Expanding the Frontier of Economic Statistics Using Alternative Data: A Case Study of Regional Employment

Abe Dunn (BEA, Presenting), Kyle Hood (BEA), Brian Quistorff (BEA), Lowell Mason (BLS), Eric English (Census)

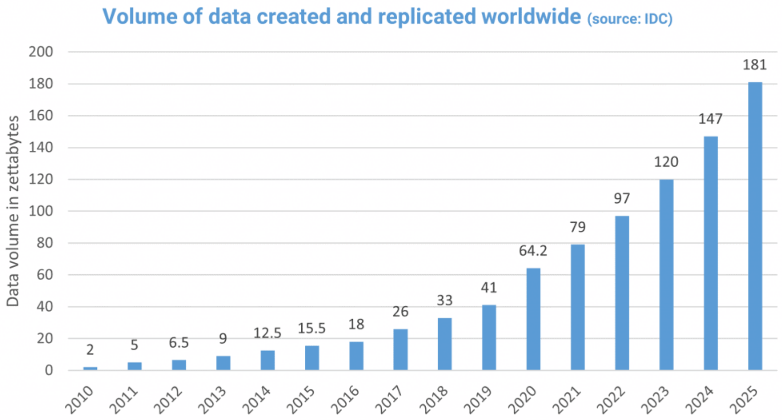
FCSM 2023 Research and Policy Conference
October 26th, 2023

Bureau of Economic Analysis
U.S. DEPARTMENT OF COMMERCE

Disclaimer

The views expressed here are those of the authors and do not represent those of the Bureau of Labor Statistics, the Bureau of Economic Analysis, the Census Bureau, the Department of Commerce, or the Department of Labor

- Exponential growth of data:



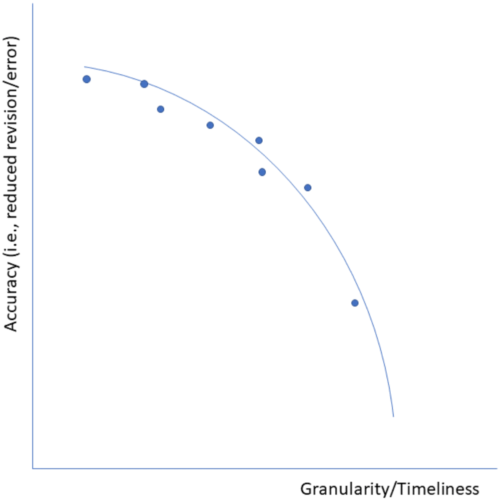
Motivation and Background

- Motivation
  - Statistical agencies are expected to:
    - Leverage increasingly available data to produce timely, high-frequency and granular statistics
    - Continue producing accurate and reliable estimates
  - Main question: How to attain these objectives considering the pros- and cons- of big data sources?
- Background
  - Several applications of third-party data by academic and institutional researchers to obtain timely estimates: Chetty et al. (2020), Aladangady et al. (2019), Dunn et al. (2020), Autor et al. (2022), Cox et al. (2020), Cajner et al. (2020)
  - Most quality checks are qualitative, quality checks typically compare to official measures, and trade-offs are not usually quantified

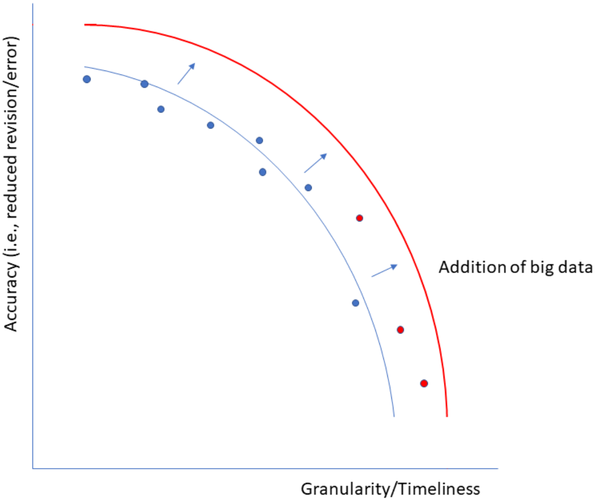
What's the plan?

