Journal of Econometrics 191 (2016) 384–397

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

Improving *GDP* measurement: A measurement-error perspective

S. Boragan Aruoba [a,*], Francis X. Diebold [b], Jeremy Nalewaik [c], Frank Schorfheide [b], Dongho Song [d]

[a] *University of Maryland, United States*
[b] *University of Pennsylvania, United States*
[c] *Federal Reserve Board, United States*
[d] *Boston College, United States*

ARTICLE INFO

Article history:
Available online 7 December 2015

JEL classification:
E01
E32

Keywords:
Income
Output
Expenditure
Business cycle
Expansion
Contraction
Recession
Turning point
State-space model
Dynamic factor model
Forecast combination

ABSTRACT

We provide a new measure of historical U.S. *GDP* growth, obtained by applying optimal signal-extraction techniques to the noisy expenditure-side and income-side *GDP* estimates. The quarter-by-quarter values of our new measure often differ noticeably from those of the traditional measures. Its dynamic properties differ as well, indicating that the persistence of aggregate output dynamics is stronger than previously thought.

1. Introduction

Aggregate real output is surely the most fundamental and important concept in macroeconomic theory. Surprisingly, however, significant uncertainty still surrounds its historical measurement. In the U.S., in particular, two often-divergent *GDP* estimates exist, a widely-used expenditure-side version, GDP_E, and a much less widely-used income-side version, GDP_I.[1] Nalewaik (2010) and Fixler and Nalewaik (2009) make clear that, at the very least, GDP_I deserves serious attention and may even have properties in certain respects superior to those of GDP_E.[2] That is, if forced to choose between GDP_E and GDP_I, a surprisingly strong case exists for GDP_I. But of course one is *not* forced to choose between GDP_E and GDP_I, and a *GDP* estimate based on *both* GDP_E and GDP_I may be superior to either one alone. In this paper we propose and implement a framework for obtaining such a blended estimate.

Our work is related to, and complements, (Aruoba et al., 2012). There we took a forecast-error perspective, whereas here we take a measurement-error perspective.[3] In particular, we work with a dynamic factor model in the tradition of Geweke (1977) and Sargent and Sims (1977), as used and extended by Watson and Engle (1983), Edwards and Howrey (1991), Harding and Scutella (1996), Jacobs and van Norden (2011), Kishor and Koenig (2012), and Fleischman and Roberts (2011), among others.[4] That is, we view "true *GDP*" as a latent variable on which we have several

* Corresponding author.
E-mail addresses: aruoba@econ.umd.edu (S.B. Aruoba), fdiebold@sas.upenn.edu (F.X. Diebold), jeremy.j.nalewaik@frb.gov (J. Nalewaik), schorf@ssc.upenn.edu (F. Schorfheide), dongho.song@bc.edu (D. Song).

1 Indeed we will focus on the U.S. because it is a key egregious example of unreconciled GDP_E and GDP_I estimates.

2 For additional informative background on GDP_E, GDP_I, the statistical discrepancy, and the national accounts more generally, see BEA (2006), McCulla and Smith (2007), Landefeld et al. (2008), and Rassier (2012).

3 Hence the pair of papers roughly parallels the well-known literature on "forecast error" and "measurement error" properties of data revisions; see for example Mankiw et al. (1984), Mankiw and Shapiro (1986), Faust et al. (2005), and Aruoba (2008).

4 See also Smith et al. (1998), who take a different but related approach, and the independent work of Greenaway-McGrevy (2011), who takes a closely-related approach but unfortunately estimates a model that we show to be unidentified in Section 2.3.

http://dx.doi.org/10.1016/j.jeconom.2015.12.009
0304-4076/

indicators, the two most obvious being GDP_E and GDP_I, and we then extract true GDP using optimal filtering techniques.

The measurement-error approach is time honored, intrinsically compelling, and very different from the forecast-combination perspective of Aruoba et al. (2012), for several reasons.[5] First, it enables extraction of latent true GDP using a model with parameters *estimated* with exact likelihood or Bayesian methods, whereas the forecast-combination approach forces one to use calibrated parameters. Second, it delivers not only point extractions of latent true GDP but also interval extractions, enabling us to assess the associated uncertainty. Third, the state-space framework in which the measurement-error models are embedded facilitates exploration of the relationship between GDP measurement errors and the economic environment, such as stage of the business cycle, which is of special interest.

We proceed as follows. In Section 2 we consider several measurement-error models and assess their identification status, which turns out to be challenging and interesting in the most realistic and hence compelling case. In Section 3 we discuss the data, estimation framework and estimation results. In Section 4 we explore the properties of our new GDP series. Finally, we conclude with both a summary and a caveat in Section 5, where the caveat refers to the potential limitations of GDP_I (relative to GDP_E) for real-time analysis.

2. Five measurement-error models of *GDP*

We use dynamic-factor measurement-error models, which embed the idea that both GDP_E and GDP_I are noisy measures of latent true GDP. We work throughout with growth rates of GDP_E, GDP_I and GDP (hence, for example, GDP_E denotes a growth rate).[6] We assume throughout that true GDP growth evolves with simple $AR(1)$ dynamics, and we entertain several measurement structures, to which we now turn.

2.1. (Identified) 2-equation model: Σ diagonal

We begin with the simplest 2-equation model; the measurement errors are orthogonal to each other and to transition shocks at all leads and lags.[7] The model has a natural state-space structure, and we write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{1}$$
$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where GDP_{Et} and GDP_{It} are expenditure- and income-side estimates, respectively, GDP_t is latent true GDP, and all shocks are Gaussian and uncorrelated at all leads and lags. That is, $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma^2_{GG} & 0 & 0 \\ 0 & \sigma^2_{EE} & 0 \\ 0 & 0 & \sigma^2_{II} \end{bmatrix}. \tag{2}$$

The Kalman smoother will deliver optimal extractions of GDP_t conditional upon observed expenditure- and income-side measurements. We will refer to measures of GDP obtained this way as GDP_M throughout the paper. Moreover, the model can be easily extended, and some of its restrictive assumptions relaxed, with no fundamental change. We now proceed to do so.

2.2. (Identified) 2-equation model: Σ block-diagonal

The first extension is to allow for correlated measurement errors. This is surely important, as there is roughly a 25% overlap in the counts embedded in GDP_E and GDP_I, and moreover, the same deflator is used for conversion from nominal to real magnitudes.[8] We write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{3}$$
$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where now ϵ_{Et} and ϵ_{It} may be correlated contemporaneously but are uncorrelated at all other leads and lags, and all other definitions and assumptions are as before; in particular, ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})'$ are uncorrelated at all leads and lags. That is, $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, where

$$\Sigma = \begin{bmatrix} \sigma^2_{GG} & 0 & 0 \\ 0 & \sigma^2_{EE} & \sigma^2_{EI} \\ 0 & \sigma^2_{IE} & \sigma^2_{II} \end{bmatrix}. \tag{4}$$

Nothing is changed, and the Kalman filter retains its optimality properties.

2.3. (Unidentified) 2-equation model, Σ unrestricted

The second key extension is motivated by Fixler and Nalewaik (2009) and Nalewaik (2010), who document cyclicality in the statistical discrepancy ($GDP_E - GDP_I$), which implies failure of the assumption that $(\epsilon_{Et}, \epsilon_{It})'$ and ϵ_{Gt} are uncorrelated at all leads and lags. Of particular concern is contemporaneous correlation between ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})'$. Hence we allow the measurement errors $(\epsilon_{Et}, \epsilon_{It})'$ to be correlated with GDP_t, or more precisely, correlated with GDP_t innovations, ϵ_{Gt}. We write

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \tag{5}$$
$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It})' \sim iid\, N(\underline{0}, \Sigma)$, with

$$\Sigma = \begin{bmatrix} \sigma^2_{GG} & \sigma^2_{GE} & \sigma^2_{GI} \\ \sigma^2_{EG} & \sigma^2_{EE} & \sigma^2_{EI} \\ \sigma^2_{IG} & \sigma^2_{IE} & \sigma^2_{II} \end{bmatrix}. \tag{6}$$

In this environment the standard Kalman filter is rendered suboptimal for extracting GDP, due to correlation between ϵ_{Gt} and $(\epsilon_{Et}, \epsilon_{It})$, but appropriately-modified optimal filters are available.

Of course in what follows we will be concerned with *estimating* our measurement-equation models, so we will be concerned with identification. The diagonal-Σ model (1)–(2) and the block-diagonal-Σ model (3)–(4) are identified. Identification of less-restricted dynamic factor models, however, is a very delicate matter. In particular, it is not obvious that the unrestricted-Σ model (5)–(6) is identified. Indeed it is not, as we prove in Appendix A. Hence we now proceed to determine minimal restrictions that achieve identification.

2.4. (Identified) 2-equation model: Σ restricted

The identification problem with the general model (5)–(6) stems from the fact that we can make true GDP more volatile (in-

[5] On the time-honored aspect, see, for example, Gartaganis and Goldberger (1955).

[6] We will elaborate on the reasons for this choice later in Section 3.

[7] Here and throughout, when we say "N-equation" state-space model, we mean that the measurement equation is an N-variable system.

[8] See Aruoba et al. (2012) for more. Many of the areas of overlap are particularly poorly measured, such as imputed financial services, housing services, and government output.

crease σ^2_{GG}) and make the measurement errors more volatile (increase σ^2_{EE} and σ^2_{II}), but reduce the covariance between the fundamental shocks and the measurement errors (reduce σ^2_{EG} and σ^2_{IG}), without changing the distribution of observables.

2.4.1. Restricting the original parameterization

We can achieve identification by slightly restricting parameterization (5)–(6). In particular, as we show in Appendix A, the unrestricted system (5)–(6) is unidentified because the Σ matrix has six free parameters with only five moment conditions to determine them. Hence we can achieve identification by restricting any single element of Σ. Imposing any such restriction would seem challenging, however, as we have no strong prior views directly on any single element of Σ. Fortunately, however, a simple re-parameterization exists about which we have a more natural prior view, to which we now turn.

2.4.2. A useful re-parameterization

Let

$$\zeta = \frac{\frac{1}{1-\rho^2}\sigma^2_{GG}}{\frac{1}{1-\rho^2}\sigma^2_{GG} + 2\sigma^2_{GE} + \sigma^2_{EE}}, \tag{7}$$

the variance of latent true GDP relative to the variance of expenditure-side measured GDP_E. Then, rather than fixing an element of Σ to achieve identification, we can fix ζ, about which we have a more natural prior view. In particular, at first pass we might take $\sigma^2_{GE} \approx 0$, in which case $0 < \zeta < 1$. Or, put differently, $\zeta > 1$ would require a *very* negative σ^2_{GE}, which seems unlikely. All told, we view a ζ value less than, but close to, 1.0 as most natural. We take $\zeta = 0.80$ as our benchmark in the empirical work that follows, although we explore a wide range of ζ values both below and above 1.0.

2.5. (Identified) 3-equation model: Σ unrestricted

Thus far we showed how to achieve identification by fixing a parameter, ζ, and we noted that our prior is centered around $\zeta = 0.80$. It is also of interest to know whether we can get some complementary data-based guidance on choice of ζ. The answer turns out to be yes, by adding a third measurement equation with a certain structure.

