from three to five years in July 2019, we consider three annual revisions, which are the only ones available for most of our sample.

We can use the posterior distribution of $\mu_x, \rho_x, \lambda_x, \sigma_x$ to assess whether comprehensive revisions have modified the static and dynamic properties of economic activity. In this respect, a noteworthy observation is that the unconditional means of the growth rates of the five different benchmark versions of US aggregate economic activity that the BEA has produced so far are remarkably similar, even though the comprehensive revision process has certainly affected the levels of US GDP, as we saw in Figure 1. In contrast, its persistence seems to have become somewhat smaller more recently, which is perhaps not surprising in view of the unusual nature of the 2020 COVID-19 recession. We will study the potential effects of this change in section 5.4 below.

Table C.1 also suggests that the common shock to the different elements of x_t is more important than their idiosyncratic shocks in explaining the variance of the innovations in the signals, as one would expect from the strong cross-sectional dependence between the different comprehensive revisions observed in the same figure.

5.2 Precision gains from using all releases for a given comprehensive revision

The root mean square error (RMSE)

$$\sqrt{V_t^\tau} = \sqrt{\text{Var}\left(\mathbb{E}\left[x_{ct} \middle| \mathcal{Y}^\tau\right] - x_{ct}\right)},$$

where $\mathcal{Y}^\tau$ denotes the σ-algebra generated by all measurements available until month τ, measures the precision of our signal extraction procedure. Figure 3 reports this RMSE for $c = 5$ and a fixed t as a function of τ for a sequence of 39 months starting in October of year t, which is when the advance GDE estimate for the third quarter becomes available, under the assumption that no comprehensive revision takes place during those three years and a quarter.[16]

As expected, the release of the advance GDE figure almost halves the RMSE of the prediction of the third quarter growth rate made at the end of September. Nevertheless, substantial precision gains also occur when the second and third estimates of GDE and GDI are released. Moreover, there are further gains when the annual estimates become available in July of the following three years. Still, the non-singular nature of our dynamic model, combined with the fact that the BEA does not attempt to reconcile the GDE and GDI figures, implies that there is a positive floor to the RMSE, which will not go to zero regardless of the number of subsequent annual revisions.

Exactly the same pattern arises if we repeat this exercise for the first and second quarters of year t in April and July, respectively, but not for the fourth quarter, which shows a slightly different initial pattern (not reported here) because there is no second GDI release in February.

5.3 Effects of combining all comprehensive revisions

To assess the effect of using data from all comprehensive revisions simultaneously, we have also estimated the single signal version of the model in Section 3 using only the data from most recent comprehensive revision. Figure 4 reports the posterior medians of GDP growth generated by our MCMC estimation and filtering procedure and their point-wise 90% credible sets based on both datasets for the period 2017Q1 to 2019Q4.[17] As can be seen, the use of the five comprehensive revisions results in not only significantly tighter bands around the smoothed estimates of economic activity but also a smoother temporal evolution for those estimates.

[16] In computing this figure we maintain the joint posterior distribution of the model parameters fixed at its estimate in September 2018 to focus on the precision gains of the smoother as new data becomes available. Consequently, the annual revisions correspond to July 2019, 2020 and 2021.

[17] In this case, we maintain the joint posterior distributions of the parameters of the models with either one or five signals fixed at their estimates in January 2022.

5.4 Solera releases

Figure 5 reports the smoothed estimates for US GDP growth from six different solera releases, which we have recursively estimated as follows. The first series uses data until January 2012 to provide estimates up to 2011Q4. Similarly, the second series provides estimates up to 2013Q4 using data until January 2014, and so forth, until the sixth series, which represents estimates of GDP growth until 2021Q4 using the data available at the BEA website at the end of January 2022. As can be seen in panel (a), which depicts the six series starting in 2004Q1, all estimates display close paths until 2010Q1.

Still, the growth rates estimates for the last few quarters of each series are somewhat different from the corresponding estimates in the next ones, an effect that it is very likely due to the smoothing embedded in our filtering algorithm, which systematically reassesses the past after observing the future.

Additionally, the two most recent solera releases that we display in green present a different pattern from the others in the second quarters of 2011 and 2012. These differences can be explained by the fact that the data underlying those last two series incorporate modifications to the GDP definition resulting from the comprehensive revision the BEA released in July 2018.

Indeed, panel (b), which only reports the two most recent series in panel (a), shows an extremely similar pattern between them even though the most recent version of *GDPsolera* includes data from the pandemic. Therefore, the post pandemic estimates for the pre-pandemic period are remarkably stable to the inclusion of the large outliers in 2020 data, which affect not only the output of the simulation smoother for fixed parameter estimates but also the posterior distribution of the parameter estimates.

5.5 Analysis of some specific quarters

Next, we shed some light on the effect of data revisions as well as the arrival of information for subsequent periods on the estimates of US GDP growth rate both through the smoother and the re-estimation of the model parameters in three specific quarters of interest: 2001Q1, 2008Q4 and 2019Q2.

We chose the first one because of the political controversy surrounding what at the time some Republican politicians called the "Clinton recession", in marked contrast to the NBER Business Cycle Dating Committee, which officially dated the peak of the previous ten year expansionary phase in March 2001. Although the BEA only publishes vintage data from September 2002 onwards, Figure 6, which uses blue crosses and red diamonds to represent GDE and GDI estimates, respectively, shows that the data initially available suggested that GDP growth had already turned negative in the first quarter of 2001. However, the comprehensive revision that became available in December 2003 is more ambiguous, with GDE and GDI growth rates having different signs. If anything, the subsequent annual revision released in July 2004 increases the degree of ambiguity. Not surprisingly, when one looks at the solid and dashed lines in that figure, which represent the posterior medians with and without parameter re-estimation, and the shaded areas, which display the corresponding 90% point-wise credible bands, the only conclusion that one can draw is that the uncertainty is too large to determine the sign of the GDP growth rate unequivocally.

Our next example focuses on 2008Q4, the worst quarter of the Great Recession, which we analyze in Figure 7. Although the advanced GDE estimate the BEA released initially pointed to a serious but not dramatic recession, subsequent releases justify the adjective ''Great''. Nevertheless, this figure also shows the adjustment of the posterior median of our solera GDP growth estimate as soon as we process the third releases of GDE and GDI, which is in line with the evidence we observed in Figure 3. In addition, Figure 7 also shows the effect that the comprehensive revision of July 2009 had on the precision of the estimates, and especially the annual revision of July 2010, which reduced further not only the growth rate but also the width of the credible sets.

Our third and final example focuses on 2019Q2, a relatively normal quarter despite the fact that some Federal Reserve officials had previously expressed concerns about a potential deceleration of the economy. This quarter is also interesting because it allows us to explicitly assess the effect of the pandemic data on our parameter estimates. As Figure 8 shows, the estimates of economic growth were noticeably revised downwards after the annual update that the BEA released in July 2020. However, a substantial part of this reduction was reversed following the July 2021 annual revision. Interestingly, the width of the credible sets goes down fairly slowly, which probably reflects the fact that the unprecedented GDP fluctations in 2020Q2 and 2020Q3 increased the uncertainty of the parameter estimates.

5.6 Comparison with *GDPplus*

We also compare our measure of economic activity – *GDPsolera* – with the *GDPplus* initially proposed by Aruoba et al. (2016), and released on a monthly basis by the Federal Reserve Bank of Philadelphia since the end of August 2013.

To begin with, we look at the smoothed estimates of GDP between the first quarter of 1985 and the fourth quarter of 2021. To construct our solera measure, we use the data released by the BEA by the end of January 2022, while for *GDPplus* we use the release that uses the same dataset to level the playing field.[18] We plot the estimated annualized growth rates in panel (a) of Figure 9. As can be seen, the two series are quite close to each other with a contemporaneous correlation of 0.86, and an average annualized growth rate of 2.61% for *GDPplus* and 2.54% for *GDPsolera* over the entire sample period. Nevertheless, our solera estimates are clearly more volatile, with a standard deviation that is 40% larger. The smoothness of *GDPplus* results in relatively more conservative estimates of the large fall and rise of economic activity after the start of the COVID-19 outbreak.

To shed further light on this, we report in panel (b) of Figure 9 the two real-time estimates of economic activity for 2020Q1 and 2020Q2 using the data available at the time. Perhaps not suprisingly, for 2020Q1 both estimators of GDP are in agreement, and remain quite stable as new information became available. In contrast, the estimators for 2020Q2 are very different and this difference increased in October 2020 when the BEA published the advance GDE estimate for 2020Q3. Interestingly, the most recent figures produced by the BEA for the COVID-19 recession are closer to the *GDPsolera* series. Nevertheless, it must be acknowledged that the extremely atypical size of the pandemic shock is a challenge to linear Gaussian state-space models, which makes the comparison difficult.

In our last exercise, we compare the concurrent online estimates of GDP growth rates generated by *GDPplus* and our procedure. Specifically, we consider estimates for each quarter based on the information available one month after the end of that quarter, by which time only the "advance" GDE estimate is available. In addition, we also look at the estimates of the same GDP growth rates obtained three months after the end of the quarter, which also make use of the "second" and "third" estimates of GDE and GDI released by the BEA. Panels (a) and (b) of Figure 10 displays these two set of results. Interestingly, the real time *GDPsolera* and *GDPplus* estimates appear to be more similar than the historical ones we saw in Figure 9. Still, we can observe a few differences in the first two quarters of 2015 afected by the 2018 comprehensive revision, and at the end of the sample, starting after the 2020Q2 drop.

[18] Given that *GDPplus* is based on the mostly recently available estimates of GDE and GDI rather than on multiple vintages, it can use data from 1960Q1. Nevertheless, this should not affect too much their estimates in recent years.

5.7 Allowing for correlation across expenditure and income measures

As mentioned earlier, there are two reasons for imposing zero correlation between the shocks to the true GDP and the GDI and GDE measurement errors. Primarily, it allows us to achieve non-parametric identification. And second, the empirical evidence that there is a sufficiently strong cyclical pattern in the statistical discrepancy is subject to debate (see, Nalewaik (2010) and the subsequent discussion).

Nevertheless, we have explicitly re-estimated a generalised version of our model in which we allow for non-zero correlation between the shocks to the common factor of the signals and the common factors in the measurement errors of the expenditure and income measures. In a set up with multiple measurements, this assumption is analogous to the one made in Aruoba et al. (2016). To conduct this exercise, and given that the identification information for those correlations comes from their priors, we have decided to use a grid of degenerate priors ranging from 0 to 30% to assess the sensitivity of our smoothed estimates to the values of that parameter.

