An Update on Developing Response Metrics for the Economic Census

FedCASIC 2015

by
Joanna Fane Lineback
Eric B. Fink

United States Census Bureau
U.S. Department of Commerce
Economics and Statistics Administration
U.S. CENSUS BUREAU
census.gov

Outline

- Motivation for Developing Response Metrics
- Where We Left Off
- Economic Census Overview
- Response Metrics
- Examples
- Discussion

Reasons for Developing New Response Metrics

- Comply with Census Bureau and Office of Management and Budget standards and guidelines
- Produce response metrics similar to those used by other economic survey programs
- Produce more meaningful information about the "quality" of Economic Census (EC) data
- Improved (near) real-time measures of data collection performance

Where We Left Off

- Last time we presented, we had introduced the standard response metrics we intended to produce for the 2017 EC. (We knew a lot about the 2007 EC, not much about the 2012 EC, and nothing about the 2017 EC.)

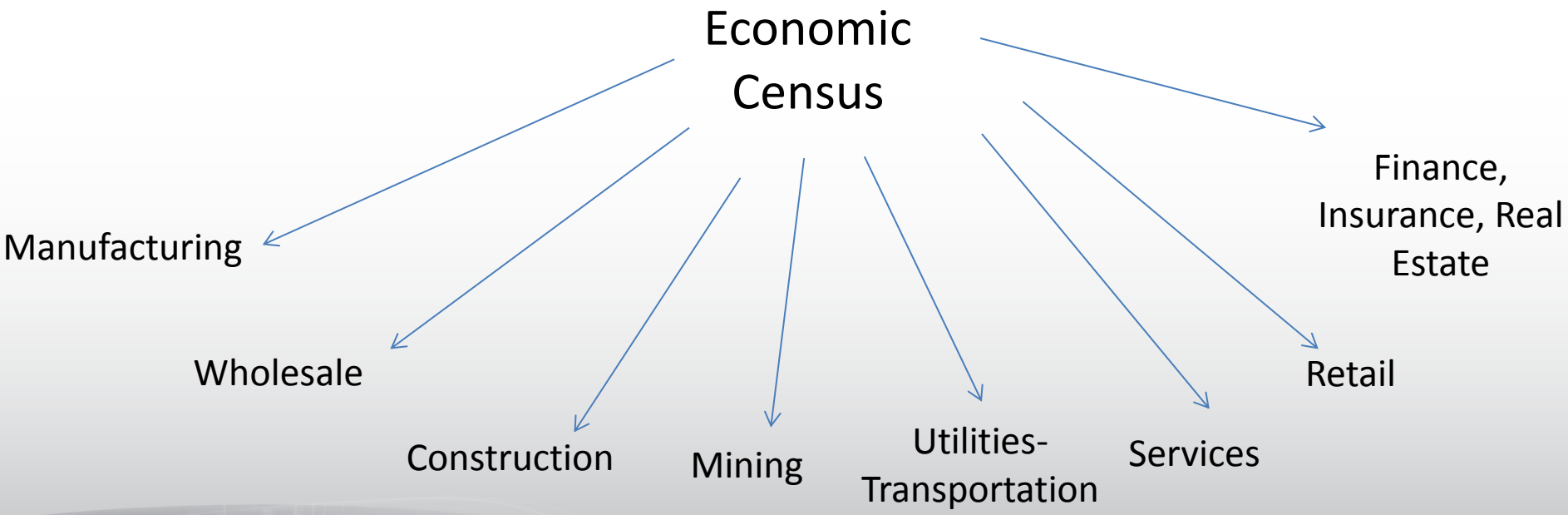
- We also discussed the need to produce metrics in “real time” to assist in targeted nonresponse follow-up.

Economic Census Overview

- Conducted every 5 years for reference years ending in 2 and 7
- Response required by law
- Data collected from business establishments in the 50 states and the District of Columbia
- In 2012, data were collected through self-administered paper and electronic questionnaires. In 2017, data will be collected electronically.
- Provides the most detailed benchmark data available about the U.S. economy with statistics down to the ZIP code level
- Supports the estimation of Gross Domestic Product
- Data users include policymakers, industry analysts, economic development commissions, academic researchers, businesses and other survey programs (e.g., as a benchmark)

Economic Census Overview, contd.

