Can GDP measurement be further improved? Data revision and reconciliation

Jan P.A.M. Jacobs*

University of Groningen, University of Tasmania, CAMA and CIRANO

Samad Sarferaz

KOF Swiss Economic Institute, ETH Zurich, Switzerland

Jan-Egbert Sturm

KOF Swiss Economic Institute, ETH Zurich, Switzerland and CESifo, Germany

Simon van Norden

HEC Montréal, CIRANO and CIREQ

August 2018

Abstract

Recent years have seen many attempts to combine expenditure-side estimates of U.S. real output (GDE) growth with income-side estimates (GDI) to improve estimates of real GDP growth. We show how to incorporate information from multiple releases of noisy data to provide more precise estimates while avoiding some of the identifying assumptions required in earlier work. This relies on a new insight: using multiple data releases allows us to distinguish news and noise measurement errors in situations where a single vintage does not.

Our new measure, GDP^{++}, fits the data better than GDP^{+}, the GDP growth measure of Aruoba et al. (2016) published by the Federal Reserve Bank of Philadephia. Historical decompositions show that GDE releases are more informative than GDI, while the use of multiple data releases is particularly important in the quarters leading up to the Great Recession.
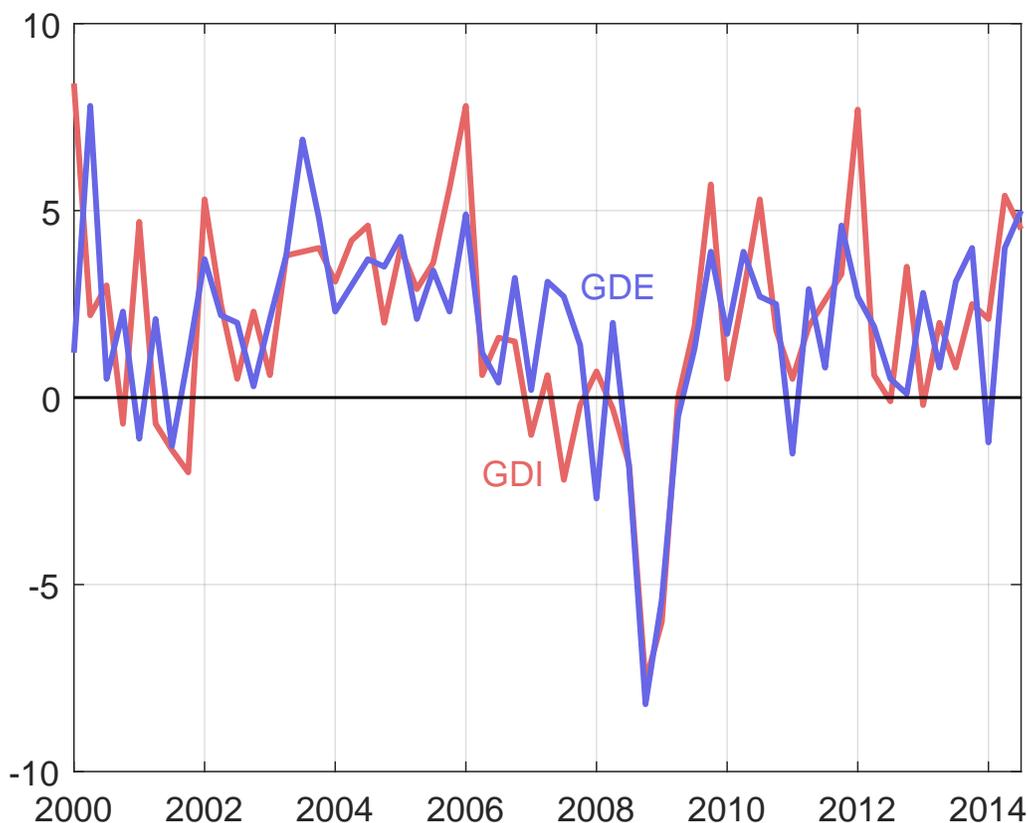
1 Introduction

Unlike many other nations, U.S. national accounts feature distinct estimates of real output based on the expenditure approach (GDE) and the income approach (GDI), see Figure 1. As pointed out by Stone, Champernowne and Meade (1942), while in theory these two approaches should give identical estimates, measurement errors cause discrepancies to arise.[1] These discrepancies are sometimes important. Chang and Li (2015) examine the impact of using GDI rather than GDE in nearly two dozen recent empirical papers published in major economic journals; they find substantive differences in roughly 15% of them. Nalewaik (2012) finds that GDI leads to quicker detection of U.S. recessions than GDE.

Figure 1: U.S. GDP growth: Expenditure side vs. income side

---

[1]The same applies to the production-based estimate of output. See e.g. the study of Rees, Lancaster and Finlay (2015) on Australian GDP.

While several studies have tried to determine which measure should be preferred in various contexts, Weale (1992) and Diebold (2010) argue that reconciling them is a more useful response as it should incorporate more information. Fixler and Nalewaik (2009) point out, however, that reconciliation traditionally relies on the assumption that measurement errors are "noise", which in turn forces the reconciled estimate of the latent variable ("true" GDP in this case) to be less variable than any of the individual series being reconciled. They instead propose that measurement errors may also include a "news" component. While this causes a loss of identification, they glean information from the revision of GDE and GDI to place bounds on relative contributions of news and noise in a least-squares framework. Aruoba et al. (2012) consider the problem from a forecast combination perspective, assuming "news" errors and imposing priors in lieu of identification without revisions, while Aruoba et al. (2016) consider alternative identifying assumptions and propose the addition of an instrumental variable. Almuzara et al. (2018) investigate a dynamic factor model with cointegration restrictions.

Aruoba et al. (2016) is the basis for the GDP^+ measure published by the Federal Reserve Bank of Philadelphia.[2] However, while their approach ignores the possibility of data revision, Figure 2 shows that the published series is subject to important revisions, which complicates its interpretation and use in policy decisions. Separately, Jacobs and van Norden (2011) and Kishor and Koenig (2012) propose state-space frameworks that allow estimation of both news- and noise-type measurement errors in data revision, but do not consider problems of data reconciliation. In this paper we extend Jacobs and van Norden (2011, henceforth JvN) to consider the problem of reconciliation and identification in which there are multiple estimates of the common underlying variable, all of which are subject to revision. Allowing for both news and noise measurement errors, the result is a modeling framework substantially more general than those previously proposed. We show that identification of these two types of measurement errors is made possible by modeling data revisions as well as the dynamics of the series. We provide a historical decomposition of GDE and GDI into news and noise shocks, and we compare those series to our improved GDP estimate, GDP^{++}. We find that GDP^{++} is more persistent than either GDE or

---

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

Figure 2: GDP^+ in real-time
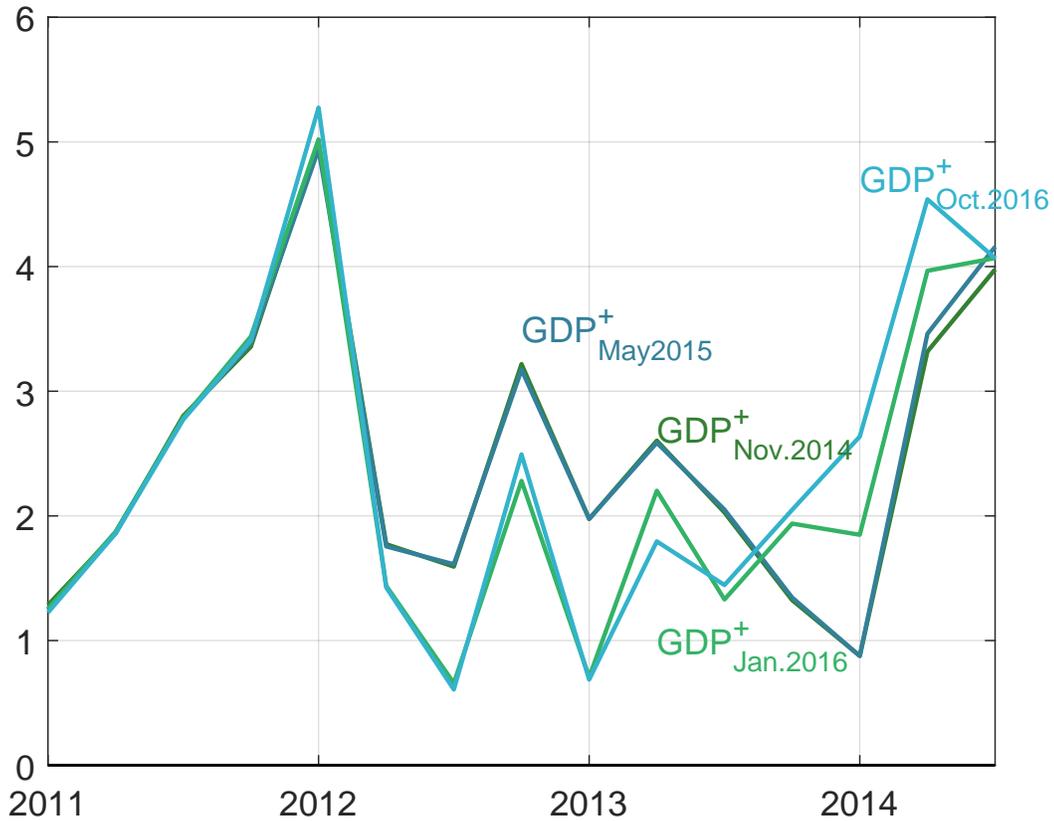
Various vintages of GDP^+ according to the estimates of the Federal Reserve Bank of Philadelphia.

GDI. While both series appear to contain both news and noise shocks, news shocks have a larger share in GDE than in GDI.

The paper is structured as follows. In Section 2 we present our econometric framework. We show that our system is identified using real-time data and news-noise assumptions. In Section 3 we describe our data and estimation method. Results are shown in Section 4 and Section 5 concludes. Formal proofs of some results related to identification are presented in an Appendix.

2 Econometric Framework

In this section, after establishing some notation, we describe our econometric framework. We begin by briefly reviewing the univariate news and noise model of JvN before generalizing it to the problem of data reconciliation. We then compare the results to the GDP^+ model of Aruoba et al. (2016) and discuss their differences for the identification of news and noise measurement errors.

