

Consumer Expenditure Surveys (CE) Data Quality Profile

Gray Jones, BLS
International Total Survey Error Workshop
Washington, D.C.
September 19, 2024



Disclaimer

- The information is being released for statistical purposes, to inform interested parties, and to encourage discussion of work in progress.



Presentation Agenda

- Background on the CE Surveys
- Creating a Data Quality Profile (DQP) in CE
 - ▶ Why we needed a DQP
 - ▶ Defining data quality
- DQP implementation
 - ▶ Quality metrics used
 - ▶ Relation to the Total Survey Error (TSE) Framework
- DQP Successes & Challenges
- Moving forward



Consumer Expenditure Surveys (CE) Background



Consumer Expenditure Surveys (CE)

- Collects spending data on the U.S. Population
- Provide expenditure weights for the U.S. Consumer Price Index (CPI)
- Sponsored by the Bureau of Labor Statistics, collected by the Census Bureau
- Survey participants report dollar amounts for all non-investment purchases. Business expenses and reimbursements are excluded.
- Two independent surveys:
 - ▶ Interview Survey: 4 waves, 3-month recall of large and regular expenses
 - ▶ Diary Survey: Household characteristic survey and two consecutive one-week expenditure diaries



Creating a Data Quality Profile (DQP) in CE

■ Why do we need a DQP?

- ▶ OMB Statistical Policy Directive No. 1: Fundamental Responsibilities of Federal Statistical Agencies
- ▶ Part of BLS Strategic Plan
 - 2.3.3 — Develop and implement standards for providing measures of data reliability to users.
- ▶ CE Redesign Project
- ▶ Ongoing improvements

Implementing a DQP in CE

CE Data Quality Definition Team Charted (2009)

- Proposed quality definitions and indicators.
- Considered Total Quality Management (TQM) and Total Survey Error (TSE) Frameworks.

First CE Prototype DQP Published (2013)

- Included response rates, expenditure edit rates, and income imputation rates.

First Annual CE DQP Published (2018)

- Began the standard of publishing a data quality profile alongside the CE annual data release in September of each year.
- Contained four quality metrics (added record use).

The CE DQP Today (2024)

- Two DQP's published per year.
 - Midyear DQP and Annual DQP accompanying both CE data releases.
- Contains eight quality metrics.

The CE DQP Today (2024)

- Two DQPs published per year.
 - ▶ Midyear DQP and Annual DQP accompanying both CE data releases.
- Metrics updated on a quarterly basis internally.
 - ▶ Available to stakeholders in a user-friendly dashboard application
- Contains eight quality metrics.



Quality Metrics Used

1. Final Disposition Rates
2. Record Use (Interview Survey Only)
3. Information Booklet Use
4. Expenditure Edit Rate
5. Income Imputation Rate
6. Reported respondent burden (Interview Survey Only)
7. Survey Mode
8. Survey Time