- Key data items (receipts, payroll, no. of employees, no. of establishments)
- Key subgroups (e.g., geography, NAICS, legal form of organization)



Sampling Methodology

- Sampling frame derived from the U.S. Business Register (BR)
- Neither a census, nor a single sample survey; sampling component to some trade areas
- For sampling purposes, distinguish between multi-establishment companies and single-establishment companies
- Some trades fill small, single-unit cases with administrative data in lieu of mailing a form
- Approximately 3.9 million out of 6.5 million establishments received a 2012 Economic Census form

2012 Economic Census Sampling Scheme By Trade Area*

Wholesale	Manufacturing and Mining	Retail, Services, and Utilities-Transportation	Finance, Insurance, Real Estate	Construction
Take all units with certainty	Take all units with certainty. Based on payroll cutoffs, unit receives a long form, short form, or no form. Impute units that receive no form with administrative data	Take those with the largest payroll with certainty Sort remaining cases on 8-digit NAICS code and payroll and take a systematic sample. Fill key items** of non-selected units with administrative data. Adjust other items using an expansion method	In select industries, take all units with certainty In all other industries: •Take those with largest payroll with certainty Sort remaining cases on 8-digit NAICS code and payroll and take a systematic sample. Fill key items** of non-selected units with administrative data. Adjust other items using an expansion method	Take all multi-unit businesses with certainty For single-unit businesses with complete 6-digit NAICS, stratify by state and NAICS. Take small strata with certainty. Within remaining strata, take a probability-proportional-to-size (payroll) sample For single-unit businesses with incomplete 6-digit NAICS, take those with largest payroll with certainty. Take a simple random sample of remaining cases

* MU's are taken with certainty across all trade areas

**Key items include receipts, payroll, and number of employees

Nonresponse Methodology

- Nonresponse follow-up
  - As many as four mailings, plus phone calls for large establishments.
- Secondary data source/administrative data substitution
  - Where possible, missing data were imputed using administrative data.
- For some trade areas, a factor adjustment was used to adjust for missing product lines data

Response Metrics and the Economic Census

- Historically, Economic Census has used check-in rate as its only metric for response. We have been developing response rates that meet quality standards or can be used as additional "real-time" performance metric.
  - Standard measures – to be computed at the end of collection.
  - Proxy measures – to be computed during data collection.

- As the Economic Census covers eight major trade areas, with different designs, and over 500 different forms, this has been a several year process. (e.g., database updates to accommodate response rates)

- We are currently testing these metrics using 2012 Economic Census data. They will go into production for the 2017 Economic Census.

Standard Response Metrics

- Check-in rate
  - The check-in rate represents the percentage of mailed forms that were returned and undeliverable as addressed.
- Unit response rate
  - The *URR* represents the percentage of reporting units eligible for data collection or of unknown eligibility that provided a valid survey response.
- Total quantity response rate
  - For a data item *t*, the *TQRR* is the percentage of the estimated total obtained from directly reported and secondary source data.

Standard Response Metrics (contd.)

- Quantity response rate
  - For data item *t*, the *QRR* is the percentage of the estimated total obtained from directly reported data.
- Administrative data rate
  - For data item *t*, the *ADR* is the percentage of the estimated total obtained from administrative data.
- Imputation rate
  - For a data item *t*, the *IR* is the percentage of the estimated total that is not obtained from directly reported or secondary source data.