We follow the standard notation in this literature by letting y_t^{t+j} be an estimate published at time $t + j$ of some real-valued scalar variable y at time t. We define $\boldsymbol{y}_t$ as a $l \times 1$ vector of l different vintage estimates of y_t^{t+i}, $i = 1, \ldots, l$ so $\boldsymbol{y}_t \equiv \left[y_t^{t+1}, y_t^{t+2}, \ldots, y_t^{t+l} \right]'$. For state-space models, we follow the notation of Durbin and Koopman (2001)

$$\boldsymbol{y}_t = \boldsymbol{Z} \cdot \boldsymbol{\alpha}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

$$\boldsymbol{\alpha}_{t+1} = \boldsymbol{T} \cdot \boldsymbol{\alpha}_t + \boldsymbol{R} \cdot \boldsymbol{\eta}_t \tag{2}$$

where $\boldsymbol{y}_t$ is $l \times 1$, $\boldsymbol{\alpha}_t$ is $m \times 1$, $\boldsymbol{\varepsilon}_t$ is $l \times 1$ and $\boldsymbol{\eta}_t$ is $r \times 1$; $\boldsymbol{\varepsilon}_t \sim N(0, \boldsymbol{H})$ and $\boldsymbol{\eta}_t \sim N(0, \boldsymbol{I}_r)$. Both error terms are i.i.d. and orthogonal to one another.[3]

2.1 A State-Space model of Measurement Error with News and Noise

JvN denote the unobserved "true" value of a variable as $\tilde{y}_t$, so that its measurement error $\boldsymbol{u}_t \equiv \boldsymbol{y}_t - \boldsymbol{\iota}_l \cdot \tilde{y}_t$, where $\boldsymbol{\iota}_l$ is an $l \times 1$ vector of ones. They model these measurement errors as the sum of "news" and "noise" measurement errors. Measurement errors are said to be noise $\left(\zeta_t^{t+i} \right)$ when they are orthogonal to the true values $\tilde{y}_t$, so that

$$y_t^{t+i} = \tilde{y}_t + \zeta_t^{t+i}, \qquad \text{cov}(\tilde{y}_t, \zeta_t^{t+i}) = 0. \tag{3}$$

---

[3]For more detailed assumptions, see Durbin and Koopman (2001, Section 3.1 and 4.1. For convenience we omit constants from the model in this exposition.

Noise implies that revisions $(y_t^{t+i+1} - y_t^{t+i})$ are generally forecastable. Measurement errors are described as news (ν_t^{t+i}) if and only if

$$\tilde{y}_t = y_t^{t+i} + \nu_t^{t+i}, \qquad \text{cov}(y_t^{t+j}, \nu_t^{t+i}) = 0 \qquad \forall j \leq i \tag{4}$$

If data revisions are pure news errors, current and past vintages of the series will be of no use in forecasting future data revision.

In their state-space model JvN impose $\boldsymbol{\varepsilon}_t \equiv \mathbf{0}_{l \times 1}$ and partition the state vector $\boldsymbol{\alpha}_t$ into four components

$$\boldsymbol{\alpha}_t = [\tilde{y}_t, \boldsymbol{\phi}_t', \boldsymbol{\nu}_t', \boldsymbol{\zeta}_t']', \tag{5}$$

of length 1, b, l and l respectively, where $\boldsymbol{\phi}_t$ is used to capture the dynamics of the true values while $\boldsymbol{\nu}_t$ and $\boldsymbol{\zeta}_t$ are the news and noise measurement errors, respectively. They similarly partition

$$\boldsymbol{Z} = [\boldsymbol{Z}_1, \boldsymbol{Z}_2, \boldsymbol{Z}_3, \boldsymbol{Z}_4] \tag{6}$$

where $\boldsymbol{Z}_1 = \boldsymbol{\iota}_l$ (a $l \times 1$ vector of 1's), $\boldsymbol{Z}_2 = \mathbf{0}_{l \times b}$ (an $l \times b$ matrix of zeros), $\boldsymbol{Z}_3 = \boldsymbol{I}_l$, and $\boldsymbol{Z}_4 = \boldsymbol{I}_l$ (both $l \times l$ identity matrices). Their measurement equation (1) then simplifies to

$$\boldsymbol{y}_t = \boldsymbol{Z} \cdot \boldsymbol{\alpha}_t = \tilde{y}_t + \boldsymbol{\nu}_t + \boldsymbol{\zeta}_t = \text{`Truth'} + \text{`News'} + \text{`Noise'}. \tag{7}$$

They conformably partition the matrix $\boldsymbol{T}$ as

$$\boldsymbol{T} = \begin{bmatrix} T_{11} & \boldsymbol{T}_{12} & \mathbf{0} & \mathbf{0} \\ \boldsymbol{T}_{21} & \boldsymbol{T}_{22} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{T}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_4 \end{bmatrix}, \tag{8}$$

where T_{11} is a scalar, and $\{\boldsymbol{T}_{12}, \boldsymbol{T}_{21}, \boldsymbol{T}_{22}, \boldsymbol{T}_3, \boldsymbol{T}_4\}$ are $1 \times b$, $b \times 1$, $b \times b$, $l \times l$ and $l \times l$; $\mathbf{0}$ is a conformably defined matrix of zeros. The $(b+1) \times (b+1)$ block in the upper left simply captures the dynamics of $\tilde{y}_t$ while $\boldsymbol{T}_3$ and $\boldsymbol{T}_4$ capture the dynamics of the news and noise shocks. If measurement errors are independent across time periods (but not vintages), then $\boldsymbol{T}_3 \equiv \boldsymbol{T}_4 \equiv \mathbf{0}_{l \times l}$. As we will see below, in the special case where $\tilde{y}_t$ is assumed to follow

an $AR(p)$ process, this will impose $p = b + 1$, the row vector $\begin{bmatrix} T_{11} & T_{12} \end{bmatrix}$ will contain the autoregressive coefficients and the remainder of the upper left $(b+1) \times (b+1)$ part will be composed of zeros and ones.[4]

The essential difference between news and noise errors is captured in the $(1 + b + 2l) \times (1 + 2l)$ matrix $\boldsymbol{R}$, which is partitioned as follows

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_3 & \boldsymbol{0} \\ \boldsymbol{R}_2 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{U}_l \cdot \operatorname{diag}(\boldsymbol{R}_3) & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_4 \end{bmatrix}, \tag{9}$$

where $\boldsymbol{U}_l$ is a $l \times l$ matrix with zeros below the main diagonal and ones everywhere else, $\boldsymbol{R}_3 = [\sigma_{\nu 1}, \sigma_{\nu 2}, \dots, \sigma_{\nu l}]$, where $\sigma_{\nu i}$ is the standard error of the measurement error associated with i-th estimate y_t^{t+i}, $\operatorname{diag}(\boldsymbol{R}_3)$ is a $l \times l$ matrix with elements of $\boldsymbol{R}_3$ on its main diagonal, and $\boldsymbol{R}_4$ is an $l \times l$ matrix. Finally, the error term is partitioned as $\boldsymbol{\eta}_t = \begin{bmatrix} \boldsymbol{\eta}'_{et}, & \boldsymbol{\eta}'_{\nu t}, & \boldsymbol{\eta}'_{\zeta t} \end{bmatrix}'$, where $\boldsymbol{\eta}_{et}$ refers to errors associated with the true values, and $\boldsymbol{\eta}_{\nu t}$ and $\boldsymbol{\eta}_{\zeta t}$ are the errors for news and noise, respectively.

JvN note that (if the model is identified, a question we deal with below) this framework permits conventional techniques to be used to estimate the model parameters, allow for missing observations, estimate and forecast the unobserved true values $\widetilde{y}_t$ together with their confidence intervals, and test hypotheses.

2.2 Data Reconciliation

We now show how the above framework may be adapted to the case where we have two alternative estimates of the same underlying true value $\widetilde{y}_t$, both of which are subject to revision. We define $\boldsymbol{Y}_t$ as a $2l \times 1$ vector of l different vintage estimates for the 2 variables $y1_t^{t+i}$ and $y2_t^{t+i}$, $i = 1, \dots, l$, for a particular observation t, so $\boldsymbol{Y}_t \equiv \begin{bmatrix} y1_t^{t+1}, y1_t^{t+2} \dots, y1_t^{t+l}, y2_t^{t+1}, y2_t^{t+2}, \dots, y2_t^{t+l} \end{bmatrix}'$, a vector of length $2l$. Our state-space model

---

[4]For details, see Jacobs and van Norden (2011).

now becomes

$$Y_t = Z \cdot \alpha_t \tag{10}$$

$$\alpha_{t+1} = T \cdot \alpha_t + R \cdot \eta_t \tag{11}$$

We again partition the state vector α_t into four components

$$\alpha_t = [\tilde{y}_t', \phi_t', \nu_t', \zeta_t']', \tag{12}$$

which are now of length 1, b, $2l$ and $2l$ respectively, and we similarly partition

$$Z = [Z_1, Z_2, Z_3, Z_4] \tag{13}$$

where $Z_1 = \iota_{2l}$ (a $2l$ vector of ones), $Z_2 = 0_{2l \times b}$ (a $2l \times b$ matrix of zeros), and $Z_3 = Z_4 = I_{2l}$ (both are $2l \times 2l$ identity matrices). The measurement equation (10) therefore again simplifies to

$$Y_t = Z \cdot \alpha_t = \tilde{y}_t + \nu_t + \zeta_t = \text{`Truth'} + \text{`News'} + \text{`Noise'}.$$

The matrix T is partitioned much as before

$$T = \begin{bmatrix} T_{11} & T_{12} & 0 & 0 \\ T_{21} & T_{22} & 0 & 0 \\ 0 & 0 & T_3 & 0 \\ 0 & 0 & 0 & T_4 \end{bmatrix}, \tag{14}$$

The upper left block (consisting of T_{11}, T_{12}, T_{21} and T_{22}) is precisely the same as in (8) above; this is because it solely determines the dynamics of $\tilde{y}_t$, which are unchanged. However, the addition of a new series increases the dimension of T_3 and T_4 from $l \times l$ to $2l \times 2l$.

R is now a $(1 + b + 4l) \times (1 + 4l)$ matrix where we separate the news and noise

7

measurement errors for the two variables

$$
\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_3 & \boldsymbol{R}_4 & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{R}_2 & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & -\boldsymbol{U}_l \cdot \operatorname{diag}(\boldsymbol{R}_3) & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & -\boldsymbol{U}_l \cdot \operatorname{diag}(\boldsymbol{R}_4) & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_5 & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{R}_6 \end{bmatrix}, \tag{15}
$$

where the row vector $\boldsymbol{R}_3 = \left[\sigma_{\nu_1^1}, \sigma_{\nu_2^1}, \dots, \sigma_{\nu_l^1} \right]$ corresponds to the news in $y1$ while $\boldsymbol{R}_4 = \left[\sigma_{\nu_1^2}, \sigma_{\nu_2^2}, \dots, \sigma_{\nu_l^2} \right]$ corresponds to the news in $y2$. $\operatorname{diag}(\boldsymbol{R}_3)$ and $\operatorname{diag}(\boldsymbol{R}_4)$ are $l \times l$ diagonal matrices with the elements of $\boldsymbol{R}_3$ and $\boldsymbol{R}_4$ on their main diagonals, while $\boldsymbol{R}_5$ and $\boldsymbol{R}_6$ are $l \times l$ diagonal matrices.

