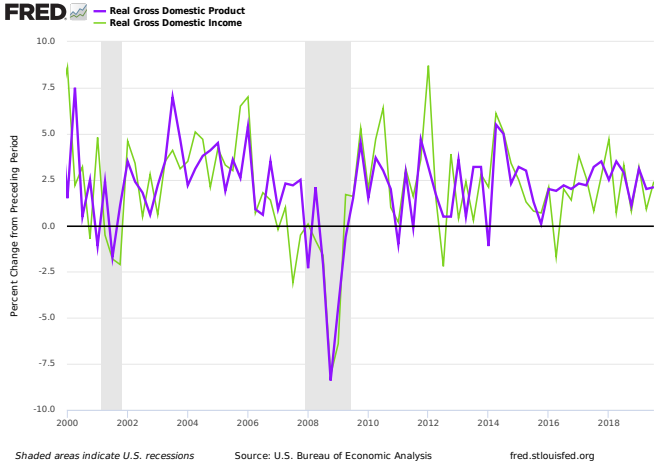
# Can GDP measurement be further improved?
# Data revision and reconciliation

Jan P.A.M. Jacobs, Samad Sarferaz, Jan-Egbert Sturm and Simon van Norden

Joint Montreal Macro BrownBag
CIRANO, November 2019
(paper @ *www.svannorden.org)*

# Motivation I




U.S. Real GDP growth: Expenditure vs Income measures (Nov. 27, 2019)

# Motivation II

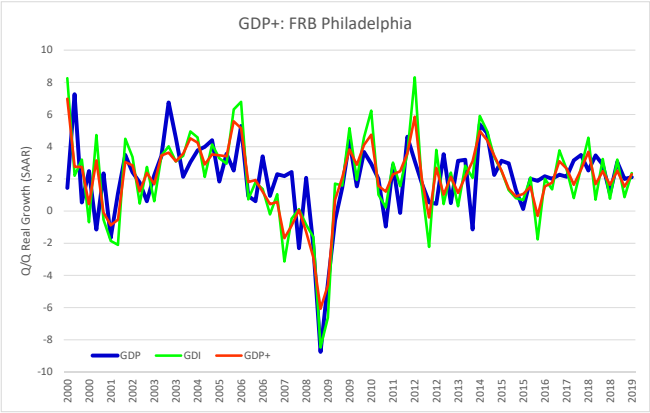
Which is the better measure of GDP?
Expenditure (GDE) or Income (GDI)?

- Nalewaik (2012)
- Chang and Li (2015)

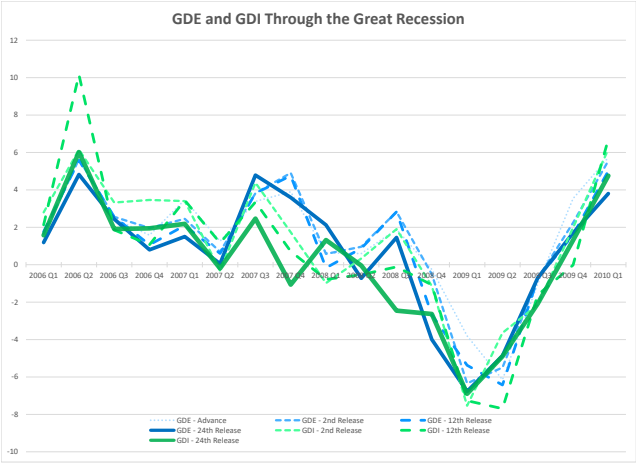
Reconciliation:

- Stone, Champernowne and Meade (1942)
- Weale (1992)
- Diebold (2010)
- ADNSS (2013) & (2016)
  - FRB Philadephia publishes GDP$^+$
- Anesti et al. (2018) (UK: **U**ncertain **K**ingdom)