- Questions about third party data and current practice
  - How do statistical agencies currently weigh the advantages and disadvantages of official data sources and data releases?
  - Third-party data source can be used to produce more granular or timely estimates than currently available, but are they an improvement?
- Steps for measuring improvement:
  1. Use official data sources to measure the current "tolerance" for error
     - Timeliness and granularity vs accuracy (i.e., less error/revisions)
  2. Use alternative data to produce *new/improved* estimates
  3. Measure error of new/improved estimates using cross-validation
  4. Errors in the new/improved estimates can be compared with tolerance levels to determine the value of new statistics

Production possibility frontier for economic measurement



Production possibility frontier for economic measurement



Application to employment data

- Data sources
  - Quarterly Census of Employment and Wages (QCEW) — monthly census of employment — lags about 5 months
  - Current Employment Survey (CES) — monthly state/MSA employment estimates — lags about 1 month
  - Paychex — granular data from payroll processor (we use county/state by 2-digit industry) — lag is customizable

- Reason for focus
  - Detailed and closely watched data, but CES lags in geographic detail (i.e., county)
  - Paychex offers the possibility to improve statistics/produce new statistics relative to CES alone
  - "True" estimate from QCEW is known on a regular basis
  - BLS, Census Bureau, and BEA all report regional employment statistics

What we do

Application to county- and state-level, 2-digit industry employment

1. Assess current tolerance from official sources (CES and QCEW)
2. Use Paychex data to produce alternative estimates (state- and county-level, two digit industry), prior to the availability of QCEW
3. Use cross-validation to evaluate the accuracy (i.e., error/revision) of alternative estimates using Paychex data
4. Compare tolerance levels from official sources to cross-validation error from new estimates to determine the value of the new estimates

What we find

- Properties of data sources:
  - Implicit tolerance for error changes with granularity (i.e., geographic & industry)
  - Construction of Paychex employment is critical to extracting a meaningful information (e.g., full sample vs continuing employers)
- For currently targeted estimates, solely using Paychex produces errors higher than tolerance levels $\Rightarrow$ use CES and Paychex together
- Paychex data (combined with CES) shows some improvement in state-level CES relative to tolerance level
- Paychex data (combined with CES) produces new timely county-level estimates in range of tolerance level for this level of granularity

Outline

- Data
- Descriptive statistics
- Framework for evaluating statistics
- Tolerance level form existing official statistics
- Evaluate input data from Paychex
- Apply Paychex data to look at new/improved statistics
  - State-level
  - County-level
- Conclusion/discussion

Data

- Time period covered
  - QCEW: Monthly [01/2017 to 06/2021]
  - CES: Monthly [01/2013 to 06/2021]
  - Paychex: Monthly [01/2016 to 12/2022]

Paychex data

- Micro data from payroll processor
  - Detailed information on employment and wages for mostly small- and medium-sized business
- About 4 percent of private sector employment
- Approx 400,000 employers

- Our adjustments
  - 2 samples: Continuing and full
  - Timing: 1-3 weeks lag or allow all future adjustments (note: We assess this but it is omitted here for time)

Framework for evaluating statistics

0. Defining evaluation criteria based on ground-truth (typically QCEW)
   - Mean absolute error (MAE) (L_1): Very understandable
   - $1 - R^2$ (L_2): More sensitive to large errors and normalizes by variability of the outcome. *Fraction of the variance unexplained*
1. Assess tolerance of existing statistics across types of granularity (County, MSA, State or National monthly employment growth at NAICS 3-digit, 2-digit, 0-digit industry level - CES compared to QCEW)
2. Produce new or improved statistics (i.e., county/state-level employment growth)
3. Out-of-sample cross-validation of new statistics using evaluation criteria
4. Evaluate performance of new statistics compared to tolerance level

CES differences (CES initial vs CES final)

Geo	Naics digits	Mean Abs. Error	MASE: MAE/MAD(Y)	1-R2	N
MSA	3d	0.0129	0.7726	0.2998	94172
MSA	2d	0.0119	0.7814	0.3873	157528
MSA	0d	0.0068	0.5828	0.3856	29024
State	3d	0.0111	0.6605	0.2136	136207
State	2d	0.0086	0.5580	0.1636	86712
State	0d	0.0030	0.2615	0.0339	4876
USA	3d	0.0048	0.3221	0.0951	6900
USA	2d	0.0032	0.2347	0.0342	2024
USA	0d	0.0014	0.1306	0.0052	92