Finally, we partition $\boldsymbol{\eta}_t = \left[\boldsymbol{\eta}'_{et}, \ \boldsymbol{\eta}'_{\nu_1 t}, \ \boldsymbol{\eta}'_{\nu_2 t}, \ \boldsymbol{\eta}'_{\zeta_1 t}, \ \boldsymbol{\eta}'_{\zeta_2 t} \right]'$, where $\boldsymbol{\eta}_{et}$ refers to errors associated with the true values, and $\boldsymbol{\eta}_{\nu it}$ and $\boldsymbol{\eta}_{\zeta it}$ are the errors for news and noise measurement errors in variable i.

To illustrate, consider the following very simple case. Let $y1 \equiv GDE$ (the growth rate of real gross domestic expenditure), $y2 \equiv GDI$ (the growth rate of real gross domestic income), $l = 2$ (we only consider two vintages, the 1st and 2nd releases) and we'll assume that the growth rate of "true" real output $\widetilde{y}$ follows an $AR(1)$. Then (10) becomes

$$
\begin{bmatrix} GDE_t^{1st} \\ GDE_t^{2nd} \\ GDI_t^{1st} \\ GDI_t^{2nd} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{y}_t \\ \widetilde{y}_{t-1} \\ \boldsymbol{\nu}_t \\ \boldsymbol{\zeta}_t \end{bmatrix}
$$

$$
= \begin{bmatrix} \widetilde{y}_t \\ \widetilde{y}_t \\ \widetilde{y}_t \\ \widetilde{y}_t \end{bmatrix} + \begin{bmatrix} \nu_t^{GDE,1} & 0 & 0 & 0 \\ 0 & \nu_t^{GDE,2} & 0 & 0 \\ 0 & 0 & \nu_t^{GDI,1} & 0 \\ 0 & 0 & 0 & \nu_t^{GDI,2} \end{bmatrix} + \begin{bmatrix} \zeta_t^{GDE,1} & 0 & 0 & 0 \\ 0 & \zeta_t^{GDE,2} & 0 & 0 \\ 0 & 0 & \zeta_t^{GDI,1} & 0 \\ 0 & 0 & 0 & \zeta_t^{GDI,2} \end{bmatrix}
$$

$$
= \ \text{'Truth'} + \text{'News'} + \text{'Noise'}.
$$

8

and (11) becomes

$$
\begin{bmatrix} \tilde{y}_{t+1} \\ \tilde{y}_t \\ \boldsymbol{\nu}_{t+1} \\ \boldsymbol{\zeta}_{t+1} \end{bmatrix} = \begin{bmatrix} \rho & 0 & \mathbf{0} & \mathbf{0} \\ 1 & 0 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{T}_3 & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \boldsymbol{T}_4 \end{bmatrix} \cdot \begin{bmatrix} \tilde{y}_t \\ \tilde{y}_{t-1} \\ \boldsymbol{\nu}_t \\ \boldsymbol{\zeta}_t \end{bmatrix} + \boldsymbol{R} \cdot \boldsymbol{\eta}_t,
$$

where

$$
\boldsymbol{R} = \begin{bmatrix}
\sigma_e & \sigma_\nu^{GDE1} & \sigma_\nu^{GDE2} & \sigma_\nu^{GDI1} & \sigma_\nu^{GDI2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -\sigma_\nu^{GDE1} & -\sigma_\nu^{GDE2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -\sigma_\nu^{GDE2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -\sigma_\nu^{GDI1} & -\sigma_\nu^{GDI2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\sigma_\nu^{GDI2} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \sigma_\zeta^{GDE1} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \sigma_\zeta^{GDE2} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_\zeta^{GDI1} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma_\zeta^{GDI2}
\end{bmatrix}
$$

$$
\boldsymbol{\eta}_t = \left[e_t, \nu_t^{GDE1}, \nu_t^{GDE2}, \nu_t^{GDI1}, \nu_t^{GDI2}, \zeta_t^{GDE1}, \zeta_t^{GDE2}, \zeta_t^{GDI1}, \zeta_t^{GDI2} \right]'
$$

2.3   Identification and GDP$^+$

Aruoba et al. (2016) consider the problem of identification in a special case of the GDE/GDI example considered above where only a single vintage is available ($l = 1$). Their unrestricted

model may be written as[5]

$$
\begin{bmatrix} GDE_t \\ GDI_t \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{y}_t \\ \widetilde{y}_{t-1} \\ \eta_t^E \\ \eta_t^I \end{bmatrix} \tag{16}
$$

$$
\begin{bmatrix} \widetilde{y}_{t+1} \\ \widetilde{y}_t \\ \eta_{t+1}^E \\ \eta_{t+1}^I \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \widetilde{y}_t \\ \widetilde{y}_{t-1} \\ \eta_t^E \\ \eta_t^I \end{bmatrix} + \begin{bmatrix} \sigma_{yy} & \sigma_{yE} & \sigma_{yI} \\ 0 & 0 & 0 \\ \sigma_{Ey} & \sigma_{EE} & \sigma_{EI} \\ \sigma_{Iy} & \sigma_{IE} & \sigma_{II} \end{bmatrix} \cdot \begin{bmatrix} e_t^y \\ e_t^E \\ e_t^I \end{bmatrix} \tag{17}
$$

and they show that it is not identified. They propose adding a third (instrumental) variable which is correlated with $\widetilde{y}_t$ but not with η_t^E or η_t^I, suggesting that household survey data may be suitable for this purpose. We argue that the model may be identified instead by increasing the number of vintages analysed and assuming that measurement errors are the sum of news and noise measurement errors as characterized above. We explore this point in the remainder of this section by comparing the available number of sample moments to the number of free parameters in the model. In the Appendix we provide a more rigorous proof of identification in a slightly simpler model using the methods of Komunjer and Ng (2011).

The essential insight comes from the form of the $\boldsymbol{R}$ matrix in (15). News and noise measurement errors have tightly constrained behaviour across successive data vintages; Noise errors are assumed to be uncorrelated across vintages and with innovations in true values, while news errors must be correlated with one another, with innovations in true values, and their variances must be decreasing as series are revised.

If we have two series to reconcile (here GDE and GDI) and l vintages of each, we have $2 \cdot l \cdot (2 \cdot l + 1)/2$ observable cross moments as well as $2 \cdot l$ first-order autocorrelation coefficients, for a total of $l \cdot (2 \cdot l + 3)$ moments. The only free parameters in the above model, however, are the autocorrelation coefficient ρ and the $(1 + 4 \cdot l)$ non-zero elements of

---

[5]See Aruoba et al. (2016), equations (A.1) and (A.2). Their model further differs from the model above in that (a) they model only the sum of news and noise shocks, and (b) they assume that $\boldsymbol{T}_3 = \boldsymbol{T}_4 = 0$, a condition that we will also impose, below.

R, for a total of $2 \cdot (1 + 2 \cdot l)$. This implies that the number of available moments increases with l^2 while the number of free parameters increases only with l.[6]

In the special case where we use only a single data release, $l = 1$, we have $2 \cdot (1 + 2 \cdot 1) = 6$ free parameters to estimate, but only $1 \cdot (2 \cdot 1 + 3) = 5$ available moments with which to do so. This is consistent with the lack of identification noted by Aruoba et al. (2016). However, if we use $l = 2$ data vintages, we have $2 \cdot (1 + 2 \cdot 2) = 10$ free parameters and $2 \cdot (2 \cdot 2 + 3) = 14$ moments with which to identify them. For $l = 3$ we have 27 moments with which to estimate 14 parameters and for $l = 4$ (the case we consider below) we have 44 moments with which to estimate 18 parameters.

This suggests that as we add more data releases, we potentially have the ability to generalize the model further still. The univariate data revision model of JvN envisages two such types of generalization.

1. We may wish to relax some of the zero restrictions on $\boldsymbol{R}$. In particular, it may be desirable to allow for news shocks to be correlated across the two variables, or to allow for noise shocks to be correlated across data releases.

2. We may wish to relax some of the zero restrictions on the transition matrix in (11) to allow for measurement errors to be correlated across calendar periods. (JvN refer to these as "spillover" effects.)

In the Appendix, we briefly explore the possibilities for identification with some of these generalizations. We now turn to consider the revisions in the available data.

3 Data and Estimation

3.1 Data

We use monthly vintages of quarterly expenditure-based and income-based estimates of GDP from the Bureau of Economic Analysis (BEA) covering the period 2003Q1–2014Q3.

---

[6]Note that we have ignored any free parameters in $\boldsymbol{T}_3$ and $\boldsymbol{T}_4$ in these calculations. We return to this, below. One must also keep in mind that identification by data revision requires that the data are in fact revised. If not, we effectively return to the underidentified case of $l = 1$.

For GDE we employ the Advance, the Third, the 12th and the 24th releases and Second/Third, 12th and the 24th releases for GDI. Due to a lag in source data availability the BEA does not prepare Advance estimates for GDI. The initial estimates for GDI are presented with the Second GDI estimate. Estimates for fourth quarter GDI are presented in the Third estimate only.[7]

3.2   Estimation

We employ Gibbs Sampling methods to obtain posterior simulations for our model's parameters (see, e.g., Kim and Nelson 1999). We use conjugate and diffuse priors for the coefficients and the variance covariance matrix, resulting in a multivariate normal posterior for the coefficients and an inverted Wishart posterior for the variance covariance matrix. For the prior for the coefficients restricted to zero we assume the mean to be zero and variance to be close to zero.