# Motivation III

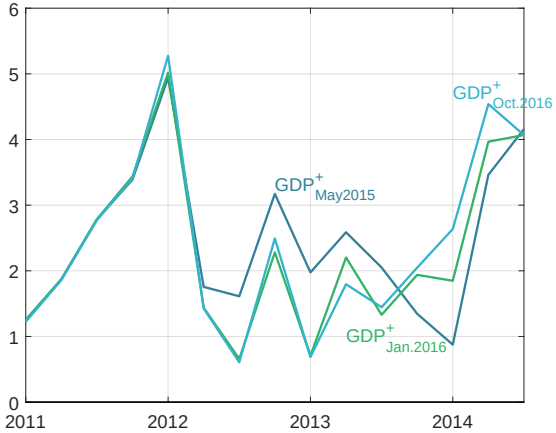


# Motivation IV




Both series have important revisions

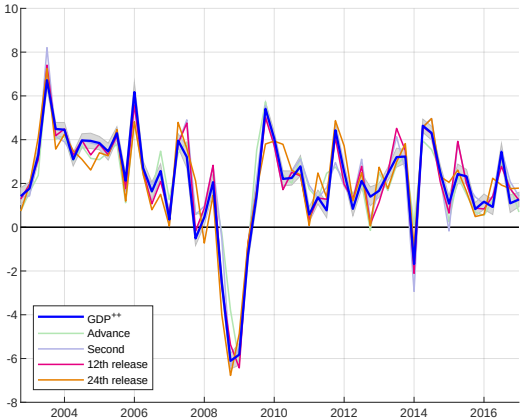
# Motivation V




...and so does $GDP^{+}$!

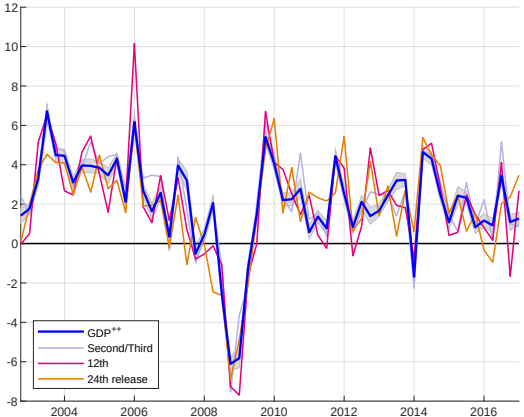
# Our Contribution

1. We model the reconciliation problem in a standard state-space framework.
2. We show how to allow for
   - multiple data releases (cf Jacobs and van Norden JoE 2011)
   - varying precision
   - series dynamics
   - news and noise errors, possibly correlated across the two series
3. We show it solves ADNSS (2013) identification problem
4. Compare our new measure ($GDP^{++}$) to real $GDE$ and $GDI$ growth and provide confidence intervals for the new estimates.
5. Decompose initial estimates of $GDE$ and $GDI$ growth into news and noise shocks




$GDP^{++}$ vs $GDE$

# $GDP^{++}$ II




$GDP^{++}$ vs $GDI$

# Problem

Reconciliation relies on assumptions about the errors in the series being reconciled.

- which is more precise?
- lead/lag relationships?
- News or Noise?

These relationships vary depending on which release(s) we consider.

- Important for producing efficient estimates.
- Important for understanding reliability of estimates.

# Revision properties

## News and Noise

Let $y_t^i$ be the $i$-th release of $y$ in period $t$ and $\tilde{y}_t \equiv$ 'true' value of $y_t$

1. **Noise:**

$$y_t^i = \tilde{y}_t + \zeta_t^i, \qquad \text{cov}(\tilde{y}_t, \zeta_t^i) = 0 \quad \forall i$$

   $\Rightarrow$ revisions (partly) forecastable
   $\Rightarrow$ vintages **more** volatile than 'true' values

2. **News:**

$$\tilde{y}_t = y_t^i + \nu_t^i, \qquad \text{cov}(y_t^i, \nu_t^i) = 0 \quad \forall i$$

   Linked to rational forecasts (De Jong 1987)
   rational statistical agency (Sargent 1989)
   $\Rightarrow$ revisions *cannot* be forecast
   $\Rightarrow$ vintages **less** volatile than "true" values

# Notation

**Data** (vectors)

| | |
|---|---|
| $GDE_t$ | real GDP growth (Expenditure measure) |
| $GDI_t$ | real GDP growth (Income measure) |

**Some scalars**

| | |
|---|---|
| $GDE_t^i$,$GDI_t^i$ | superscript $^i$ indicates release $(1, ..., L)$ |
| $GDP_t$ | real GDP growth ("Truth" - unobserved) |
| $GDP^{50/50}$ | BEA: Average of $(GDE_t^i, GDI_t^i)$ |
| $GDP_t^+$ | real GDP growth - FRB Philadelphia measure |
| $GDP_t^{++}$ | our real GDP growth measure |