As can be seen in Figure 11, the posterior means are hardly affected, except in the third quarter of 2020. As a consequence, the identifying assumption of zero correlation does not seem to explain on its own the fact that *GDPsolera* is more volatile than *GDPplus*.

6 Conclusion

We make the best use of the information in the different vintages of GDE and GDI from the current comprehensive revision to obtain an improved timely measure of US aggregate output by imposing cointegration between the different measures and taking seriously their monthly release calendar. We also combine overlapping comprehensive revisions to achieve further improvements.

We express our model in linear state-space form, and use Bayesian methods of inference for both parameters and latent variables. Specifically, we develop a Gibbs sampling algorithm that tackles estimation and signal extraction simultaneously, allowing for an efficient and conceptually simple integration of uncertainty coming from different sources. Thus, we obtain a posterior distribution for the underlying GDP measure, whence we can obtain not only point estimates but also measures of dispersion.

The estimated parameters of our dynamic state-space model suggest that comprehensive revisions have not changed the long-run growth rate of US GDP, but they have somewhat lowered the persistence of its shocks.

Our results suggest that noticeable precision gains in signal extraction occur not only when the advance, second and third estimates of GDE and GDI are released but also when the annual estimates become available in July of the subsequent years. We also observe that the use of the five comprehensive revisions not only results in significantly tighter bands around the smoothed estimates of economic activity, but also a smoother temporal evolution for those estimates.

In addition, we pay particular attention to certain recent episodes, like the Great Recession or the COVID-19 pandemic, which, despite producing dramatic fluctuations, does not generate noticeable revisions in previous growth rates.

We also find that the real time GDPsolera and GDPplus estimates appear to be remarkably similar, with small exceptions at the end of the sample after the 2020Q2 drop.

Although the objective of our analysis is not the creation of a real time activity index (see e.g. Lewis, Mertens, Stock, and Trivedi (Forthcoming) and the references therein), combining our approach with either high frequency data or additional quarterly variables constitutes a promising avenue for further research. Assessing the effect of incorporating the seasonally unadjusted GDE and GDI data that the BEA has released since 2018 to our empirical results would also provide a valuable addition.

Similarly, the potential forecasting improvements of the model we propose in this paper for the early releases of GDE and GDI would be worth investigating, as they would provide an external validity check on our modelling approach. In this respect, another potential extension would allow for a more flexible autocorrelation structure, as well as conditional heteroskedasticity and non-normal shocks, although the latter would require replacing the analytical Kalman filter by a numerical non-linear one.

Finally, it would interesting to apply our *Solera* approach to the different components of GDE and GDI, as well as other macroeconomic series subject to revisions, like the Non-farm Payroll Employment figures or the Chained Consumer Price Index for All Urban Consumers released by the US Bureau of Labor Statistics.

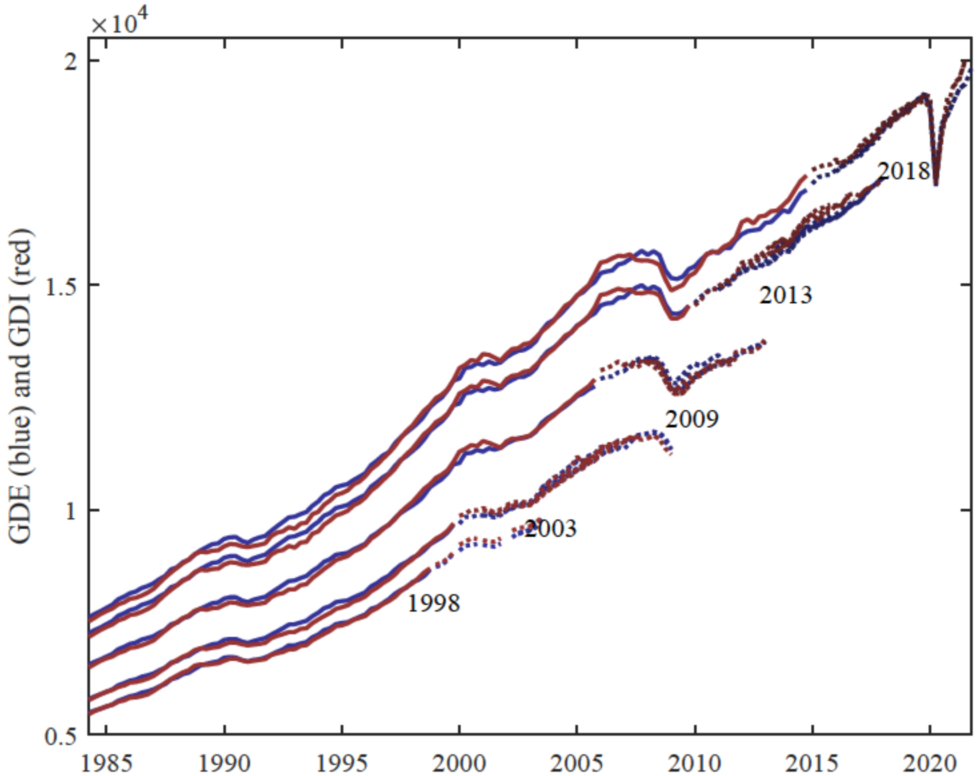
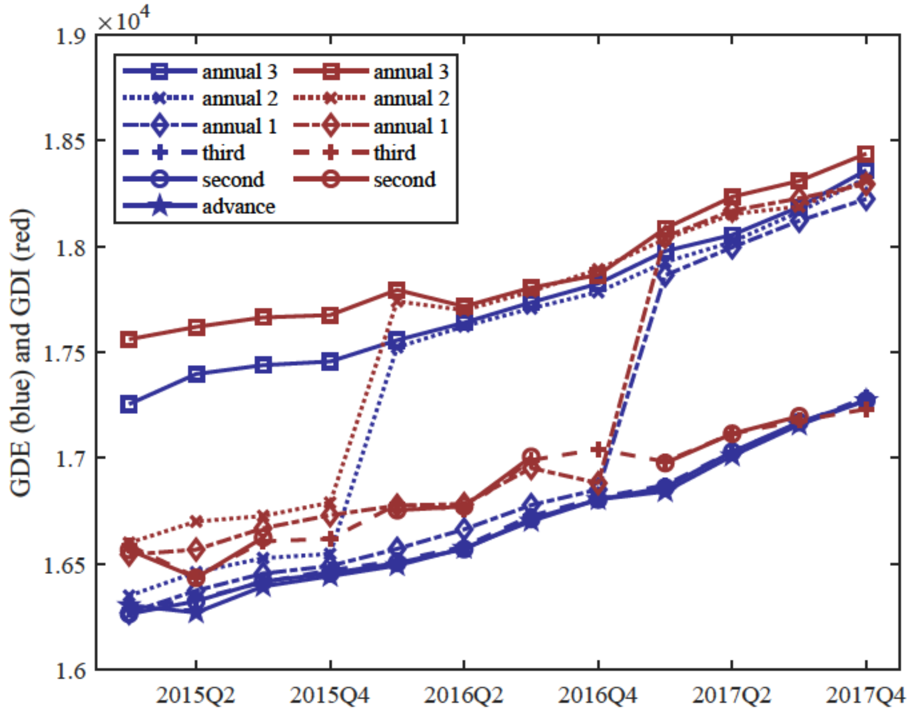
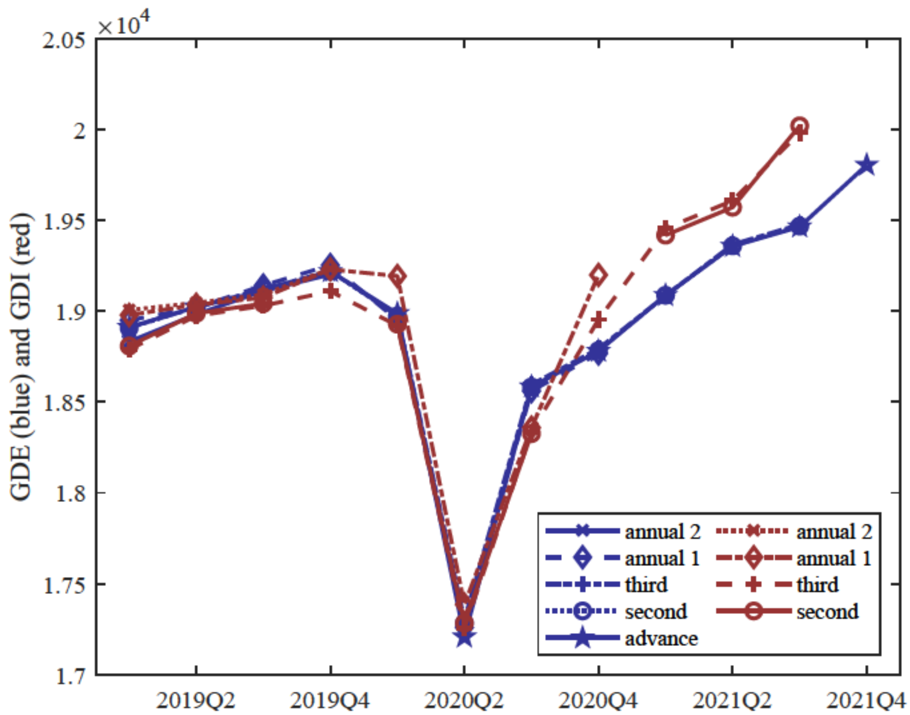


FIGURE 1. **GDE and GDI data from the BEA** (1984Q1 to 2021Q4; MM US$). Solid lines represent data released under comprehensive revisions while dashed lines represent data produced by early and annual revisions.



(a) Measurements 2015Q1-2017Q4



(b) Measurements 2019Q1-2021Q4

FIGURE 2. **GDE and GDI data from the BEA** (MM US$). Each subplot reports levels for a different version of economic activity.

