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. Borağan 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.