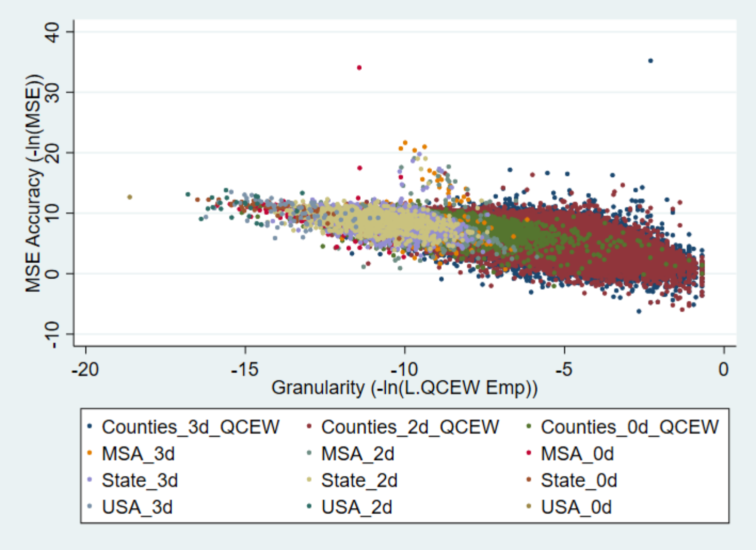
Notes: MAD is Mean Absolute Deviation, $N^{-1}\sum_i |y_i - \bar{y}|$. $1 - R^2 = MSE/Var(Y)$. Observations weighted by lagged QCEW employment.

CES differences (CES initial vs QCEW)

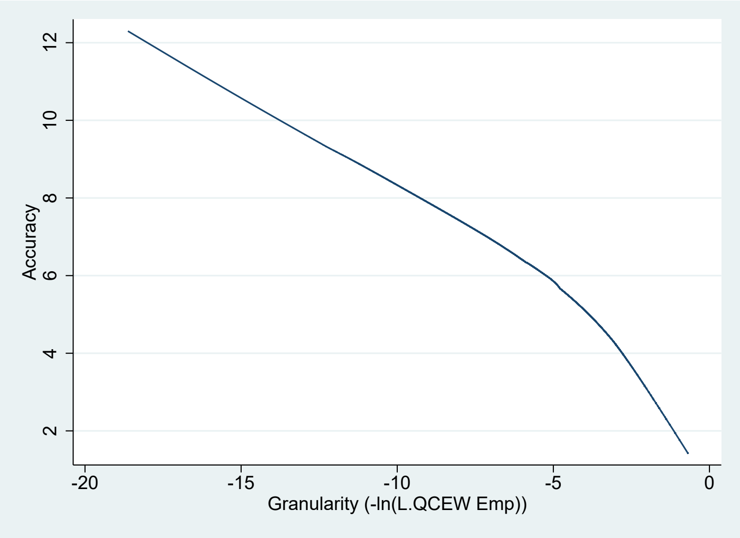
Geo	Naics digits	Mean Abs. Error	MASE: MAE/MAD(Y)	1-R2	N
Counties	3d	0.0214	0.8907	0.6684	4844287
Counties	2d	0.0180	0.8779	0.6240	3124127
Counties	0d	0.0102	0.7511	0.5234	179114

Notes: MAD is Mean Absolute Deviation, $N^{-1}\sum_i |y_i - \bar{y}|$. $1 - R^2 = MSE/Var(Y)$. Observations weighted by lagged QCEW employment.