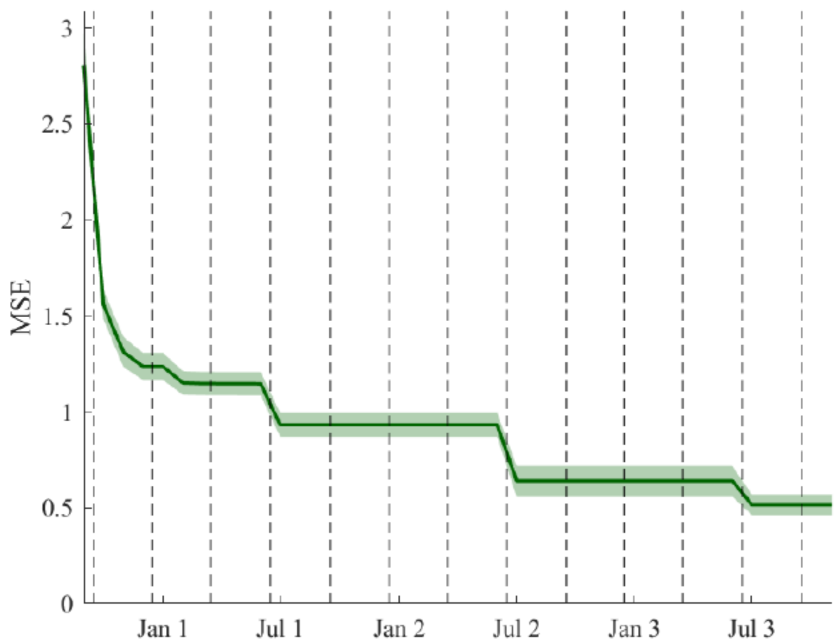
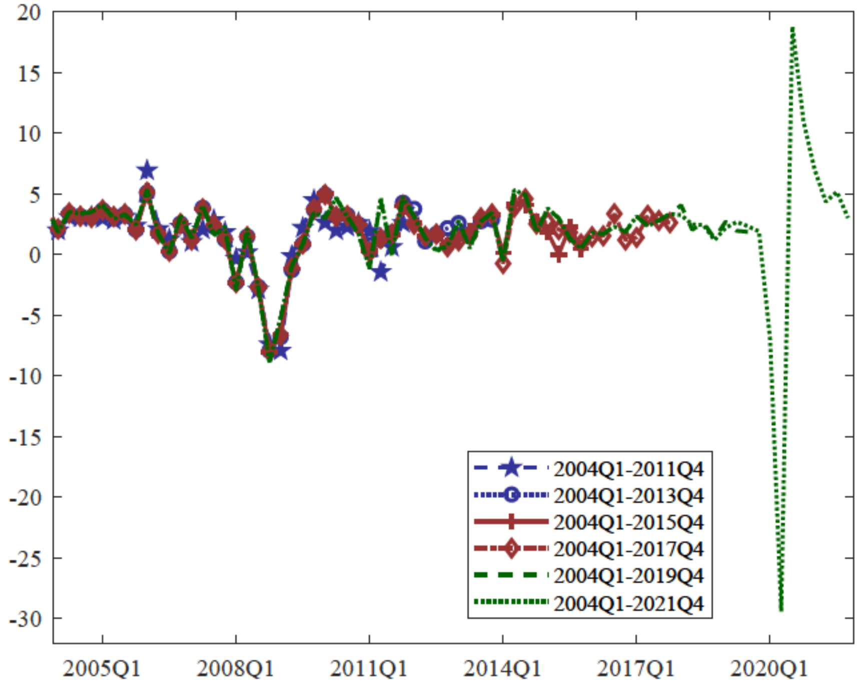
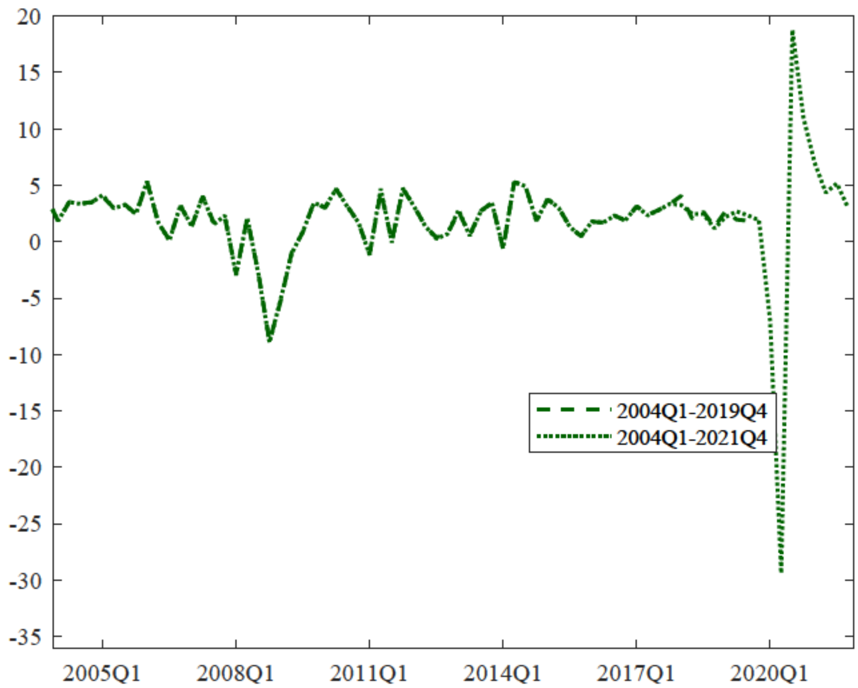


FIGURE 3. **Root mean square error improvements from using all releases for a given comprehensive revision** The solid green line is the posterior median of the root-MSEs while the shaded areas are point-wise 90% probability intervals.

FIGURE 4. **Signal extraction** for Δx_{Ct}. The solid line is the median of Δx_{Ct} given $y_{1:T}$ when all signals are used while the dashed line refers to its conditional median when only the most recent signal is used. Shaded areas represent point-wise 90%-probability intervals.



(a) Stability with respect to 2018 comprehensive revision and pandemic data



(b) Stability with respect to pandemic data

FIGURE 5. ***GDPsolera* releases.** The first release uses data until January of 2012 to provide estimates up to the 2011Q4. Similarly, the second one provides estimates up to 2013Q4 using data until January of 2014; following in that manner until the sixth one which, using data until January of 2022, delivers estimates of GDP growth until 2021Q4.

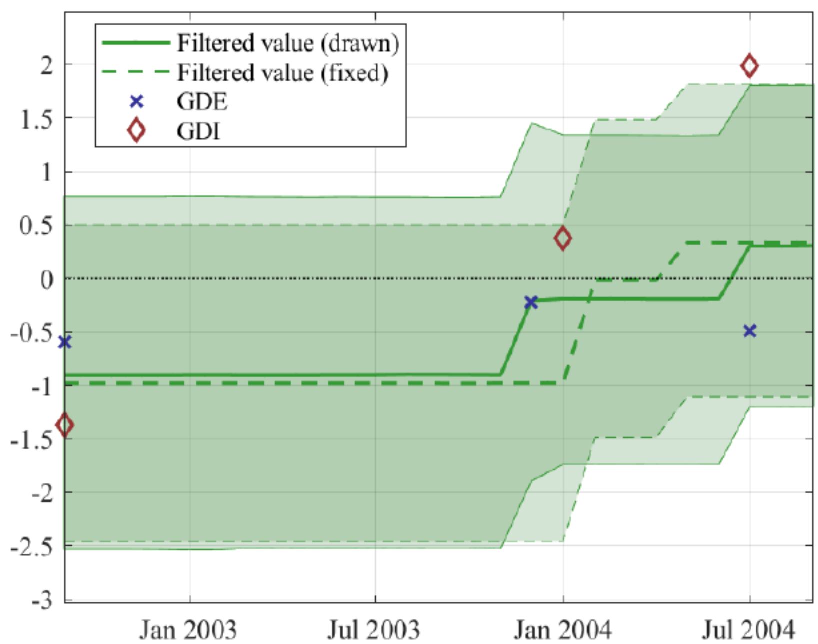


FIGURE 6. **Real time filtering of Δx_{2001Q1}.** The dashed and solid lines are posterior medians with and without parameter re-estimation while the shaded areas represent 90%-point-wise credible bands. Data releases for GDE (blue crosses) and GDI (red diamonds) are displayed too.

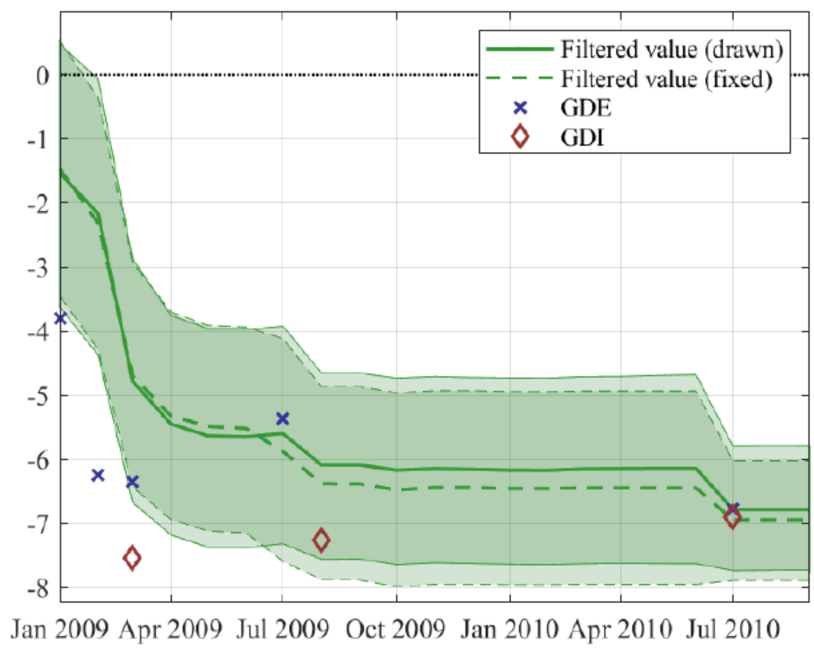


FIGURE 7. **Real time filtering of Δx_{2008Q4}.** The dashed and solid lines are posterior medians with and without parameter re-estimation while the shaded areas represent 90%-point-wise credible bands. Data releases for GDE (blue crosses) and GDI (red diamonds) are displayed too.