Suppose, in particular, that we have an additional observable variable U_t that loads on true GDP_t with measurement error orthogonal to those of GDP_I and GDP_E. In particular, consider the 3-equation model

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \kappa \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \lambda \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \tag{8}$$

$$GDP_t = \mu(1-\rho) + \rho GDP_{t-1} + \epsilon_{Gt},$$

where $(\epsilon_{Gt}, \epsilon_{Et}, \epsilon_{It}, \epsilon_{Ut})' \sim iid\, N(\underline{0}, \Omega)$, with

$$\Omega = \begin{bmatrix} \sigma^2_{GG} & \sigma^2_{GE} & \sigma^2_{GI} & \sigma^2_{GU} \\ \sigma^2_{EG} & \sigma^2_{EE} & \sigma^2_{EI} & 0 \\ \sigma^2_{IG} & \sigma^2_{IE} & \sigma^2_{II} & 0 \\ \sigma^2_{UG} & 0 & 0 & \sigma^2_{UU} \end{bmatrix}. \tag{9}$$

Note that the upper-left 3×3 block of Ω is just Σ, which is now unrestricted. Nevertheless, as we prove in Appendix B, the 3-equation model (8)–(9) is identified. Of course some of the remaining elements of the overall 4×4 covariance matrix Ω *are* restricted, which is how we achieve identification in the 3-equation model, but the economically interesting sub-matrix, which the 3-equation model leaves completely unrestricted, is Σ.

Depending on the application, of course, it is not obvious that an identifying variable U_t with measurement errors orthogonal to those of GDP_E and GDP_I (i.e., with stochastic properties that satisfy (9)), is available. Hence it is not obvious that estimation of the 3-equation model (8)–(9) is feasible in practice, despite the model's appeal in principle. Indeed, much of the data collected from business surveys is used in the BEA's estimates, invalidating use of that data as U_t since any measurement error in that data appears directly in either GDP_E or GDP_I, producing correlation across the measurement errors. Moreover, variables drawn from business surveys similar to those used to produce GDP_E and GDP_I, even if they are not used directly in the estimation of GDP_E and GDP_I, might still be invalid identifying variables if the survey methodology itself produces similar measurement errors.[9]

Fortunately, however, some important macroeconomic data is collected not from surveys of businesses, but from samples of households. A sample of data drawn from a universe of households seems likely to have measurement errors that are different than those contaminating a data sample drawn from a universe of businesses, especially when the "universes" of businesses and households are not complete census counts, as is the case here. For example, the universe of business surveys is derived from tax records, so businesses not paying taxes will not appear on that list, but individuals working at that business may appear in the universe of households.

Importantly, very little data collected from household surveys are used to construct GDP_E and GDP_I, so a U_t variable computed from a household survey seems most likely to meet our identification conditions. The change in the unemployment rate is a natural choice (hence our notational choice U_t). U_t arguably loads on true GDP with a measurement error orthogonal to those of GDP_E and GDP_I, because the U_t data is being produced independently (by the BLS rather than BEA) from different types of surveys. In addition, virtually all of the GDP_E and GDP_I data are estimated in nominal dollars and then converted to real dollars using a price deflator, whereas U_t is estimated directly with no deflation.

All told, we view "3-equation identification" as a useful complement to the "ζ-identification" discussed earlier in Section 2.4. All identifications involve assumptions. ζ-identification involves introspection about likely values of ζ, given its structure and components, and that introspection is of course subject to error. 3-equation identification involves introspection about various measurement-error correlations involving the newly-introduced third variable, which is of course also subject to error. Indeed the two approaches to identification are usefully used in tandem, and compared.

One can even view the 3-equation approach as a device for implicitly selecting ζ. In particular, we can find the ζ implied by the 3-equation model estimate, that is, find the ζ that minimizes the divergence between $\hat{\Sigma}_\zeta$ and $\hat{\Sigma}_3$, in an obvious notation.[10] For example, using the Frobenius matrix-norm to measure divergence, we obtain an optimum of $\zeta^* = 0.82$. The minimum is sharp and unique. The implied ζ^* of 0.82 is of course quite close to the directly-assessed value of 0.80 at which we arrived earlier, which lends additional credibility to the earlier assessment. See (online) Appendix C.2.1 for details.

[9] For example, if the business surveys used to produce GDP_E and GDP_I tend to oversample large firms, variables drawn from a business survey that also oversamples large firms may have measurement errors that are correlated with those in GDP_E and GDP_I, absent appropriate corrections.

[10] We will discuss subsequently the estimation procedure used to obtain $\hat{\Sigma}_\zeta$ and $\hat{\Sigma}_3$.

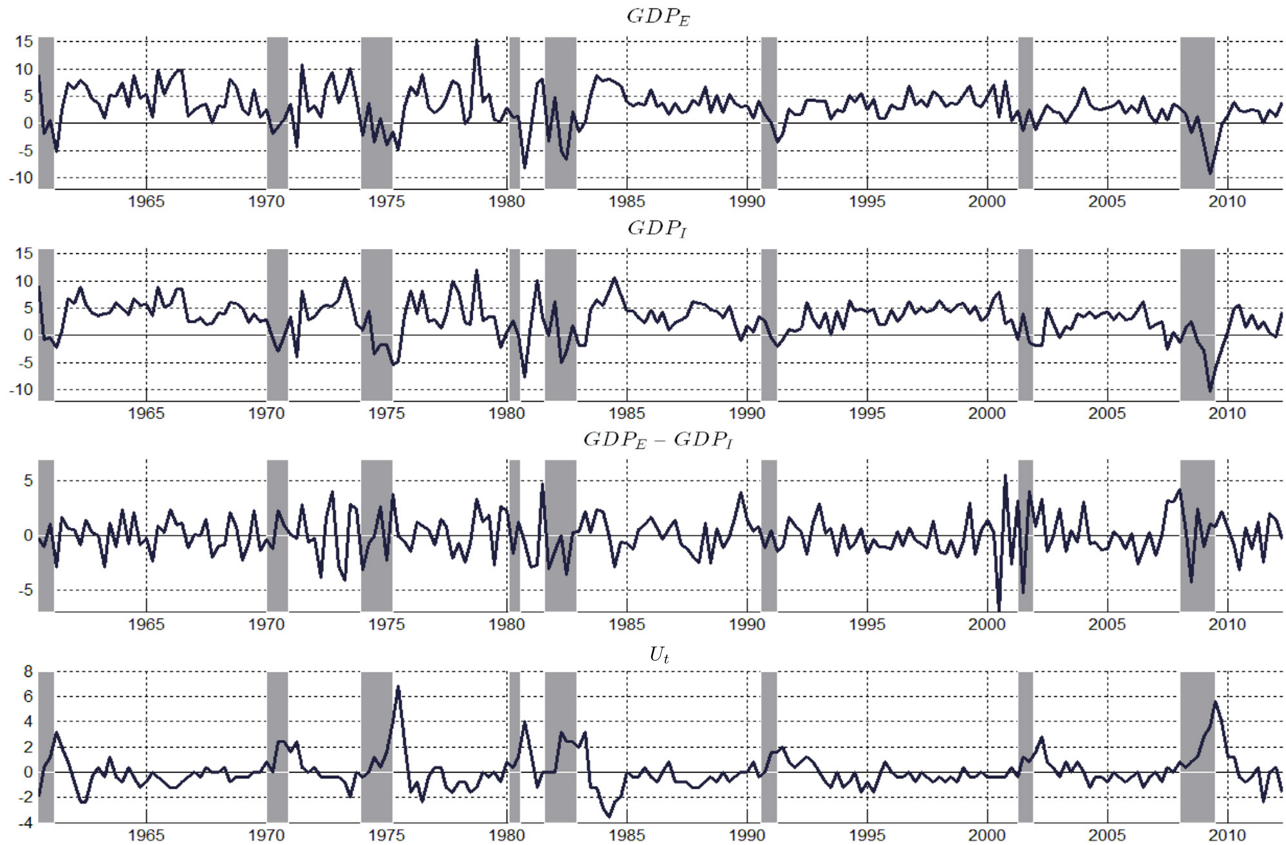

Fig. 1. *GDP* and unemployment data. GDP_E and GDP_I are in growth rates and U_t is in changes. All are measured in annualized percent.

3. Data and estimation

We intentionally work with a stationary system in growth rates, because we believe that measurement errors are best modeled as *iid* in growth rates rather than in levels, due to BEA's devoting maximal attention to estimating the "best change". [11] In its above-cited "Concepts and Methods . . ." document, for example, the BEA emphasizes that:

> Best change provides the most accurate measure of the period-to-period movement in an economic statistic using the best available source data. In an annual revision of the NIPAs, data from the annual surveys of manufacturing and trade are generally incorporated into the estimates on a best-change basis. In the current quarterly estimates, most of the components are estimated on a best-change basis from the annual levels established at the most recent annual revision.

The monthly source data used to estimate GDP_E (such as retail sales) and GDP_I (such as nonfarm payroll employment) are generally produced on a best-change basis as well, using a so-called "link-relative estimator". This estimator computes growth rates using firms in the sample in both the current and previous months, in contrast to a best-level estimator, which would generally use all the firms in the sample in the current month regardless of whether or not they were in the sample in the previous month. For example, for retail sales the BEA notes that[12]:

> Advance sales estimates for the most detailed industries are computed using a type of ratio estimator known as the link-relative estimator. For each detailed industry, we compute a ratio of current-to-previous month weighted sales using data from units for which we have obtained usable responses for both the current and previous month.

Indeed the BEA produces estimates on a best-level basis only at 5-year benchmarks. These best-level benchmark revisions should drive only the very-low frequency variation in GDP_E, and thus probably matter very little for the quarterly growth rates estimated on a best-change basis.

3.1. Descriptive statistics

We show time-series plots of the "raw" GDP_E and GDP_I data in Fig. 1, and we show summary statistics for the raw series in the top panel of Table 1. Not captured in the table but also true is that the raw data are highly correlated; the simple correlations are $corr(GDP_E, GDP_I) = 0.85$, $corr(GDP_E, U) = -0.67$, and $corr(GDP_I, U) = -0.73$. Median GDP_I growth is a bit higher than that of GDP_E, and GDP_I growth is noticeably more persistent than that of GDP_E. Related, GDP_I also has smaller $AR(1)$ innovation variance and greater predictability as measured by the predictive R^2. Fig. 1 also depicts the sample paths of changes in the unemployment rate, which we use to estimate the 3-equation model, and the discrepancy between the growth rates GDP_E and GDP_I. According to our state-space models, the discrepancy equals the measurement error difference $\epsilon_{Et} - \epsilon_{It}$. The mean of the discrepancy series is zero, and its variance is approximately 30% of the variance of GDP_E. The first-order autoregressive coefficient is slightly negative, but the R^2 associated with an $AR(1)$ regression is only about 4%.

[11] For example, see "Concepts and Methods in the U.S. National Income and Product Accounts", available at http://www.bea.gov/national/pdf/methodology/chapters1-4.pdf.