CES initial vs CES final (USA, State, MSA) and QCEW (County)



CES initial vs CES final: Lowess



Prediction equation

We wanted to compare (unmodelled) CES with models that include CES and Paychex data.

$$\text{Monthly QCEW Employment Growth} = f(CES, Paychex) + Error \tag{1}$$

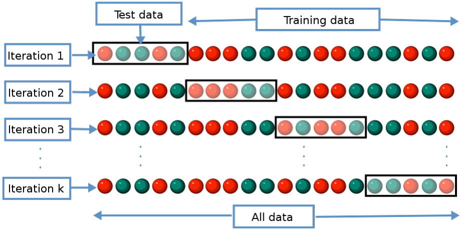
- $f()$ — linear functional form
- We try various specifications (CES only, Paychex only, CES + Paychex)

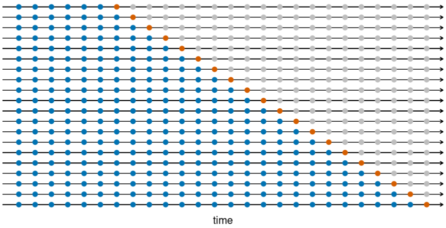
Evaluation criteria

Two out-of-sample approaches:

- Cross-fold estimation:
  - Advantage – Uses all time-series variation in the data
  - Disadvantage – Does not match how estimation will work in practice

- Rolling one-step-ahead estimation:
  - Advantage – Closer to how estimation would work in practice
  - Disadvantage – Does not use the full time series variation





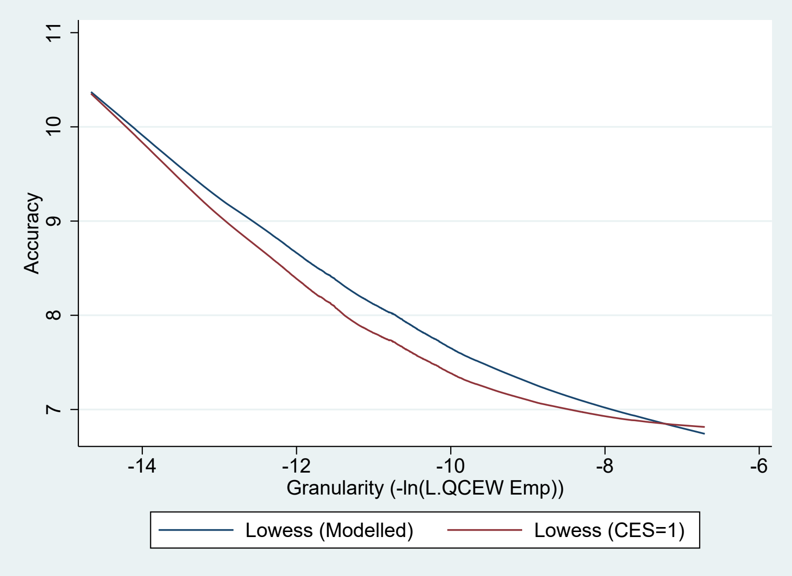
What Paychex input data to use for this application?

- Continuing sample vs full sample?
  - Full sample provides noisy signal $\Rightarrow$ *Use continuing sample*
- Use Paychex alone or in combination with CES data?
  - Paychex alone provides a weak signal $\Rightarrow$ *Use in combination with CES*

Improved state-level estimates?

	(1)	(2)	(3)	(4)
CES Emp Gr	1	0.853***	0.755***	0.716***
	(.)	(438.58)	(315.00)	(292.91)
Pay Emp Gr (cont; t=3)			0.0925***	0.0612***
			(64.58)	(38.20)
Pay (cont; t3) Coverage x Emp GR				-0.000896
				(-1.38)
Paychex Emp Gr (cnty-Agg; t3)				0.163***
				(51.64)
Observations	41938	41938	41938	41938
$1 - R^2$ OOS (rolling)	0.177	0.216	0.196	0.171
$1 - R^2$ OOS (cf)	0.179	0.180	0.163	0.154
MAE OOS (rolling)	0.00876	0.00865	0.00840	0.00806
MAE OOS (cf)	0.00853	0.00803	0.00787	0.00774

Expanding the PPF for state estimates



New county-level estimates?