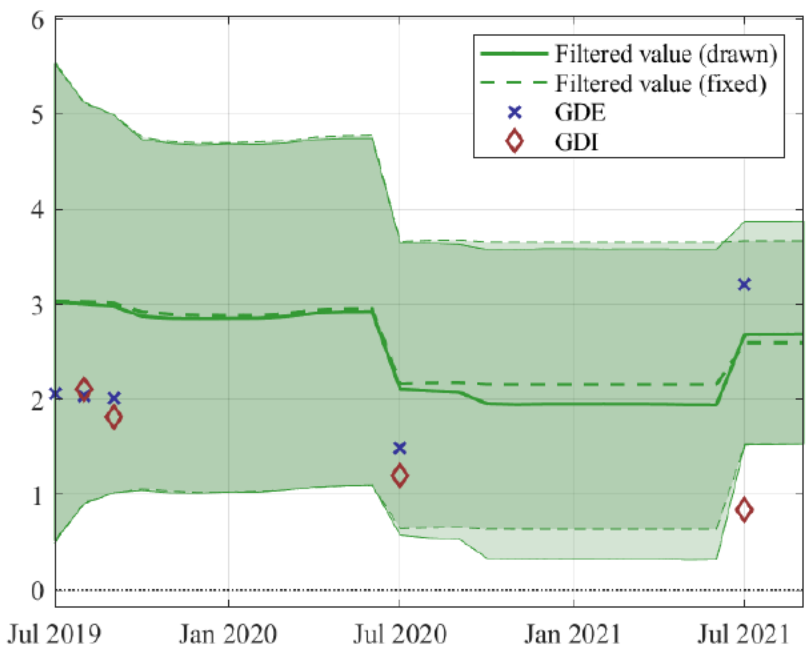


FIGURE 8. **Real time filtering of Δx_{2019Q2}.** The dashed and solid lines are posterior medians with and without parameter re-estimation while the shaded areas represent 90%-point-wise credible bands. Data releases for GDE (blue crosses) and GDI (red diamonds) are displayed too.

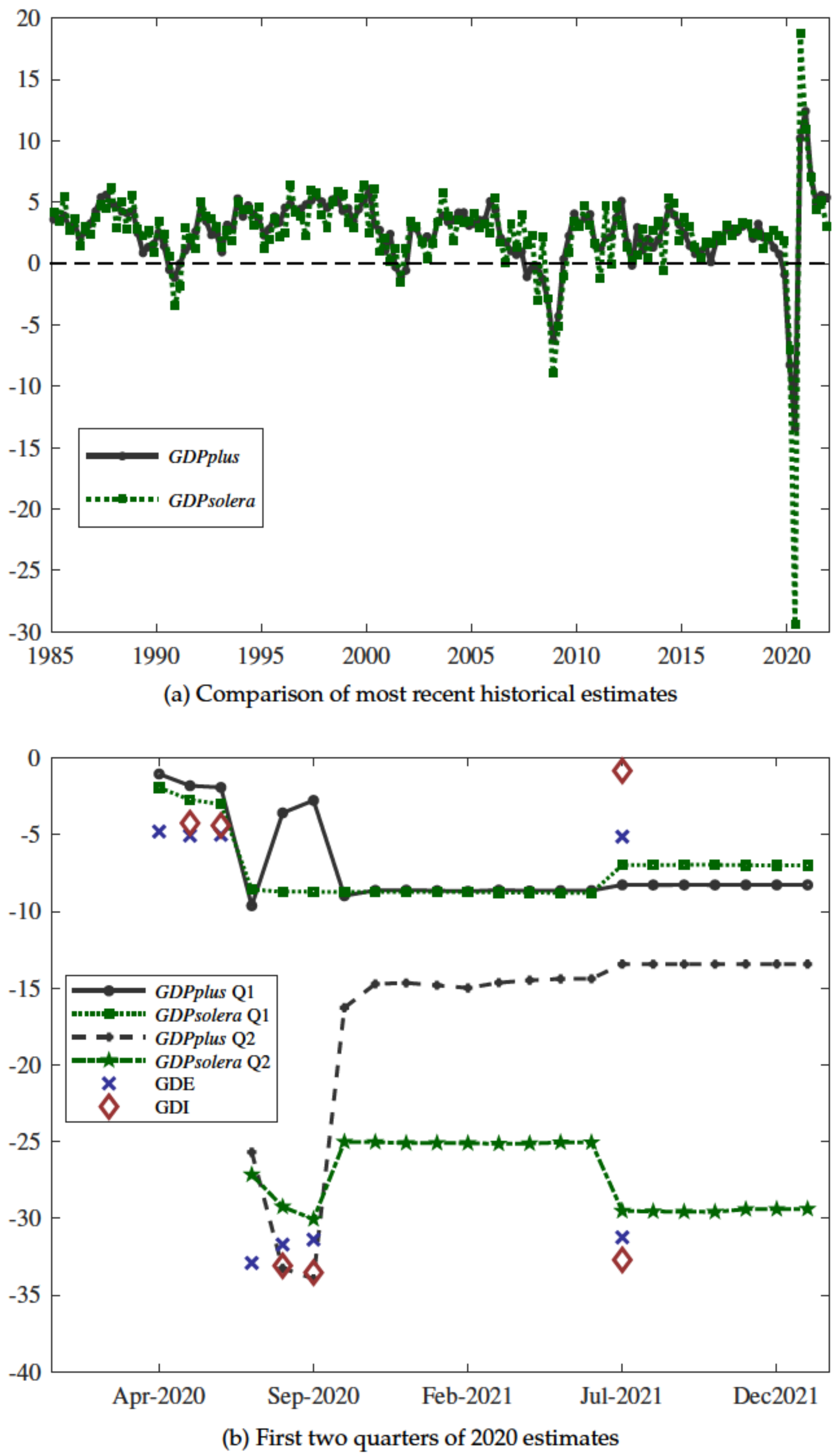
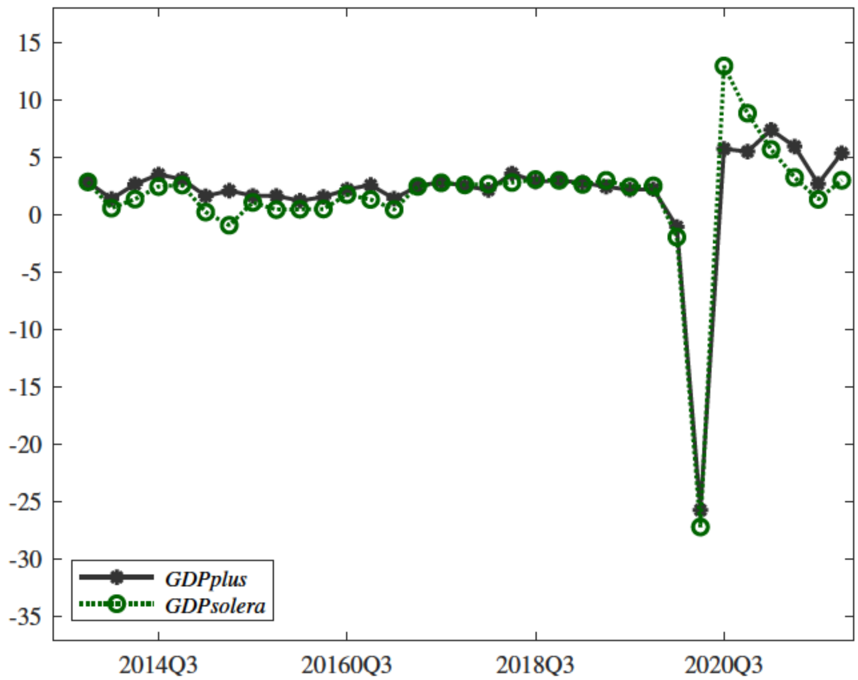
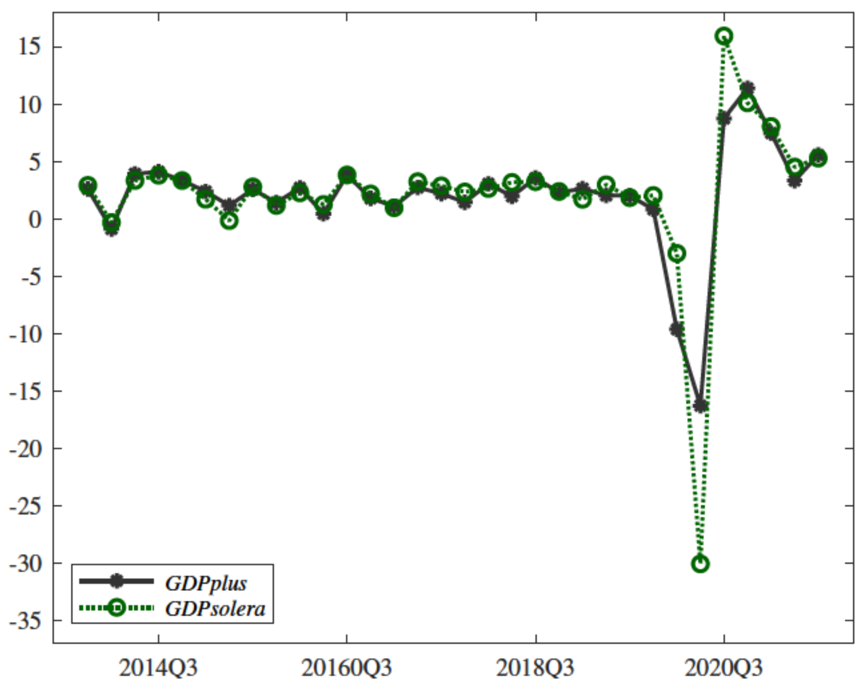


FIGURE 9. ***GDPplus* versus *GDPsolera*.** Panel (a) displays *GDPplus* and *GDPsolera* series estimated using data until January 2022. Panel (b) displays GDP estimates at the beginning of the COVID-19 outbreak revised from April 2020 to January 2022.



(a) GDP growth estimates one month after the end of the quarter



(b) GDP growth estimates three months after the end of the quarter

FIGURE 10. **Nowcast: *GDPplus* versus *GDPsolera*.** Panel (a) displays *GDPplus* and *GDPsolera* series estimated with information available one month after the end of the quarter when only advance of GDE is available for the most recent quarter. *GDPplus* for 2018Q4 was released in February 2019. Panel (b) displays *GDPplus* and *GDPsolera* series estimated with information available three months after the end of the quarter.