[12] See http://www.census.gov/retail/marts/how_surveys_are_collected.html.

Table 1
Descriptive statistics for various *GDP* series.

	$\bar{x}$	50%	$\hat{\sigma}$	*Sk*	$\hat{\rho}_1$	$\hat{\rho}_2$	$\hat{\rho}_3$	$\hat{\rho}_4$	Q_{12}	$\hat{\sigma}_e$	R^2	$\hat{V}_e$
GDP_E	3.03	3.04	3.49	−0.31	.33	.27	.08	.09	47.07	3.28	.06	12.12
GDP_I	3.02	3.39	3.40	−0.55	.47	.27	.22	.08	81.60	2.99	.12	11.43
GDP_M 2-eqn, Σ diag	3.02	3.22	3.00	−0.56	.56	.34	.21	.09	108.25	2.48	.18	8.92
GDP_M 2-eqn, Σ block	3.02	3.35	2.64	−0.64	.70	.45	.28	.13	170.08	1.89	.29	6.90
GDP_M 2-eqn, $\zeta = 0.65$	3.02	3.32	2.61	−0.64	.67	.43	.27	.12	157.56	1.92	.26	6.73
GDP_M 2-eqn, $\zeta = 0.75$	3.02	3.30	2.77	−0.63	.65	.41	.26	.11	148.23	2.08	.25	7.60
GDP_M 2-eqn, $\zeta = 0.80$	3.02	3.29	2.87	−0.62	.64	.39	.25	.11	141.14	2.19	.24	8.16
GDP_M 2-eqn, $\zeta = 0.85$	3.02	3.31	2.89	−0.64	.66	.41	.28	.12	153.27	2.15	.25	8.29
GDP_M 2-eqn, $\zeta = 0.95$	3.02	3.26	3.02	−0.64	.66	.40	.28	.12	149.61	2.27	.25	9.07
GDP_M 2-eqn, $\zeta = 1.05$	3.01	3.22	3.12	−0.65	.67	.40	.28	.12	155.60	2.30	.26	9.69
GDP_M 2-eqn, $\zeta = 1.15$	3.04	3.34	3.07	−0.67	.76	.47	.31	.15	201.15	1.99	.35	9.46
GDP_M 3-eqn	3.02	3.37	3.02	−1.14	.63	.37	.21	.03	141.79	2.33	.23	9.03
GDP_F	3.02	3.29	3.30	−0.51	.46	.29	.19	.07	78.28	2.92	.12	10.80

Notes: The sample period is 1960Q1–2011Q4. In the top panel we show statistics for the raw data. In the middle panel we show statistics for various posterior-median measurement-error-based ("*M*") estimates of true *GDP*, where all estimates are smoothed extractions. In the bottom panel we show statistics for the forecast-error-based estimate of true *GDP* produced by Aruoba et al. (2012), GDP_F. $\bar{x}$, 50%, $\hat{\sigma}$ and *Sk* are sample mean, median, standard deviation and skewness, respectively, and $\hat{\rho}_\tau$ is a sample autocorrelation at a displacement of τ quarters. Q_{12} is the Ljung–Box serial correlation test statistic calculated using $\hat{\rho}_1, \ldots, \hat{\rho}_{12}$. $R^2 = 1 - \frac{\hat{\sigma}_e^2}{\hat{\sigma}^2}$, where $\hat{\sigma}_e$ denotes the estimated disturbance standard deviation from a fitted $AR(1)$ model, is a predictive R^2. $\hat{V}_e$ is the unconditional variance implied by a fitted $AR(1)$ model, $\hat{V}_e = \frac{\hat{\sigma}_e^2}{1-\hat{\rho}^2}$.

3.2. Estimation

Bayesian estimation involves parameter estimation and latent state smoothing. First, we generate draws from the posterior distribution of the model parameters using a Random-Walk Metropolis–Hastings algorithm. Next, we apply the simulation smoother of Durbin and Koopman (2001) to obtain draws of the latent states conditional on the parameters. See (online) Appendix C for details.

Here we present and discuss estimation results for our various models. In Table 2 we show details of parameter prior and posterior distributions, as well as statistics describing the overall posterior and likelihood, for various 2-equation models, and in Table 3 we provide the same information for the 3-equation model.

The complete estimation information in the tables can be difficult to absorb fully, however, so here we briefly present aspects of the results in a more revealing way. For the 2-equation models, the parameters to be estimated are those in the transition equation and those in the covariance matrix Σ, which includes variances and covariances of both transition and measurement shocks. Hence we simply display the estimated transition equation and the estimated Σ matrices. For the 3-equation model, we also need to estimate a factor loading in the measurement equation, so we display the estimated measurement equation as well. Below each posterior median parameter estimate, we show the posterior interquartile range in brackets.

For the 2-equation model with Σ diagonal, we have

$$GDP_t = \underset{[2.81,3.33]}{3.07}(1-0.53) + \underset{[0.48,0.57]}{0.53}\, GDP_{t-1} + \epsilon_{Gt}, \tag{10}$$

$$\Sigma = \begin{bmatrix} \underset{[6.39,7.44]}{6.90} & 0 & 0 \\ 0 & \underset{[2.12,2.55]}{2.32} & 0 \\ 0 & 0 & \underset{[1.52,1.85]}{1.68} \end{bmatrix}. \tag{11}$$

For the 2-equation model with Σ block-diagonal, we have

$$GDP_t = \underset{[2.77,3.34]}{3.06}(1-0.62) + \underset{[0.57,0.68]}{0.62}\, GDP_{t-1} + \epsilon_{Gt}, \tag{12}$$

$$\Sigma = \begin{bmatrix} \underset{[4.39,5.95]}{5.17} & 0 & 0 \\ 0 & \underset{[3.34,4.48]}{3.86} & \underset{[0.96,1.95]}{1.43} \\ 0 & \underset{[0.96,1.95]}{1.43} & \underset{[2.25,3.22]}{2.70} \end{bmatrix}. \tag{13}$$

For the 2-equation model with benchmark $\zeta = 0.80$, we have

$$GDP_t = \underset{[2.79,3.35]}{3.08}(1-0.57) + \underset{[0.51,0.62]}{0.57}\, GDP_{t-1} + \epsilon_{Gt}, \tag{14}$$

$$\Sigma = \begin{bmatrix} \underset{[6.54,7.70]}{7.09} & \underset{[-1.15,-0.29]}{-0.69} & \underset{[-0.74,-0.04]}{-0.38} \\ \underset{[-1.15,-0.29]}{-0.69} & \underset{[3.14,4.77]}{3.90} & \underset{[0.80,1.85]}{1.29} \\ \underset{[-0.74,-0.04]}{-0.38} & \underset{[0.80,1.85]}{1.29} & \underset{[1.98,2.82]}{2.36} \end{bmatrix}. \tag{15}$$

Finally, for the 3-equation model, we have

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \underset{[1.53,1.71]}{1.62} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ \underset{[-0.55,-0.50]}{-0.52} \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \tag{16}$$

$$GDP_t = \underset{[2.60,2.95]}{2.78}(1-0.58) + \underset{[0.54,0.63]}{0.58}\, GDP_{t-1} + \epsilon_{Gt}, \tag{17}$$

$$\begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \underset{[6.73,7.35]}{6.96} & \underset{[-1.27,-0.84]}{-1.10} & \underset{[-1.03,-0.59]}{-0.82} & \underset{[1.27,1.66]}{1.46} \\ \underset{[-1.27,-0.84]}{-1.10} & \underset{[4.17,4.79]}{4.57} & \underset{[1.70,2.12]}{1.95} & 0 \\ \underset{[-1.03,-0.59]}{-0.82} & \underset{[1.70,2.12]}{1.95} & \underset{[2.54,3.27]}{3.07} & 0 \\ \underset{[1.27,1.66]}{1.46} & 0 & 0 & \underset{[0.50,0.71]}{0.59} \end{bmatrix} \right). \tag{18}$$

Many aspects of the results are noteworthy; here we simply mention a few. First, every posterior interval in every model reported above excludes zero. Hence the diagonal and block diagonal models do not appear satisfactory.

Second, the Σ estimates are qualitatively similar across specifications. Covariances are always negative, as per our conjecture based on the counter-cyclicality in the statistical discrepancy ($GDP_E - GDP_I$) documented by Fixler and Nalewaik (2009) and Nalewaik (2010). Shock variances always satisfy $\hat{\sigma}^2_{GG} > \hat{\sigma}^2_{EE} > \hat{\sigma}^2_{II}$.

Finally, GDP_M is highly serially correlated across all specifications ($\rho \approx .6$), much more so than the current "consensus" based on GDP_E ($\rho \approx .3$). We shall have more to say about these and other results in Section 4.

Table 2
Priors and posteriors, 2-equation models, 1960Q1–2011Q4.

	Prior	Posterior			Posterior		
	(Mean, Std.Dev)	25%	50%	75%	25%	50%	75%
		Diagonal			Block diagonal		
μ	$N(3, 10)$	2.81	3.07	3.33	2.77	3.06	3.34
ρ	$N(0.3, 1)$	0.48	0.53	0.57	0.57	0.62	0.68
σ^2_{GG}	$IG(10, 15)$	6.39	6.90	7.44	4.39	5.17	5.95
σ^2_{GE}	$N(0, 10)$	–	–	–	–	–	–
σ^2_{GI}	$N(0, 10)$	–	–	–	–	–	–
σ^2_{EE}	$IG(10, 15)$	2.12	2.32	2.55	3.34	3.86	4.48
σ^2_{EI}	$N(0, 10)$	–	–	–	0.96	1.43	1.95
σ^2_{II}	$IG(10, 15)$	1.52	1.68	1.85	2.25	2.70	3.22
Posterior	–	−984.57	−983.46	−982.60	−986.23	−985.00	−984.01
Likelihood	–	−951.68	−950.41	−949.43	−950.70	−949.49	−948.60
		$\zeta = 0.75$			$\zeta = 0.80$		
μ	$N(3, 10)$	2.75	3.03	3.31	2.79	3.08	3.35
ρ	$N(0.3, 1)$	0.53	0.59	0.64	0.51	0.57	0.62
σ^2_{GG}	$IG(10, 15)$	5.78	6.31	6.92	6.54	7.09	7.70
σ^2_{GE}	$N(0, 10)$	−0.76	−0.29	0.15	−1.15	−0.69	−0.29
σ^2_{GI}	$N(0, 10)$	−0.34	0.01	0.34	−0.74	−0.38	−0.04
σ^2_{EE}	$IG(10, 15)$	3.08	3.88	4.75	3.14	3.90	4.77
σ^2_{EI}	$N(0, 10)$	0.73	1.23	1.78	0.80	1.29	1.85
σ^2_{II}	$IG(10, 15)$	1.94	2.30	2.76	1.98	2.36	2.82
Posterior	–	−982.50	−980.99	−979.87	−982.48	−981.05	−979.91
Likelihood	–	−950.93	−949.55	−948.40	−950.85	−949.44	−948.41
		$\zeta = 0.85$			$\zeta = 0.95$		
μ	$N(3, 10)$	2.72	2.96	3.14	2.84	3.03	3.25
ρ	$N(0.3, 1)$	0.51	0.56	0.60	0.49	0.54	0.60
σ^2_{GG}	$IG(10, 15)$	6.67	7.19	7.76	7.69	8.43	9.28
σ^2_{GE}	$N(0, 10)$	−2.17	−1.98	−1.77	−2.88	−2.73	−2.50
σ^2_{GI}	$N(0, 10)$	−0.97	−0.80	−0.53	−1.99	−1.58	−1.22
σ^2_{EE}	$IG(10, 15)$	5.36	5.79	6.28	5.64	6.10	6.39
σ^2_{EI}	$N(0, 10)$	2.04	2.33	2.63	2.43	2.64	2.93
σ^2_{II}	$IG(10, 15)$	2.36	2.65	3.04	2.45	3.22	3.81
Posterior	–	−982.62	−981.40	−980.48	−984.09	−982.80	−981.57
Likelihood	–	−949.42	−948.25	−947.49	−950.19	−948.84	−947.81
		$\zeta = 1.05$			$\zeta = 1.15$		
μ	$N(3, 10)$	2.85	3.07	3.33	2.55	2.89	3.21
ρ	$N(0.3, 1)$	0.48	0.53	0.58	0.52	0.56	0.61
σ^2_{GG}	$IG(10, 15)$	8.92	9.57	10.25	9.07	9.88	10.73
σ^2_{GE}	$N(0, 10)$	−4.04	−3.88	−3.70	−5.61	−5.50	−5.22
σ^2_{GI}	$N(0, 10)$	−3.09	−2.65	−2.29	−4.38	−4.21	−4.01
σ^2_{EE}	$IG(10, 15)$	6.74	7.13	7.41	8.51	9.07	9.30
σ^2_{EI}	$N(0, 10)$	3.23	3.46	4.13	5.29	5.52	5.89
σ^2_{II}	$IG(10, 15)$	3.27	3.66	4.43	5.68	6.00	6.31
Posterior	–	−984.89	−983.63	−982.49	−988.63	−987.18	−986.32
Likelihood	–	−949.31	−948.30	−947.53	−949.82	−948.51	−947.67