**Errors** (vectors)

| | |
|---|---|
| $\nu_t$ | News measurement error |
| $\zeta_t$ | Noise measurement error |

# State-Space Model I

**Measurement Equation:**

$$\begin{bmatrix} GDE_t^L \\ GDI_t^L \end{bmatrix} = GDP_t + \begin{bmatrix} \nu_{Et}^L \\ \nu_{It}^L \end{bmatrix} + \begin{bmatrix} \zeta_{Et}^L \\ \zeta_{It}^L \end{bmatrix} \tag{1}$$

where

$$GDE_t^L = [GDE_t^1, \ldots, GDE_t^l]', \qquad GDI_t^L = [GDI_t^1, \ldots, GDI_t^l]',$$
$$\nu_{Et}^L = [\nu_{Et}^1, \ldots, \nu_{Et}^l]', \qquad \nu_{It}^L = [\nu_{It}^1, \ldots ..., \nu_{It}^l]'$$
$$\zeta_{Et}^L = [\zeta_{Et}^1, ..., \zeta_{Et}^l]', \qquad \zeta_{It}^L = [\zeta_{It}^1, ..., \zeta_{It}^l]',$$

**Noise:** $E[\zeta_{Et}^L | GDP_t] = 0 = E[\zeta_{It}^L | GDP_t]$
**News:** $E[\nu_{Et}^j | GDE_t^k] = 0 = E[\nu_{It}^j | GDI_t^k] \; \forall j > k$

# State-Space Model II

**Transition Equation:**

Let $\boldsymbol{\alpha}_t = [GDP_t, \nu_{Et}^{L\prime}, \nu_{It}^{L\prime}, \zeta_{Et}^{L\prime}, \zeta_{It}^{L\prime}]'$

The transition equation may be compactly written as

$$\boldsymbol{\alpha}_t = \begin{bmatrix} \rho & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{\alpha}_{t-1} + \boldsymbol{R} \cdot \boldsymbol{\eta}_t, \tag{2}$$

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{R}_1 & \boldsymbol{R}_2 + \boldsymbol{R}_3 & 0 & 0 \\ -\boldsymbol{V}_l \cdot diag(\boldsymbol{R}_1) & -\boldsymbol{V}_l \cdot diag(\boldsymbol{R}_3) & 0 & 0 \\ 0 & -\boldsymbol{V}_l \cdot diag(\boldsymbol{R}_2) & 0 & 0 \\ 0 & 0 & \boldsymbol{R}_4 & \boldsymbol{R}_6 \\ 0 & 0 & 0 & \boldsymbol{R}_5 \end{bmatrix} \tag{3}$$

$\boldsymbol{V}_l$ is upper triangular matrix of ones,

$\boldsymbol{R}_1, ..., \boldsymbol{R}_3$ are $1 \times l$ vectors and $\boldsymbol{R}_4, ..., \boldsymbol{R}_6$ are $l \times l$ diagonal matrices.

$\boldsymbol{\eta}_t = [\eta_{E\nu t}^{i\prime}, \eta_{I\nu t}^{i\prime}, \eta_{E\zeta t}^{i\prime}, \eta_{I\zeta t}^{i\prime}]' \sim N(0, \boldsymbol{I}_{4l})$

# Identification

How can we hope to distinguish News and Noise measurement errors?

(2) implies that all persistence comes through $GDP_t$.

- News shocks are part of $GDP_t$, and so have a persistent effect. Also, they do not decay over subsequent releases.
- Noise shocks are have no persistence across time or releases.

2 series with $l$ releases each give us $2l \cdot (2l+1)/2$ observable cross moments and $2l$ first-order autocorrelations $= l \cdot (2l+3)$ moments to identify $1+6l$ parameters.

- $\rho$ and $\boldsymbol{R}_1, ..., \boldsymbol{R}_6$ with $l$ elements each.
- For $l=1$, we have 5 moments for 7 unknown parameters.
- For $l=2$, we have 14 moments for 13 unknown parameters.
- For $l=3$, we have 27 moments for 19 unknown parameters.

We give a formal proof of identification based on Komunjer and Ng (2011).