Our Gibbs sampler has the following structure. We first initialize the sampler with values for the coefficients and the variance covariance matrix. Conditional on data, the most recent draw for the coefficients and for the variance covariance matrix, we draw the latent state variables $\boldsymbol{\alpha}_t$ for $t = 1, ..., T$ using the procedure described in Carter and Kohn (1994). In the next step, we condition on data, the most recent draw for the latent variable $\boldsymbol{\alpha}_t$ and for the variance covariance matrix, drawing the coefficients from a multivariate normal distribution. Finally, conditional on data, the most recent draw for the latent variables and the coefficients, we draw the variance covariance matrix from an inverted Wishart distribution. We cycle through 100K Gibbs iterations, discarding the first 90K as burn-in. Of those 10K draws we save only every 10th draw, which gives us in total 1000 draws on which we base our inference. Convergence of the sampler was checked by studying recursive mean plots and by varying the starting values of the sampler and comparing results.

---

[7]See Fixler et al. (2014) for a more detailed discussion of the GDE-GDI vintage history.

4 Results

Here we compare our measure of GDP to releases of GDE and GDI in four different ways: (i) in graphs, (ii) looking at historical decompositions, (iii) by investigating dynamics, and (iv) by calculating relative contributions. To distinguish between the true unknown values of GDP and our model's estimates of these values, we refer to our model's estimates as GDP^{++}.

Figure 3: GDP^{++} vs. GDE
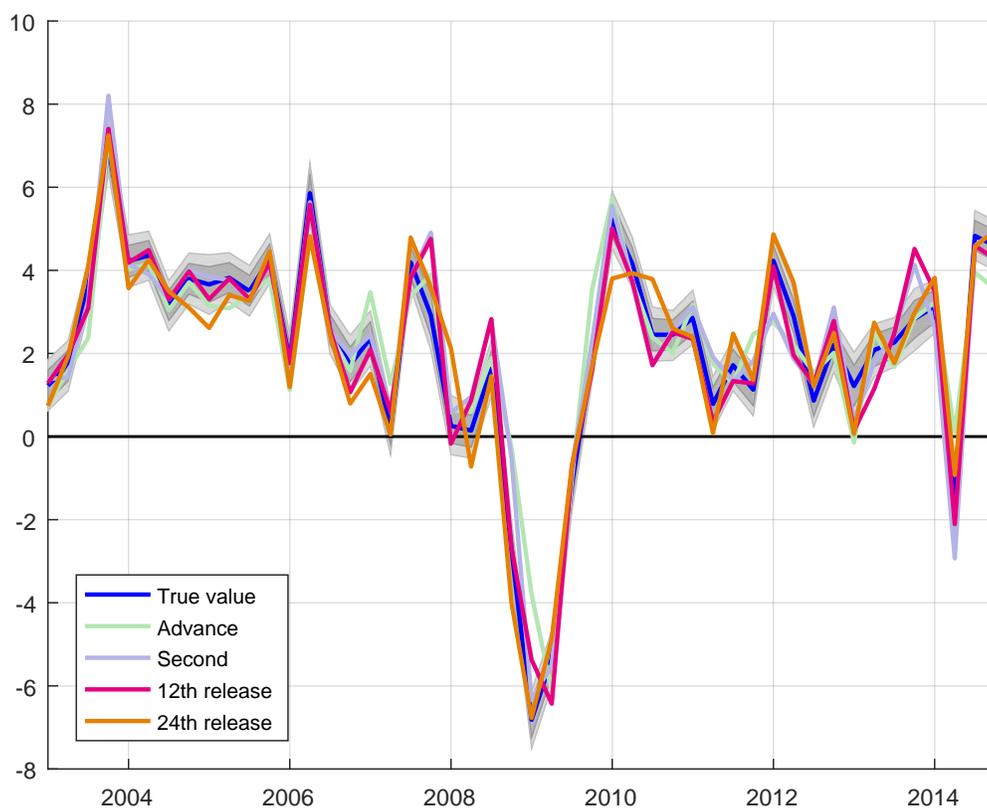
The blue line represents the posterior mean of GDP (the "true" value) and the shaded area around the blue line indicates 90% of posterior probability mass. The green line represents the advance estimate, the purple line is the second estimate, the red line the 12th release and the orange line the 24th release of expenditure side GDP growth.

4.1 Comparison of GDP^{++} and releases of GDE and GDI

In Figure 3 we compare GDP^{++} and its shaded posterior ranges (90% of probability mass) to the four releases of GDE we employed in the estimation, the Advance, third, the 12th and the 24th release. There is some evidence that the releases are more volatile than the true values of GDP. We observe that the releases are outside the posterior bounds for some periods. This observation holds especially for the Advance release and the 24th release; in some periods, like e.g. 2010Q1, the Advance release and the 24th release are on different sides of the posterior range.
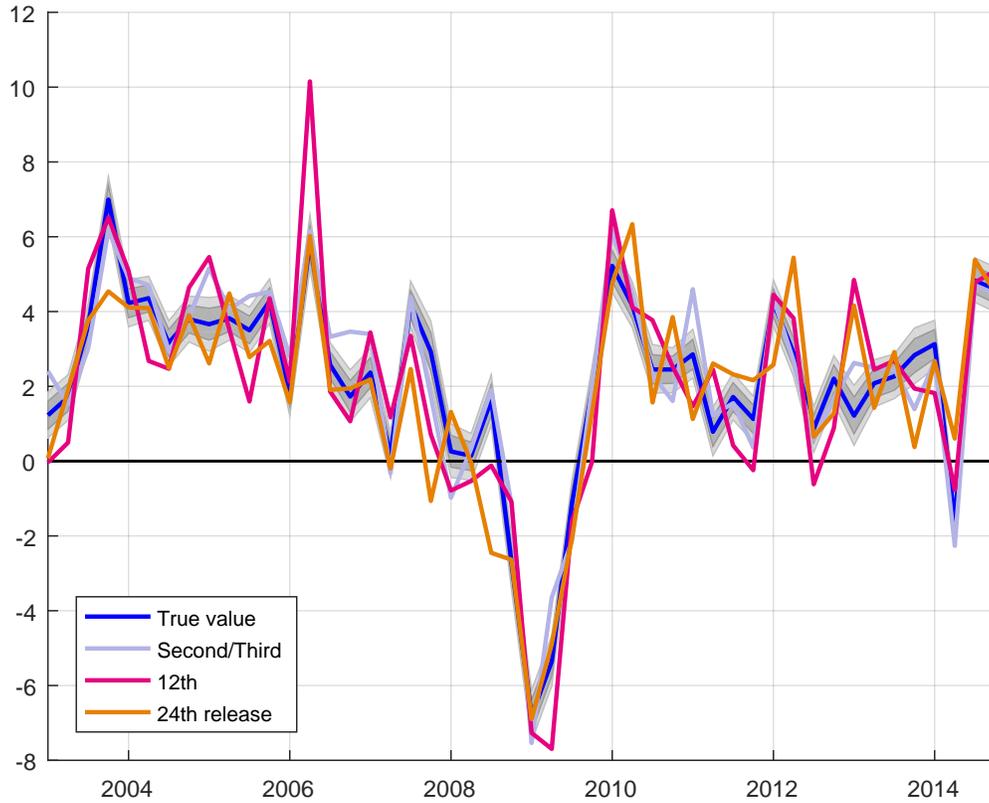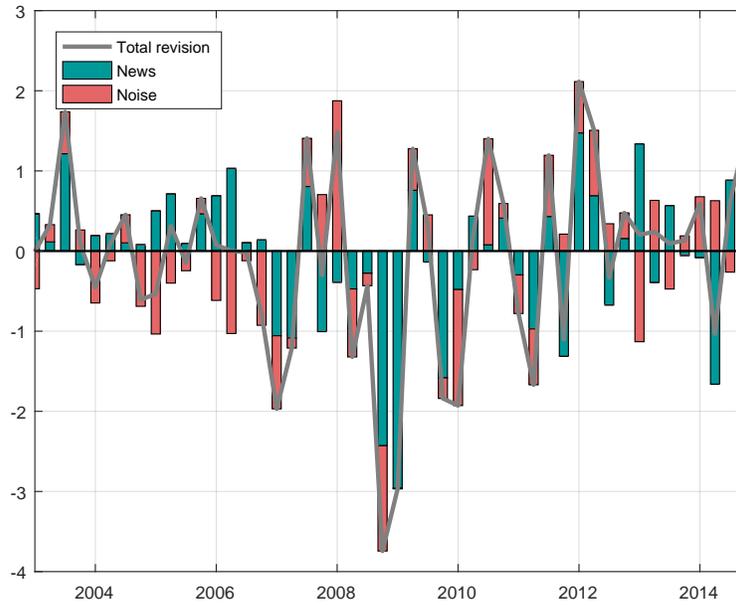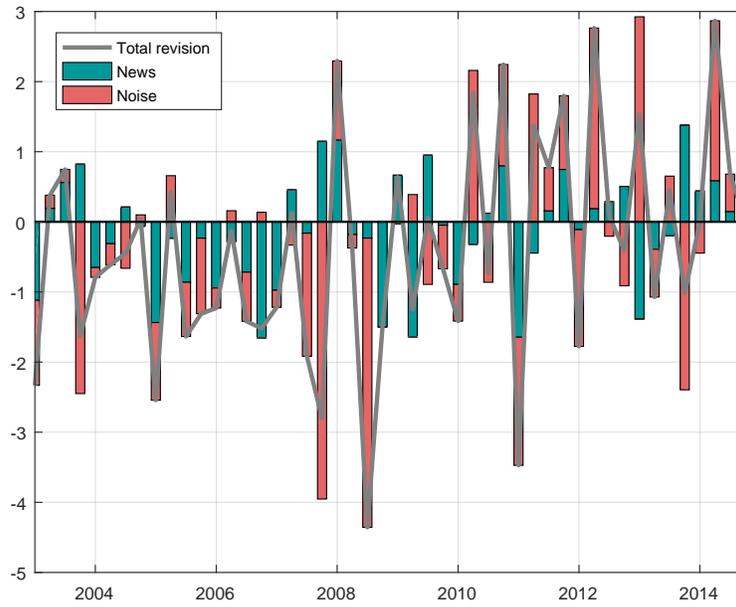
Figure 4: GDP^{++} vs. GDI

The blue line represents the posterior mean of GDP, the "true" value, and the shaded area around the blue line indicates 90% of posterior probability mass. The purple line is the second/third estimate, the red line the 12th release and the orange line the 24th release of income side GDP growth.

Figure 4 shows GDP^{++} together with shaded posterior ranges (90% of probability mass) and the three releases of GDI we employed in the estimation, the Second/Third, the 12th and the 24th release. The releases fluctuate around the posterior bounds of the true values. The GDI releases are more volatile than our estimates GDP^{++}. The releases of GDI are also much more volatile than the releases of GDE. Note that the sample paths of GDP_M and GDE and GDI in Aruoba et al. (2016, Figure 3) show a different picture than our Figures 3 and 4. GDE differs more from their GDP measure than GDI.