3.3. Diagnostic checks

We have assumed throughout that all shocks are Gaussian white noise. As regards normality, we feel that it is an obvious benchmark. The recent severe recession does not necessarily invalidate the normality assumption, as occasional extreme draws will occur even under normality, and moreover our Kalman filtering remains BLUE even under non-normality. Nevertheless it is of course interesting and important to check the validity of the normality assumption.

We report diagnostic normality checks in Fig. 2 for the three model shocks, ϵ_E, ϵ_I and ϵ_G. In the top panel we show the distributions of residual skewness across our 25,000 posterior draws. All are tightly and symmetrically distributed around zero, providing strong support for symmetry. In the middle panel we show the distributions of residual kurtosis. Those for the measurement errors ϵ_E and ϵ_I are tightly and symmetrically distributed around three, consistent with normality. The distribution of residual kurtosis for ϵ_G again appears consistent with normality, although less strongly than for the distributions of ϵ_E and ϵ_I. It is centered around a median slightly greater than three, and it is skewed slightly rightward.

As regards the white noise assumption, we show the interquartile ranges of our 25,000 posterior residual autocorrelation function draws in the bottom panel of Fig. 2, again for each of ϵ_E, ϵ_I and ϵ_G. They are tightly centered around zero and reveal no evidence of serial correlation in measurement errors or true *GDP* innovations.

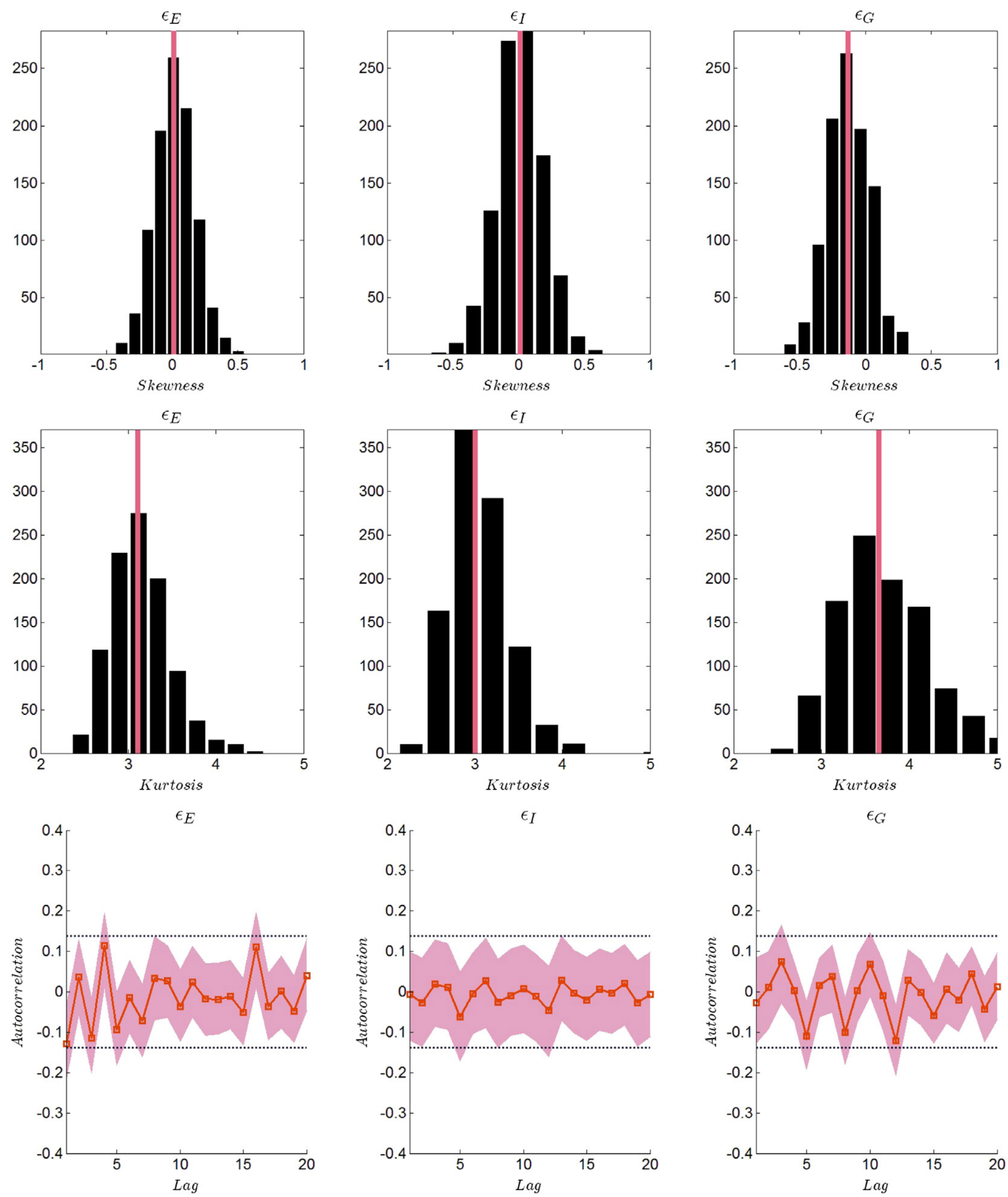

Fig. 2. Distributions of residual skewness, kurtosis and autocorrelations across 25,000 posterior draws. In the skewness and kurtosis plots the solid vertical lines denote posterior medians. The shaded region in the autocorrelation plots denotes the posterior interquartile range.

All told, then, the GDP_E and GDP_I data appear to accord quite well with our benchmark dynamic factor model.

4. New perspectives on the properties of *GDP*

Our various extracted GDP_M series differ in fundamental ways from other measures, such as GDP_E and GDP_I. Here we discuss some of the most important differences.

4.1. GDP sample paths

Let us begin by highlighting the sample-path differences between our GDP_M and the obvious competitors GDP_E and GDP_I. We make those comparisons in Fig. 3. In each panel we show the sample path of GDP_M together with a shaded posterior interquartile range, and we show one of the competitor series.[13] In the top panel we show GDP_M vs. GDP_E. There are often wide divergences, with GDP_E well outside the posterior interquartile range of GDP_M. Indeed GDP_E is substantially more volatile than GDP_M. In the bottom panel of Fig. 3 we show GDP_M vs. GDP_I. Noticeable divergences again appear often, with GDP_I also outside the posterior interquartile range of GDP_M. The divergences are not as pronounced, however, and the "excess volatility" apparent in GDP_E is less apparent in GDP_I. That is because, as we will show later, GDP_M loads relatively more heavily on GDP_I.

[13] For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

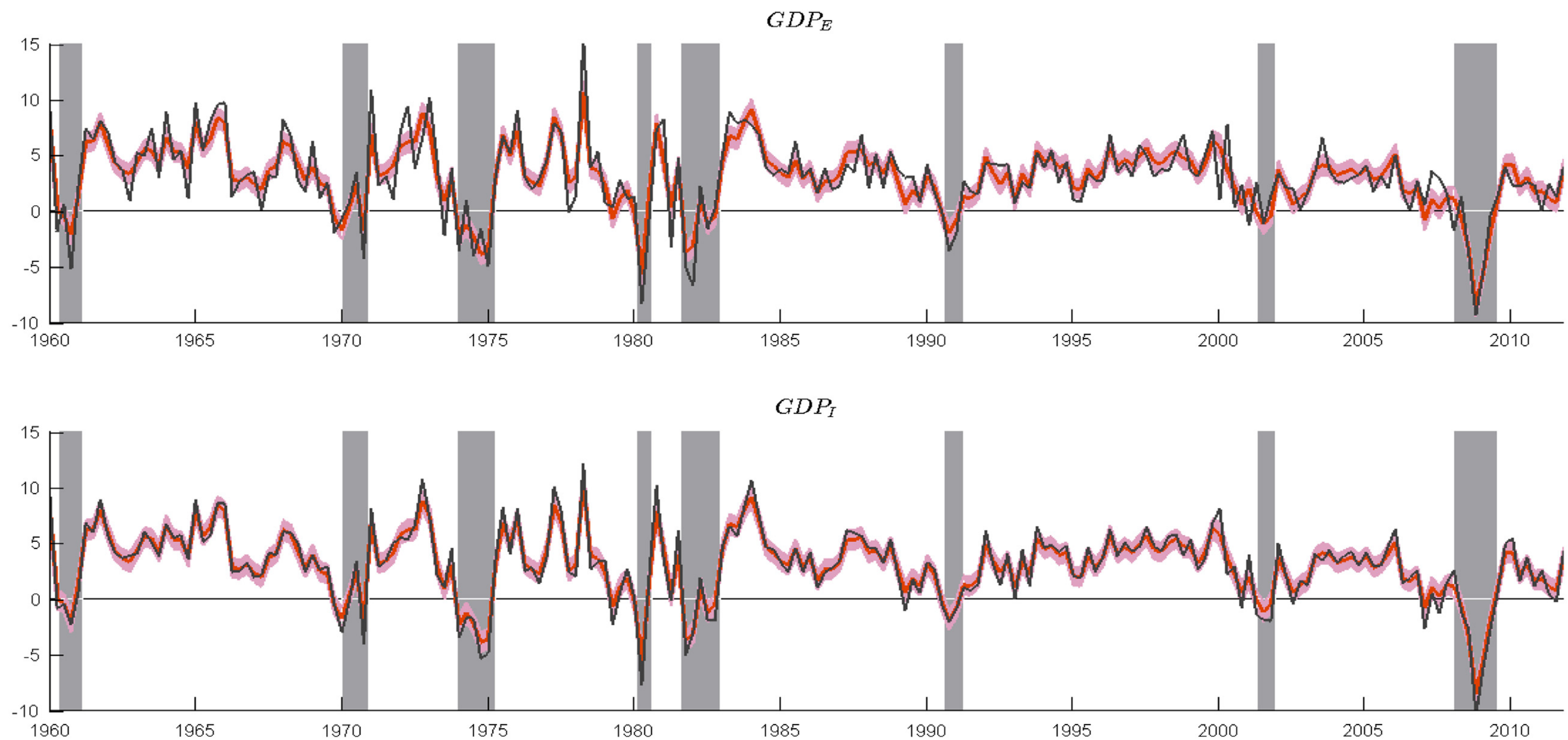

Fig. 3. *GDP* sample paths, 1960Q1–2011Q4. In each panel we show the sample path of GDP_M (light color) together with posterior interquartile range with shading and we show one of the competitor series (dark color). For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

