

# Data Warehouse and Access Initiative

Michael D. Levi

BLS Data Users Advisory Committee

Nov. 9, 2017



# Background

- BLS Program Reviews
  - ▶ Scrutiny by outside experts (from BLS)
  - ▶ Issue findings and recommendations for program improvement
  - ▶ Highlight program best practices
- Opportunity to identify corporate issues beyond the control of an individual program



# Corporate Findings

- Multiple program reviews over past several years identified corporate issues related to
  - ▶ Availability of data
  - ▶ Accessibility of data
  - ▶ Presentation of data



# Causes

- Limitations of public database and query tools mostly designed 25 years ago
- Demise of a commercial table-producing software package
  - ▶ Exacerbated by agency stove pipes
- Optimizing for a program vs. optimizing across programs

# In Response

- Data Repository and Access Working team (DRAW)
  - ▶ Chartered by BLS Executive Oversight Board (EOB) in mid-October, 2017



# Problems We Want to Solve

- Not all data are directly available to customers
- Some of the data that are available to customers are hard to find on the website
- It is difficult for customers to combine data from more than one program.



# Problems to Solve (cont.)

- There is a proliferation of data access tools customized for individual programs
- Current practices at BLS make maintenance and enhancement of data dissemination capabilities inefficient
- It is inefficient to load very large data sets



# DRAW Charge

- Develop the business requirements for
  - ▶ An adaptable BLS-wide repository to house all BLS publishable data and a relevant set of metadata in one location
  - ▶ The user interface(s) that will allow BLS staff and customers to easily find and access BLS data and create desired products.



# Repository

## ■ BLS publishable data

- ▶ All estimates that meet individual programs' publishability criteria

- Not microdata / public use data files

- No expectation that programs will change their series or estimates

## ■ Other significant data?

- ▶ Experimental measures, vintage data, etc.

# Repository (cont.)

- Statistical measures?
  - ▶ Significance, relative importance, etc.
- Metadata?
  - ▶ Inform customers of important characteristics and potential differences in the data they are selecting
- Ability to connect data across programs

# User Interfaces

- Query multiple dimensions of data
  - ▶ Across programs, time series, etc.
- Consistent “layouts”
  - ▶ Time series, tables, data visualizations
- Enable customers to create layouts
  - ▶ Appropriate set of BLS products
  - ▶ Empower users to create their own



# User Interfaces (cont.)

- Consistent delivery formats
  - ▶ HTML, CSV, Excel, print-friendly, etc.



# Time Frame

- DRAW deliverables due end of June, 2018
- Report to Executive Oversight Board (EOB)
- EOB decides on next steps
  - ▶ When and how BLS will proceed with development and implementation
  - ▶ Where resources will come from



# Questions for DUAC

- Is BLS asking itself the right questions?
- Are there additional important shortcomings in current data availability and access that are not being addressed?
- Should BLS be prioritizing this initiative high on the list of all desirable agency improvements?



# Contact Information

**Michael D. Levi**

Associate Commissioner

Office of Publications and Special Studies

202-691-5100

[levi.michael@bls.gov](mailto:levi.michael@bls.gov)