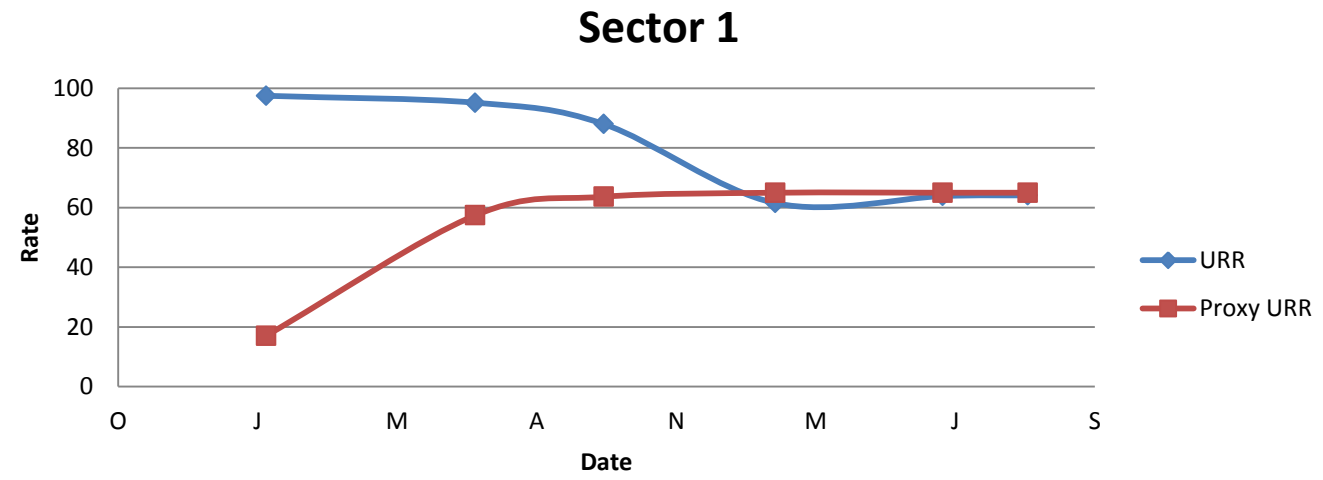
Plans for Economic Census "Real-Time" Response Metrics

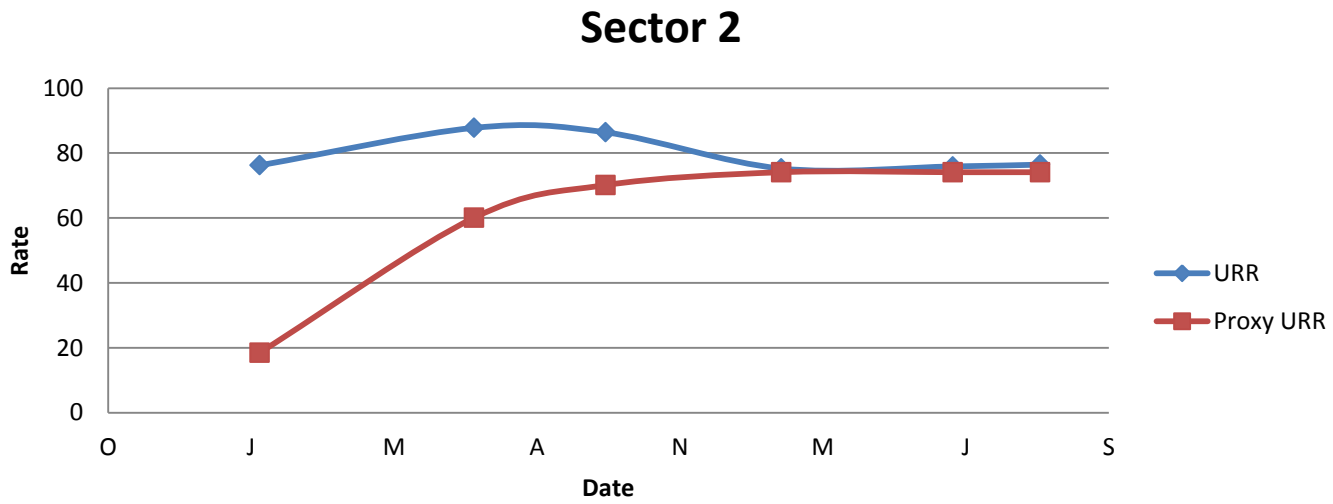
- Will allow survey managers to monitor data collection performance in real-time.
- The proxy URR differs from the standard URR in that we make some simplifying assumptions about eligibility and response (use mail and check-in status as proxies).
- Proxy item-level response rates differ from their standard counterparts in that they use administrative data (use mail status, check-in status and administrative payroll, which is highly correlated with all the key items of primary interest).