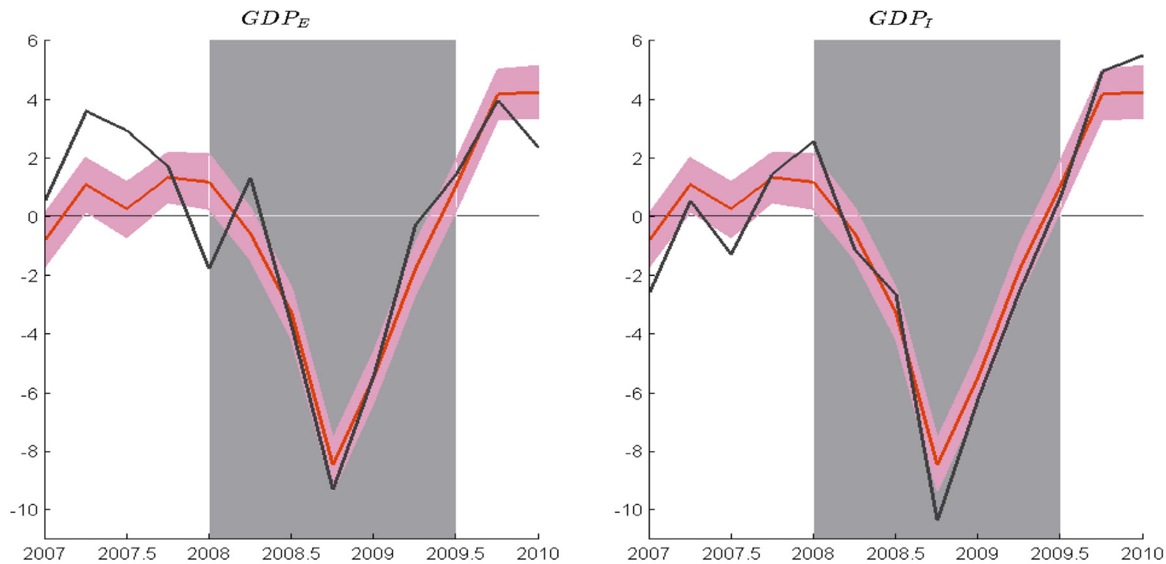

Fig. 4. *GDP* sample paths, 2007Q1–2009Q4. In each panel we show the sample path of GDP_M (light color) together with posterior interquartile range with shading and we show one of the competitor series (dark color). For GDP_M we use our benchmark estimate from the 2-equation model with $\zeta = 0.80$.

Table 3
Priors and posteriors, 3-equation model, 1960Q1–2011Q4.

Parameter	Prior	Posterior		
	(Mean, Std)	25%	50%	75%
μ	$N(3, 10)$	2.60	2.78	2.95
ρ	$N(0.3, 1)$	0.54	0.58	0.63
σ^2_{GG}	$IG(10, 15)$	6.73	6.96	7.35
σ^2_{GE}	$N(0, 10)$	−1.27	−1.10	−0.84
σ^2_{GI}	$N(0, 10)$	−1.03	−0.82	−0.59
σ^2_{EE}	$IG(10, 15)$	4.17	4.57	4.79
σ^2_{EI}	$N(0, 10)$	1.70	1.95	2.12
σ^2_{II}	$IG(10, 15)$	2.54	3.07	3.27
σ^2_{GU}	$N(0, 10)$	1.27	1.46	1.66
σ^2_{UU}	$IG(0.3, 10)$	0.50	0.59	0.71
κ	$N(0, 10)$	1.53	1.62	1.71
λ	$N(-0.5, 10)$	−0.55	−0.52	−0.50
Posterior	–	−1251.1	−1249.6	−1248.3
Likelihood	–	−1199.0	−1197.5	−1196.2

To emphasize the economic importance of the differences in competing real activity assessments, in Fig. 4 we focus on the tumultuous period 2007Q1–2009Q4. The figure makes clear not only that *both* GDP_E and GDP_I can diverge substantially from *GDP*, but also that the timing and nature of their divergences can be very different. In 2007Q3, for example, GDP_E growth was strongly positive and GDP_I growth was negative.

4.2. GDP dynamics

In our linear framework, the data-generating process for true GDP_t is completely characterized by the pair, (σ^2_{GG}, ρ).[14] In Fig. 5 we show those pairs across MCMC draws for all of our measurement-error models, and for comparison we show (ρ, σ^2) values corresponding to $AR(1)$ models fit to GDP_E alone and GDP_I alone. In addition, in Table 1 we show a variety of statistics quantifying the sample properties of our various optimally extracted GDP_M measures vs. those of GDP_E, GDP_I and GDP_F, the forecast-error-based estimate of true *GDP* produced by Aruoba et al. (2012).

A key result of our analysis is the strong serial correlation (persistence, forecastability, etc.) of true *GDP* and our extracted GDP_M, regardless of the particular specification. First consider the (ρ, σ^2_{GG}) draws, which determine the population autocovariance function of the true *GDP* process, depicted in Fig. 5. Depending on the specification of the measurement error model, the posterior mean estimates of ρ lie in the interval of 0.5–0.6. For comparison, the estimated $AR(1)$ coefficient for GDP_E is only 0.33. The large ρ values are accompanied by relatively small innovation variances σ^2_{GG}.

Now consider the sample statistics of the extracted GDP_M series summarized in Table 1. As expected from the parameter estimates depicted in Fig. 5, the GDP_M series is robustly more serially correlated than GDP_E, GDP_I, GDP_F. More specifically, if we fit an $AR(1)$ model to GDP_M we find that the shock persistence is roughly double that of GDP_E (ρ of roughly 0.60 for GDP_M vs. 0.30 for GDP_E). Simultaneously, the estimated innovation variances of the GDP_M series are much smaller than those associated with the raw data. This translates into much higher predictive R^2's for GDP_M. Indeed GDP_M is twice as predictable as GDP_I or GDP_F, which in turn are twice as predictable as GDP_E. Table 1 also reveals that the various GDP_M series are all less volatile than each of GDP_E, GDP_I and GDP_F, and a bit more skewed left.

To appreciate these results, consider the 2-equation model with block-diagonal Σ. A straightforward analysis of the implied autocovariances implies that in population both GDP_E and GDP_I have to be more volatile than true *GDP*. Moreover, due to the presence of measurement errors that are independent of the *GDP* innovations, the first-order autocorrelations of GDP_E and GDP_I always provide downward-biased estimates of ρ, the autocorrelation of true *GDP*.

Once we allow for the measurement errors to be correlated with ϵ_{Gt}, the volatility ranking and the sign of the bias are ambiguous. We can express the first-order autocorrelation of GDP_E as

$$Corr(GDP_{Et}, GDP_{E,t-1}) = \rho \, \frac{V(GDP_t) + \sigma^2_{GE}}{V(GDP_t) + 2\sigma^2_{GE} + \sigma^2_{EE}}. \tag{19}$$

Thus the autocorrelation of GDP_E provides an upward-biased estimate of ρ if

$$\sigma^2_{GE} > 2\sigma^2_{GE} + \sigma^2_{EE}. \tag{20}$$

[14] We provide complementary nonlinear Markov-switching results in (online) Appendix C.2.3.

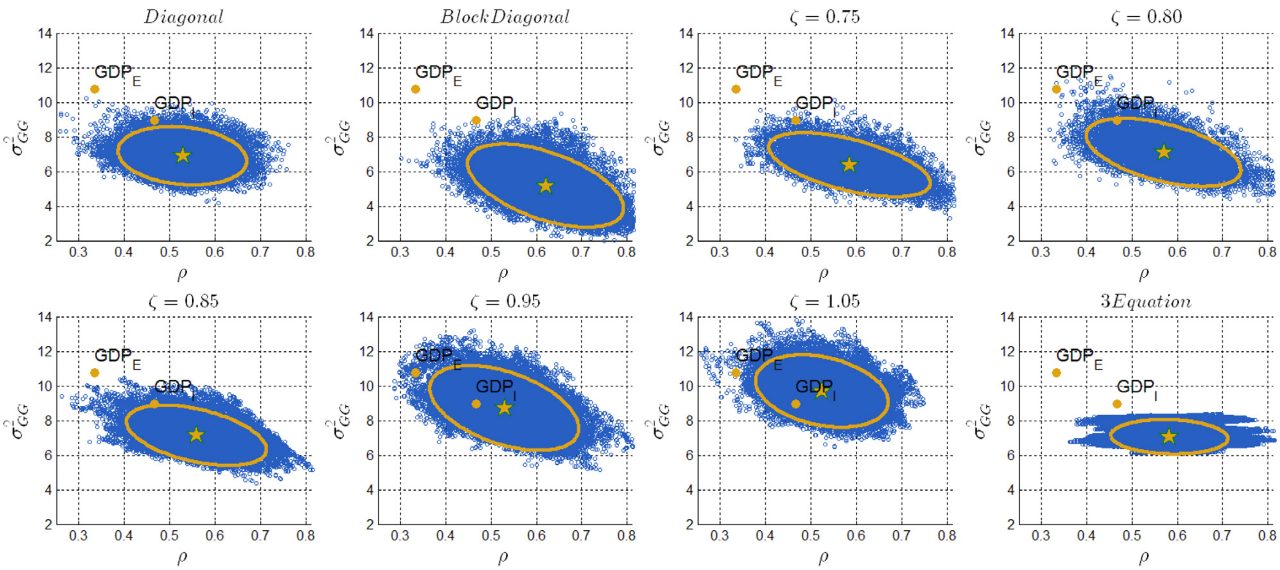

Fig. 5. (ρ, σ^2_{GG}) Pairs across MCMC draws. Solid lines indicate 90% (σ^2_{GG}, ρ) posterior coverage ellipsoids for the various models. Stars indicate posterior median values. The sample period is 1960Q1-2011.Q4. For comparison we show (σ^2, ρ) values corresponding to $AR(1)$ models fit to GDP_E alone and GDP_I alone.