4.2 Historical decomposition

Our econometric framework (10-11) allows the historical decomposition of GDE and GDI in terms of news and noise measurement errors. We illustrate the decomposition for GDE.

Suppose, we have l releases of GDE_t

$$GDE_t^1 = \rho GDP_{t-1}^{++} + \eta_{Gt} + \eta_{E\zeta t}^1$$

$$GDE_t^2 = \rho GDP_{t-1}^{++} + \eta_{Gt} + \eta_{E\nu t}^1 + \eta_{E\zeta t}^2$$

$$\vdots \quad = \qquad \vdots$$

$$GDE_t^l = \rho GDP_{t-1}^{++} + \eta_{Gt} + \eta_{E\nu t}^1 + ... + \eta_{E\nu t}^{l-1} + \eta_{E\zeta t}^l.$$

Then the total revision of GDE can be written as

$$GDE_t^l - GDE_t^1 = \underbrace{\eta_{E\nu t}^1 + ... + \eta_{E\nu t}^{l-1}}_{\text{News}} + \underbrace{\eta_{E\zeta t}^l - \eta_{E\zeta t}^1}_{\text{Noise}} \tag{18}$$

where every element on the right-hand side of the equation is part of the state vector whose estimates may be recovered using standard techniques.

Figure 5: Historical Decomposition

GDE

GDI

Historical decomposition of the total revision (24th release minus second estimate) into news and noise.
The red bars depict the share of news and the green bars the share of noise in total revision (grey line).
The historical decomposition is based on the decomposition described in (18).

The outcomes of the historical decompositions are shown in Figure 5. The top panel shows total revisions in GDE with news and noise shares, the bottom panel total GDE revisions with news and noise shares. We observe that total revisions in GDI, the bottom panel, are larger than total revisions in GDE, a stylized fact which can also be distilled from the previous two figures. The two panels suggest that the news share in total GDE revisions is larger than the noise share while the opposite seems to hold for total revisions in GDI. This observation is consistent with Fixler and Nailewaik (2009), who also reject the pure noise assumption in GDI. It also appears that GDI was particularly noisy around the start of 2008 and after 2012.

4.3   Dynamics of GDP^{++} and other GDP measures

In Figure 6 we depict the (ρ, σ^2) pairs summarizing the dynamics of our true GDP estimate across all draws. We contrast the (ρ, σ^2) pairs corresponding to our GDP^{++} estimate to the (ρ, σ^2) pairs obtained when using a news measurement error only or a noise measurement error only version of our model, the benchmark model estimated in Aruoba et al. (2016) and when fitting an AR(1) model to GDE and GDI.

Figure 6 reveals that our real-time data based estimate of GDP is somewhat less persistent than the GDP^+ measure of Aruoba et al. (2016), but exhibits a higher persistence than the estimates for GDE and GDI.[8] We also find that the posterior mean of the innovation variance of our GDP^{++} is much smaller than the innovation variances of GDE, GDI and the benchmark model of Aruoba et al. (2016). The innovation variance of GDP^{++} is also smaller than the innovation variance of the models estimated with news and noise measurement errors only, which in turn are higher than the innovation variance of GDP^+. The combination of a ρ that is close to those implied by the various models estimated in Aruoba et al. (2016) and a σ^2 that is much smaller than the ones implied by Aruoba et al. (2016) leads to a higher forecastability of the GDP^{++} measure.

---

[8]We thank Dongho Song for making his Matlab code available online.

Figure 6: GDP Dynamics
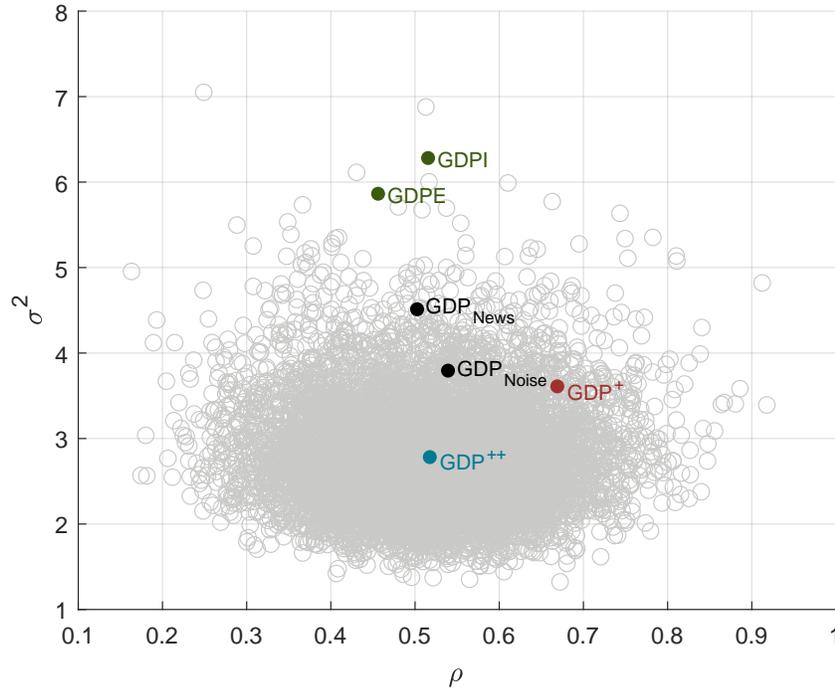
The grey shaded area consists of (ρ, σ^2) pairs across draws from our sampler and the blue dot is the posterior mean of the (ρ, σ^2) pairs across draws. The black dots represent the posterior mean of the (ρ, σ^2) pairs of the news only and noise model, respectively. The red dot is the posterior mean of the (ρ, σ^2) pairs of GDP^+ using the benchmark specification ($\zeta = 0.8$) described in Aruoba et al. (2016). The green dots are (ρ, σ^2) pairs, resulting from AR(1) models fitted to expenditure side and income side GDP growth, respectively. The sampling period for re-estimating the Aruoba et al. (2016) model and for fitting the AR(1) models to the two GDP measures is 2003Q3–2014Q3 (released on October 28, 2016).

4.4 Relative contributions of GDE and GDI to GDP^{++}

To assess the relative importance of GDI and GDE at different releases, we use the Kalman gains. They represent the weight that the estimated value places on estimates of various releases. The outcomes are listed in Table 1.

The results show that the weights assigned to different releases vary greatly as we change the assumed structure of the measurement error. When they are assumed to be pure News, the second panel of the table shows that 98% of the weight is put on the last release of GDE. Once we allow for the possibility of noise errors, however, more weight is assigned to GDI and weights are spread over more releases. The earliest releases of GDE receive

Table 1: Kalman Gains

Weight on	GDE	GDI
News and Noise		
Advance	0.0474	
Second/Third	0.0402	0.2624
12th	0.2879	0.0473
24th Release	0.2824	0.079
News Only		
Advance	0.0073	
Second/Third	0.0000	0.0073
12th	0.0000	0.0000
24th Release	0.9786	0.0142
Noise Only		
Advance	0.1421	
Second/Third	0.2823	0.0688
12th	0.3309	0.0303
24th Release	0.0997	0.0283

less weight than the later releases, while the opposite is true for GDI. In all cases, we also find that GDE releases are more important for explaining GDP than GDI releases, in contrast to Aruoba et al. (2016).

5 Conclusion

We have described a new approach to data reconciliation that exploits multiple data releases on each series. This helps both with the identification of measurement errors and with optimally extracting information from multiple noisy series. We used this to propose a new measure of U.S. GDP growth using real-time data on GDE and GDI. Our measure GDP^{++} is shown to be more persistent than GDE and GDI and has smaller residual variance. In addition it has a similar autoregressive coefficient but smaller residual variance than the GDP measure GDP^+ of Aruoba et al. (2016). Historical decompositions of GDE and GDI measurement errors reveal a larger news share in GDE than in GDI.

Appendix

This appendix first analyzes the identification of the univariate state space system in Jacobs and van Norden (2011), using the procedure described in Komunjer and Ng (2011) and used by Aruoba et al. (2016). Thereafter we discuss the possibilities for identification in more general reconciliation models by comparing the number of free parameters to the number of available moment conditions.

Identification in the univariate, two vintage JvN framework

The state space form of the Jacobs and van Norden (JvN) model with two vintages and no spillovers can be expressed as

$$
\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{y}_t \\ \nu_t^1 \\ \nu_t^2 \\ \zeta_t^1 \\ \zeta_t^2 \end{bmatrix}, \tag{A.1}
$$

$$
\begin{bmatrix} \tilde{y}_t \\ \nu_t^1 \\ \nu_t^2 \\ \zeta_t^1 \\ \zeta_t^2 \end{bmatrix} = \begin{bmatrix} \rho & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \tilde{y}_{t-1} \\ \nu_{t-1}^1 \\ \nu_{t-1}^2 \\ \zeta_{t-1}^1 \\ \zeta_{t-1}^2 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & -1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \eta_{t,\tilde{y}} \\ \eta_{t,\nu}^1 \\ \eta_{t,\nu}^2 \\ \eta_{t,\zeta}^1 \\ \eta_{t,\zeta}^2 \end{bmatrix}, \tag{A.2}
$$

where y_t^i for $i = 1, 2$ denotes the different releases, $\tilde{y}_t$ is the "true" value of the variable of interest, ν_t^i and ζ_t^i for $i = 1, 2$ are the news and the noise components and $\eta_{t,\nu}^i$ and $\eta_{t,\zeta}^i$ for $i = 1, 2$ are the news and the noise shocks, $[\eta_{t,\tilde{y}} \ \eta_{t,\nu}^1 \ \eta_{t,\nu}^2 \ \eta_{t,\zeta}^1 \ \eta_{t,\zeta}^2]' \sim N(\mathbf{0}, \boldsymbol{H})$ with $\boldsymbol{H} = \mathrm{diag}(\sigma_{\tilde{y}}^2, \sigma_{\nu 1}^2, \sigma_{\nu 2}^2, \sigma_{\zeta 1}^2, \sigma_{\zeta 2}^2)$, where diag denotes a diagonal matrix.