Real-Time Rates

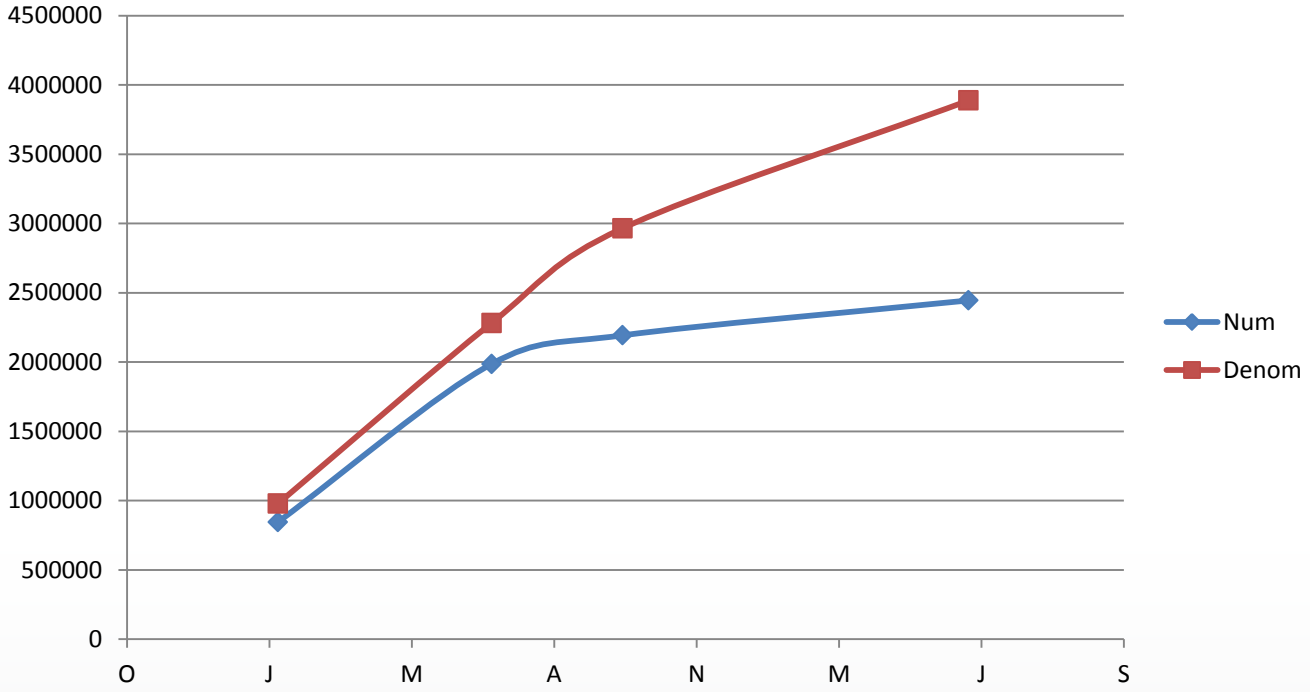
- Check-in rate
- Proxy URR
- Proxy TQRR (payroll)
- Proxy QRR (payroll)
- Proxy Administrative Date Rate (payroll)
- Proxy Imputation Rate (payroll)