Because the measurement error variance σ^2_{EE} is always non-negative, an upward bias only arises if GDP innovation and measurement error are negatively correlated and the measurement error is small. Consider, for instance, the estimated 3-equation model. Although $\hat{\sigma}^2_{GE} < 0$, the inequality (20) is not satisfied: $\hat{\sigma}^2_{GE} = -1.10$ and $2\hat{\sigma}^2_{GE} + \hat{\sigma}^2_{EE} = 2.37$. Thus, we emphasize that the high serial correlation of GDP_M is not a spurious artifact of our signal-extraction approach. In view of the flexibility of our measurement-error model, it is a genuine empirical finding that is a reflection of estimated size of the measurement error and its correlation with the innovation to true GDP.

4.3. On the relative contributions of GDP_E and GDP_I to GDP_M

It is of interest to know how the observed indicators GDP_E and GDP_I contribute to our extracted true GDP. We do this in two ways, by examining the Kalman gains, and by finding the convex combination of GDP_E and GDP_I closest to our extracted GDP.

The Kalman gains associated with GDP_E and GDP_I govern the amount by which news about GDP_E and GDP_I, respectively, causes the optimal extraction of GDP_t (conditional on time-t information) to differ from the earlier optimal prediction of GDP_t (conditional on time-$(t-1)$ information). Put more simply, the Kalman gain of GDP_E (resp. GDP_I) measures its importance in influencing GDP_M, and hence in informing our views about latent true GDP.

We summarize the posterior distributions of Kalman gains in Fig. 6. Posterior median GDP_I Kalman gains are large in absolute terms, and most notably, very large relative to those for GDP_E. Indeed posterior median GDP_E Kalman gains are zero in several specifications. In any event, it is clear that GDP_I plays a larger role in informing us about GDP than does GDP_E. For our benchmark ζ-model with $\zeta = 0.80$, the posterior median GDP_I and GDP_E Kalman gains are 0.59 and 0.23, respectively.

The Kalman filter extractions average not only over space, but also over time. Nevertheless, we can ask what contemporaneous convex combination of GDP_E and GDP_I, $\lambda GDP_E + (1-\lambda)GDP_I$, is closest to the extracted GDP_M. That is, we can find $\lambda^* = argmin_\lambda L(\lambda)$, where $L(\lambda)$ is a loss function. Under quadratic loss we have

$$\lambda^* = \underset{\lambda}{\operatorname{argmin}} \sum_{t=1}^{T} [(\lambda GDP_{Et} + (1-\lambda)GDP_{It}) - GDP_{Mt}]^2 ,$$

where GDP_{Mt} is our smoothed extraction of true GDP_t. Over our sample of 1960Q1–2011Q4, the optimum under quadratic loss is $\lambda^* = 0.29$. The minimum is quite sharp, and it puts more than twice as much weight on GDP_I than on GDP_E.[15] That weighting accords closely with both the Kalman gain results discussed above and the forecast-combination calibration results in Aruoba et al. (2012). It does not, of course, mean that time series of GDP_M will "match" time series of GDP_F, because the Kalman filter does much more than simple contemporaneous averaging of GDP_E and GDP_I in its extraction of latent true GDP.

5. Conclusions, caveats, and future research

We produce several estimates of GDP that blend both GDP_E and GDP_I. All estimates feature GDP_I prominently, and our blended GDP estimate has properties quite different from those of the "traditional" GDP_E (as well as GDP_I). In a sense we build on the literature on "balancing" the national income accounts, which extends back almost as far as national income accounting itself, as for example in Stone et al. (1942). We do not, however, advocate that the U.S. publishes only GDP_M, as there may at times be value in being able to see the income and expenditure sides separately. But we would certainly advocate the additional calculation of GDP_M and using it as the benchmark GDP estimate.[16]

A caveat is in order, however, as GDP_I is released in less-timely fashion than GDP_E, and moreover, early releases of GDP_I may be inferior to corresponding releases of GDP_E. A key reason is the simple fact that it takes time for the tax returns underlying much of GDP_I to be filed and processed. Hence if one is interested in real-time tracking of real activity (during the most-recent four quarters, say), GDP_M is not likely to add much relative to GDP_E.[17] On the other

15 See Appendix C.2.2 for a plot of the entire surface.

16 The Federal Reserve Bank of Philadelphia recently began doing this; see their "GDPplus" series at http://www.philadelphiafed.org/research-and-data/real-time-center/gdpplus/.

17 Of course one would surely also not want to use GDP_E alone. Instead, for real-time analysis GDP_E should be blended with other higher-frequency (monthly, weekly) indicators as in Aruoba et al. (2009) and Aruoba and Diebold (2010), implemented in real time by the Federal Reserve Bank of Philadelphia at http://www.philadelphiafed.org/research-and-data/real-time-center/business-conditions-index.

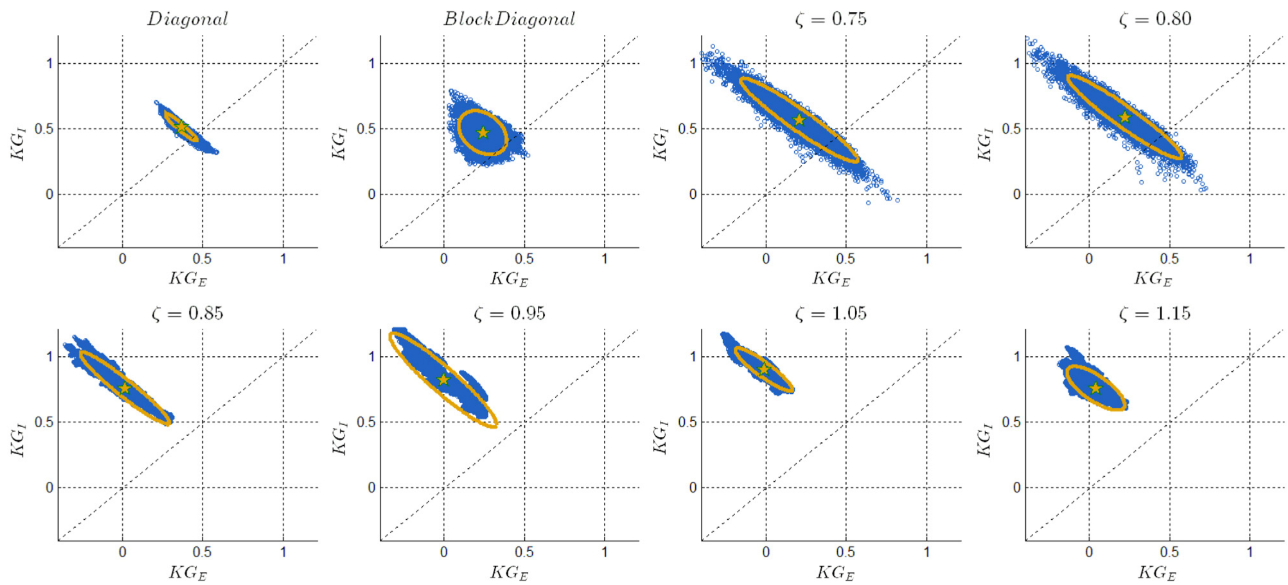

Fig. 6. (KG_E, KG_I) Pairs across MCMC draws. Solid lines indicate 90% posterior coverage ellipsoids. Stars indicate posterior median values.

hand, whether one uses up-to-the-instant *GDP* data, as opposed to up-to-a-year-ago data, is typically irrelevant to the research work for which we seek to contribute a superior input.

Interesting extensions of our framework and methods are possible. Consider, for example, forecasting. When forecasting a "traditional" *GDP* series such as GDP_E, we must take it as given (i.e., we must ignore measurement error). The analogous procedure in our framework would take GDP_M as given, modeling and forecasting it directly, ignoring the fact that it is only an estimate. Fortunately, however, in our framework we need not do that. Instead we can estimate and forecast directly from the dynamic factor model, accounting for all sources of uncertainty, which should translate into superior interval and density forecasts. Related, it would be interesting to calculate directly the point, interval and density forecast functions corresponding to our measurement-error model.

Acknowledgments

We are grateful to the Editor (William Barnett) and two anonymous referees for insightful comments that improved the paper significantly. For helpful comments we are also grateful to Bob Chirinko, Rob Engle, Graham Elliott, Don Harding, John Haltiwanger, Andrew Harvey, Jan Jacobs, Greg Mankiw, Adrian Pagan, John Roberts, Matt Shapiro, Neil Shephard, Minchul Shin, Chris Sims, Mark Watson, Justin Wolfers and Simon van Norden. For research support we thank the National Science Foundation and the Real-Time Data Research Center at the Federal Reserve Bank of Philadelphia.

Appendix

Here we report various details of theory, establishing identification results for the two- and three-variable models in Appendices A and B, respectively. The identification analysis is based on Komunjer and Ng (2011).

Appendix A. Identification in the two-equation model

The constants in the state-space model can be identified from the means of GDP_{Et} and GDP_{It}. To simplify the subsequent exposition we now set the constant terms to zero:

$$GDP_t = \rho GDP_{t-1} + \epsilon_{Gt} \tag{A.1}$$

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

and the joint distribution of the errors is

$$\epsilon_t = \begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \end{bmatrix} \sim iidN\left(0, \Sigma\right), \quad \text{where } \Sigma = \begin{bmatrix} \Sigma_{GG} & \cdot & \cdot \\ \Sigma_{EG} & \Sigma_{EE} & \cdot \\ \Sigma_{IG} & \Sigma_{IE} & \Sigma_{II} \end{bmatrix}.$$

Using the notation in Komunjer and Ng (2011), we write the system as

$$s_{t+1} = A(\theta)s_t + B(\theta)\epsilon_{t+1} \tag{A.3}$$

$$y_{t+1} = C(\theta)s_t + D(\theta)\epsilon_{t+1}, \tag{A.4}$$

where

$$s_t = GDP_t, \qquad y_t = \begin{bmatrix} GDP_{Et} \\ GDP_{It} \end{bmatrix} \tag{A.5}$$

$$A(\theta) = \rho, \qquad B(\theta) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

$$C(\theta) = \begin{bmatrix} \rho \\ \rho \end{bmatrix}, \qquad D(\theta) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

and $\theta = [\rho, vech(\Sigma)']'$. Note that only $A(\theta)$ and $C(\theta)$ are non-trivial functions of θ.

Assumption 1. The parameter vector θ satisfies the following conditions: (i) Σ is positive definite; (ii) $0 \leq \rho < 1$.

Because the rows of D are linearly independent, Assumption 1(i) implies that $D\Sigma D'$ is non-singular. In turn, we deduce that Assumptions 1, 2, and 4-NS of Komunjer and Ng (2011) are satisfied.