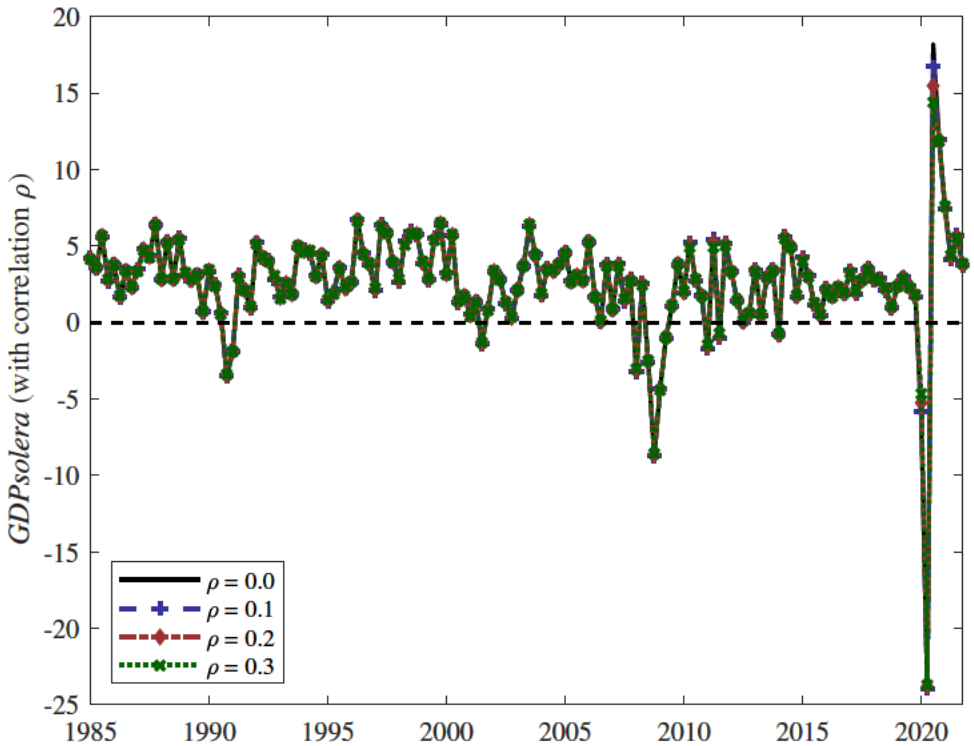


FIGURE 11. ***GDPsolera*** estimated using data until January 2022 allowing for correlation between the different shocks.

Appendix A Identification

A.1 Proof of proposition 1

Let f_ω denote the spectrum of a time series $\{\omega_t\}$. Identification of the autocovariance function of $\{\omega_t\}$ is equivalent to identification of f_ω. Hence, an alternative statement to proposition 1 is that under assumption 1, if $N > 1$, $f_{\Delta x}$ and $f_{v_1}, \ldots, f_{v_I}$ are nonparametrically identified from $f_{\Delta y}$.

To see why the proposition holds, let us write

$$f_{\Delta y}(\lambda) = 1_{M\times M} f_{\Delta x}(\lambda) + |1 - e^{i\lambda}|^2 \operatorname{diag}\Big(f_{v_1}(\lambda), \ldots, f_{v_N}(\lambda)\Big), \quad 0 \le \lambda \le 2\pi.$$

If E_i is the $M_i \times M$ matrix such that $y_{it} = E_i y_t$, we get $E_{i_1} f_{\Delta y}(\lambda) E'_{i_2} = 1_{M_{i_1}\times M_{i_2}} f_{\Delta x}(\lambda)$ for $i_1 \neq i_2$—such a pair i_1, i_2 exists only if $N > 1$. With $f_{\Delta x}$ pinned down, one then recovers

$$f_{v_i}(\lambda) = |1 - e^{i\lambda}|^{-2} E_i \left(f_{Dy}(\lambda) - 1_{M\times M} f_{Dx}(\lambda) \right) E'_i,$$

dealing with the removable singularity at $\lambda = 0$ by using that each entry f_{v_i} is holomorphic over the unit circle. ■

It follows from the proof of proposition 1 that if in addition to $N > 1$ we have $M_i > 1$ for at least one i, the model imposes overidentifying restrictions and is, therefore, testable. This is the case in our empirical analysis, although we do not pursue such tests. If the spectra $f_{\Delta x}, f_{v_1}, \ldots, f_{v_N}$ belong to a particular parametric class, an indirect approach to testing the overidentifying restrictions is to use dynamic specification tests as in Fiorentini and Sentana (2019).

A.2 Proof of proposition 2

By condition (b) in the proposition, $D_{i_1} = E_{i_1} \delta_t$ and $D_{i_2} = E_{i_2} \delta_t$ are time-invariant. By assumption 1 and condition (a), moreover, $E_{i_1} v_t$ and $E_{i_2} v_t$ are uncorrelated at all lags and leads. Ergo,

$$E_{i_1} f_{\Delta y}(\lambda) E'_{i_2} = D_{i_1} f_{\Delta x}(\lambda) D'_{i_2}, \quad 0 \le \lambda \le 2\pi.$$

Now, by condition (c), $\operatorname{rank}(D_{i_1}) = \operatorname{rank}(D_{i_2}) = C$. In that case,

$$f_{\Delta x} = (D'_{i_1} D_{i_1})^{-1} D'_{i_1} f_{\Delta y} D_{i_2} (D'_{i_2} D_{i_2})^{-1}.$$

Identification of $f_{v_1}, \ldots, f_{v_N}$ then follows by an analogous argument to that in proposition 1. ■

Appendix B Details of estimation algorithm

The algorithm updates unknowns by drawing iteratively from the following distributions:

(1) $p(X_{1:T}\big|\theta, y_{1:T})$: using the state-space representation of the model, $X_{1:T}$ is obtained from the simulation smoother proposed by Durbin and Koopman (2002).

(2) $p(\theta_x\big|\theta_1, \ldots, \theta_N, X_{1:T}, y_{1:T})$: first notice that $(\Delta x_{1:T}, \eta_{x,1:T})$ are sufficient for θ_x, i.e.,

$$p(\theta_x\big|\theta_1, \ldots, \theta_N, X_{1:T}, y_{1:T}) = p\Big(\theta_x\big|\Delta x_{1:T}, \eta_{x,1:T}\Big),$$

and because of the conjugacy of the prior we recover $\mu_x, \rho_x, \lambda_x, \sigma_x$ from

(i) $\pi_x = 1/\sigma_x^2|\Delta x_{1:T}, \eta_{x,1:T} \sim \Gamma_C(\tilde{d}_x/2, \tilde{p}_x/\tilde{d}_x)$ where

$$\begin{aligned}
\tilde{d}_x &= d_x + T - 1, \\
\frac{\tilde{d}_x}{\tilde{p}_x} &= \frac{d_x}{p_x} + \sum_{t=2}^{T} (\Delta x_t - \mu_x - \operatorname{diag}(\rho_x)(\Delta x_{t-1} - \mu_x) - \lambda_x \eta_{xt})^2 ;
\end{aligned}$$

(ii) $\beta_x = ((I_C - \operatorname{diag}(\rho_x)\mu_x, \rho_x, \lambda_x)|\sigma_x, \Delta x_{1:T}, \eta_{x,1:T} \sim N(\tilde{b}_x, \tilde{R}_x \otimes \operatorname{diag}(\sigma_x^2))$ where

$$\begin{aligned}
\tilde{R}_x &= R_x + \sum_{t=2}^{T} \begin{pmatrix} 1 & \Delta x_{t-1} & \eta_{xt} \\ \Delta x_{t-1} & \Delta x_{t-1}^2 & \Delta x_{t-1}\eta_{xt} \\ \eta_{xt} & \Delta x_{t-1}\eta_{xt} & \eta_{xt}^2 \end{pmatrix}, \\
\tilde{R}_x \tilde{b}_x &= R_x b_x + \sum_{t=2}^{T} \begin{pmatrix} \Delta x_t \\ \Delta x_{t-1}\Delta x_t \\ \eta_{xt}\Delta x_t \end{pmatrix}.
\end{aligned}$$

(3) $p\Big(\theta_i\big|\theta_x, (\theta_j)_{j\neq i}, X_{1:T}, y_{1:T}\Big)$ for each i: first notice that $(v_{i,1:T}, \eta_{i,1:T})$ are sufficient for θ_i, i.e.,

$$p\Big(\theta_i\big|\theta_x, (\theta_j)_{j\neq i}, X_{1:T}, y_{1:T}\Big) = p\Big(\theta_i\big|v_{i,1:T}, \eta_{i,1:T}\Big),$$

and because of the conjugacy of the prior we recover $\rho_i, \lambda_i, \sigma_i$ from

(i) $\pi_i = 1/\sigma_i^2|v_{i,1:T}, \eta_{i,1:T} \sim \Gamma_{M_i}(\tilde{d}_i/2, \tilde{p}_i/\tilde{d}_i)$ where

$$\tilde{d}_i = d_i + T - 1,$$

$$\frac{\tilde{d}_i}{\tilde{p}_i} = \frac{d_i}{p_i} + \sum_{t=2}^{T} \left(v_{it} - \mathrm{diag}(\rho_i) v_{i,t-1} - \lambda_i \eta_{it} \right)^2 ;$$

(ii) $\beta_i = (\rho_i, \lambda_i) | \sigma_i, v_{i,1:T}, \eta_{i,1:T} \sim N(\tilde{b}_i, \tilde{R}_i \otimes \mathrm{diag}(\sigma_i^2))$ where

$$\tilde{R}_i = R_i + \sum_{t=2}^{T} \begin{pmatrix} v_{i,t-1}^2 & v_{i,t-1}\eta_{it} \\ v_{i,t-1}\eta_{it} & \eta_{it}^2 \end{pmatrix},$$

$$\tilde{R}_i \tilde{b}_i = R_i b_i + \sum_{t=2}^{T} \begin{pmatrix} v_{i,t-1} v_{it} \\ \eta_{it} v_{it} \end{pmatrix}.$$

A small comment is that the choice of hyperparameters $d_x = 0_{C\times 1}$, $R_x = 0_{3\times 3}$, $d_i = 0_{M_i \times 1}$, and $R_i = 0_{2\times 2}$, despite implying improper priors, still leads to a well-defined algorithm and a proper posterior distribution.