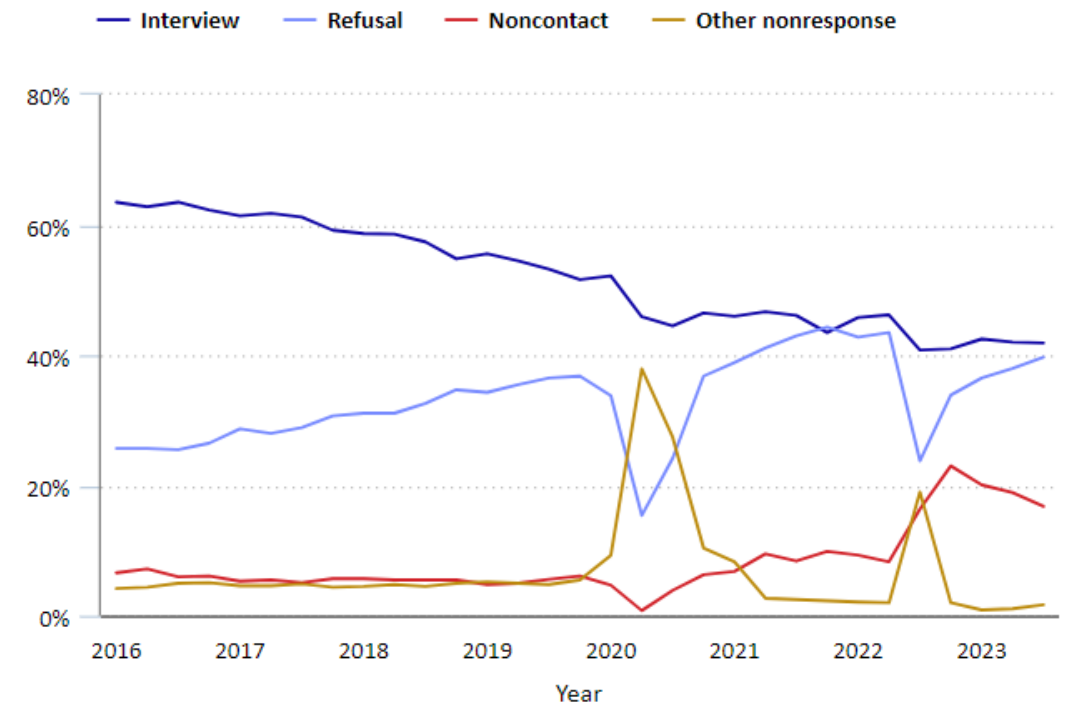
Metric Relation to TSE Framework

	Total Survey Error Dimensions Associated with the Metrics		
	Measurement	Nonresponse	Processing
1. Final Disposition Rates		X	
2. Record Use	X		
3. Information Booklet Use	X		
4. Expenditure Edit Rate	X	X	X
5. Income Imputation Rate	X	X	X
6. Respondent burden	X		
7. Survey Mode	X		
8. Survey Time	X		

1. Final Disposition Rates

- Final disposition rates of eligible sample units represent the final participation outcomes of CE field staff's survey recruitment efforts.
- Beneficial for illustrating the impacts of survey fielding changes.
 - ▶ COVID-19 Pandemic
 - ▶ Cost containment measures

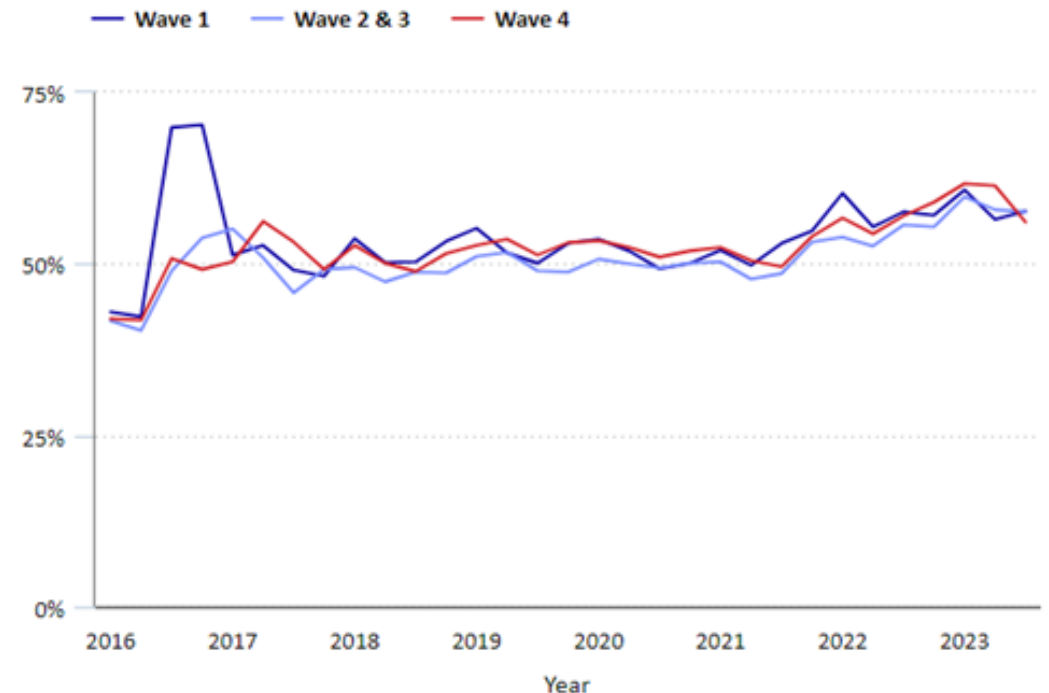
Chart 1.2 Interview Survey: Distribution of final dispositions for eligible sample units (unweighted) (in percent)



Record Use (Interview Survey Only)

- Measures the proportion of respondents who refer to records while answering the Interview Survey questions, as reported by Census field staff.
 - ▶ Records include, but are not limited to: receipts, bills, checkbooks, and bank statements.
 - ▶ Increased record use has been linked to improved data quality.
 - ▶ Assists in the evaluation of incentives test (2016).