We now express the state-space system in terms of its innovation representation

$$s_{t+1|t+1} = A(\theta)s_{t|t} + K(\theta)a_{t+1} \tag{A.6}$$

$$y_{t+1} = C(\theta)\hat{s}_{t|t} + a_{t+1},$$

where a_{t+1} is the one-step-ahead forecast error of the system whose variance we denote by $\Sigma_a(\theta)$. The innovation representation is obtained from the Kalman filter as follows. Suppose that conditional on time t information $Y_{1:t}$ the distribution of $s_t|Y_{1:t} \sim N(s_{t|t}, P_{t|t})$. Then the joint distribution of $[s_{t+1}, y'_{t+1}]'$ is

$$\begin{bmatrix} s_{t+1} \\ y_{t+1} \end{bmatrix} \Bigg| Y_{1:T} \sim \left(\begin{bmatrix} As_{t|t} \\ Cs_{t|t} \end{bmatrix}, \begin{bmatrix} AP_{t|t}A' + B\Sigma B' & AP_{t|t}C' + B\Sigma D' \\ CP_{t|t}A' + D\Sigma B' & CP_{t|t}C' + D\Sigma D' \end{bmatrix} \right).$$

In turn, the conditional distribution of $s_{t+1}|Y_{1:t+1}$ is

$$s_{t+1}|Y_{1:t+1} \sim N\big(s_{t+1|t+1}, P_{t+1|t+1}\big),$$

where

$$\begin{aligned} s_{t+1|t+1} &= As_{t|t} + (AP_{t|t}C + B\Sigma D')(CP_{t|t}C' + D\Sigma D')^{-1}(y_t - Cs_{t|t}) \\ P_{t+1|t+1} &= AP_{t|t}A' + B\Sigma B' - (AP_{t|t}C' + B\Sigma D') \\ &\quad \times (CP_{t|t}C' + D\Sigma D')^{-1}(CP_{t|t}A' + D\Sigma B'). \end{aligned}$$

Now let P be the matrix that solves the Riccati equation,

$$\begin{aligned} P &= APA' + B\Sigma B' - (APC' + B\Sigma D')(CPC' + D\Sigma D')^{-1} \\ &\quad \times (CPA' + D\Sigma B'), \end{aligned} \tag{A.7}$$

and let K be the Kalman gain matrix

$$K = (APC' + B\Sigma D')(CPC' + D\Sigma D')^{-1}. \tag{A.8}$$

Then the one-step-ahead forecast error matrix is given by

$$\Sigma_a = CPC' + D\Sigma D'. \tag{A.9}$$

Eqs. (A.7)–(A.9) determine the matrices that appear in the innovation-representation of the state-space system (A.6).

In order to be able to apply Proposition 1-NS of Komunjer and Ng (2011) we need to express P, K, and Σ_a in terms of θ. While solving Riccati equations analytically is in general not feasible, our system is scalar, which simplifies the calculation considerably. Replacing A by ρ and P by p such that scalars appear in lower case, and defining

$$\Sigma_{BB} = B\Sigma B', \qquad \Sigma_{BD} = B\Sigma D', \quad \text{and} \quad \Sigma_{DD} = D\Sigma D',$$

we can write (A.7) as

$$\begin{aligned} p &= p\rho^2 + \Sigma_{BB} - (p\rho C' + \Sigma_{BD})(pCC' + \Sigma_{DD})^{-1} \\ &\quad \times (p\rho C + \Sigma_{DB}). \end{aligned} \tag{A.10}$$

Likewise,

$$\begin{aligned} &K = (p\rho C' + \Sigma_{BD})(pCC' + \Sigma_{DD})^{-1} \quad \text{and} \\ &\Sigma_a = pCC' + \Sigma_{DD}. \end{aligned} \tag{A.11}$$

Because $\Sigma_{BB} - \Sigma_{BD}\Sigma'_{DD}\Sigma_{DB} > 0$ we can deduce that $p > 0$. Moreover, because $A = \rho \geq 0$ and $C \geq 0$, we deduce that $K \neq 0$ and therefore Assumption 5-NS of Komunjer and Ng (2011) is satisfied. According to Proposition 1-NS in Komunjer and Ng (2011), two vectors θ and θ_1 are observationally equivalent if and only if there exists a scalar $\gamma \neq 0$ such that

$$A(\theta_1) = \gamma A(\theta)\gamma^{-1} \tag{A.12}$$

$$K(\theta_1) = \gamma K(\theta) \tag{A.13}$$

$$C(\theta_1) = C(\theta)\gamma^{-1} \tag{A.14}$$

$$\Sigma_a(\theta_1) = \Sigma_a(\theta). \tag{A.15}$$

Define $\theta = [\rho, vech(\Sigma)']'$ and $\theta_1 = [\rho_1, vech(\Sigma_1)']'$. Using the definition of the scalar $A(\theta)$ in (A.5) we deduce from (A.12) that $\rho_1 = \rho$. Since $C(\theta)$ depends on θ only through ρ we can deduce from (A.14) that $\gamma = 1$. Thus, given θ and ρ, the elements of the vector $vech(\Sigma_1)$ have to satisfy conditions (A.13) and (A.15), which, using (A.11), can be rewritten as

$$\Sigma_a = \Sigma_{a1} = p_1CC' + \Sigma_{DD1} \tag{A.16}$$

$$K = K_1 = (p_1\rho C' + \Sigma_{BD1})\Sigma_a^{-1}. \tag{A.17}$$

Moreover, p_1 has to solve the Riccati equation (A.10):

$$p_1 = p_1\rho^2 + \Sigma_{BB1} - K_0(p_1\rho C + \Sigma_{BD}). \tag{A.18}$$

Eqs. (A.16)–(A.18) are satisfied if and only if

$$pCC' + \Sigma_{DD} = p_1CC' + \Sigma_{DD1} \tag{A.19}$$

$$p\rho C' + \Sigma_{BD} = p_1\rho C' + \Sigma_{BD1} \tag{A.20}$$

$$p(1-\rho^2) - \Sigma_{BB} = p_1(1-\rho^2) - \Sigma_{BB1}. \tag{A.21}$$

We proceed by deriving expressions for the Σ_{xx} matrices that appear in (A.19)–(A.21):

$$\Sigma_{BB} = \Sigma_{GG}$$
$$\Sigma_{BD} = \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} \end{bmatrix}$$
$$\Sigma_{DD} = \begin{bmatrix} \Sigma_{GG} + \Sigma_{EE} + 2\Sigma_{EG} & \cdot \\ \Sigma_{GG} + \Sigma_{GE} + \Sigma_{GI} + \Sigma_{EI} & \Sigma_{GG} + \Sigma_{II} + 2\Sigma_{GI} \end{bmatrix}.$$

Without loss of generality let

$$\Sigma_{GG1} = \Sigma_{GG} + (1-\rho^2)\delta, \tag{A.22}$$

which implies that

$$\Sigma_{BB1} = \Sigma_{BB} + (1-\rho^2)\delta.$$

We now distinguish the cases $\delta = 0$ and $\delta \neq 0$.

Case 1: $\delta = 0$. (A.21) implies $p_1 = p$. It follows from (A.20) that $\Sigma_{BD1} = \Sigma_{BD}$. In turn, $\Sigma_{GE1} = \Sigma_{GE}$ and $\Sigma_{GI1} = \Sigma_{GI}$. Finally, to satisfy (A.19) it has to be the case that $\Sigma_{DD1} = \Sigma_{DD}$, which implies that the remaining elements of Σ and Σ_1 are identical. We conclude that $\theta_1 = \theta$.

Case 2: $\delta \neq 0$. (A.21) implies $p_1 = p + \delta$. Now consider (A.20):

$$\begin{aligned} p\rho C' + \Sigma_{BD} &= p\rho^2 \begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} \end{bmatrix} \\ &\overset{!}{=} p\rho^2 \begin{bmatrix} 1 & 1 \end{bmatrix} + \delta\rho^2 \begin{bmatrix} 1 & 1 \end{bmatrix} \\ &\quad + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE1} & \Sigma_{GG} + \Sigma_{GI1} \end{bmatrix} \\ &\quad + \delta(1-\rho^2)\begin{bmatrix} 1 & 1 \end{bmatrix}. \end{aligned}$$

We deduce that

$$\Sigma_{GE1} = \Sigma_{GE} - \delta, \qquad \Sigma_{GI1} = \Sigma_{GI} - \delta. \tag{A.23}$$

Finally, consider (A.19), which can be rewritten as

$$0 = \Sigma_{DD1} - \Sigma_{DD} + \delta CC'.$$

Using the previously derived expressions for Σ_{DD} and Σ_{DD1} we obtain the following three conditions

$$0 = (1-\rho^2)\delta + (\Sigma_{EE1} - \Sigma_{EE}) - 2\delta + \rho^2\delta = \Sigma_{EE1} - \Sigma_{EE} - \delta$$
$$0 = (1-\rho^2)\delta - 2\delta + (\Sigma_{EI1} - \Sigma_{EI}) + \rho^2\delta = \Sigma_{EI1} - \Sigma_{EI} - \delta$$
$$0 = (1-\rho^2)\delta + (\Sigma_{II1} - \Sigma_{II}) - 2\delta + \rho^2\delta = \Sigma_{II1} - \Sigma_{II} - \delta.$$

Thus, we deduce that

$$\begin{aligned} &\Sigma_{EE1} = \Sigma_{EE} + \delta, \qquad \Sigma_{EI1} = \Sigma_{EI} + \delta, \quad \text{and} \\ &\Sigma_{II1} = \Sigma_{II} + \delta. \end{aligned} \tag{A.24}$$

Combining (A.22)–(A.24) we find that

$$\Sigma_1 = \begin{bmatrix} \Sigma_{GG} + \delta(1-\rho^2) & \Sigma_{GE} - \delta & \Sigma_{GI} - \delta \\ \Sigma_{GE} - \delta & \Sigma_{EE} + \delta & \Sigma_{EI} + \delta \\ \Sigma_{GI} - \delta & \Sigma_{EI} + \delta & \Sigma_{II} + \delta \end{bmatrix}. \tag{A.25}$$

Thus, we have proved the following theorem:

$$\Sigma_{BB} = \Sigma_{GG}$$
$$\Sigma_{BD} = \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} & \lambda \Sigma_{GG} + \Sigma_{GU} \end{bmatrix}$$
$$\Sigma_{DD} = \begin{bmatrix} \Sigma_{GG} + \Sigma_{EE} + 2\Sigma_{GE} & \cdot & \cdot \\ \Sigma_{GG} + \Sigma_{GE} + \Sigma_{GI} + \Sigma_{EI} & \Sigma_{GG} + \Sigma_{II} + 2\Sigma_{GI} & \cdot \\ \lambda \Sigma_{GG} + \lambda \Sigma_{GE} + \Sigma_{GU} & \lambda \Sigma_{GG} + \lambda \Sigma_{GI} + \Sigma_{GU} & \lambda^2 \Sigma_{GG} + 2\lambda \Sigma_{GU} + \Sigma_{UU} \end{bmatrix}$$

Box I.