The system in (A.1) and (A.2) can also be written as

$$
\begin{bmatrix} y_t^1 \\ y_t^2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tilde{y}_t + \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \omega_{t,\tilde{y}} \\ \omega_{t,\nu}^1 \\ \omega_{t,\nu}^2 \\ \omega_{t,\zeta}^1 \\ \omega_{t,\zeta}^2 \end{bmatrix}, \tag{A.3}
$$

$$
\tilde{y}_t = \rho \tilde{y}_{t-1} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \omega_{t,\tilde{y}} \\ \omega_{t,\nu}^1 \\ \omega_{t,\nu}^2 \\ \omega_{t,\zeta}^1 \\ \omega_{t,\zeta}^2 \end{bmatrix}, \tag{A.4}
$$

where $\omega_{t,\tilde{y}} = \eta_{t,\tilde{y}} + \eta_{t,\nu}^1 + \eta_{t,\nu}^2$, $\omega_{t,\nu}^1 = -\eta_{t,\nu}^1 - \eta_{t,\nu}^2$, $\omega_{t,\nu}^2 = -\eta_{t,\nu}^2$, $\omega_{t,\zeta}^1 = \eta_{t,\zeta}^1$, $\omega_{t,\zeta}^2 = \eta_{t,\zeta}^2$ and $[\omega_{t,\tilde{y}} \ \omega_{t,\nu}^1 \ \omega_{t,\nu}^2 \ \omega_{t,\zeta}^1 \ \omega_{t,\zeta}^2]' \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ with variance-covariance matrix $\boldsymbol{\Sigma}$ defined as

$$
\boldsymbol{\Sigma} = \begin{bmatrix} \Sigma_{\tilde{y}\tilde{y}} & \Sigma_{\tilde{y}\nu 1} & \Sigma_{\tilde{y}\nu 2} & 0 & 0 \\ \Sigma_{\nu 1 \tilde{y}} & \Sigma_{\nu 1 \nu 1} & \Sigma_{\nu 1 \nu 2} & 0 & 0 \\ \Sigma_{\nu 2 \tilde{y}} & \Sigma_{\nu 2 \nu 1} & \Sigma_{\nu 2 \nu 2} & 0 & 0 \\ 0 & 0 & 0 & \Sigma_{\zeta 1 \zeta 1} & 0 \\ 0 & 0 & 0 & 0 & \Sigma_{\zeta 2 \zeta 2} \end{bmatrix}, \tag{A.5}
$$

where

$$
\Sigma_{\tilde{y}\tilde{y}} = \sigma_{\tilde{y}}^2 + \sigma_{\nu 1}^2 + \sigma_{\nu 2}^2, \quad \Sigma_{\tilde{y}\nu 1} = -\sigma_{\nu 1}^2 - \sigma_{\nu 2}^2,
$$
$$
\Sigma_{\tilde{y}\nu 2} = -\sigma_{\nu 2}^2, \quad \Sigma_{\nu 1 \nu 1} = \sigma_{\nu 1}^2 + \sigma_{\nu 2}^2, \quad \Sigma_{\nu 1 \nu 2} = \sigma_{\nu 2}^2, \tag{A.6}
$$
$$
\Sigma_{\nu 2 \nu 2} = \sigma_{\nu 2}^2, \quad \Sigma_{\zeta 1 \zeta 1} = \sigma_{\zeta 1}^2, \quad \Sigma_{\zeta 2 \zeta 2} = \sigma_{\zeta 2}^2,
$$

which implies

$$\Sigma_{\tilde{y}\nu1} = -\Sigma_{\nu1\nu1},$$
$$\Sigma_{\tilde{y}\nu2} = -\Sigma_{\nu2\nu2}, \tag{A.7}$$
$$\Sigma_{\nu2\nu2} = \Sigma_{\nu1\nu2}.$$

Moreover, consider the following restriction

$$\sigma_{\nu2}^2 = 0, \tag{A.8}$$

which is justified due to the fact that the second release news shock corresponds to information outside the sample and is thus not needed.

Aruoba et al. (2016) show that a state space system described in Equations (A.3) and (A.4) is not identified with $\boldsymbol{\Sigma}$ unrestricted and identified under certain restrictions on elements of $\boldsymbol{\Sigma}$. We now investigate whether the restrictions implied by JvN lead to an identified system following the procedure described in Aruoba et al. (2016).

Theorem 1. *Suppose that Assumptions 1, 2, 4-NS and 5-NS of Komunjer and Ng (2011) hold. Then according to Proposition 1-NS of Komunjer and Ng (2011), the state space model described in (A.1) and (A.2) is identified given the restrictions implied by (A.1), (A.2) and (A.8).*

Proof of Theorem 1. We begin by rewriting the state space model in (A.3) and (A.4) to match the notation used in Komunjer and Ng (2011)

$$x_{t+1} = A(\theta)x_t + \boldsymbol{B(\theta)}\boldsymbol{\epsilon}_{t+1} \tag{A.9}$$
$$\boldsymbol{z}_{t+1} = \boldsymbol{C(\theta)}x_t + \boldsymbol{D(\theta)}\boldsymbol{\epsilon}_{t+1}, \tag{A.10}$$

where $x_t = \tilde{y}_t$, $\boldsymbol{z}_t = [y_t^1 \ y_t^2]'$, $\boldsymbol{\epsilon}_t = [\omega_{t,\tilde{y}} \ \omega_{t,\nu}^1 \ \omega_{t,\nu}^2 \ \omega_{t,\zeta}^1 \ \omega_{t,\zeta}^2]'$, $A(\theta) = \rho$, $\boldsymbol{B(\theta)} = [1 \ 0 \ 0 \ 0 \ 0]$, $\boldsymbol{C(\theta)} = [\rho \ \rho]'$,

$$\boldsymbol{D(\theta)} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{bmatrix}$$

and $\boldsymbol{\theta} = [\rho \ \sigma_{\tilde{y}}^2 \ \sigma_{\nu1}^2 \ \sigma_{\nu2}^2 \ \sigma_{\zeta1}^2 \ \sigma_{\zeta2}^2]'$.

Given that $\boldsymbol{\Sigma}$ is positive definite and $0 \leq \rho < 1$, Assumption 1 and 2 of Komunjer and Ng (2011) are satisfied. Given that $\boldsymbol{D}(\boldsymbol{\theta})\boldsymbol{\Sigma}\boldsymbol{D}(\boldsymbol{\theta})'$ is nonsingular also Assumption 4-NS of Komunjer and Ng (2011) is satisfied. Rewriting the state space model in (A.9) and (A.10) into its innovation representation gives

$$\hat{x}_{t+1|t+1} = A(\theta)\hat{x}_{t|t} + \boldsymbol{K}(\boldsymbol{\theta})\boldsymbol{a}_{t+1} \tag{A.11}$$

$$\boldsymbol{z}_{t+1} = \boldsymbol{C}(\boldsymbol{\theta})\hat{x}_{t|t} + \boldsymbol{a}_{t+1}, \tag{A.12}$$

where $\boldsymbol{K}(\boldsymbol{\theta})$ is the Kalman gain and $\boldsymbol{a}_{t+1}$ is the one-step ahead forecast error of z_{t+1} with variance $\boldsymbol{\Sigma}_a(\boldsymbol{\theta})$. The Kalman gain and the variance of the one-step ahead forecast error for this system can be expressed as

$$\boldsymbol{K}(\boldsymbol{\theta}) = (p\rho\boldsymbol{C}' + \boldsymbol{\Sigma_{BD}})(p\boldsymbol{CC}' + \boldsymbol{\Sigma_{DD}})^{-1} \tag{A.13}$$

$$\boldsymbol{\Sigma}_a(\boldsymbol{\theta}) = p\boldsymbol{CC}' + \boldsymbol{\Sigma_{DD}}, \tag{A.14}$$

where p is the variance of the state vector, solving the following Riccati equation

$$p = p\rho^2 + \Sigma_{BB} - (p\rho\boldsymbol{C}' + \boldsymbol{\Sigma_{BD}})(p\boldsymbol{CC}' + \boldsymbol{\Sigma_{DD}})^{-1}(p\rho\boldsymbol{C} + \boldsymbol{\Sigma_{DB}}). \tag{A.15}$$

and $\Sigma_{BB} = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{B}'$, $\boldsymbol{\Sigma_{BD}} = \boldsymbol{B}\boldsymbol{\Sigma}\boldsymbol{D}'$, $\boldsymbol{\Sigma_{DD}} = \boldsymbol{D}\boldsymbol{\Sigma}\boldsymbol{D}'$ with

$$
\begin{aligned}
\Sigma_{BB} &= \Sigma_{\tilde{y}\tilde{y}}, \\
\boldsymbol{\Sigma_{BD}} &= \begin{bmatrix} \Sigma_{\tilde{y}\tilde{y}} + \Sigma_{\tilde{y}\nu 1} & \Sigma_{\tilde{y}\tilde{y}} + \Sigma_{\tilde{y}\nu 2} \end{bmatrix}, \\
\boldsymbol{\Sigma_{DD}} &= \begin{bmatrix} \Sigma_{\tilde{y}\tilde{y}} + 2\Sigma_{\tilde{y}\nu 1} + \Sigma_{\nu 1\nu 1} + \Sigma_{\zeta 1\zeta 1} & \cdot \\ \Sigma_{\tilde{y}\tilde{y}} + \Sigma_{\tilde{y}\nu 1} + \Sigma_{\nu 2\tilde{y}} + \Sigma_{\nu 2\nu 1} & \Sigma_{\tilde{y}\tilde{y}} + 2\Sigma_{\tilde{y}\nu 2} + \Sigma_{\nu 2\nu 2} + \Sigma_{\zeta 2\zeta 2} \end{bmatrix}.
\end{aligned} \tag{A.16}
$$

By using the definitions in (A.6), the expressions in (A.16) can also be written as

$$\Sigma_{BB} = \sigma_{\tilde{y}}^2 + \sigma_{\nu 1}^2 + \sigma_{\nu 2}^2, \qquad \boldsymbol{\Sigma_{BD}} = \begin{bmatrix} \sigma_{\tilde{y}}^2 & \sigma_{\tilde{y}}^2 + \sigma_{\nu 1}^2 \end{bmatrix},$$

$$\boldsymbol{\Sigma_{DD}} = \begin{bmatrix} \sigma_{\tilde{y}}^2 + \sigma_{\zeta 1}^2 & \cdot \\ \sigma_{\tilde{y}}^2 & \sigma_{\tilde{y}}^2 + \sigma_{\nu 1}^2 + \sigma_{\zeta 2}^2 \end{bmatrix}. \tag{A.17}$$

Assumption 5-NS of Komunjer and Ng (2011) relates to the *controllability* and *observability* of state space systems. The state space system in (A.3) and (A.4) is *controllable* if matrix $[\boldsymbol{K(\theta)}\ A(\theta)\boldsymbol{K(\theta)}]$ has full row rank and it is *observable* if the matrix $[\boldsymbol{C(\theta)}'\ A(\theta)'\boldsymbol{C(\theta)}']$ has full column rank and is thus said to be minimal.