Chart 2.1 Interview Survey: Prevalence of records use among respondents (in percent)

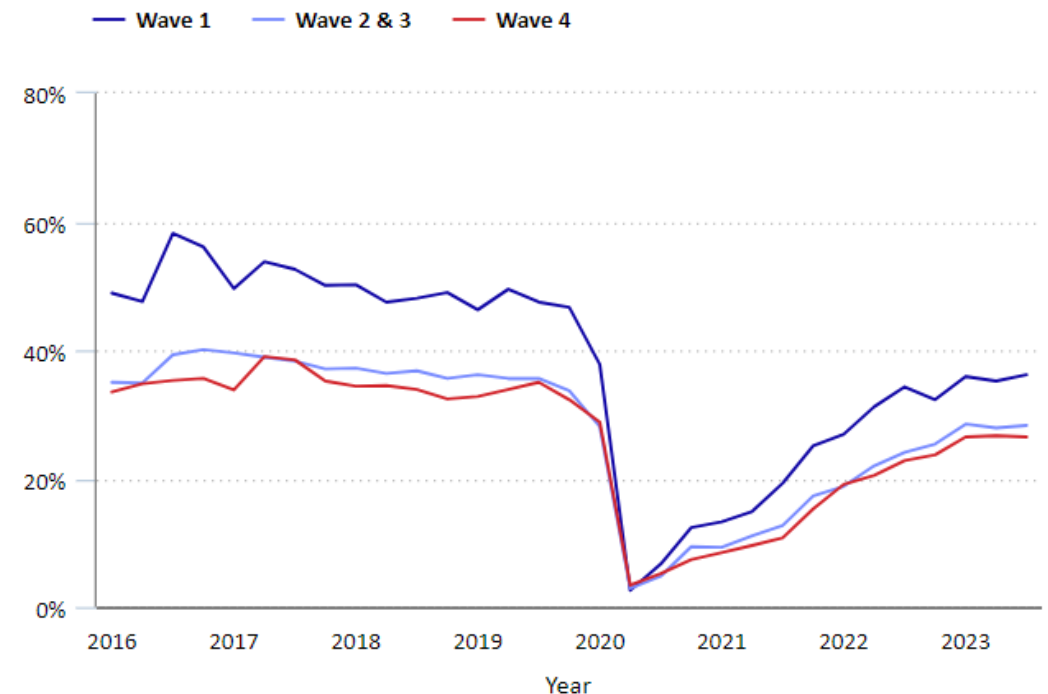


Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Information Booklet Use

- Identifies the prevalence of Information Booklet use among respondents during their interviews, according to Census field staff.
- Information Booklet use has been associated with higher quality expenditure reporting.
- Ability to track use as new or updated Information Booklets are introduced.
 - ▶ Disposable booklets during pandemic

Chart 3.2 Interview Survey: Prevalence of information booklet use among respondents (in percent)