Theorem A.1. *Suppose Assumption 1 is satisfied. Then the two-variable model is*

(i) *identified if Σ is diagonal as in Section 2.1;*
(ii) *identified if Σ is block-diagonal as in Section 2.2;*
(iii) *not identified if Σ is unrestricted as in Section 2.3;*
(iv) *identified if Σ is restricted as in Section 2.4.*

Appendix B. Identification in the three-equation model

The identification analysis of the three-variable is similar to the analysis of the two-variable model in the previous section. The system is given by

$$GDP_t = \rho GDP_{t-1} + \epsilon_{Gt} \tag{B.26}$$

$$\begin{bmatrix} GDP_{Et} \\ GDP_{It} \\ U_t \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \lambda \end{bmatrix} GDP_t + \begin{bmatrix} \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix}, \tag{B.27}$$

and the joint distribution of the errors is

$$\epsilon_t = \begin{bmatrix} \epsilon_{Gt} \\ \epsilon_{Et} \\ \epsilon_{It} \\ \epsilon_{Ut} \end{bmatrix} \sim iidN\big(0, \Sigma\big),$$

$$\text{where } \Sigma = \begin{bmatrix} \Sigma_{GG} & \cdot & \cdot & \cdot \\ \Sigma_{EG} & \Sigma_{EE} & \cdot & \cdot \\ \Sigma_{IG} & \Sigma_{IE} & \Sigma_{II} & \cdot \\ \Sigma_{UG} & \Sigma_{UE} & \Sigma_{UI} & \Sigma_{UU} \end{bmatrix}.$$

The matrices $A(\theta)$, $B(\theta)$, $C(\theta)$, and $D(\theta)$ are now given by

$$A(\theta) = \rho, \qquad B(\theta) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

$$C(\theta) = \begin{bmatrix} \rho \\ \rho \\ \lambda\rho \end{bmatrix}, \qquad D(\theta) = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ \lambda & 0 & 0 & 1 \end{bmatrix},$$

where $\theta = [\rho, \lambda, vech(\Sigma)']'$.

Assumption 2. The parameter vector θ satisfies the following conditions: (i) Σ is positive definite; (ii) $0 < \rho < 1$; (iii) $\lambda \neq 0$; (iv) $\Sigma_{UE} = \Sigma_{UI} = 0$.

Condition (A.12) implies that $\rho_1 = \rho$. Moreover, (A.14) implies that $\gamma = 1$ and that $\lambda_1 = \lambda$ provided that $\rho \neq 0$. As for the two-variable model, we have to verify that (A.19)–(A.21) are satisfied. The matrices Σ_{xx} that appear in these equations are given in Box I.

Without loss of generality, let

$$\Sigma_{GG,1} = \Sigma_{GG} + (1-\rho^2)\delta,$$

which implies that

$$\Sigma_{BB,1} = \Sigma_{BB} + (1-\rho^2)\delta.$$

Case 1: $\delta = 0$. (A.21) implies $p_1 = p$. It follows from (A.20) that $\Sigma_{BD,1} = \Sigma_{BD}$. In turn, $\Sigma_{GE,1} = \Sigma_{GE}$, $\Sigma_{GI,1} = \Sigma_{GI}$, and $\Sigma_{GU,1} = \Sigma_{GU}$. Finally, to satisfy (A.17) it has to be the case that $\Sigma_{DD,1} = \Sigma_{DD}$, which implies that the remaining elements of Σ and Σ_1 are identical for the two parameterizations. We conclude that it has to be the case that $\theta_1 = \theta$.

Case 2: $\delta \neq 0$. (A.21) implies $p_1 = p + \delta$. Now consider (A.20):

$$\begin{aligned} &p\rho C' + \Sigma_{BD} \\ &\quad = p\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE} & \Sigma_{GG} + \Sigma_{GI} & \lambda\Sigma_{GG} + \Sigma_{GU} \end{bmatrix} \\ &\quad \overset{!}{=} p\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} + \delta\rho^2 \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix} \\ &\qquad + \begin{bmatrix} \Sigma_{GG} + \Sigma_{GE,1} & \Sigma_{GG} + \Sigma_{GI,1} & \lambda\Sigma_{GG} + \Sigma_{GU,1} \end{bmatrix} \\ &\qquad + (1-\rho^2)\delta \begin{bmatrix} 1 & 1 & \lambda \end{bmatrix}. \end{aligned}$$

We deduce that

$$\Sigma_{GE,1} = \Sigma_{GE} - \delta, \qquad \Sigma_{GI,1} = \Sigma_{GI} - \delta, \qquad \Sigma_{GU,1} = \Sigma_{GU} - \delta.$$

Finally, consider (A.19), which can be rewritten as

$$0 = \Sigma_{DD,1} - \Sigma_{DD} + \delta CC'.$$

Using the previously derived expressions for Σ_{DD} and Σ_{DD1} we obtain the following five conditions

$$\begin{aligned} 0 &= (1-\rho^2)\delta + (\Sigma_{EE1} - \Sigma_{EE}) - 2\delta + \rho^2\delta = \Sigma_{EE1} - \Sigma_{EE} - \delta \\ 0 &= (1-\rho^2)\delta - 2\delta + (\Sigma_{EI1} - \Sigma_{EI}) + \rho^2\delta = \Sigma_{EI1} - \Sigma_{EI} - \delta \\ 0 &= (1-\rho^2)\delta + (\Sigma_{II1} - \Sigma_{II}) - 2\delta + \rho^2\delta = \Sigma_{II1} - \Sigma_{II} - \delta \\ 0 &= \lambda(1-\rho^2)\delta - \lambda\delta - \delta + \lambda\rho^2\delta = \delta \\ 0 &= \lambda^2(1-\rho^2)\delta - 2\lambda\delta + (\Sigma_{UU1} - \Sigma_{UU}) + \lambda^2\rho^2\delta \\ &= \Sigma_{UU1} - \Sigma_{UU} - \lambda(2-\lambda)\delta. \end{aligned}$$

Thus, we deduce that

$$\delta = 0, \qquad \Sigma_{EE1} = \Sigma_{EE}, \qquad \Sigma_{EI1} = \Sigma_{EI}, \qquad \Sigma_{II1} = \Sigma_{II}, \qquad \text{and} \quad \Sigma_{UU1} = \Sigma_{UU}.$$

This proves the following theorem:

Theorem B.1. *Suppose Assumption 2 is satisfied. Then the three-variable model is identified.*

Appendix C. Supplementary data

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2015.12.009.

References

Aruoba, B., 2008. Data revisions are not well-behaved. J. Money Credit Bank. 40, 319–340.
Aruoba, S.B., Diebold, F.X., 2010. Real-time macroeconomic monitoring: Real activity, inflation, and interactions. Am. Econ. Rev. 100, 20–24.
Aruoba, S.B., Diebold, F.X., Nalewaik, J., Schorfheide, F., Song, D., 2012. Improving GDP measurement: A forecast combination perspective. In: Chen, X., Swanson, N. (Eds.), Recent Advances and Future Directions in Causality, Prediction, and Specification Analysis: Essays in Honour of Halbert L. White Jr., Springer, pp. 1–26.
Aruoba, S.B., Diebold, F.X., Scotti, C., 2009. Real time measurement of business conditions. J. Bus. Econom. Statist. 27, 417–427.
BEA 2006. A Guide to the National Income and Product Accounts of the United States. Bureau of Economic Analysis, Washington, DC. http://www.bea.gov/national/pdf/nipaguid.pdf.
Durbin, J., Koopman, S.J., 2001. Time Series Analysis by State Space Methods. Oxford University Press, Oxford.
Edwards, C.L., Howrey, E.P., 1991. A 'true' time series and its indicators: An alternative approach. J. Amer. Statist. Assoc. 86, 878–882.

Faust, J., Rogers, J.H., Wright, J.H., 2005. News and noise in G-7 GDP announcements. J. Money Credit Bank. 37, 403–417.
Fixler, D.J., Nalewaik, J.J., 2009. News, Noise, and Estimates of the True Unobserved State of the Economy. Bureau of Labor Statistics and Federal Reserve Board, Manuscript.
Fleischman, C.A., Roberts, J.M., 2011. A Multivariate Estimate of Trends and Cycles. Federal Reserve Board, Manuscript.
Gartaganis, A.J., Goldberger, A.S., 1955. A note on the statistical discrepancy in the national accounts. Econometrica 23, 166–173.
Geweke, J.F., 1977. The dynamic factor analysis of economic time series models. In: Aigner, D., Goldberger, A. (Eds.), Latent Variables in Socioeconomic Models. North Holland, pp. 365–383.
Greenaway-McGrevy, Ryan, 2011. Is GDP or GDI a better measure of output? A statistical approach. Bureau of Economic Analysis Manuscript.
Harding, D., Scutella, R., (1996). Efficient Estimates of GDP, Unpublished Seminar Notes, La Trobe University, Australia.
Jacobs, J.P.A.M., van Norden, S., 2011. Modeling data revisions: Measurement error and dynamics of true values. J. Econometrics 161, 101–109.
Kishor, N.K., Koenig, E.F., 2012. VAR estimation and forecasting when data are subject to revision. J. Bus. Econom. Statist. 30 (2), 181–190.
Komunjer, I., Ng, S., 2011. Dynamic identification of dynamic stochastic general equilibrium models. Econometrica 79, 1995–2032.
Landefeld, J.S., Seskin, E.P., Fraumeni, B.M., 2008. Taking the pulse of the economy: Measuring GDP. J. Econ. Perspect. 22, 193–216.
Mankiw, N.G., Runkle, D.E., Shapiro, M.D., 1984. Are preliminary announcements of the money stock rational forecasts? J. Monetary Econ. 14, 15–27.
Mankiw, N.G., Shapiro, M.D., 1986. News or noise: An analysis of GNP revisions. Survey of Current Business 20–25.
McCulla, S.H., Smith, S., 2007. Measuring the Economy: A Primer on GDP and the National Income and Product Accounts. Bureau of Economic Analysis, Washington, DC. http://www.bea.gov/national/pdf/nipa_primer.pdf.
Nalewaik, J.J., 2010. The income- and expenditure-side estimates of U.S. output growth. Brookings Papers on Economic Activity 1, 71–127 (with discussion).
Rassier, D.G., 2012. The role of profits and income in the statistical discrepancy. Survey of Current Business 8–22.
Sargent, T.J., Sims, C.A., 1977. Business cycle modeling without pretending to have too much a priori theory. In: Sims, C.A. (Ed.), New Methods in Business Cycle Research: Proceedings from a Conference. Federal Reserve Bank of Minneapolis, pp. 45–109.
Smith, R.J., Weale, M.R., Satchell, S.E., 1998. Measurement error with accounting constraints: Point and interval estimation for latent data with an application to U.K. gross domestic product. Rev. Econom. Stud. 65, 109–134.
Stone, R., Champernowne, D.G., Meade, J.E., 1942. The precision of national income estimates. Rev. Econom. Stud. 9, 111–125.
Watson, M.W., Engle, R.F., 1983. Alternative algorithms for the estimation of dynamic factor, MIMIC and varying coefficient regression models. J. Econometrics 23, 385–400.