To show that Assumption 5-NS is satisfied, first note that $\Sigma_{BB} - \boldsymbol{\Sigma_{BD}\Sigma_{DD}^{-1}\Sigma_{DB}}$ is the Schur complement of $\boldsymbol{\Omega}$, the variance covariance matrix of the joint distribution of x_{t+1} and z_{t+1}, with respect to $\boldsymbol{\Sigma_{DD}}$ where

$$\boldsymbol{\Omega} = \begin{bmatrix} \Sigma_{BB} & \boldsymbol{\Sigma_{BD}} \\ \boldsymbol{\Sigma_{DB}} & \boldsymbol{\Sigma_{DD}} \end{bmatrix}.$$

Because $\boldsymbol{\Omega}$ is a positive definite matrix, its Schur complement is also positive definite thus leading to $\Sigma_{BB} - \boldsymbol{\Sigma_{BD}\Sigma_{DD}^{-1}\Sigma_{DB}} > 0$. Now to show that this inequality leads to $p > 0$, we use the following lemma

Lemma 1. *Assume $\boldsymbol{A}$ and $(\boldsymbol{A} + \boldsymbol{B})$ are invertible and that $rank(\boldsymbol{B}) = 1$, then*

$$(\boldsymbol{A} + \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} - \frac{1}{1 + tr(\boldsymbol{BA}^{-1})}\boldsymbol{A}^{-1}\boldsymbol{BA}^{-1}.$$

We can now use Lemma 1 to rewrite Equation (A.15) as

$$p = p\rho^2 + \Sigma_{BB} - (p\rho\boldsymbol{C}' + \boldsymbol{\Sigma_{BD}})\boldsymbol{\Sigma_{DD}^{-1}}(p\rho\boldsymbol{C} + \boldsymbol{\Sigma_{DB}})$$

$$+ \frac{p}{g}(p\rho\boldsymbol{C}' + \boldsymbol{\Sigma_{BD}})\boldsymbol{\Sigma_{DD}^{-1}}\boldsymbol{CC}'\boldsymbol{\Sigma_{DD}^{-1}}(p\rho\boldsymbol{C} + \boldsymbol{\Sigma_{DB}}), \tag{A.18}$$

where $g = 1 + ptr(\boldsymbol{CC}'\boldsymbol{\Sigma_{DD}^{-1}})$. After some manipulations we find the following quadratic equation

$$ap^2 + bp + c = 0, \tag{A.19}$$

24

with

$$a = -tr(\boldsymbol{CC'}\boldsymbol{\Sigma}_{DD}^{-1}),$$

$$b = (\rho - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB})^2 + tr(\boldsymbol{CC'}\boldsymbol{\Sigma}_{DD}^{-1})(\Sigma_{BB} - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB}) - 1,$$

$$c = \Sigma_{BB} - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB}.$$

The necessary and sufficient conditions for $p > 0$ are $\sqrt{b^2 - 4ac} > 0$ and $\frac{-b - \sqrt{b^2 - 4ac}}{2a} > 0$. The first condition leads to $b^2 + 4tr(\boldsymbol{C'}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{C})(\Sigma_{BB} - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB}) > 0$ and the second to $tr(\boldsymbol{C'}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{C})(\Sigma_{BB} - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB}) > 0$ Since $\boldsymbol{\Sigma}_{DD}$ is positive definite (thus $tr(\boldsymbol{C'}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{C}) > 0$) both conditions are satisfied if $\Sigma_{BB} - \boldsymbol{\Sigma}_{BD}\boldsymbol{\Sigma}_{DD}^{-1}\boldsymbol{\Sigma}_{DB} > 0$.

Given also that $A(\theta) = \rho \geq 0$ and $\boldsymbol{C}(\boldsymbol{\theta}) \geq \mathbf{0}$, we obtain $\boldsymbol{K}(\boldsymbol{\theta}) \neq 0$ and thus Assumption 5-NS is satisfied.

Now Proposition 1-NS of Komunjer and Ng (2011) can be applied, which implies that two vectors

$$\boldsymbol{\theta}_0 = [\rho \ \sigma_{\tilde{y},0}^2 \ \sigma_{\nu 1,0}^2 \ \sigma_{\nu 2,0}^2 \ \sigma_{\zeta 1,0}^2 \ \sigma_{\zeta 2,0}^2]'$$

and

$$\boldsymbol{\theta}_1 = [\rho \ \sigma_{\tilde{y},1}^2 \ \sigma_{\nu 1,1}^2 \ \sigma_{\nu 2,1}^2 \ \sigma_{\zeta 1,1}^2 \ \sigma_{\zeta 2,1}^2]'$$

are observationally equivalent iff there exists a scalar $\tau \neq 0$ such that

$$A(\theta_1) = \tau A(\theta_0)\tau^{-1} \tag{A.20}$$

$$\boldsymbol{K}(\boldsymbol{\theta_1}) = \tau \boldsymbol{K}(\boldsymbol{\theta_0}) \tag{A.21}$$

$$\boldsymbol{C}(\boldsymbol{\theta_1}) = \boldsymbol{C}(\boldsymbol{\theta_0})\tau^{-1} \tag{A.22}$$

$$\boldsymbol{\Sigma}_a(\boldsymbol{\theta_1}) = \boldsymbol{\Sigma}_a(\boldsymbol{\theta_0}). \tag{A.23}$$

Given that $A(\theta) = \rho$, it follows from Equation (A.20) that $\rho_0 = \rho_1$ and thus we can deduce from (A.22) that $\gamma = 1$. Hence, by using Equations (A.13) and (A.14), the conditions

(A.21) and (A.23) can be expressed as

$$\boldsymbol{K}_1 = \boldsymbol{K}_0 = (p_0 \rho \boldsymbol{C}' + \boldsymbol{\Sigma_{BD0}})(p \boldsymbol{C} \boldsymbol{C}' + \boldsymbol{\Sigma_{DD0}})^{-1} \tag{A.24}$$

$$\boldsymbol{\Sigma_{a1}} = \boldsymbol{\Sigma_{a0}} = p_0 \boldsymbol{C} \boldsymbol{C}' + \boldsymbol{\Sigma_{DD0}}, \tag{A.25}$$

where p_0 solves the following Riccati equation

$$p_0 = p_0 \rho^2 + \Sigma_{BB0} - \boldsymbol{K_0}(p_0 \rho \boldsymbol{C} + \boldsymbol{\Sigma_{DB0}}). \tag{A.26}$$

Equations (A.24) to (A.26) are satisfied if and only if

$$p_1(1 - \rho^2) - \Sigma_{BB1} = p_0(1 - \rho^2) - \Sigma_{BB0} \tag{A.27}$$

$$p_1 \rho \boldsymbol{C}' + \boldsymbol{\Sigma_{BD1}} = p_0 \rho \boldsymbol{C}' + \boldsymbol{\Sigma_{BD0}} \tag{A.28}$$

$$p_1 \boldsymbol{C} \boldsymbol{C}' + \boldsymbol{\Sigma_{DD1}} = p_0 \boldsymbol{C} \boldsymbol{C}' + \boldsymbol{\Sigma_{DD0}}. \tag{A.29}$$

Without loss of generality let

$$\Sigma_{\tilde{y}\tilde{y},1} = \Sigma_{\tilde{y}\tilde{y},0} + \delta(1 - \rho^2) \tag{A.30}$$

leading to

$$\Sigma_{BB,1} = \Sigma_{BB,0} + \delta(1 - \rho^2). \tag{A.31}$$

We now proceed by splitting the analysis into two cases.

Case 1: $\delta = 0$. From (A.27) we obtain $p_1 = p_0$. (A.28) hence implies $\sigma_{\tilde{y},1}^2 = \sigma_{\tilde{y},0}^2$ and $\sigma_{\nu 1,1}^2 = \sigma_{\nu 1,0}^2$ and given that $\Sigma_{\tilde{y}\tilde{y},1} = \Sigma_{\tilde{y}\tilde{y},0}$ it follows $\sigma_{\nu 2,1}^2 = \sigma_{\nu 2,0}^2$. (A.29) implies that $\boldsymbol{\Sigma_{DD1}} = \boldsymbol{\Sigma_{DD0}}$ and thus $\sigma_{\zeta 1,1}^2 = \sigma_{\zeta 1,0}^2$ and $\sigma_{\zeta 2,1}^2 = \sigma_{\zeta 2,0}^2$, leading to the fact that $\theta_1 = \theta_0$.

Case 2: $\delta \neq 0$. From (A.27) we obtain $p_1 = p_0 + \delta$. From (A.28) it follows

$$\sigma_{\tilde{y},1}^2 = \sigma_{\tilde{y},0}^2 - \delta\rho^2 \quad \text{and} \quad \sigma_{\nu 1,1}^2 = \sigma_{\nu 1,0}^2. \tag{A.32}$$

Moreover, (A.27) gives

$$\sigma_{\nu 2,1}^2 = \sigma_{\nu 2,0}^2 + \delta. \tag{A.33}$$

26

Finally, the equations in (A.29) lead to

$$\sigma^2_{\zeta 1,1} = \sigma^2_{\zeta 1,0} \quad \text{and} \quad \sigma^2_{\zeta 2,1} = \sigma^2_{\zeta 2,0}. \tag{A.34}$$

Note that (A.6) and (A.32) to (A.34) result into

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} \Sigma_{\tilde{y}\tilde{y},0} + \delta(1-\rho^2) & \Sigma_{\tilde{y}\nu 1,0} - \delta & \Sigma_{\tilde{y}\nu 2,0} - \delta & 0 & 0 \\ \Sigma_{\nu 1\tilde{y},0} - \delta & \Sigma_{\nu 1\nu 1,0} + \delta & \Sigma_{\nu 1\nu 2,0} + \delta & 0 & 0 \\ \Sigma_{\nu 2\tilde{y},0} - \delta & \Sigma_{\nu 2\nu 1,0} + \delta & \Sigma_{\nu 2\nu 2,0} + \delta & 0 & 0 \\ 0 & 0 & 0 & \Sigma_{\zeta 1\zeta 1} & 0 \\ 0 & 0 & 0 & 0 & \Sigma_{\zeta 2\zeta 2} \end{bmatrix}. \tag{A.35}$$

Finally, from (A.33) and (A.34) it follows that $\delta = 0$. $\qquad\square$

Identification in generalized reconciliation models

We now return to the use of moment conditions to discuss possible approaches to identification in data reconciliation models with multiple data releases. We only consider the reconciliation of exactly two data series, but otherwise consider linear dynamic data generating processes more general than any that we have seen used in the literature.