Appendix C Posterior distributions

Parameter	Post. mean	90%-CI	MC s.e.
${\mu_x}^{(1)}$	2.579	[2.231, 2.935]	0.0029
${\mu_x}^{(2)}$	2.639	[2.334, 2.966]	0.0028
${\mu_x}^{(3)}$	2.623	[2.317, 2.949]	0.0030
${\mu_x}^{(4)}$	2.679	[2.390, 2.985]	0.0028
${\mu_x}^{(5)}$	2.711	[2.406, 3.035]	0.0030
${\rho_x}^{(1)}$	0.514	[0.466, 0.562]	0.0003
${\rho_x}^{(2)}$	0.512	[0.470, 0.554]	0.0003
${\rho_x}^{(3)}$	0.511	[0.471, 0.552]	0.0003
${\rho_x}^{(4)}$	0.501	[0.462, 0.539]	0.0003
${\rho_x}^{(5)}$	0.448	[0.405, 0.491]	0.0003
${\lambda_x}^{(1)}$	2.688	[2.462, 2.911]	0.0019
${\lambda_x}^{(2)}$	2.664	[2.487, 2.846]	0.0015
${\lambda_x}^{(3)}$	2.751	[2.574, 2.929]	0.0014
${\lambda_x}^{(4)}$	2.796	[2.634, 2.957]	0.0013
${\lambda_x}^{(5)}$	3.227	[3.055, 3.410]	0.0015
${\sigma_x}^{(1)}$	0.710	[0.613, 0.831]	0.0005
${\sigma_x}^{(2)}$	0.622	[0.543, 0.720]	0.0004
${\sigma_x}^{(3)}$	0.618	[0.541, 0.712]	0.0004
${\sigma_x}^{(4)}$	0.567	[0.501, 0.643]	0.0003
${\sigma_x}^{(5)}$	0.651	[0.564, 0.759]	0.0006
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(1)}$	0.055	[0.014, 0.095]	0.0002
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(2)}$	0.055	[0.027, 0.082]	0.0002
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(3)}$	0.059	[0.029, 0.090]	0.0002
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(4)}$	0.115	[−0.008, 0.239]	0.0005
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(5)}$	0.092	[0.005, 0.185]	0.0004
${\rho_{\mathrm{GDE}}^{\mathrm{nc}}}^{(6)}$	0.068	[−0.009, 0.159]	0.0006

${\rho^{c}_{GDE}}^{(1)}$	0.036	[−0.003, 0.079]	0.0002
${\rho^{c}_{GDE}}^{(2)}$	0.010	[−0.021, 0.043]	0.0002
${\rho^{c}_{GDE}}^{(3)}$	−0.006	[−0.036, 0.025]	0.0001
${\rho^{c}_{GDE}}^{(4)}$	−0.002	[−0.034, 0.033]	0.0002
${\rho^{c}_{GDE}}^{(5)}$	0.014	[−0.024, 0.056]	0.0002
${\lambda^{nc}_{GDE}}^{(1)}$	1.638	[1.534, 1.744]	0.0012
${\lambda^{nc}_{GDE}}^{(2)}$	1.604	[1.520, 1.690]	0.0012
${\lambda^{nc}_{GDE}}^{(3)}$	1.580	[1.494, 1.670]	0.0012
${\lambda^{nc}_{GDE}}^{(4)}$	0.581	[0.390, 0.777]	0.0017
${\lambda^{nc}_{GDE}}^{(5)}$	−0.148	[−0.298, 0.003]	0.0017
${\lambda^{nc}_{GDE}}^{(6)}$	−0.076	[−0.216, 0.056]	0.0014
${\lambda^{c}_{GDE}}^{(1)}$	1.196	[1.085, 1.308]	0.0011
${\lambda^{c}_{GDE}}^{(2)}$	0.950	[0.850, 1.054]	0.0011
${\lambda^{c}_{GDE}}^{(3)}$	0.763	[0.668, 0.852]	0.0011
${\lambda^{c}_{GDE}}^{(4)}$	0.624	[0.532, 0.712]	0.0010
${\lambda^{c}_{GDE}}^{(5)}$	0.707	[0.613, 0.800]	0.0010
${\sigma^{nc}_{GDE}}^{(1)}$	0.594	[0.515, 0.689]	0.0004
${\sigma^{nc}_{GDE}}^{(2)}$	0.300	[0.256, 0.358]	0.0003
${\sigma^{nc}_{GDE}}^{(3)}$	0.364	[0.308, 0.433]	0.0003
${\sigma^{nc}_{GDE}}^{(4)}$	1.574	[1.379, 1.813]	0.0011
${\sigma^{nc}_{GDE}}^{(5)}$	0.827	[0.679, 0.995]	0.0010
${\sigma^{nc}_{GDE}}^{(6)}$	0.683	[0.524, 0.855]	0.0015
${\sigma^{c}_{GDE}}^{(1)}$	0.379	[0.300, 0.492]	0.0007
${\sigma^{c}_{GDE}}^{(2)}$	0.303	[0.252, 0.374]	0.0003
${\sigma^{c}_{GDE}}^{(3)}$	0.293	[0.244, 0.360]	0.0004
${\sigma^{c}_{GDE}}^{(4)}$	0.316	[0.260, 0.393]	0.0005
${\sigma^{c}_{GDE}}^{(5)}$	0.376	[0.300, 0.484]	0.0008
${\rho^{nc}_{GDI}}^{(1)}$	0.306	[0.175, 0.427]	0.0007
${\rho^{nc}_{GDI}}^{(2)}$	0.334	[0.222, 0.444]	0.0006
${\rho^{nc}_{GDI}}^{(3)}$	0.395	[0.285, 0.500]	0.0005
${\rho^{nc}_{GDI}}^{(4)}$	0.462	[0.371, 0.550]	0.0006
${\rho^{nc}_{GDI}}^{(5)}$	0.400	[0.316, 0.486]	0.0005
${\rho^{c}_{GDI}}^{(1)}$	0.000	[−0.041, 0.044]	0.0002
${\rho^{c}_{GDI}}^{(2)}$	0.365	[0.300, 0.429]	0.0003
${\rho^{c}_{GDI}}^{(3)}$	0.226	[0.170, 0.282]	0.0003
${\rho^{c}_{GDI}}^{(4)}$	0.188	[0.150, 0.229]	0.0003
${\rho^{c}_{GDI}}^{(5)}$	0.217	[0.176, 0.260]	0.0004
${\lambda^{nc}_{GDI}}^{(1)}$	1.260	[0.993, 1.533]	0.0015
${\lambda^{nc}_{GDI}}^{(2)}$	1.236	[0.985, 1.490]	0.0013
${\lambda^{nc}_{GDI}}^{(3)}$	0.915	[0.693, 1.140]	0.0013
${\lambda^{nc}_{GDI}}^{(4)}$	1.018	[0.818, 1.218]	0.0014
${\lambda^{nc}_{GDI}}^{(5)}$	1.228	[1.029, 1.422]	0.0015
${\lambda^{c}_{GDI}}^{(1)}$	0.822	[0.713, 0.927]	0.0011
${\lambda^{c}_{GDI}}^{(2)}$	1.259	[1.111, 1.411]	0.0010
${\lambda^{c}_{GDI}}^{(3)}$	1.707	[1.566, 1.851]	0.0010
${\lambda^{c}_{GDI}}^{(4)}$	1.913	[1.820, 2.010]	0.0008
${\lambda^{c}_{GDI}}^{(5)}$	1.982	[1.885, 2.082]	0.0008
${\sigma^{nc}_{GDI}}^{(1)}$	1.661	[1.412, 1.975]	0.0017
${\sigma^{nc}_{GDI}}^{(2)}$	1.704	[1.479, 1.981]	0.0014
${\sigma^{nc}_{GDI}}^{(3)}$	1.720	[1.501, 1.988]	0.0012
${\sigma^{nc}_{GDI}}^{(4)}$	1.360	[1.184, 1.570]	0.0010

${\sigma^{\text{nc}}_{\text{GDI}}}^{(5)}$	1.270	[1.101, 1.473]	0.0012
${\sigma^{\text{c}}_{\text{GDI}}}^{(1)}$	0.392	[0.306, 0.527]	0.0011
${\sigma^{\text{c}}_{\text{GDI}}}^{(2)}$	0.853	[0.730, 1.004]	0.0007
${\sigma^{\text{c}}_{\text{GDI}}}^{(3)}$	0.753	[0.644, 0.876]	0.0007
${\sigma^{\text{c}}_{\text{GDI}}}^{(4)}$	0.402	[0.325, 0.493]	0.0006
${\sigma^{\text{c}}_{\text{GDI}}}^{(5)}$	0.429	[0.342, 0.531]	0.0007

TABLE C.1. Posterior distribution of parameters of the model

NOTES. Unconditional means μ_x, loadings $\lambda_x, \lambda_{\text{GDE}}, \lambda_{\text{GDI}}$ and standard deviations, $\sigma_x, \sigma_{\text{GDE}}, \sigma_{\text{GDI}}$ are annualized.