# Data and estimation

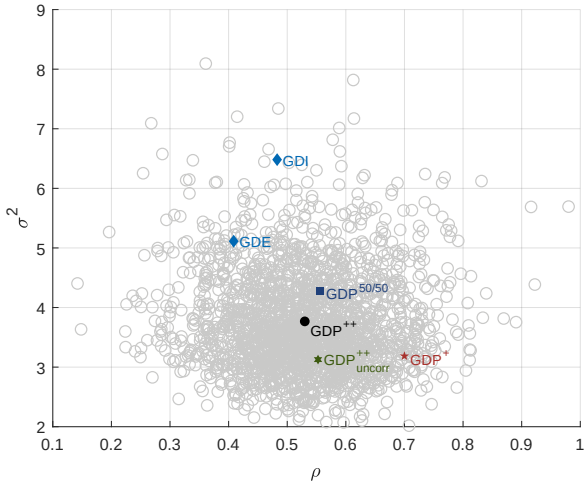
## Data

- Monthly vintages of quarterly series 2003Q1–2014Q3
  from Bureau of Economic Analysis (BEA)
- For real $GDE$ growth
  we use the advance, third, the 12th and the 24th releases
- For real $GDI$ growth
  we use the second/third, the 12th and the 24th releases

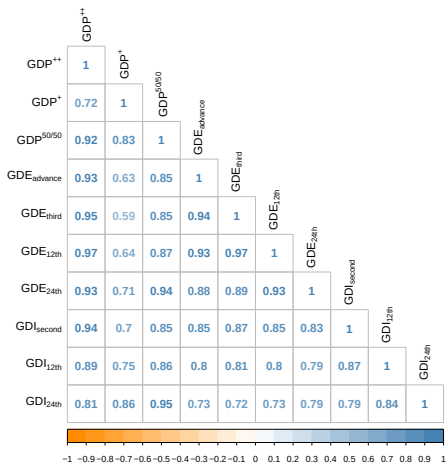
## Estimation

- Gibbs Sampling with diffuse priors
- Estimate with errors correlated or uncorrelated across the two series.
- Compare results to AR(1) for GDE, GDI, GDP$^{50/50}$ and GDP$^{+}$

# Real $GDP$ growth dynamics I



# Real $GDP$ growth dynamics II

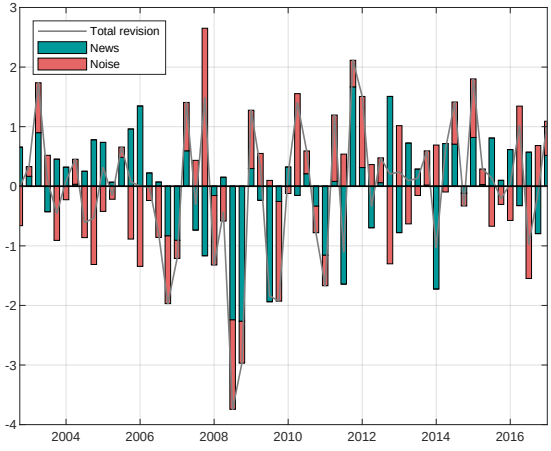



Correlations

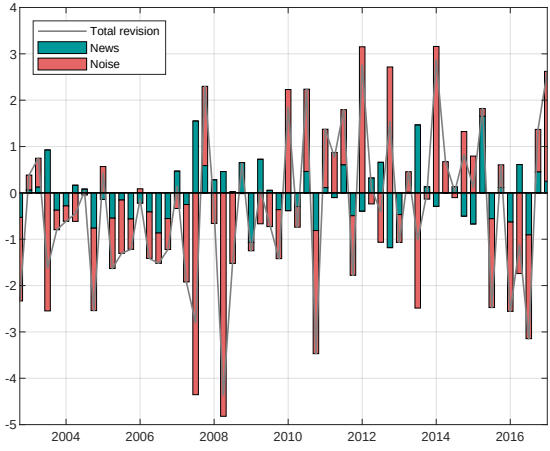
# $GDP_t^{++}$ Loadings (Kalman Gains)

