

Emerging Trends Impacting Data Sharing

Michael D. Levi

U.S. Bureau of Labor Statistics

APDU 2015 Annual Conference

Sept. 2, 2015



Outline

1. BLS data resources
2. The (arguably) biggest threat to data availability
3. The (arguably) biggest obstacle to data sharing

Agency Context

- Fundamental mission:
 - ▶ Produce and disseminate data that are
 - Accurate
 - Objective
 - Relevant
 - Timely
 - Accessible
 - ▶ ... to help businesses, households, policy-makers, and NGOs make better decisions

Access to BLS Data: Website

- www.bls.gov
- Database of published estimates
 - ▶ Suite of query tools (improving)
 - ▶ API (also improving)
- Flat files
- Tables (HTML, PDF, Excel)

Push Mechanisms

- E-mail subscriptions
- Twitter
- RSS feeds

Research Data Centers

- BLS expanding presence in Federal statistical research data centers
- Two data sets coming soon (2015)
 - ▶ SOII
 - ▶ NLS-Y
 - ▶ More under way
- BLS looking to establish an FSRDC in our building

The (arguably) Biggest Threat

- Increasingly contentious public discourse may undermine users' confidence in published data
- Few users can verify the completeness and accuracy of published data
- Reputation is essential to retain:
 - ▶ End users
 - ▶ Respondents

Social Context

Percent of American adults who trust the Federal government:

- 3% - Just about always
- 20% - Most of the time
- 56% - Only some of the time
- 19% - None of the time

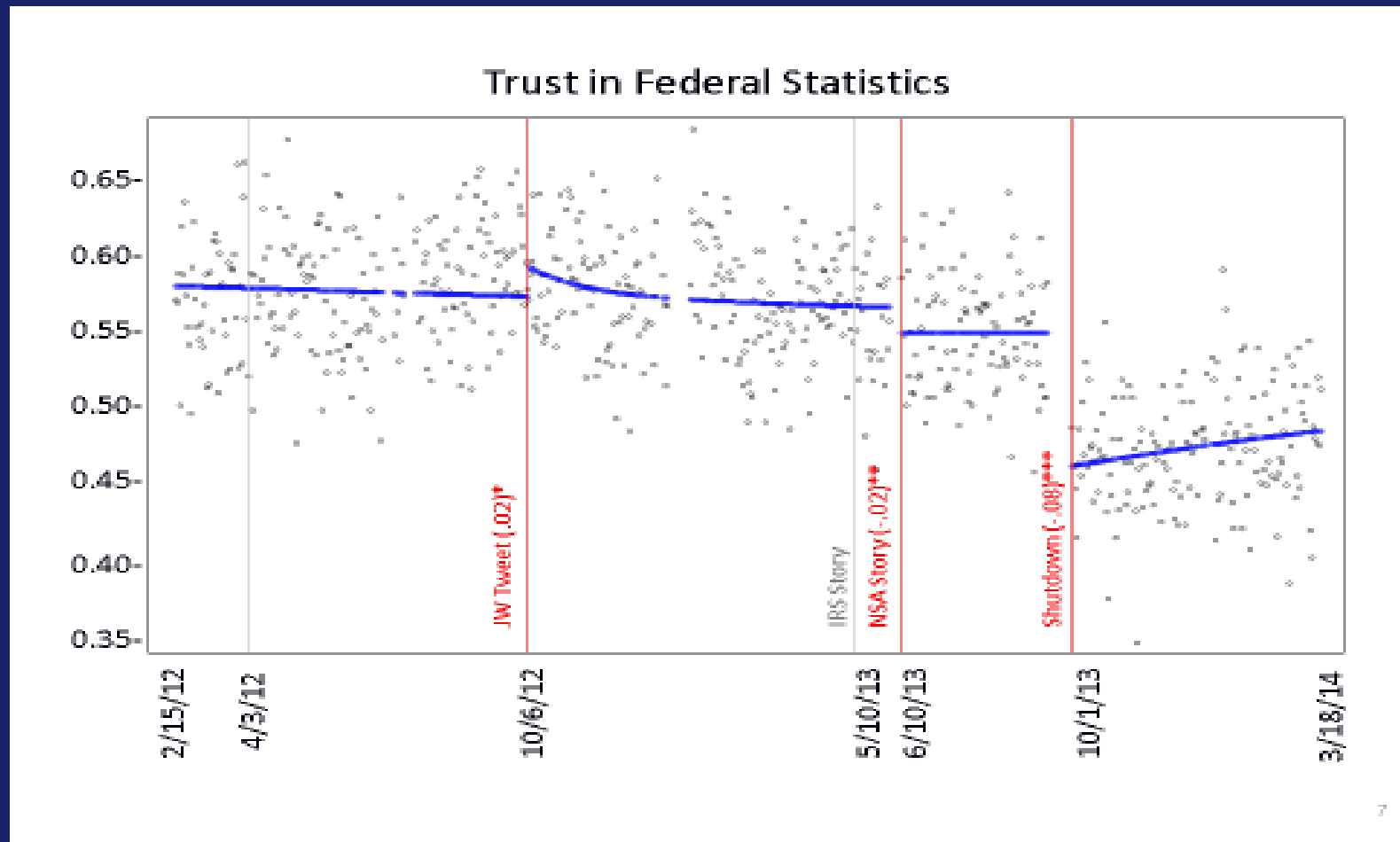
Pew Research American Trends Panel (Nov 17 to Dec 15, 2014)

FSS Trust Survey – Some Results

- Time span: Feb. 2012 - March 2014
- Four potential shocks:
 - ▶ Oct. 2012: Jack Welch tweet
 - ▶ May 2013: Internal Revenue Service story
 - ▶ June 2013: National Security Agency story
 - ▶ Oct. 2013: US Government shutdown

See Childs, J.H., King, R., Toth, D., & Earp, M. (2014). Potential Effects of Government Shutdown on the Federal Statistical System. American Association for Public Opinion Research. Anaheim, CA.