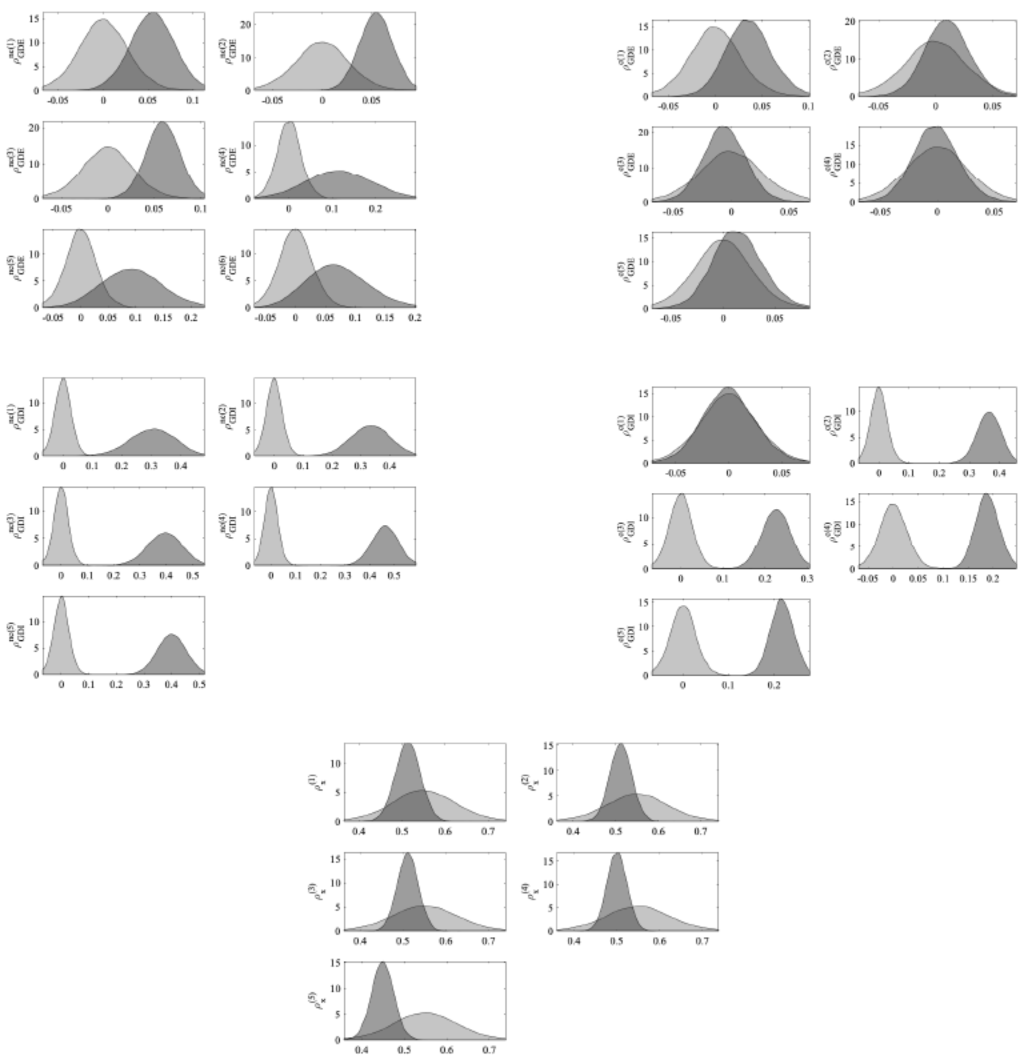


FIGURE C.1. **Parameter estimation.** Priors (light area) and posteriors (dark area) distributions of the parameters.

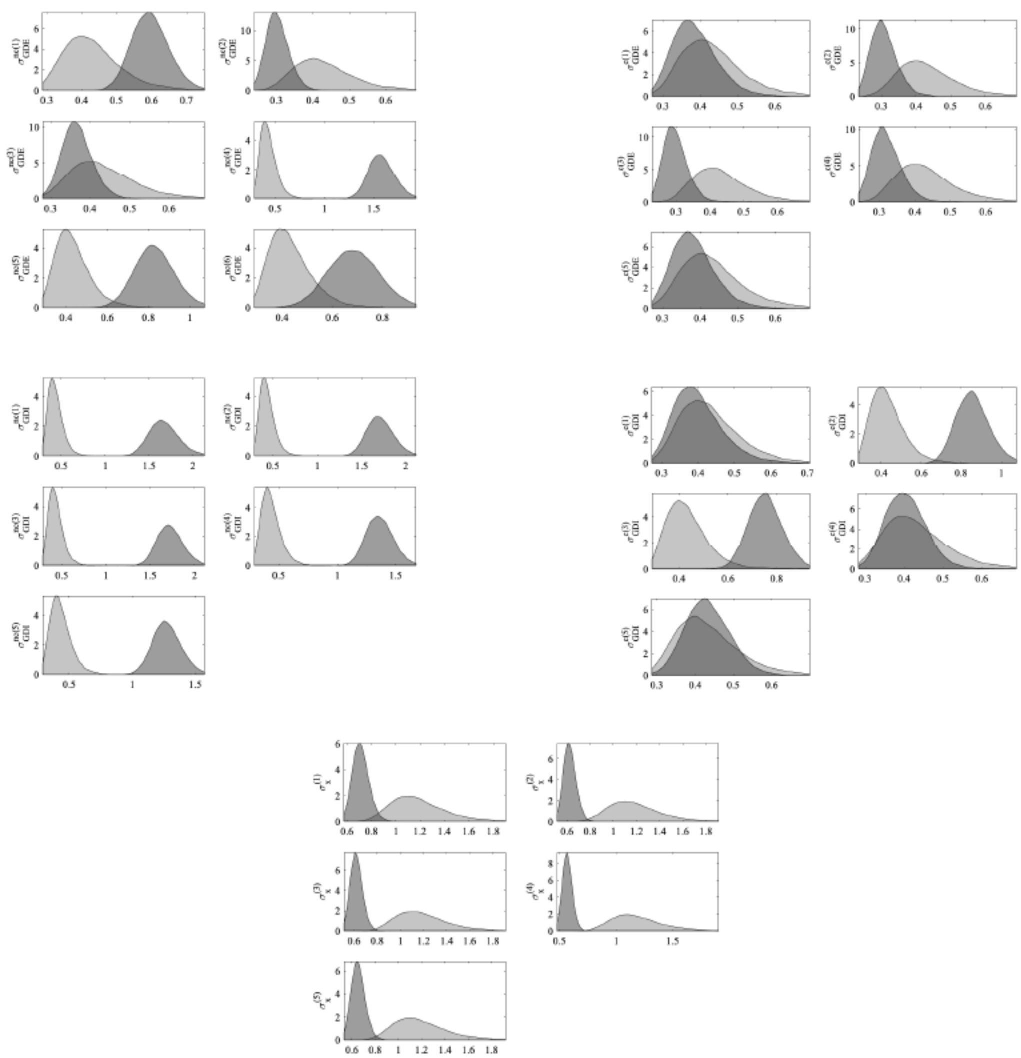


FIGURE C.2. **Parameter estimation.** Priors (light area) and posteriors (dark area) distributions of the parameters.

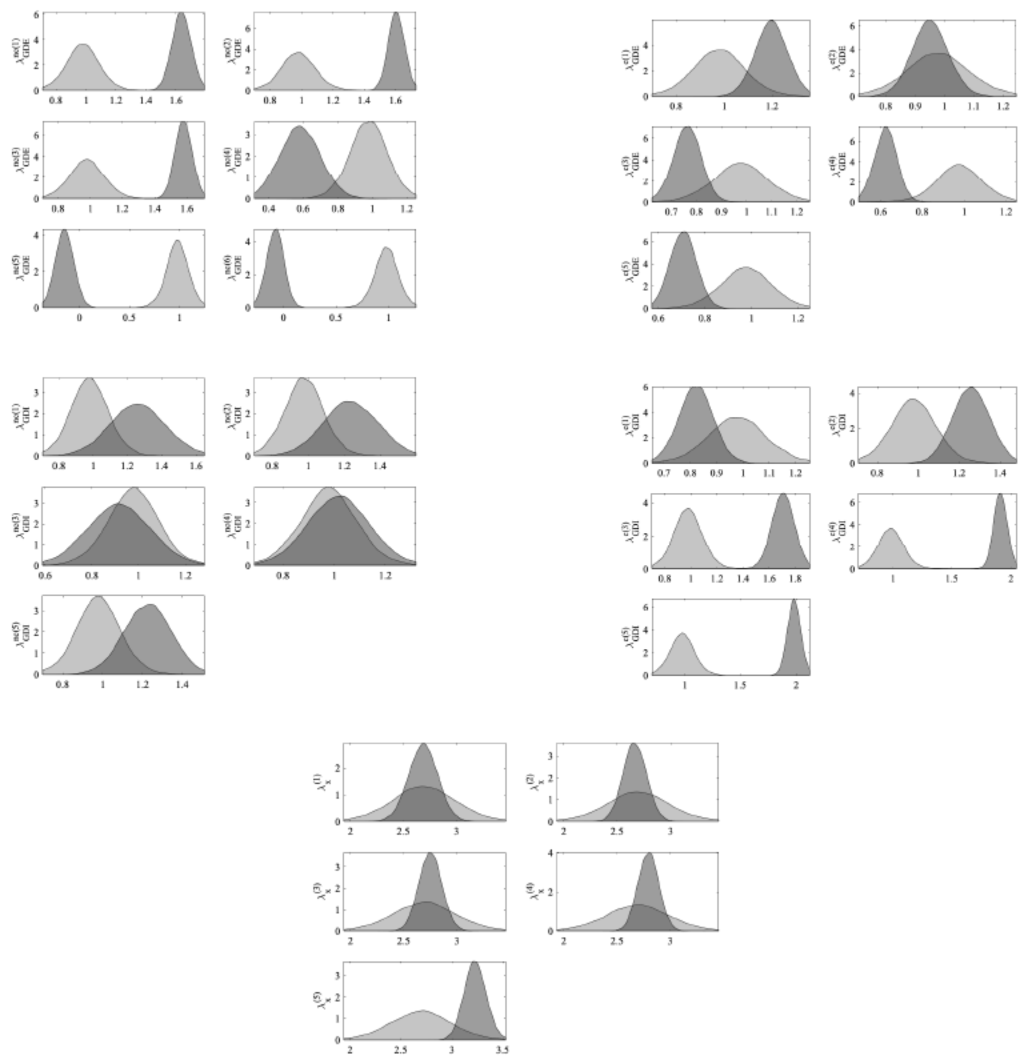


FIGURE C.3. **Parameter estimation.** Priors (light area) and posteriors (dark area) distributions of the parameters.

Appendix D Implications of L^2-optimality

Consider the following model for the release process. For each type of estimate i and quarter t, the statistical office collects inputs $\iota_{it}^1, \ldots, \iota_{it}^{J_{it}}$ – such as sectoral surveys – on which the estimates y_{it}^m are based. Our objective is to show that if estimates were produced to minimize expected square loss (i.e., the L^2-distance between the estimate and x_t), the optimal signal-extraction rule would map x_t to its most recent release. For ease of exposition, let $C = 1$ (which gives the model with no comprehensive revisions). Fix i and t and let $\sigma(\cdot)$ denote a (generated) σ-algebra. We will assume that (i) there are integers $\{J_{it}^m\}_{m=1}^{M_i}$ such that $J_{it}^m \leq J_{it}^{m+1}$ and y_{it}^m is $\mathcal{I}_{it}^m$-measurable with $\mathcal{I}_{it}^m = \sigma\{\iota_{it}^1, \ldots, \iota_{it}^{J_{it}^m}\}$ for all m, and (ii) the statistical office minimizes $L^2(y_{it}^m - x_t) = \mathbb{E}\left[|y_{it}^m - x_t|^2\right]$. Assumption (i) allows for data on past and future periods to be included among the time-t inputs. We also assume x_t has finite variance by an appropriate choice of initial conditions. From (i) we obtain $\mathcal{I}_{it}^m \subset \mathcal{I}_{it}^{m+1}$ for all m, and from (ii),

$$y_{it}^m = \mathbb{E}\left[x_t \middle| \mathcal{I}_{it}^m\right], \quad m = 1, \ldots, M_i.$$

Let $\tilde{\mathcal{I}}_{it}$ be a σ-algebra such that $\tilde{\mathcal{I}}_{it} \subset \mathcal{I}_{it}^m$ for all m. For example, if the time-t inputs include all the data needed to construct past measurements, $\tilde{\mathcal{I}}_{it}$ may be the σ-algebra generated by all past measurements. With a slight abuse of notation,

$$\mathbb{E}\left[x_t \middle| y_{it}^1, \ldots, y_{it}^m, \tilde{\mathcal{I}}_{it}\right] = \mathbb{E}\left[\mathbb{E}\left[x_t \middle| \mathcal{I}_{it}^m\right] \middle| y_{it}^1, \ldots, y_{it}^m, \tilde{\mathcal{I}}_{it}\right] = \mathbb{E}\left[y_{it}^m \middle| y_{it}^1, \ldots, y_{it}^m, \tilde{\mathcal{I}}_{it}\right] = y_{it}^m,$$

by the law of iterated expectations.