Why We Need Proxy Response Rates



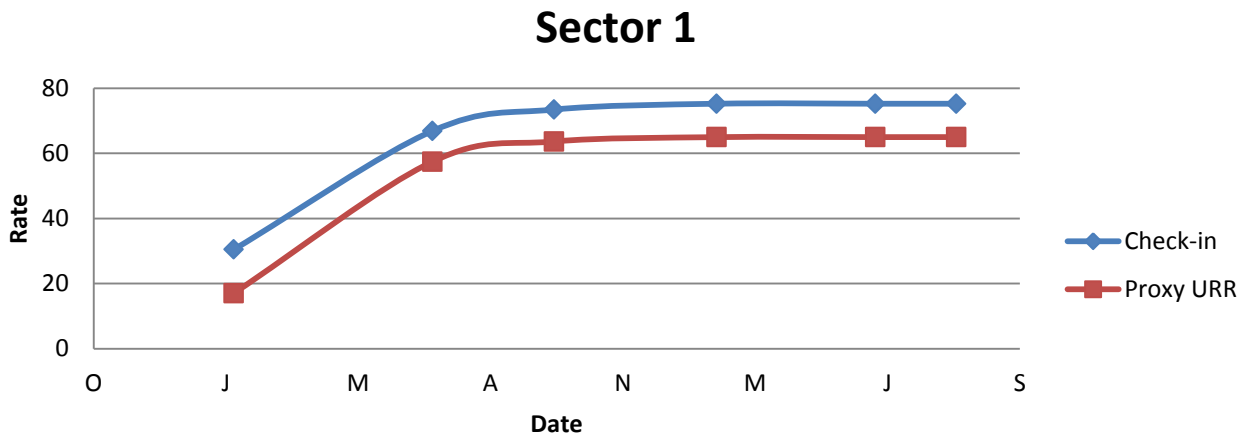


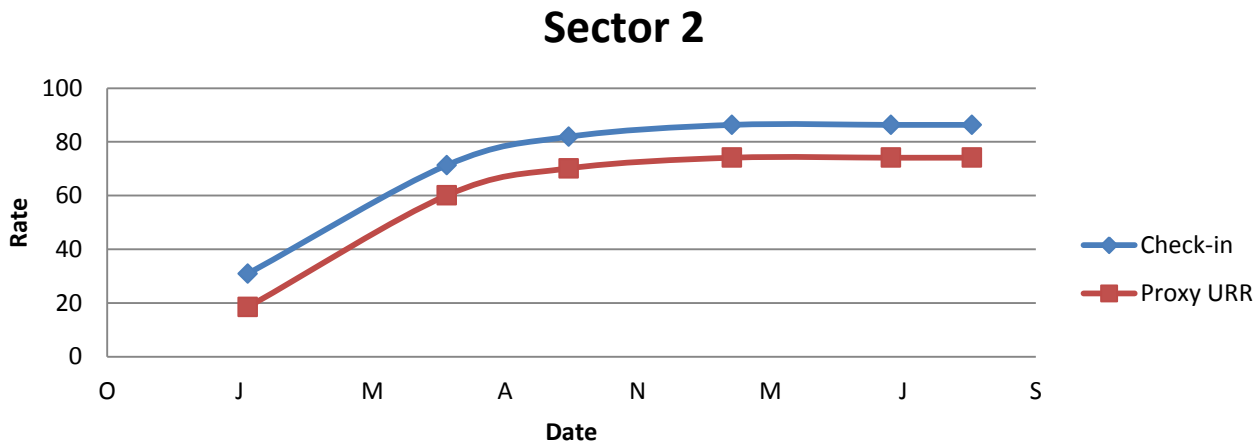
Numerator and Denominator For URR



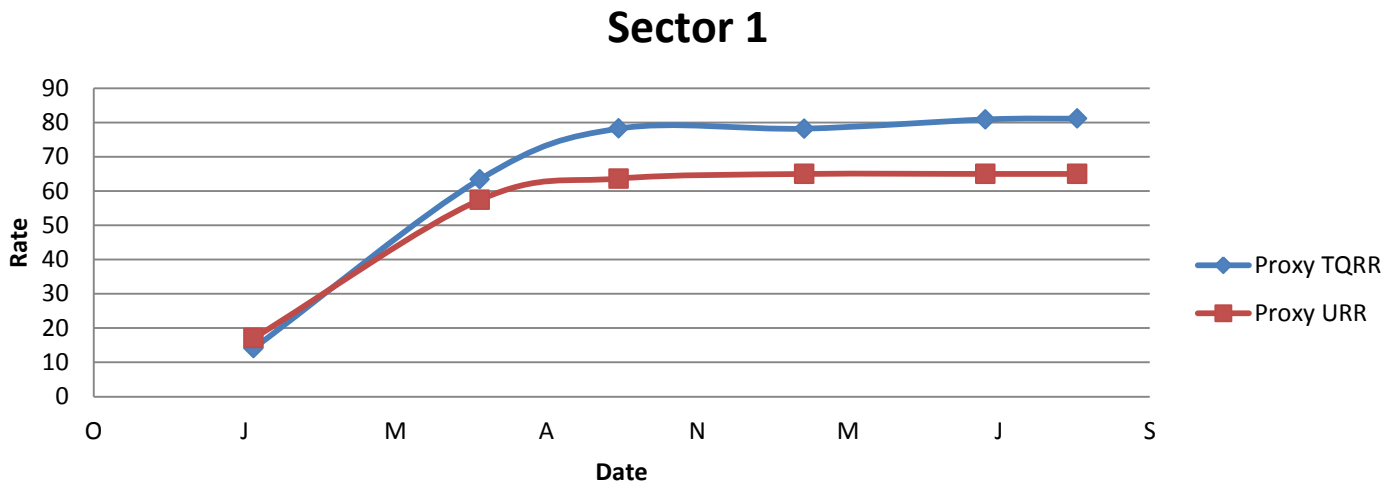
- Initially there were a “small” number of cases defined as eligible, and many of those were flagged as respondents.
- As processing continued, our pool of eligible cases expanded, but not as many eligible units actually reported.