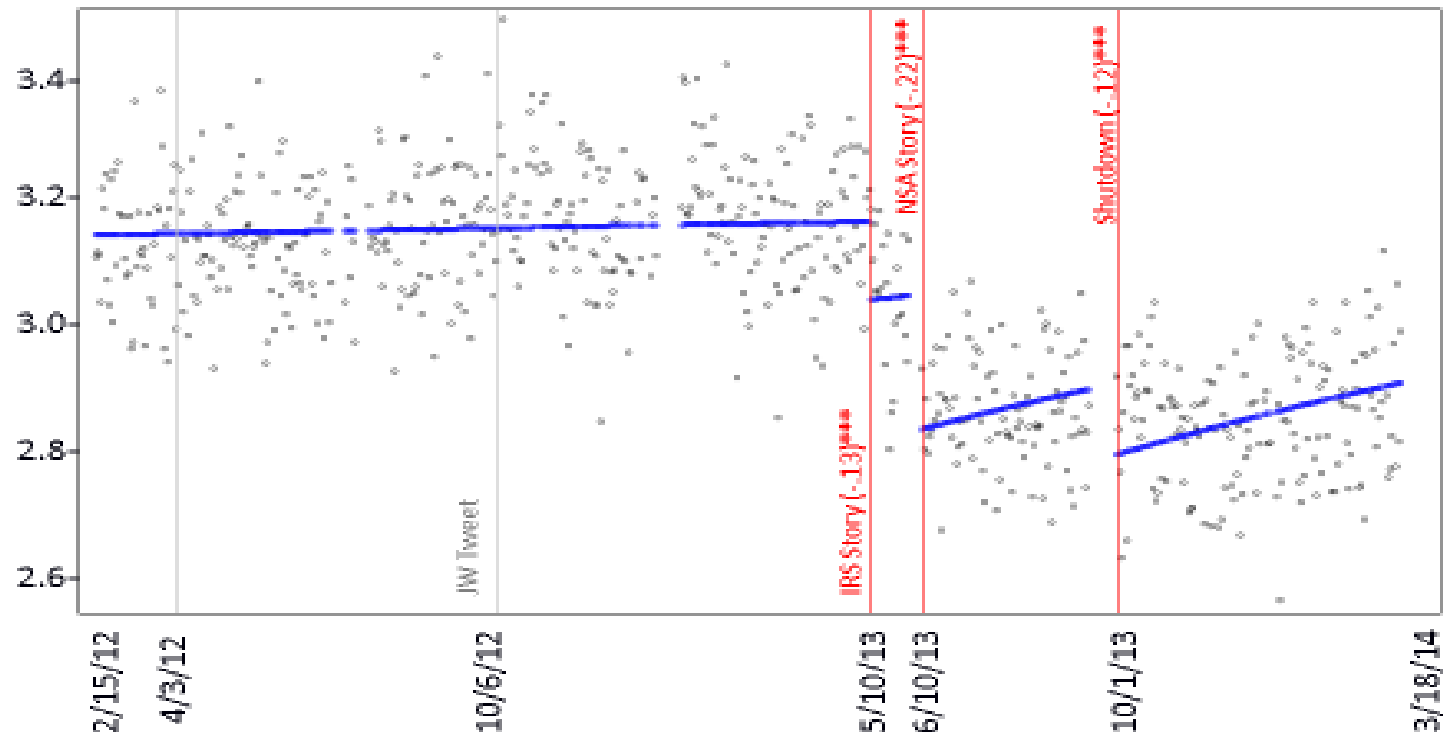
[Do] you tend to trust Federal statistics?



(Childs, King, Toth, & Earp; AAPOR, 2014)

Confidentiality

People Can Trust Federal Statistical Agencies to Keep Information About Them Confidential



11

Interpretation of Results

- Direct attack on a statistical product:
 - ▶ No change in public trust, or
 - ▶ Rise in trust.
- Damage comes from events unrelated
 - ▶ To BLS
 - ▶ To other Federal statistical agencies.
- Third-party support helped mitigate potential harm

The (arguably) Biggest Obstacle

- Matching data from diverse sources
 - ▶ Even for different programs from one data provider
 - ▶ Especially across data providers

The Problem

- Concepts, definitions, and nomenclature evolve independently in individual programs
- Similar concepts may be described using different words
- Different concepts may be described using the same words

Examples

- Earnings / income / wages / compensation
- Boston

A Solution-in-Progress

- Develop a unified BLS taxonomy and lexicon
- In plain English
- Geared towards the lay public
- To be used in
 - ▶ data query tools
 - ▶ document tagging
 - ▶ website search and navigation

Some Conceptual Issues

- Different user communities' potential need for different language
- “Lumping” vs “splitting”
- Cross-walk vs harmonized data
- Continuity breaks in time series
- Does it scale?
- Stay tuned...

Contact Information

Michael D. Levi

Associate Commissioner for Publications and
Special Studies

U.S. Bureau of Labor Statistics

www.bls.gov

202-691-5100

levi.michael@bls.gov



www.bls.gov