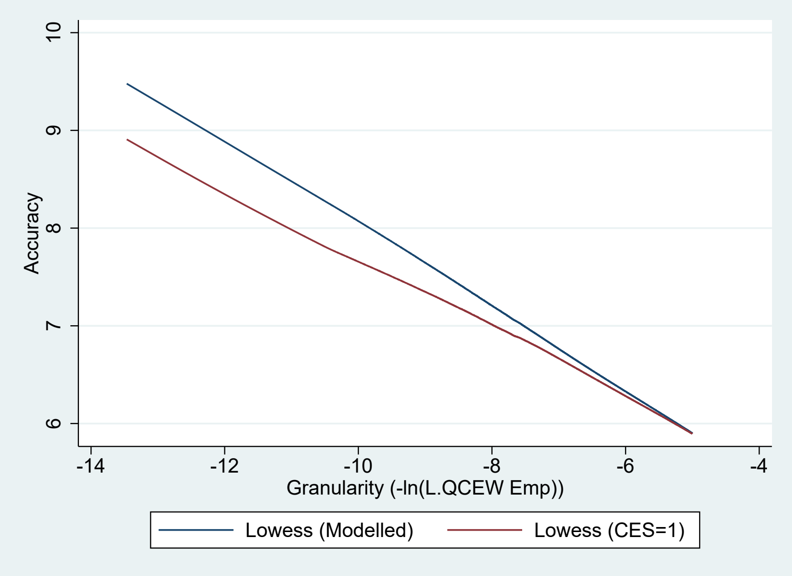
	(1)	(2)	(3)	(4)	(5)	(6)
CES Emp GR (MSA or State)	1	0.840***			0.804***	0.668***
	(.)	(741.08)			(690.87)	(528.56)
Paychex Emp Gr (cnty-naics2; t3)			0.198***	0.0457***	0.0486***	0.0175***
			(392.75)	(76.73)	(100.72)	(35.13)
Paychex Emp Gr (Cnty-naics2) x Paychex Coverage (naics2) (t3)				0.494***		0.292***
				(68.83)		(49.44)
Paychex Emp Gr (cnty-Agg; t3)				0.191***		0.109***
				(192.61)		(132.68)
Paychex Emp Gr (state-naics2; t3)				0.367***		0.132***
				(425.52)		(165.71)
Observations	897174	897174	897174	897174	897174	897174
$1 - R^2$ (rolling)	0.473	0.541	0.934	0.723	0.513	0.469
$1 - R^2$ (CrossFold)	0.478	0.472	0.856	0.650	0.448	0.426
MAE (rolling)	0.0160	0.0151	0.0187	0.0172	0.0149	0.0145
MAE (CrossFold)	0.0156	0.0144	0.0183	0.0161	0.0143	0.0138

t statistics in parentheses

Hiding non-MSA interaction terms for main GR covariates.

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Expanding the PPF for county estimates



Conclusions so far

- Current estimates provide guidance for "tolerance" for new or improved estimates
- Some evidence of improved estimates at the state level (about 10 percent reduction in error)
- County-level estimates appear promising and with reasonable tolerance levels (10-15 percent reduction in error, relative to CES alone)
- Adjustments to alternative data are critical

Aditya Aladangady, Shifrah Aron-Dine, Wendy Dunn, Laura Feiveson, Paul Lengermann, and Claudia Sahm. From transactions data to economic statistics: constructing real-time, high-frequency, geographic measures of consumer spending. Technical report, National Bureau of Economic Research, 2019.

David Autor, David Cho, Leland D Crane, Mita Goldar, Byron Lutz, Joshua Montes, William B Peterman, David Ratner, Daniel Villar, and Ahu Yildirmaz. An evaluation of the paycheck protection program using administrative payroll microdata. *Journal of Public Economics*, 211:104664, 2022.

Tomaz Cajner, Leland D Crane, Ryan A Decker, John Grigsby, Adrian Hamins-Puertolas, Erik Hurst, Christopher Kurz, and Ahu Yildirmaz. The us labor market during the beginning of the pandemic recession. Technical report, National Bureau of Economic Research, 2020.

Raj Chetty, John N Friedman, Nathaniel Hendren, Michael Stepner, et al. The economic impacts of covid-19: Evidence from a new public database built using private sector data. Technical report, national Bureau of economic research, 2020.

Natalie Cox, Peter Ganong, Pascal Noel, Joseph Vavra, Arlene Wong, Diana Farrell, Fiona Greig, and Erica Deadman. Initial impacts of the pandemic on consumer behavior: Evidence from linked income, spending, and savings data. *Brookings Papers on Economic Activity*, 2020(2):35–82, 2020.

Abe Dunn, Kyle Hood, and Alexander Driessen. *Measuring the effects of the COVID-19 pandemic on consumer spending using card transaction data*. US Department of Commerce, Bureau of Economic Analysis, 2020.