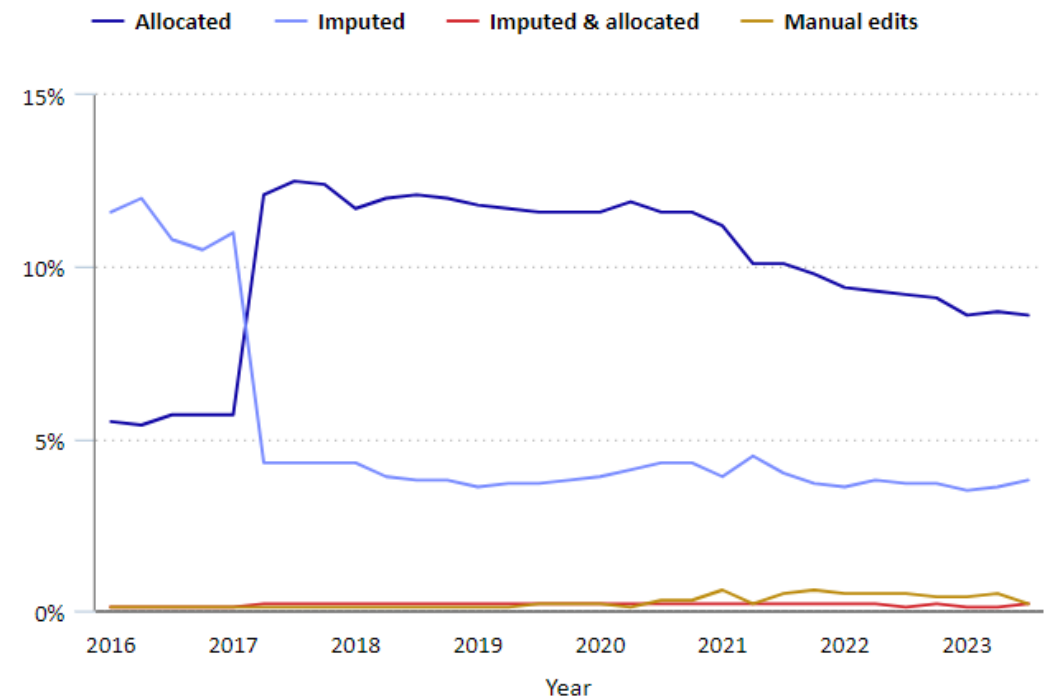


Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Expenditure Edit Rate

- Measures the proportion of reported expenditure data that are edited during CE data processing.
- Beneficial in evaluating the impact of processing changes.
 - ▶ Cable, internet, and phone allocation (Interview Survey, 2017)
 - ▶ Consolidation of classification codes (Diary Survey, 2022)

Chart 4.2 Interview Survey: Reported expenditure records (in percent)

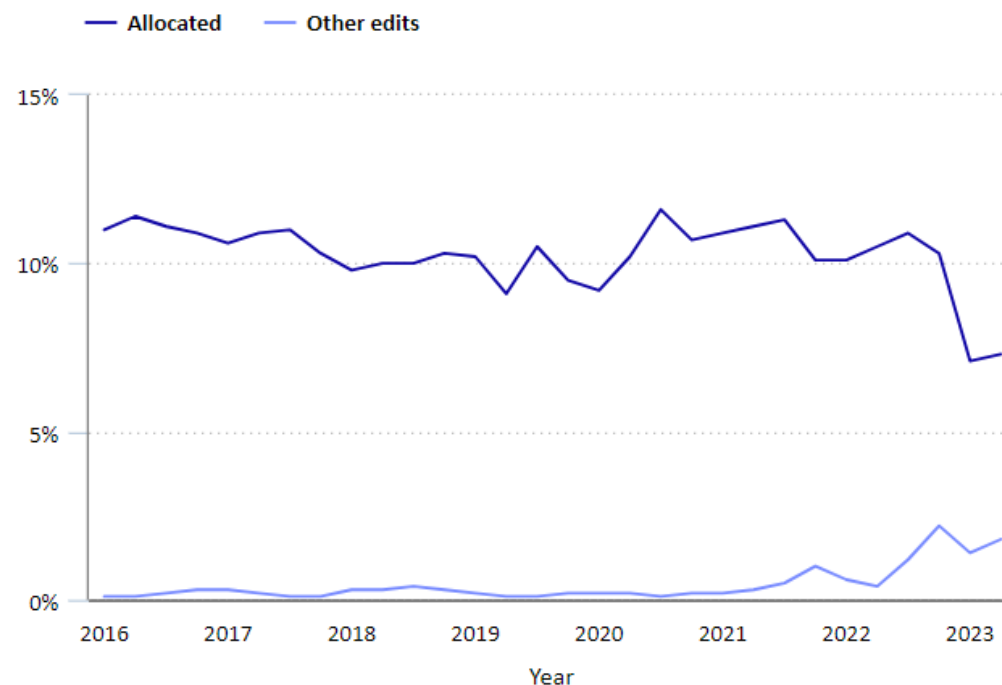


Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Expenditure Edit Rate Continued

- Measures the proportion of reported expenditure data that are edited during CE data processing.
- Beneficial in evaluating the impact of processing changes.
 - ▶ Cable, internet, and phone allocation (Interview Survey, 2017)
 - ▶ Consolidation of classification codes (Diary Survey, 2022)

Chart 4.1 Diary Survey: Reported expenditure records (in percent)



Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Survey Mode

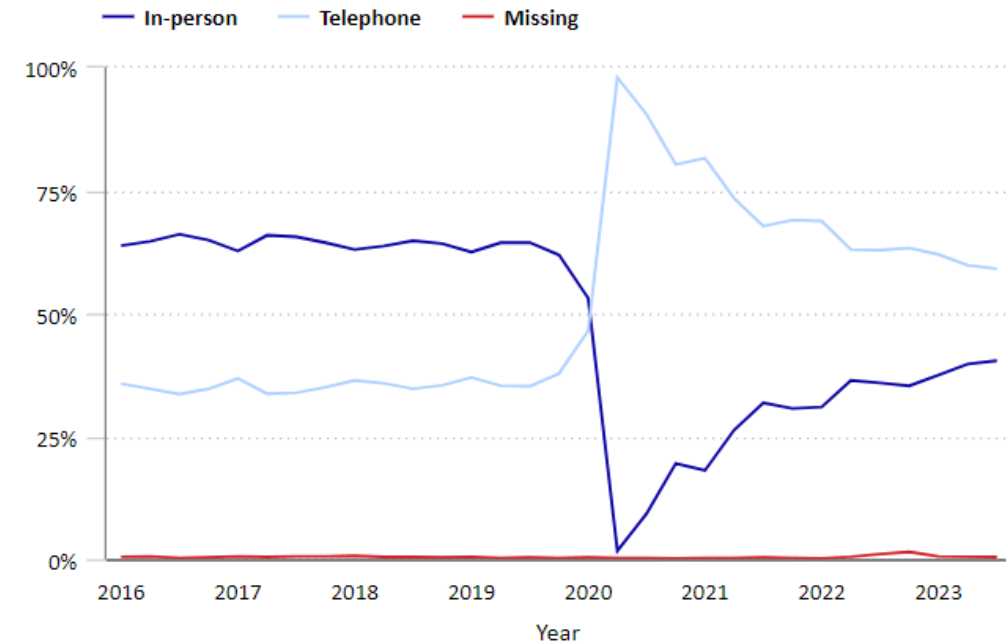
■ Measure the mode of data collection for the Diary and the Interview Surveys.

- ▶ In-Person Interviews vs Telephone Interviews
- ▶ Paper Diaries vs Online Diaries

■ Assists in evaluating:

- ▶ Data collection changes
- ▶ New modes (online diary)

Chart 7.3 Interview Survey: Survey mode (all waves) (in percent)



Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Survey Mode Continued

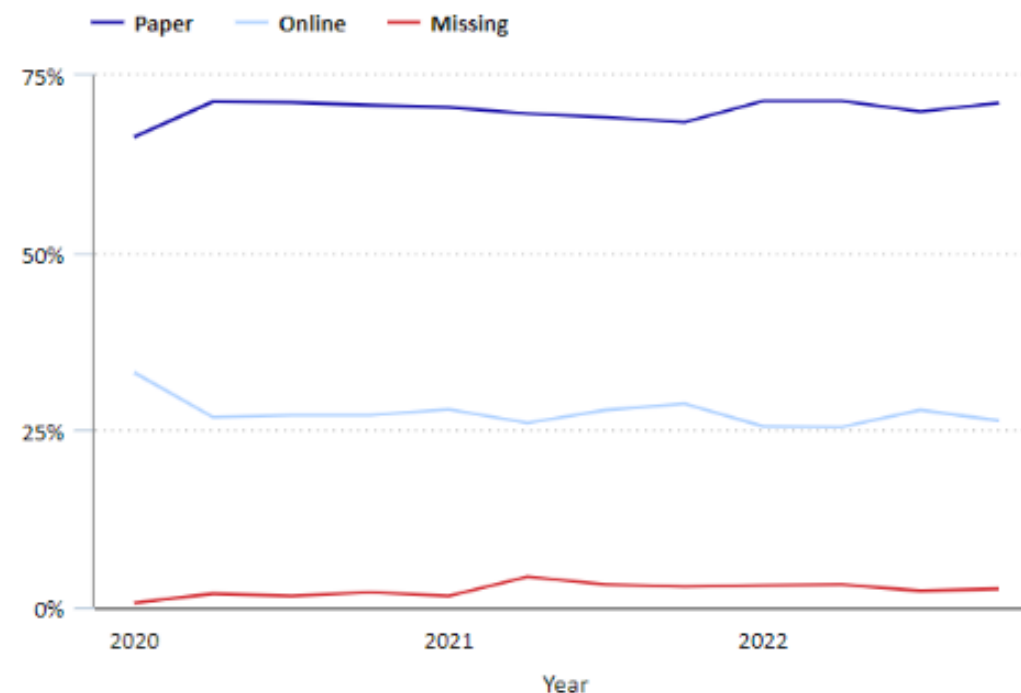
■ Measure the mode of data collection for the Diary and the Interview Surveys.

