Improving the Accuracy of Economic Measurement with Multiple Data Sources: The Case of Payroll Employment Data

CRIW Pre-Conference — July 18, 2018

Tomaz Cajner – Leland Crane – Ryan Decker – Adrian Hamins-Puertolas – Chris Kurz

Federal Reserve Board of Governors

Opinions expressed herein are those of the authors alone and do not necessarily reflect the views of the Federal Reserve System. All results have been reviewed to ensure that no confidential data are disclosed. We thank the staff of ADP Inc. for the use of their data and their support.

Question

How do we combine official and alternative data to improve estimates?

- Motivation: Government surveys have measurement error.
- Our approach: State-space model. Treat official and alternative data as noisy indicators of truth.
- Secondary questions: How should we assess, weight and benchmark alternative data?

Use CES, QCEW heavily in our application: not about replacing government statistics.

Data

Our data: records from the payroll processor ADP (2001-present).

- Weekly worker counts for each firm.
- ADP processes paychecks for **20 percent** of US workers.
- Observe all weeks, not just a reference period.
- Not a probability sample, but approximately representative.

More

Process microdata into aggregate private employment estimates.

- Methodology roughly similar to CES.

Methods: State space model

Kalman Filter assumptions:

- Unobserved state (actual employment growth) follows an AR(1):

$$\Delta EMP_t^U = \mu(1-\rho) + \rho \Delta EMP_{t-1}^U + \epsilon_t^U$$

- ADP and CES are noisy signals of truth:

$$\begin{bmatrix} \Delta EMP_t^{CES} \\ \Delta EMP_t^{ADP} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \Delta EMP_t^U + \begin{bmatrix} \epsilon_t^{CES} \\ \epsilon_t^{ADP} \end{bmatrix}$$

Feed in the data:

- Extract estimates of observation noise.
- Extract estimates of true employment growth.

Preliminary Results

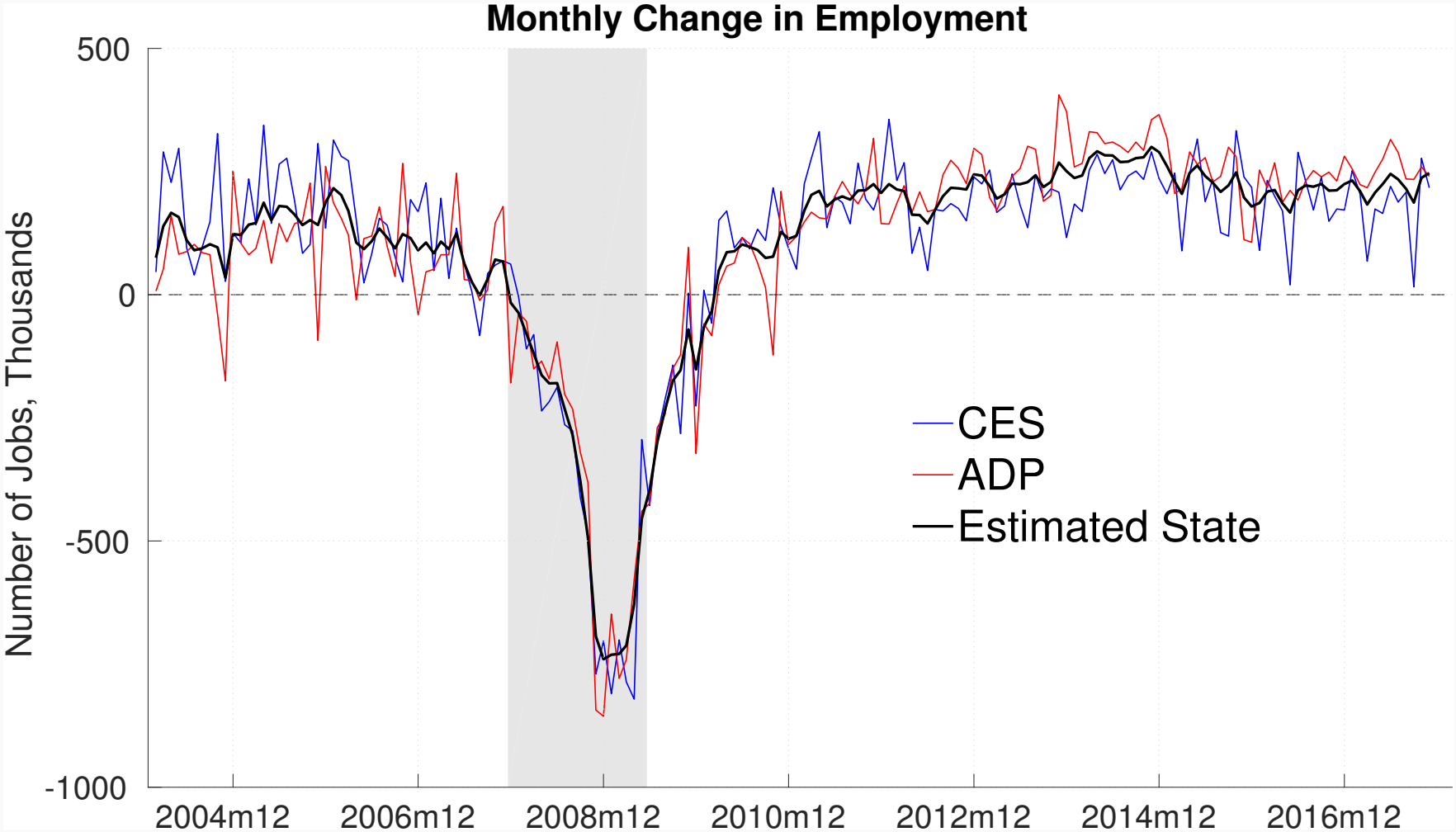
State-space estimates:

- Model puts roughly equal weight on CES, ADP.
- Robust to changes in assumptions: AR(1), random walk, correlation of observation noise.
- To do: real time estimation, forecasting.