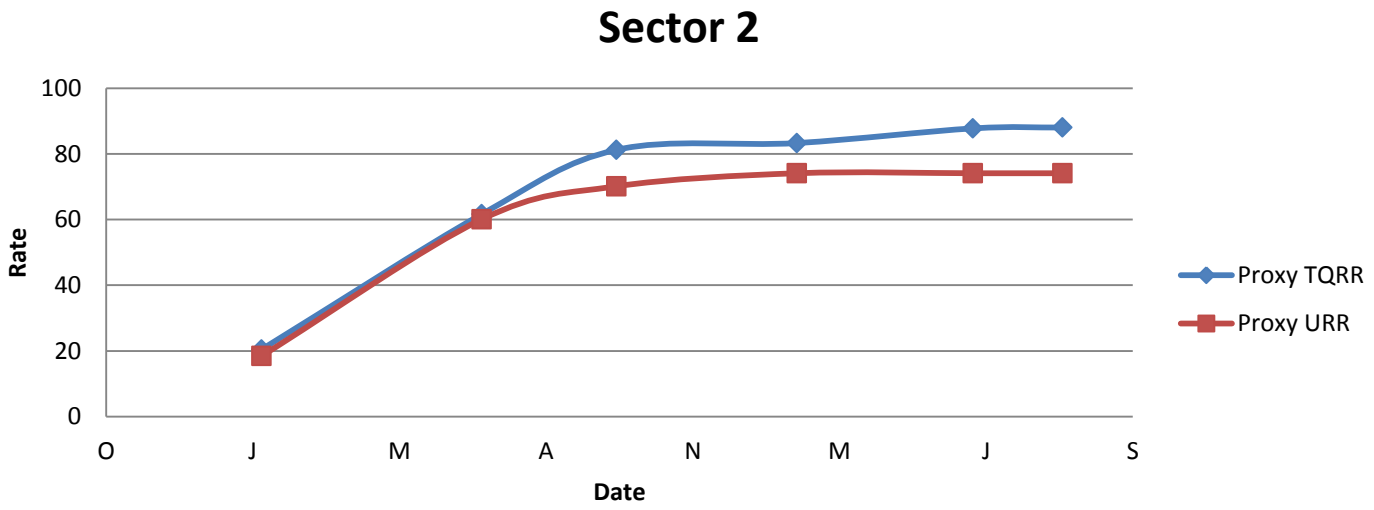
Example: Check-in Rate vs. Proxy URR





Example: Proxy URR vs. Proxy TQRR





Real-Time Response Metrics Table

		Response Metrics																	
		By Sector by Unit Type																	
		Check-in Rate			Proxy Unit Response Rate			Proxy Total Quantity Response Rate*			Proxy Quantity Response Rate			Proxy Administrative Data Rate			Proxy Imputation Rate		
		Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)
Sector 21	Overall																		
	Multi-unit																		
	Single-unit																		
22	Overall																		
	Multi-unit																		
	Single-unit																		
23	Overall																		
	Multi-unit																		
	Single-unit																		
31-33	Overall																		
	Multi-unit																		
	Single-unit																		

*Proxy rates based on 20xx Business Register payroll data

Disclosure Prohibited, 13 & 26 U.S.C.

Post-Processing Item Response Metrics

		Item Response Metrics for Payroll												
		By Sector by Unit Type												
			Total Quantity Response Rate			Quantity Response Rate			Administrative Data Rate			Imputation Rate		
		N	Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)	Num	Den	Rate (%)
Sector 21	Overall													
	Multi-unit													
	Single-unit													
22	Overall													
	Multi-unit													
	Single-unit													
23	Overall													
	Multi-unit													
	Single-unit													
31-33	Overall													
	Multi-unit													
	Single-unit													

Disclosure Prohibited, 13 & 26 U.S.C.

Future Plans

- In addition to response rate efforts, current 2017 EC plans include:
  - Variance estimation
  - Better use of survey process data (paradata)
  - “Adaptive design” strategies

- To prepare for using these in production in 2017, we hope to test them earlier (hopefully in the 2016 ASM because of the similarities).

- In the future, we will investigate the use of additional quality metrics.

Contact Information

- Fane Lineback
  joanna.fane.lineback@census.gov

- Eric B. Fink
  eric.fink@census.gov