- ▶ In-Person Interviews vs Telephone Interviews
- ▶ Paper Diaries vs Online Diaries

■ Assists in evaluating:

- ▶ Data collection changes
- ▶ New modes (online diary)

Chart 7.2 Diary Survey: Expenditure diary survey mode (in percent)

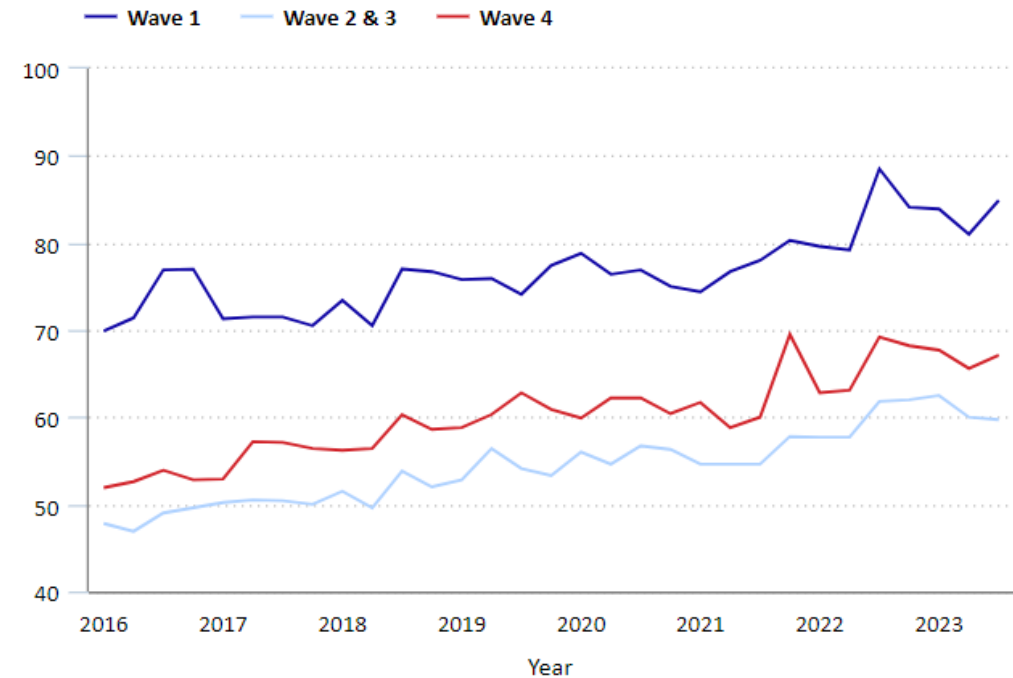


Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

Survey Time

- Defined as the number of minutes needed to complete an interview.
- Helpful for evaluating:
 - ▶ Objective burden
 - ▶ New survey content
 - ▶ Fielding procedures
 - Computer Assisted Recorded Interviewing (CARI) in 2022.

Chart 8.2 Interview Survey: Median length of time to complete survey (in minutes)



Click legend items to change data display. Hover over chart to view data.
Source: U.S. Bureau of Labor Statistics.

CE DQP Successes

- Indicates quality of published data for our users.
- Helps our internal stakeholders identify survey components in need of improvement.
- Identifying issues in the data.
- Has become a “go-to” resource for answering questions related to data quality.
- CE audit cited the CE DQP as a best practice.



CE DQP Challenges

■ Metric thresholds

- ▶ Originally intended to have thresholds for each metric.
- ▶ Only the “Interview Survey Mode” metric has a threshold, which stipulates the acceptable level of “in-person” interviews.
- ▶ More difficult to create for some metrics like “Information Booklet Use,” “Record Use,” and “Perceived Respondent Burden.”

■ Maintaining the DQP Dashboard

■ Tracking metric development

CE DQP Future Updates

- External data comparisons
- Online diary paradata metrics



Contact Information

Gray Jones, BLS jones.graham@bls.gov