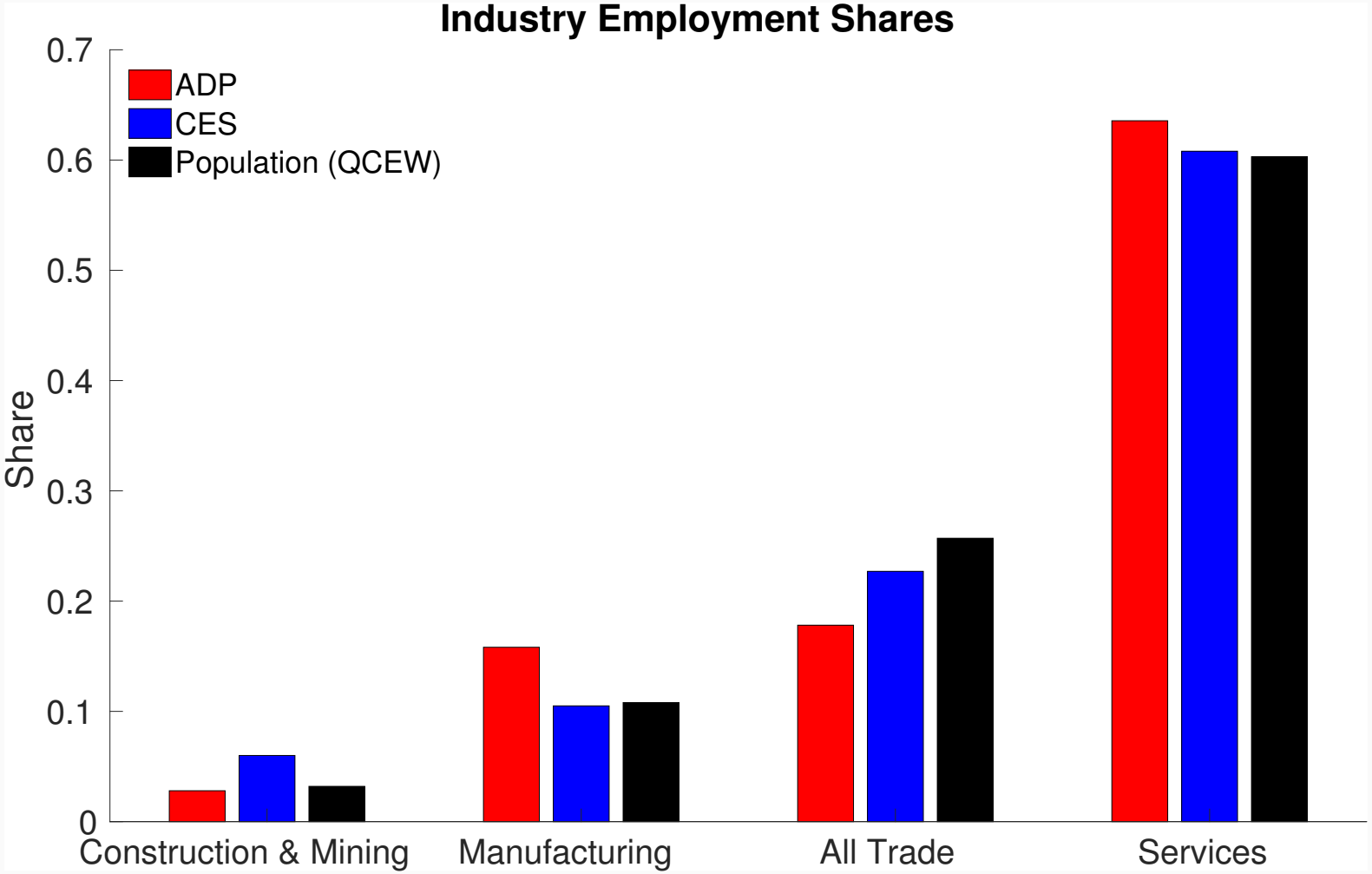
Auxiliary results:

- Business-cycle frequency correlation b/w ADP and CES growth: 0.95.
- ADP series helps forecast final CES (real time, pseudo out-of-sample).

Preliminary Results



Representativeness



Back

Representativeness

Pay frequency	ADP emp.	ADP estabs.	QCEW estabs.
Weekly	23.4	22.4	32.2
Biweekly	55.1	45.8	40.0
Semimonthly	17.5	20.6	18.5
Monthly	4.0	11.2	9.3

Back

Representativeness

Census Region	ADP emp.	ADP estabs.	QCEW emp.
Northeast	28.2	28.1	18.2
South	29.4	30.2	34.9
Midwest	20.2	16.6	20.1
West	22.2	25.2	26.8

Back