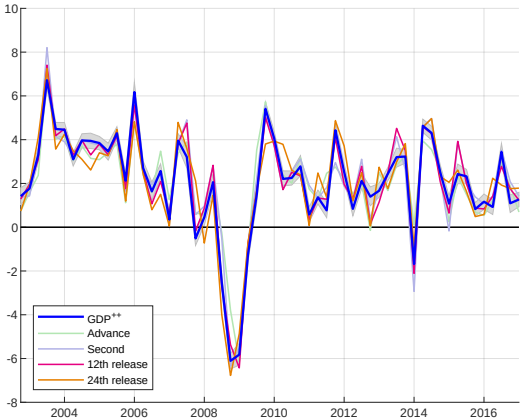
| | Balanced Sample | | Ragged-Edge Sample | |
|---|---|---|---|---|
| Weight on | $GDE$ | $GDI$ | $GDE$ | $GDI$ |
| **News and Noise** | | | | |
| Advance | 0.0272 | | 0.2311 | |
| Second/Third | -0.2103 | 0.3067 | 0.3363 | 0.4804 |
| 12th | 0.7104 | 0.1081 | 0 | 0 |
| 24th Release | 0.0479 | 0.0125 | 0 | 0 |
| **Uncorrelated News and Noise** | | | | |
| Advance | 0.0380 | | 0.1363 | |
| Second/Third | 0.1240 | 0.1672 | 0.4934 | 0.3768 |
| 12th | 0.2318 | 0.0796 | 0 | 0 |
| 24th Release | 0.2799 | 0.0826 | 0 | 0 |

# Historical decomposition of real $GDE$ growth



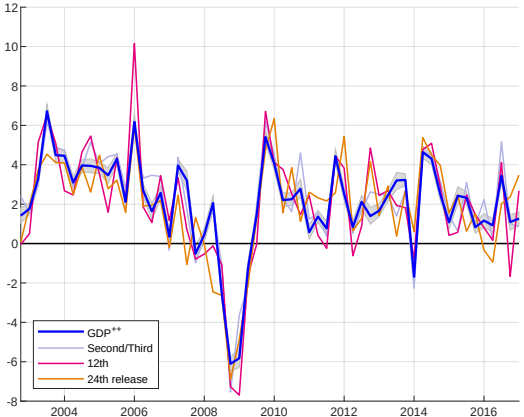
# Historical decomposition of real $GDI$ growth






$GDP^{++}$ vs $GDE$

# $GDP^{++}$ II




$GDP^{++}$ vs $GDI$

# Conclusion

We show how to reconcile series subject to revision due to news and noise.

- Identification possible due to differing impact of news and noise errors across data vintages, and persistence in "true" $GDP$

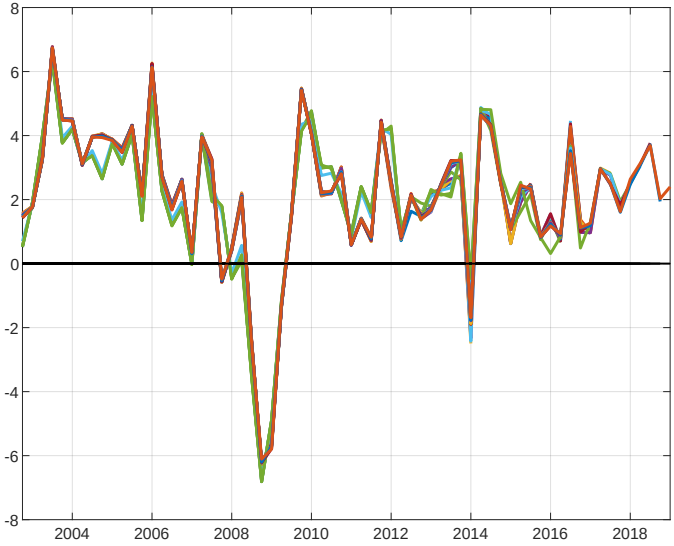
We provided a new real $GDP$ growth measure using real-time data

- More persistent and smaller residual variance than real $GDE$ growth and real $GDI$ growth
- Similar AR-coefficient but smaller residual variance than $GDP^{+}$

Computed historical decomposition of real $GDE$ growth and real $GDP$ growth measurement errors

- Higher news share in real $GDE$ growth than in real $GDI$ growth
- 2008 downturn in $GDI$ seems like noise rather than a leading indicator of recession.

# $GDP^{++}$ in Real-Time



# Revisions in $GDP^{++}$, $GDP^{+}$ and $GDP^{50/50}$