In words, if measurements minimize expected square loss, all measurements of x_t but the most recent one contain no useful information to extract x_t. A reasonable situation is one where the statistical office computes y_{it}^m using input data corresponding only to quarter-t economic activity. A measure that captures L^2-optimality in that context would compare the expected loss of y_{it}^m with that of $\mathbb{E}\left[x_t \middle| y_{it}^1, \ldots, y_{it}^m\right]$ (i.e., taking $\tilde{\mathcal{I}}_{it} = \emptyset$). For example,

$$D_{it}^m = \mathrm{Var}\left(\mathbb{E}\left[x_t \middle| y_{it}^1, \ldots, y_{it}^m\right] - x_t\right) \Big/ \mathrm{Var}\left(v_{it}^m\right).$$

We have $0 \leq D_{it}^m \leq 1$ with $D_{it}^m = 1$ indicating full L^2-optimality. Thus, $D_{it}^m < 1$ may be evidence that, for example, the measurements optimize a different loss function or the weights given to the inputs disregard the dynamic model.

Appendix E News and noise model

Consider a setup in which $N = 1$, so that we can omit the subindex indicating type, which would be 1, $M = M_1 = 3$, and there is a single comprehensive version of GDP, so that $C = 1$. Suppose the data follows the news-noise model of Jacobs and van Norden (2011) and Jacobs et al. (2022):

$$\Delta y_t = \begin{pmatrix} \Delta y_t^1 \\ \Delta y_t^2 \\ \Delta y_t^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \Delta \tilde{y}_t + \begin{pmatrix} \nu_t^1 \\ \nu_t^2 \\ \nu_t^3 \end{pmatrix} + \begin{pmatrix} \zeta_t^1 \\ \zeta_t^2 \\ \zeta_t^3 \end{pmatrix} = 1_{3\times 1} \Delta \tilde{y}_t + \nu_t + \zeta_t,$$

where ν_t^m and ζ_t^m are news and noise components. News are defined by the condition that $\text{Cov}\left(\nu_t^m, \Delta \tilde{y}_t + \nu_t^{m'}\right) = 0$ for all $m' \leq m$, while noise must satisfy $\text{Cov}(\zeta_t^m, \Delta \tilde{y}_t + \nu_t^m) = 0$. These, however, are not enough to pin down a unique decomposition of y_t in terms of $\tilde{y}_t, \nu_t, \zeta_t$ and we will further impose $\zeta_t^1, \zeta_t^2, \zeta_t^3$ are uncorrelated with each other.

To simplify the argument, we will assume that (i) $\Delta \tilde{y}_t + \nu_t^3$ follows an AR(1) process and (ii) ν_t and ζ_t are uncorrelated over time. Moreover, we note that the news-noise model is typically applied to measurements of GDP growth, as opposed to our model, which focuses on the level.

The goal is to understand how the news-noise model maps to ours, namely

$$\begin{pmatrix} y_t^1 \\ y_t^2 \\ y_t^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} x_t + \begin{pmatrix} v_t^1 \\ v_t^2 \\ v_t^3 \end{pmatrix} = 1_{3\times 1} x_t + v_t.$$

We can write

$$\Delta y_t = \begin{pmatrix} \Delta y_t^1 \\ \Delta y_t^2 \\ \Delta y_t^3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} (\Delta \tilde{y}_t + \nu_t^3) + \begin{pmatrix} (\nu_t^2 - \nu_t^3) + (\nu_t^1 - \nu_t^2) \\ (\nu_t^2 - \nu_t^3) \\ 0 \end{pmatrix} + \begin{pmatrix} \zeta_t^1 \\ \zeta_t^2 \\ \zeta_t^3 \end{pmatrix},$$

where $\nu_t^1 - \nu_t^2, \nu_t^2 - \nu_t^3, \zeta_t^1, \zeta_t^2, \zeta_t^3$ are mutually orthogonal white noise processes. If we set

$$\begin{aligned} \Delta x_t &= \Delta \tilde{y}_t + \nu_t^3, \\ \Delta v_t^1 &= (\nu_t^2 - \nu_t^3) + (\nu_t^1 - \nu_t^2) + \zeta_t^1, \\ \Delta v_t^2 &= (\nu_t^2 - \nu_t^3) + \zeta_t^2, \end{aligned}$$

$$\Delta v_t^3 = \zeta_t^3,$$

we obtain a particular case of our model in which, not surprisingly, $\rho = 1_{3\times 1}$. Measurement error are therefore white noise in first differences with a particular variance matrix,

$$\text{Var}\left(\Delta v_t\right) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} & 0 \\ \Sigma_{12} & \Sigma_{22} & 0 \\ 0 & 0 & \Sigma_{33} \end{pmatrix}.$$

If we give Δv_t the factor structure in (4) (again maintaining $\rho = 1_{M\times 1}$),

$$\Delta v_t = \begin{pmatrix} \Delta v_t^1 \\ \Delta v_t^2 \\ \Delta v_t^3 \end{pmatrix} = \begin{pmatrix} \lambda^1 \\ \lambda^2 \\ \lambda^3 \end{pmatrix} \eta_t + \begin{pmatrix} \sigma^1 \varepsilon_t^1 \\ \sigma^2 \varepsilon_t^2 \\ \sigma^3 \varepsilon_t^3 \end{pmatrix} = \lambda \eta_t + \text{diag}(\sigma)\varepsilon_t,$$

with $\eta_t \overset{iid}{\sim} N(0,1)$, $\varepsilon_t \overset{iid}{\sim} N(0_{3\times 1}, I_3)$ and η_t independent of ε_t, the news-noise model implies the restriction $\lambda_3 = 0$. The rest of the parameters, $\lambda^1, \lambda^2, \sigma^1, \sigma^2, \sigma^3$, can be recovered from Var (Δv_t).

References

Almuzara, M., D. Amengual, and E. Sentana (2019): "Normality tests for latent variables," *Quantitative Economics*, 10, 981–1017.

Almuzara, M., G. Fiorentini, and E. Sentana (2022): "US aggregate output measurement: A common trend approach," CEMFI Working Paper 2101, revised March 2022, forthcoming in Y. Chang, S. Lee and J.I. Miller (eds.) Essays in honor of Joon Y. Park, Advances in Econometrics.

Aruoba, S. B. (2008): "Data revisions are not well behaved," *Journal of Money, Credit and Banking*, 40, 319–340.

Aruoba, S. B., F. X. Diebold, J. Nalewaik, F. Schorfheide, and D. Song (2016): "Improving GDP measurement: A measurement-error perspective," *Journal of Econometrics*, 191, 384–397.

Campbell, J. and G. Mankiw (1987): "Are output fluctuations transitory?" *Quarterly Journal of Economics*, 102, 857–880.

Durbin, J. and S. J. Koopman (2002): "A simple and efficient simulation smoother for state space time series analysis," *Biometrika*, 89, 603–615.

Fiorentini, G. and E. Sentana (2019): "Dynamic specification tests for dynamic factor models," *Journal of Applied Econometrics*, 34, 325–346.

Greenaway-McGrevy, R. (2011): "Is GDP or GDI a better measure of output? A statistical approach," Working Paper 2011–08, Bureau of Economic Analysis.

Grimm, B. T. (2007): "The statistical discrepancy," Working Paper 2007-01, Bureau of Economic Analysis.

Jacobs, J. P. A. M., S. Sarferaz, J. Sturm, and S. van Norden (2022): "Can GDP measurement be further improved? Data revision and reconciliation," *Journal of Business and Economic Statistics*, 40, 423–431.

Jacobs, J. P. A. M. and S. van Norden (2011): "Modeling data revisions: Measurement error and dynamics of "true" values," *Journal of Econometrics*, 161, 101–109.

Kerry, P., S. H. McCulla, and D. B. Wasshausen (2018): "Improved estimates of the national income and product accounts: Results of the 2018 comprehensive update," *Survey of Current Business*, 98.

Landefeld, J. S., E. P. Seskin, and B. M. Fraumeni (2008): "Taking the pulse of the economy: Measuring GDP," *Journal of Economic Perspectives*, 22, 193–216.

Lewis, D. J., K. Mertens, J. H. Stock, and M. Trivedi (Forthcoming): "Measuring real activity using a weekly economic index," *Journal of Applied Econometrics*.

Nalewaik, J. (2010): "The income- and expenditure-side measures of output growth," *Brookings Papers on Economic Activity*, 1, 71–106.

——— (2011): "The income- and expenditure-side measures of output growth – An update through 2011Q2," *Brookings Papers on Economic Activity*, 2, 385–402.

Orphanides, A. (2001): "Monetary policy rules based on real-time data," *American Economic Review*, 94, 964–984.

Smith, R. J., M. R. Weale, and S. E. Satchell (1998): "Measurement error with accounting constraints: Point and interval estimation for latent data with an application to U.K. gross domestic product," *Review of Economic Studies*, 65, 109–134.

Stock, J. H. (1994): "Unit roots, structural breaks and trends," in *Handbook of Econometrics*, ed. by D. McFadden and R. Engle, Elsevier, vol. 4, chap. 46.

Stone, R., D. G. Champernowne, and J. E. Meade (1942): "The precision of national income estimates," *Review of Economic Studies*, 9, 111–125.

Weale, M. (1992): "Estimation of data measured with error and subject to linear restrictions," *Journal of Applied Econometrics*, 7, 167–174.