Specifically, we consider models of the form

$$\begin{bmatrix} \boldsymbol{Y}_t^0 \\ \boldsymbol{Y}_t^1 \end{bmatrix} = \boldsymbol{Z} \cdot \boldsymbol{\alpha}_t \tag{A.36}$$

and

$$\boldsymbol{\alpha}_t = \boldsymbol{T} \cdot \boldsymbol{\alpha}_{t-1} + \boldsymbol{R} \cdot \boldsymbol{\eta}_t \tag{A.37}$$

where

- $\boldsymbol{Y}_t^i$ is a $1 \times l$ vector containing l releases of series $i = \{0, 1\}$ estimates of the unobserved true value $\widetilde{y}_t$.

- $\boldsymbol{\alpha}_t$ is a $(p + 4 \cdot l) \times 1$ latent state vector that we may partition as $\boldsymbol{\alpha}_t' = \begin{bmatrix} \widetilde{\boldsymbol{Y}}_{t-1}^{p\prime} & \boldsymbol{\nu}_t^{0\prime} & \boldsymbol{\nu}_t^{1\prime} & \boldsymbol{\zeta}_t^{0\prime} & \boldsymbol{\zeta}_t^{1\prime} \end{bmatrix}$

- $\widetilde{\boldsymbol{Y}}_{t-1}^{p}$ is a $p \times 1$ vector $\equiv [\widetilde{y}_t, \ldots, \widetilde{y}_{t-p+1}]'$

- $\boldsymbol{\nu}_t^i$ for $i = \{0, 1\}$ is a $l \times 1$ vector of news shocks contained in $\boldsymbol{Y}_t^i$

- $\boldsymbol{\zeta}_t^i$ for $i = \{0, 1\}$ is a $l \times 1$ vector of noise shocks contained in $\boldsymbol{Y}_t^i$

- $\boldsymbol{\eta}_t \sim N(\boldsymbol{0}, \boldsymbol{\Sigma})$ is a $(1 + 4 \cdot l) \times 1$ vector of i.i.d. mean zero normally distributed shocks.with diagonal covariance matrix $\boldsymbol{\Sigma}$.

- $\boldsymbol{Z}$ is a $(2 \cdot l) \times (p + 4 \cdot l)$ matrix of the form

$$Z \equiv \begin{bmatrix} \boldsymbol{1}_{(2 \cdot l) \times 1} & \boldsymbol{0}_{(2 \cdot l) \times (p-1)} & \boldsymbol{I}_{2 \cdot l} & \boldsymbol{I}_{2 \cdot l} \end{bmatrix}$$

- $\boldsymbol{1}_{a \times b}$ is a matrix of dimension $a \times b$ composed entirely of 1's

- $\boldsymbol{0}_{a \times b}$ is a matrix of dimension $a \times b$ composed entirely of 0's

- $\boldsymbol{I}_a$ is a $a \times a$ identity matrix

- $\boldsymbol{T}$ is a $(p + 4 \cdot l) \times (p + 4 \cdot l)$ block diagonal matrix of the form

$$\boldsymbol{T} \equiv \begin{bmatrix} \boldsymbol{T}_p & \boldsymbol{0}_{p \times (4 \cdot l)} \\ \boldsymbol{0}_{(4 \cdot l) \times p} & \boldsymbol{T}_S \end{bmatrix} \tag{A.38}$$

- $\boldsymbol{T}_p$ is a $p \times p$ matrix

$$\boldsymbol{T}_p = \begin{bmatrix} \rho_1, \rho_2, \cdots & \rho_p \\ \boldsymbol{I}_{p-1} & \boldsymbol{0}_{(p-1) \times 1} \end{bmatrix} \tag{A.39}$$

- $\boldsymbol{T}_S$ is a $(4 \cdot l) \times (4 \cdot l)$ arbitrary diagonal matrix

- $\boldsymbol{R}$ is a $(p + 4 \cdot l) \times (1 + 4 \cdot l)$ matrix of the form

$$\boldsymbol{R} = \begin{bmatrix} 1 & \mathbf{1}_{1 \times l} & \mathbf{1}_{1 \times l} & \mathbf{0}_{1 \times l} & \mathbf{0}_{1 \times l} \\ \mathbf{0}_{(p-1) \times 1} & \mathbf{0}_{(p-1) \times l} & \mathbf{0}_{(p-1) \times l} & \mathbf{0}_{(p-1) \times l} & \mathbf{0}_{(p-1) \times l} \\ \mathbf{0}_{l \times 1} & -\boldsymbol{U}_l & \boldsymbol{\Psi} & \mathbf{0}_{l \times l} & \mathbf{0}_{l \times l} \\ \mathbf{0}_{l \times 1} & \mathbf{0}_{l \times l} & -\boldsymbol{U}_l & \mathbf{0}_{l \times l} & \mathbf{0}_{l \times l} \\ \mathbf{0}_{l \times 1} & \mathbf{0}_{l \times l} & \mathbf{0}_{l \times l} & \boldsymbol{I}_{l \times l} & \boldsymbol{\Phi} \\ \mathbf{0}_{l \times 1} & \mathbf{0}_{l \times l} & \mathbf{0}_{l \times l} & \mathbf{0}_{l \times l} & \boldsymbol{I}_{l \times l} \end{bmatrix} \tag{A.40}$$

- $\boldsymbol{U}_l$ is a $l \times l$ matrix with 0's below the main diagonal and 1's everywhere else

- $\boldsymbol{\Psi}, \boldsymbol{\Phi}$ are unrestricted $l \times l$ matrices.

The model estimated in the paper is the special case of the above where

1. $p = 1$

2. $\boldsymbol{T}_S = \mathbf{0}_{(4 \cdot l) \times (4 \cdot l)}$

3. $\boldsymbol{\Phi} = \boldsymbol{\Psi} = \mathbf{0}_{l \times l}$

Relaxing the first condition allows us to consider model where the dynamics of the unobserved true values follow an $AR(p)$ process rather than simply an $AR(1)$. Allowing for higher-order autocorrelations adds an additional $p - 1$ free parameters to the model, but also adds an additional $2 \cdot (p-1)$ sample autocorrelations that may be used for identification.

Relaxing the second condition allows what JvN refer to as "spillover" effects. This permits revisions to the values for calendar period t to be correlated with revisions to calendar period $t - 1$. This may occur, for example, when revisions tend to shift measured growth from one quarter to an adjascent quarter, or when the incorporation of lower frequency data sources (e.g. annual tax returns) shift multiple periods in the same direction. This adds an additional $4 \cdot l$ free parameters to the model. However, it also brings into play an additional $2 \cdot (l - 1)$ moments capturing the 1st-order autocorrelations of the revisions of our two series.

Relaxing the third condition allows for the possibility that measurement errors of either type may be correlated across the two series. Contemporaneous correlations (i.e. measurement errors that affect the same *release* of each series) are captured by the diagonals of these two matrices. Evidence that information tends to be incorporated into releases of y^0 before (after) those of y^1 implies that there should be non-zero entries of $\boldsymbol{\Psi}$ above (below) the main diagonal. Contemporaneous correlations would add an additional $2 \cdot l$ free parameters to the model, while unrestricted correlations would add an additional $2 \cdot l^2$ free parameters. However, we have already assumed the use of all $l \cdot (2 \cdot l + 1)$ contemporaneous cross-moments of the various vintages of both series, so there is no offsetting gain in the number of moments available for identification.

References

Almuzara, Tincho, Gabriele Fiorentini and Enrique Sentana (2018). *U.S. aggregate output measurement: A common trend approach*. manuscript, January 2018.

Aruoba, S. Borağan et al. (2012). 'Improving GDP Measurement: A Forecast Combination Perspective'. In: *Causality, Prediction, and Specification Analysis: Recent Advances and Future Directions: Essays in Honor of Halbert L. White Jr.* Ed. by X. Chen and N. Swanson. New York: Springer, pp. 1–26.

— (2016). 'Improving GDP Measurement: A Measurement-Error Perspective'. In: *Journal of Econometrics* 191, pp. 384–397.

Carter, C. and R. Kohn (1994). 'On Gibbs sampling for state space models'. In: *Biometrika* 81, pp. 541–553.

Chang, Andrew C. and Phillip Li (2015). *Measurement Error in Macroeconomic Data and Economics Research: Data Revisions, Gross Domestic Product, and Gross Domestic Income*. Federal Reserve Board Finance and Economics Discussion Series Working Paper 2015-102.

Diebold, Francis X. (2010). 'Comment'. In: *Brookings Papers on Economic Activity*, pp. 107–112.

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

Fixler, Dennis J., Ryan Greenaway-McGrevy and Bruce T. Grimm (2014). *The Revisions to GDP , GDI , and Their Major Components*. U.S. Bureau of Economic Analysis, Department of Commerce and Federal Reserve Board Manuscript.

Fixler, Dennis J. and Jeremy J. Nalewaik (2009). *News, Noise, and Estimates of the "True" Unobserved State of the Economy*. U.S. Bureau of Economic Analysis, Department of Commerce and Federal Reserve Board Manuscript.

Jacobs, Jan P.A.M and Simon van Norden (2011). 'Modeling Data Revisions: Measurement Error and Dynamics of "True" Values'. In: *Journal of Econometrics* 161, pp. 101–109.

Kim, Chang-Jin and Charles R. Nelson (1999). *State-Space Models with Regime Switching*. Cambridge MA and London: The MIT Press.

Kishor, N.K. and E.F. Koenig (2012). 'VAR Estimation and Forecasting When Data Are Subject to Revision'. In: *Journal of Business & Economic Statistics* 30, pp. 181–190.

Komunjer, Ivana and Serena Ng (2011). 'Dynamic Identification of Dynamic Stochastic General Equilibrium Models'. In: *Econometrica* 79, pp. 1995–2031.

Nalewaik, Jeremy J. (2012). 'Estimating probabilities of recession in real time using GDP and GDI'. In: *Journal of Money, Credit and Banking* 44, pp. 235–253.

Rees, Daniel M., David Lancaster and Richard Finlay (2015). 'A state-space approach to Australian Gross Domestic Product'. In: *Australian Economic Review* 48, pp. 133–149.

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

Weale, Martin (1992). 'Estimation of Data Measured with Error and Subject to Linear Restrictions'. In: *Journal of Applied Econometrics*, pp. 167–174.